Science
AAAS

Supplementary Materials for

Structure and mechanism of the plant RNA polymerase V

Guohui Xie *et al*.

Corresponding author: Jiamu Du, dujm@sustech.edu.cn

Science **379**, 1209 (2023)
DOI: 10.1126/science.adf8231

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Materials and Methods

Antibody, peptide, and oligos

A peptide from cauliflower (*Brassica oleracea* var. *botrytis*) NRPE1 (KKNPETELNAAAWG) (**Fig. 1A**), whose homology region in Arabidopsis NRPE1 (DKKNSETESDAAAWG) was shown to be able to raise anti-sera with sufficient quality for Pol V purification (*7, 36*), was selected as the antigen to immune mouse by the ABclonal Company (Wuhan, China). A monoclonal antibody strain 5D2D8 that showed the best behavior was chosen for further large-scale production by ABclonal Company (Wuhan, China). The protein A purified 5D2D8 antibody obtained from the company was directly used in the following purification outlined below. The 5D2D8 epitope peptide was purchased from the GL Biochem Company (Shanghai, China). All the oligos were ordered from the Sangon Company (Shanghai, China).

Purification of cauliflower nucleus

The inflorescence tissue was cut from 10 kg of fresh cauliflower and subsequently frozen by liquid nitrogen. The frozen inflorescence was homogenized by a blender. About 16 L of the Extraction Buffer (10 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 5 mM EDTA, 250 mM sucrose, 0.5% Triton-X-100, 0.2 mM PMSF, and 10 mM HEPES, pH 7.8) was added into the homogenized inflorescence and the mixture was stirred gently. The resulting curd was further filtered through two layers of Miracloth (Millipore). The cauliflower nucleus was collected by centrifugation at 3,000 g for 20 min at 4 °C and further washed by Extraction Buffer twice. About 60 g crude nucleus can be obtained from 10 kg cauliflower.

Extraction of cauliflower nucleus

About 60 g crude cauliflower nucleus was suspended in 60 ml of Low Salt Buffer (20 mM KCl, 2 mM $MgCl_2$, 0.5 mM DTT, 0.2 mM EDTA, 25% Glycerol, 0.2 mM PMSF, and 20 mM HEPES, pH 7.8) to disrupt the nucleus by osmotic shock. Then 90 ml High Salt Buffer (1.6 M KCl, 2 mM $MgCl_2$, 0.5 mM DTT, 0.2 mM EDTA, 25% Glycerol, 0.2 mM PMSF, and 20 mM HEPES, pH 7.8) and 2.5 kU SuperNuclease (SinoBiological, SSNP01) were added and the solution was gently shaken for 10 min to release the nucleic acids. The nucleic acid was further removed by sonication for 45 min. The lysate was centrifuged at 17,000 g for 20 min at 4 °C. The supernatant was filtered by 4-layer Miracloth (Millipore) and dialyzed against an RB400 buffer (400 mM NaCl, 5 mM $MgCl_2$, 2 mM DTT, 10% Glycerol, 0.1% NP40, and 50 mM Tris-HCl, pH 7.5) for 3 h, and filtered by a 0.22 μm filter (Millipore).

Purification of *Bo*Pol V

About 2 ml Protein G agarose resin (Yeasen, 36405ES) was added into the cauliflower nucleus extraction. After incubation for 1 hour, the resin was discarded by low-speed centrifuge to remove non-specific binding proteins. Following, about 200 μg antibody 5D2D8 was added into the cauliflower nucleus extraction for capturing the *Bo*NRPE1. After 3 hours incubation, 2 ml fresh Protein G resin was added and further incubated for 1.5 hours. The Protein G resin, as well as the bounded *Bo*Pol V, were collected by low-speed centrifuge. The resin was further successively washed by 200 ml Washing Buffer (400 mM NaCl, 5 mM $MgCl_2$, 2 mM DTT, 5% Glycerol, 0.2% NP40, and 20 mM Tris-HCl, pH 7.5) and 100 ml Pre-Elution Buffer (400 mM NaCl, 5 mM $MgCl_2$, 2 mM DTT, 1% Glycerol, 0.05% NP40, and 20 mM Tris-HCl, pH 7.5). The *Bo*Pol V was eluted by 5 ml 5D2D8 epitope peptide (KKNPETELNAAAWG) with a final concentration of 0.5 mg/ml. The affinity purified *Bo*Pol V was further purified using a Superose 6 Increase 3.2/300 column

(Cytiva) with an RB150 Buffer (150 mM KCl, 5 mM $MgCl_2$, 2 mM DTT, 1% Glycerol, 0.05% NP40, and 20 mM HEPES, pH 7.8). About 3 μg *Bo*Pol V can be obtained from 10 kg cauliflower.

Purification of *Bo*Pol II

The mouse monoclonal antibody strain 8WG16 against NRPB1 CTD (Santa Cruz Biotechnology, sc-56767) was used for *Bo*Pol II purification. The overall purification steps were similar as the *Bo*Pol V purification. The only difference is the EG Buffer (0.1 mM EDTA, 0.75 M ammonium sulfate, 40% ethylene glycol, and 50 mM Tris-HCl, pH 7.5) was used as elution buffer (*46*). The purified *Bo*Pol II was buffer-exchanged into the RB150 buffer for further use. Both the purified *Bo*Pols II and V were subjected to mass spectrometry analysis using an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific) with a nanospray ion source and an EASY-nLC 1000 system (Thermo Fisher Scientific) at SUSTech Core Research Facility (SUSTech CRF). The presence of all 12 subunits Pols II and V were confirmed.

In vitro transcription assays

Equal amounts of 5'-FAM labeled RNA primer and the corresponding DNA strand(s) (**table S1**) were dissolved in a buffer of 50 mM NaCl, 20 mM Tris-HCl, pH 8.0 with a final concentration of 40 μM for each oligo. To anneal the oligos, the oligo mixtures were heated to 95 °C for 5 min and then slowly cooled down to room temperature. All the annealed templates were diluted to 10 μM and stored at -80 °C for further use. The *in vitro* transcription assay was carried out in a 20 μL reaction system with a buffer of 50 mM NaCl, 5 mM $MgCl_2$, 2 mM DTT, 10 % glycerol, and 40 mM Tris-Cl, pH 8.0. In each reaction, the purified Pols and various substrates were pre-incubated at 25 °C for 25 min with a final concentration of 120 nM and 200 nM, respectively. After pre-incubation, 0.5 mM NTP (ATP, UTP, CTP, and GTP) was added to allow the reaction lasting for 2 hours. The reactions were quenched by adding equal volume RNA loading buffer containing 98% formamide, 1 mM EDTA, bromophenol blue, and xylene cyanole by heating to 95 °C for 5 min followed by cooling down to 0 °C on ice immediately. To monitor the reactions, 20 μL of each sample was loaded onto a 20% denatured PAGE gel with 8 M urea and 1 × Tris-borate-EDTA buffer. Then, the gel was run with 0.5 × TBE buffer at 7 W for 3 hours and imaged under blue fluorescence mode of a Tanon 6100C Imager. To compare the activity between Pols V and II, the reaction system was prepared according to the experiment design.

Cryo-EM specimen preparation

To obtain the *Bo*Pol V elongation conformation complex, the peptide eluted *Bo*Pol V (about 3 μg) was concentrated to 80 μL and then incubated with the designed transcription bubble (**Fig. 1D** and **table S1**) with a molar ratio of 1:2 at 25 °C in a reaction buffer of 150 mM KCl, 5 mM $MgCl_2$, 2 mM DTT, 10% Glycerol, 0.5 mM ATP, 0.5 mM UMPPNP, and 20 mM HEPES, pH 7.8. After one-hour incubation, the free nucleic acids were removed by a Superose 6 Increase 3.2/300 column (Cytiva) with the RB150 Buffer. The Pol V-transcription bubble complex was further concentrated to about 0.1 mg/ml for specimen preparation. For apo Pol V sample, the concentrated fractions collected from gel filtration were directly used. All grid preparations were conducted immediately after purification. About 4 μL Pol V or Pol V- transcription bubble complex was applied to a 300-mesh Cu R1.2/1.3 grid with 2 nm carbon (Quantifoil, Micro Tools GmbH), which was pre-glow discharged by a plasma cleaner (PDC-32G, Harrick Plasma). Then, the grids were blotted by a Vitrobot instrument (Thermo Fisher Scientific) under the condition of 2.5 s blotting time and 5 s

wait time under the condition of 6 °C temperature and 95 % humidity, and subsequently plunge-frozen in liquid ethane which was already pre-cooled by liquid nitrogen.

EM data collection and processing

All the cryo-EM data were collected on 300 kV Titan Krios electron microscopes (Thermo Fisher Scientific) equipped with K3 Summit direct electron detectors (Gatan), and the movies were automatically collected by EPU (Thermo Fisher Scientific) under super-resolution counting mode. Electron micrograph datasets were acquired at a magnification of 81,000 × (1.095 Å /physical pixel) at 20 $e^-/Å^2$ per second for a total dose of 50 $e^-/Å^2$ which fractionated into 32 frames, and the defocus values of data collection were ranged from -1.0 μm to -2.5 μm.

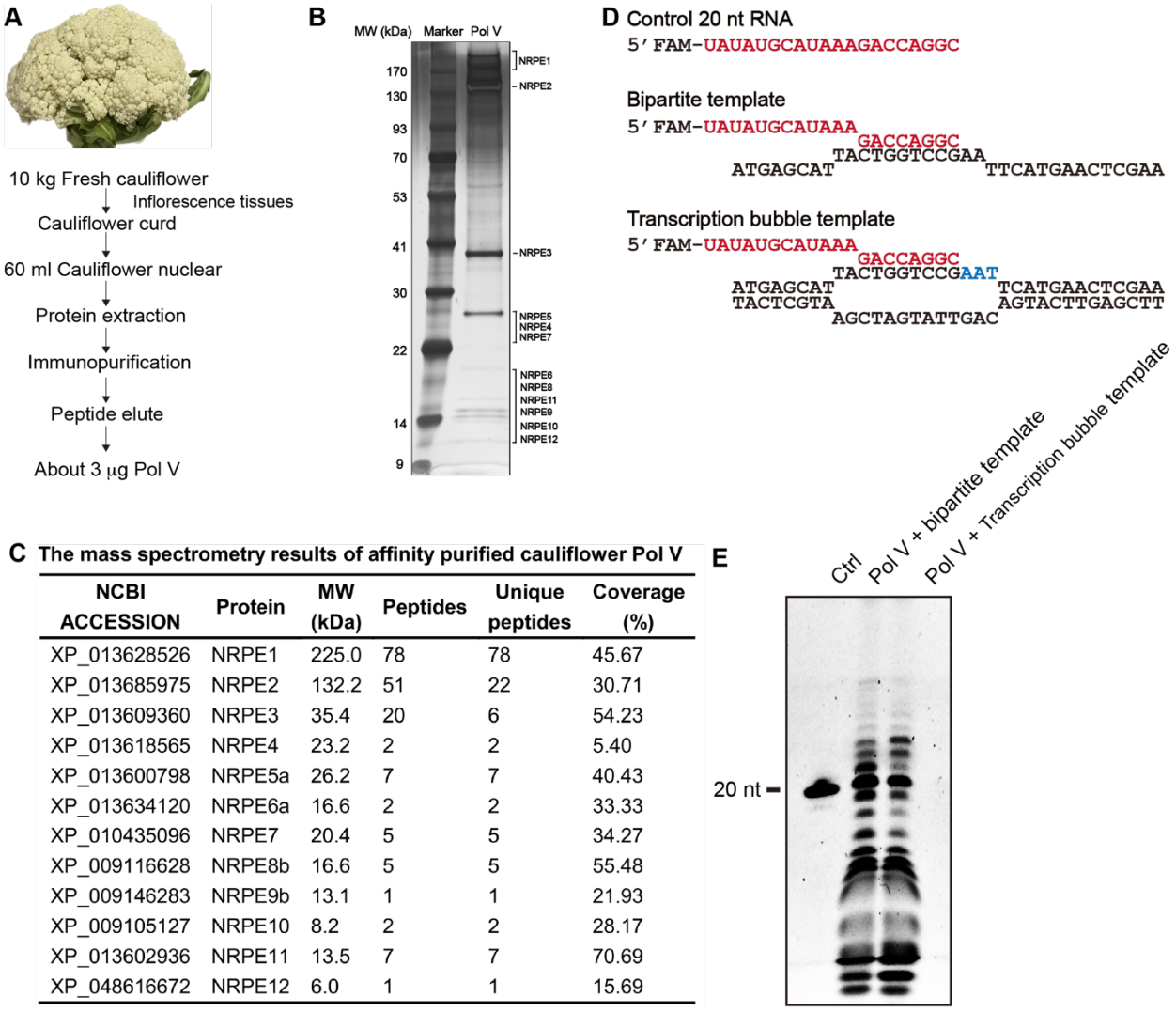
The cryo-EM data of apo-Pol V was processed using the program RELION3.1 (*47*). The movies were firstly aligned using MotionCor2 in Relion-3.1 (*48*), and the binning factor was set to 2 to recover the pixel size back to 1.095 Å. The defocus estimation and contrast transfer function (CTF) were estimated by CTFfind v4.1 (*49*). A total of 1,609,875 particles were picked from 6,805 micrographs. To accelerate the calculation, the pixel size of extracted particles was expanded to 2.19 Å with the particle box size of 150 × 150 pixels accordingly. After 3 rounds of 2D classification, 1,141,590 particles were selected and further classified by one more round of 3D classification with low quality particles excluded. The remaining 794,368 selected particles were re-centered and re-extracted at a pixel size of 1.095 Å with the box size was expanded to 300 × 300 pixels at the same time. A consensus 3D auto-refine job yielded an initial resolution of 4.5 Å. Next, 125,972 particles were selected by three rounds of 3D classification incorporation with 3D auto-refine, yielding a refined resolution of 4.2 Å. Subsequently 2 more rounds alignment-free 3D classification were conducted to further clean the particles, finally 63,603 particles were selected and auto-refined to 3.65 Å resolution. Further 3 rounds of CTF refinement in addition with 3D auto-refinement were applied to push the final refined resolution to 3.57 Å (*50*).

The cryo-EM data processing of Pol V EC was performed using RELION 4.0 (*51*). Total 5,694 movie frames were used for the data processing. Motion correction and CTF estimation were carried out using MotionCor2 and CTFFIND, respectively (*48*). Particle picking was done using Topaz implemented within RELION 4.0 (*52*). The particles with figure-of-merit (FOM) cut-off above -3 were selected and extracted with a binning factor of 8. A total of 2,750,896 particles were extracted from the two datasets and then subjected to a round of reference-free 2D classification. 2D classes that showed polymerase features were selected and subjected to a round of 3D classification using the Pol V apo structure map as reference. 3D classification was done using an angular sampling of 15° for the first 25 iterations, followed by additional 25 iterations with an angular sampling of 7.5°. Particles in selected 3D classes went through another round of reference-free 2D classification to remove sub-optimal particles. Good 2D classes were selected and combined. Particles from these classes were then re-extracted using 300 × 300 pixel box size for 3D consensus refinement which yielded a 3.04 Å resolution map. Subsequent refinements of beam tilt, anisotropic magnification, and defocus improved the resolution to 2.8 Å. One more step of 3D classification without alignment was carried out with regularization parameter *T* of 10 and a global mask covering Pol V EC. One 3D class was selected and refined, yielding a final map of 2.73 Å resolution.

Model building, structure refinement, and analysis

The manual model building was directly carried out using the program Coot (*53*) guided by homology model of Arabidopsis Pol IV (PDB code: 7EU1) (*13*). The structure refinement was

conducted using the program Phenix by global minimization with Ramachandran restraints (*54*). The 3D FSC correlation of cryo-EM maps were analyzed using the 3DFSC Processing Server (https://3dfsc.salk.edu/) (*55*). A list of the statistics of the cryo-EM data collection, structure refinement, and validation was given in **table S2**. The graphics were produced using the program PyMOL (Schrödinger) and Chimera (*56*). The sequences were aligned using the T-Coffee server and further illustrated using the ESPript server (*57, 58*). The sequences used in alignments are: *Bo*NRPE1 (NCBI: XP_013628526), *At*NRPE1 (NCBI: NP_181532), *At*NRPB1 (NCBI: NP_001329346), *At*NRPD1 (NCBI: NP_001185298), *Bo*NRPE(D)2 (NCBI: XP_013685975), *At*NRPE(D)2 (NCBI: NP_001189957), *Glycine max* NRPE(D)2 (NCBI: XP_003523670), *Solanum lycopersicum* NRPE(D)2 (NCBI: XP_004236361), *Zea mays* NRPE(D)2 (NCBI: AQK46524), *Oryza sativa* NRPE(D)2 (NCBI: XP_015636222), *Bo*NRPA2 (NCBI: XP_013597647), *Saccharomyces cerevisiae* RPA2 (NCBI: AJW03181), *Bo*NRPB2 (NCBI: XP_013598194), *Saccharomyces cerevisiae* RPB2 (NCBI: EWG93581), *Bo*NRPC2 (NCBI: XP_013598951), *Saccharomyces cerevisiae* RPC2 (NCBI: AJU02406).



C The mass spectrometry results of affinity purified cauliflower Pol V

NCBI ACCESSION	Protein	MW (kDa)	Peptides	Unique peptides	Coverage (%)
XP_013628526	NRPE1	225.0	78	78	45.67
XP_013685975	NRPE2	132.2	51	22	30.71
XP_013609360	NRPE3	35.4	20	6	54.23
XP_013618565	NRPE4	23.2	2	2	5.40
XP_013600798	NRPE5a	26.2	7	7	40.43
XP_013634120	NRPE6a	16.6	2	2	33.33
XP_010435096	NRPE7	20.4	5	5	34.27
XP_009116628	NRPE8b	16.6	5	5	55.48
XP_009146283	NRPE9b	13.1	1	1	21.93
XP_009105127	NRPE10	8.2	2	2	28.17
XP_013602936	NRPE11	13.5	7	7	70.69
XP_048616672	NRPE12	6.0	1	1	15.69

Fig. S1. The purification and characterization of *Bo*Pol V. (**A**) The purification steps of *Bo*Pol V from cauliflower. Typically, 3 μg purified Pol V can be obtained from 10 kg fresh cauliflower. (**B**) The SDS-PAGE of purified *Bo*Pol V stained by silver staining. (**C**) The mass spectrum confirming the existence of all the 12 subunits of the *Bo*Pol V. (**D**) The substrate design for the monitoring the transcription activity *Bo*Pol V. (**E**) The *in vitro* transcription activity of *Bo*Pol V with two different designed substrates. A 20 nt RNA was used as a control (ctrl).

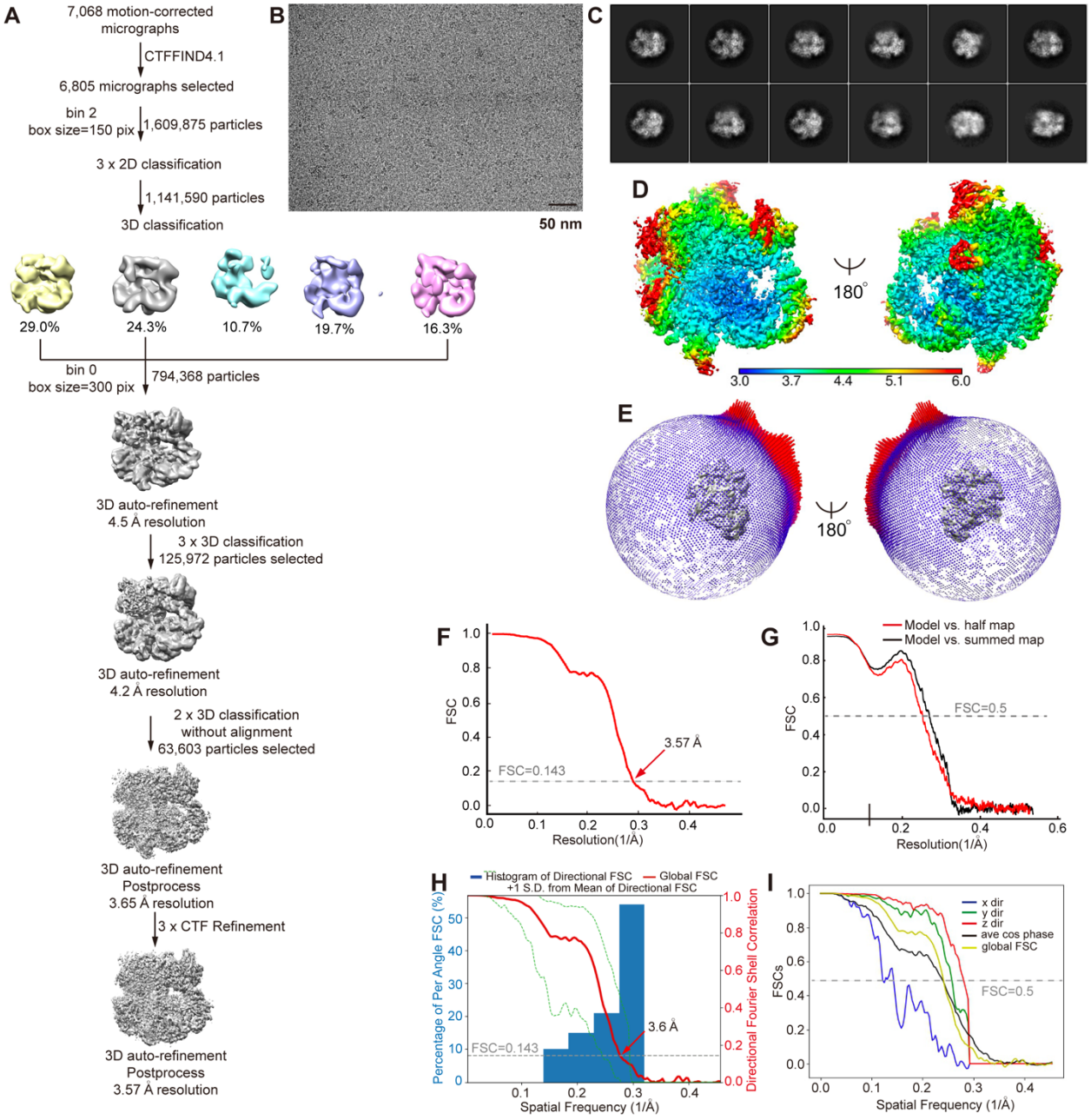


Fig. S2. Cryo-EM structural analysis of *Bo*Pol V in apo form. (**A**) The flowchart for the data processing. (**B**) A representative image of the cryo-EM micrograph. (**C**) 2D class averages. (**D-E**) Local resolution map (D) and angular distribution of particles (E) of the final 3D reconstruction in two orientations related by 180°. (**F**) The FSC curves of the final 3D reconstruction. (**G**) The FSC calculated between the refined structure and the half map used for refinement and the full map. (**H-I**) The global (H) and directional (I) FSC of the 3D auto-refinement by the 3DFSC Processing Server.

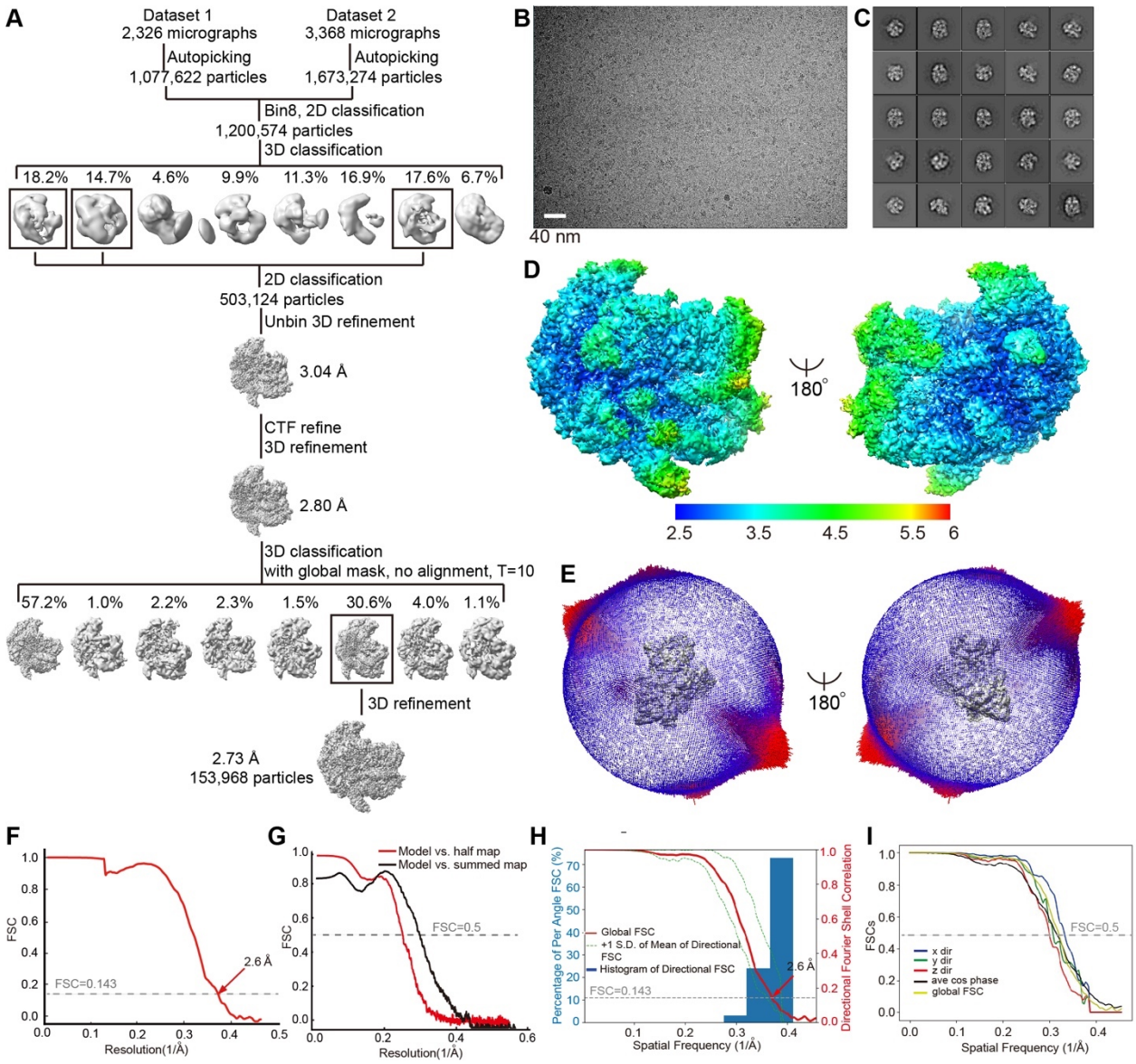


Fig. S3. Cryo-EM structural analysis of *Bo*Pol V in elongation conformation. (**A**) The flowchart of the cryo-EM data processing. (**B**) A representative cryo-EM image. (**C**) 2-D classification average. (**D**) Local resolution map in two orientations related by 180°. (**E**) Angular distribution of particles in two orientations related by 180°. (**F**) The FSC curves of the final 3D reconstruction. (**G**) The FSC calculated between the refined structure and the half map used for refinement and the full map. (**H-I**) The global (H) and directional (I) FSC of the 3D auto-refinement by the 3DFSC Processing Server.

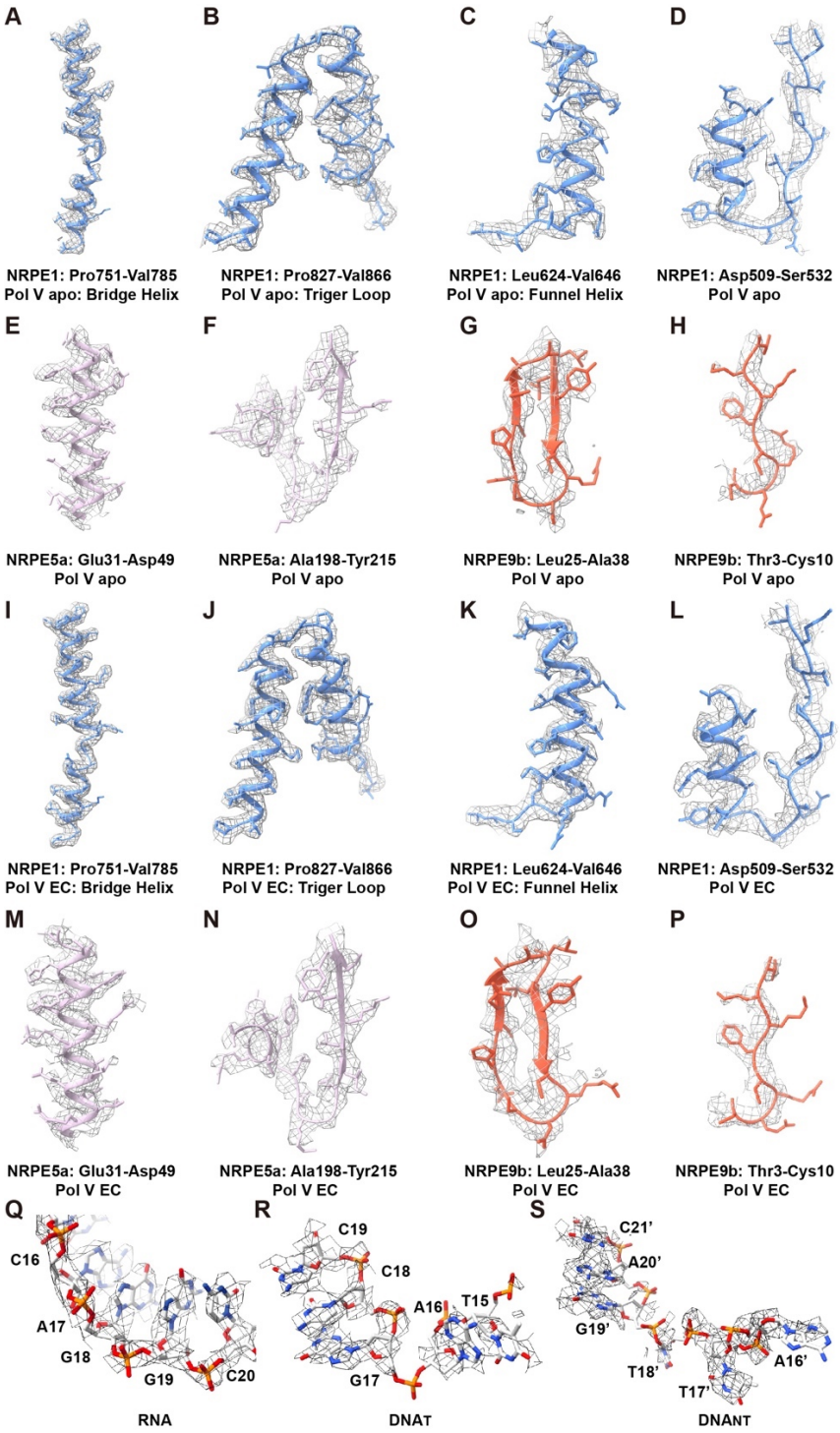


Fig. S4. Cryo-EM maps for representative regions of *Bo*Pol V. (**A-S**) Electron density maps showed the fitting of representative regions of the apo *Bo*Pol V structure (A-H), the *Bo*Pol V elongation complex (Pol V EC) structure (I-P), the nascent RNA (Q), the template DNA (DNA_T, R), and the non-template DNA (DNA_{NT}, S).

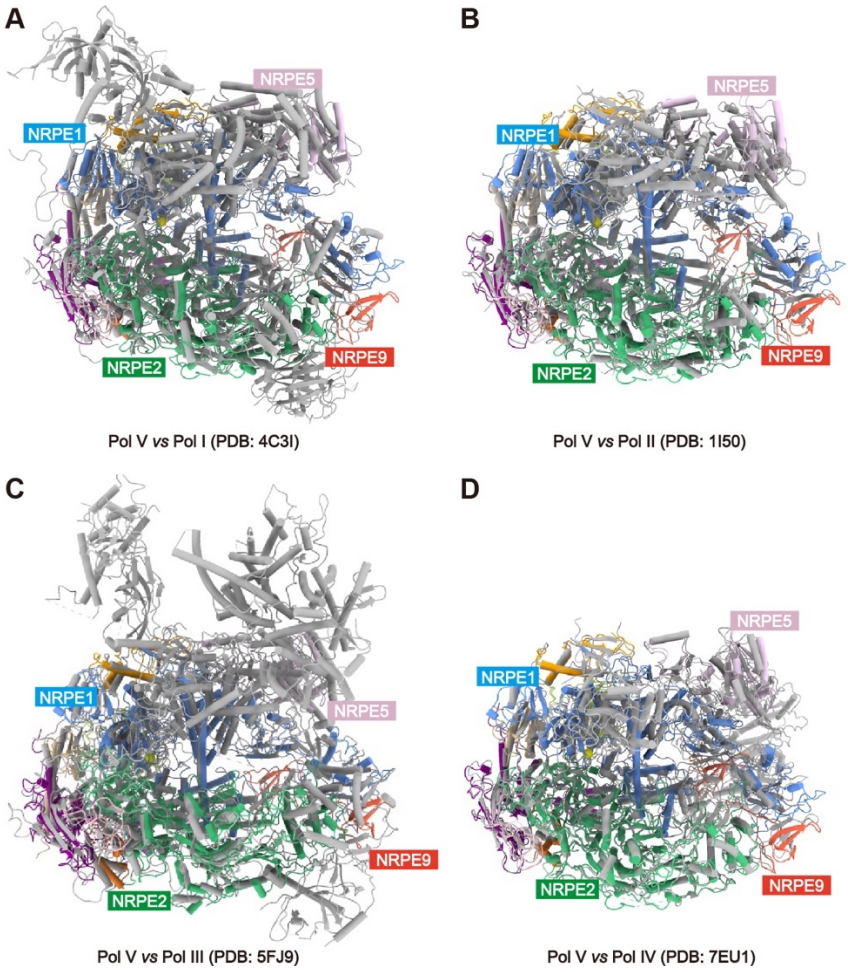


Fig. S5. The structural comparison of Pol V and other Pols. (**A-D**). The superimposition of our Pol V structure with Pol I (A), Pol II (B), Pol III (C), and Pol IV (D) with the Pol V in color and other Pols in silver.

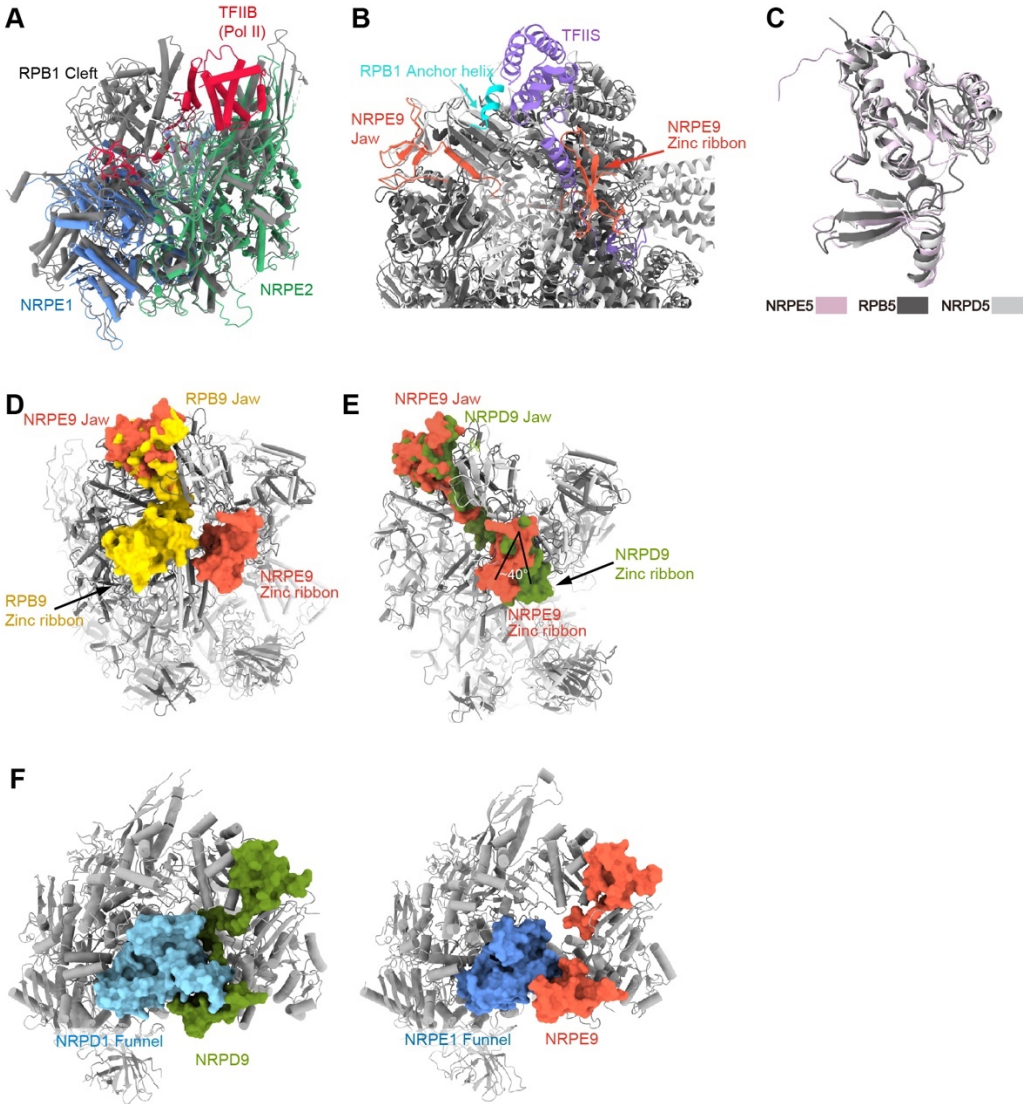


Fig. S6. Structure features of *Bo*PolV. (**A-B**) The structural superimposition of Pol V with the Pol II-TFIIB complex (A, PDB code: 4BBR) and Pol II-TFIIS (B, PDB code: 3PO3) complexes revealed steric conflicts, suggesting the Pol V lacks the binding capacity for the Pol II transcription factors. (**C**) The superimposition of the Pol V NRPE5, Pol II RPB5 (PDB code: 1I50), and Pol IV NRPD5 (PDB code: 7EU1) suggests a very similar conformation, although they are encoded by different genes. (**D-E**) Superimposition of Pol V with Pol II and Pol IV suggested conformation differences of Pol V NRPE9 with Pol II RPB9 (D, PDB code: 1I50) and Pol IV NRPD9 (E, PDB code: 7EU1). (**F**) The Funnel domains of the Pol IV NRPD1 (left panel, PDB code: 7EU1) and Pol V NRPE1 (right panel) induced different conformations of their binding partner NRPD9 and NRPE9, respectively.

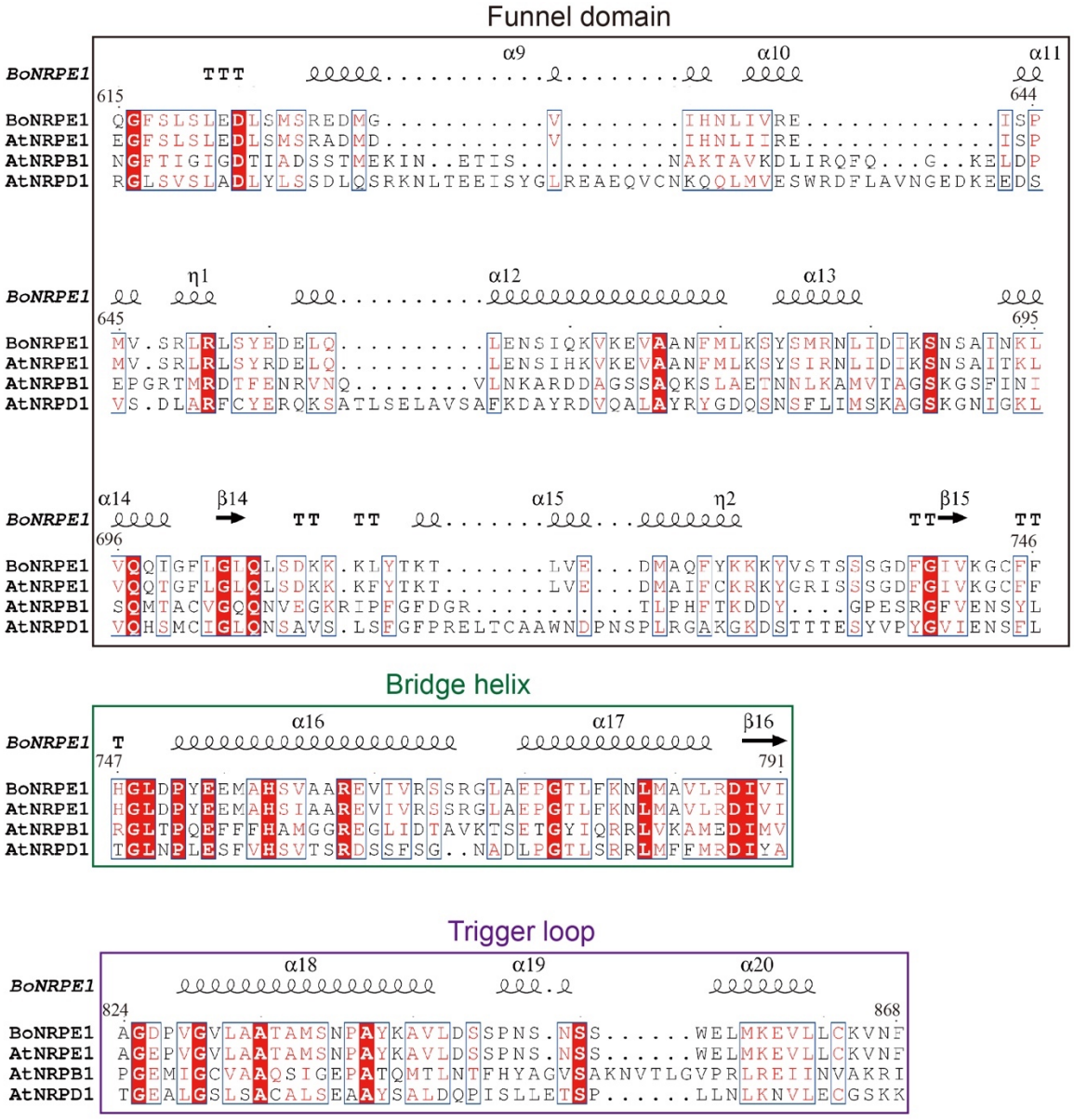


Fig. S7. A structure-based sequence alignment of the Funnel domain (top panel), the Bridge Helix (middle panel), and the Trigger Loop (lower panel) of the *Bo*Pol V NRPE1, *At*Pol V NRPE1, *At*Pol II NRPB1, and *At*Pol IV NRPD1 subunits.

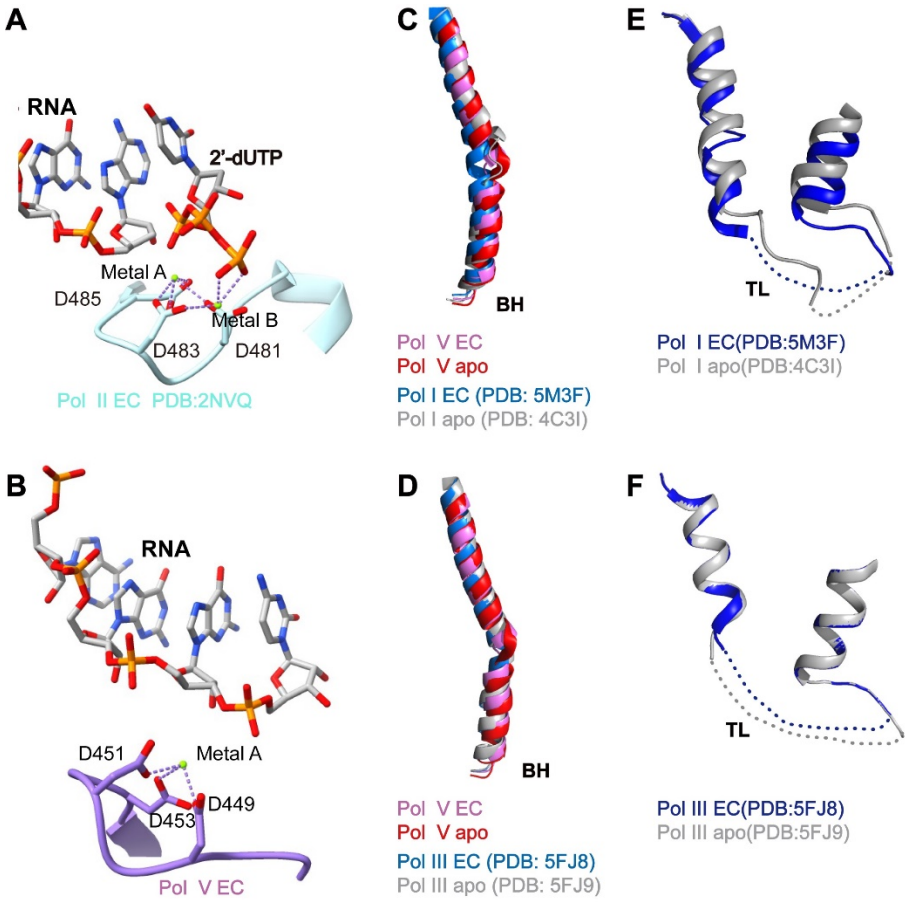


Fig. S8. The active center feature of *Bo*Pol V. (**A-B**) The active center metal binding sites of Pol II (A) and *Bo*Pol V (B). (**C-D**) The superimposition of the BH of Pol V with Pol I (C) or Pol III (D) in apo and elongation forms showing very similar conformations. (**E-F**) The superimposition of TL of Pol I (E) or Pol III (F) in apo and elongation forms showing that the TL of Pols I and III are very flexible.

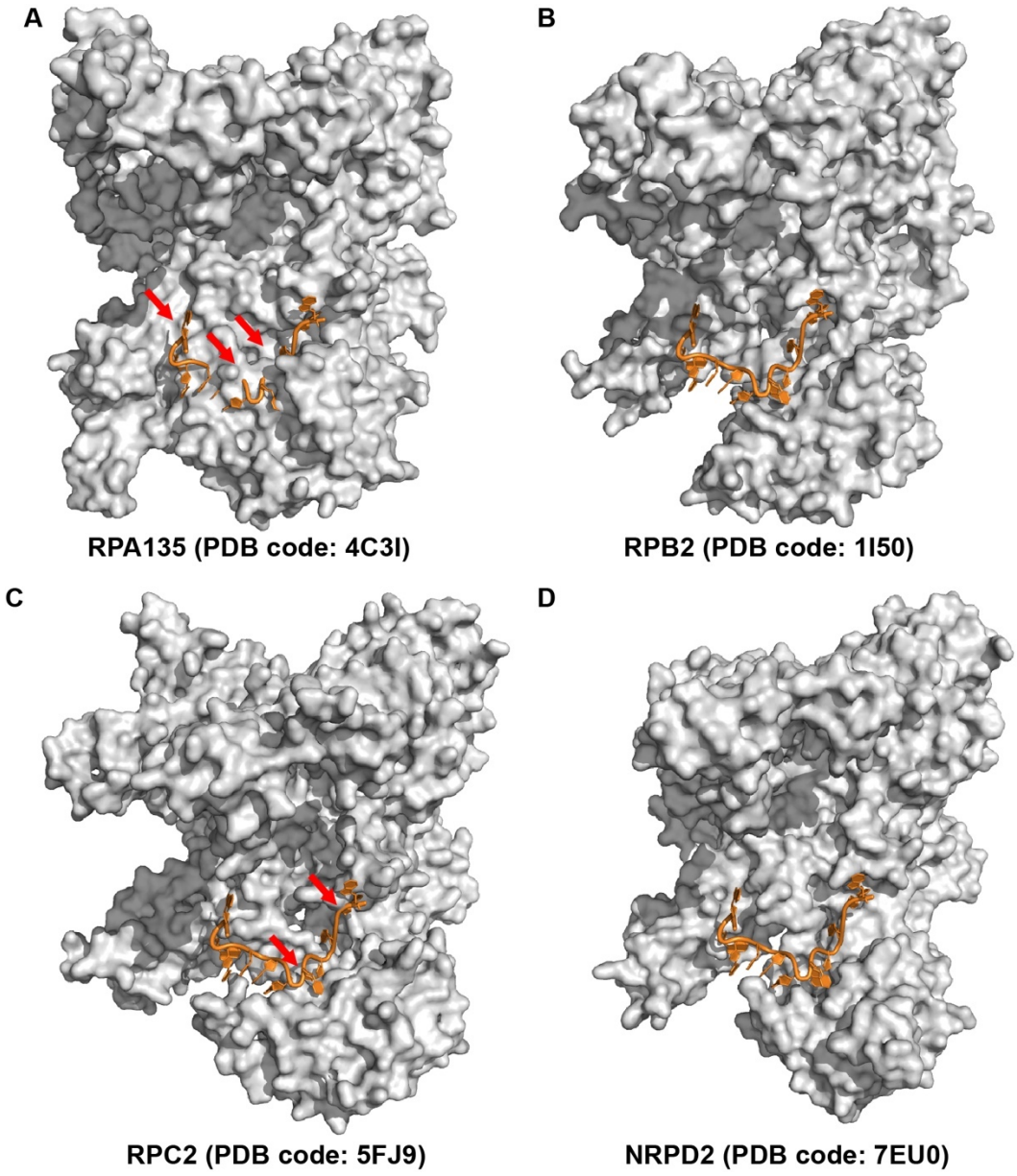
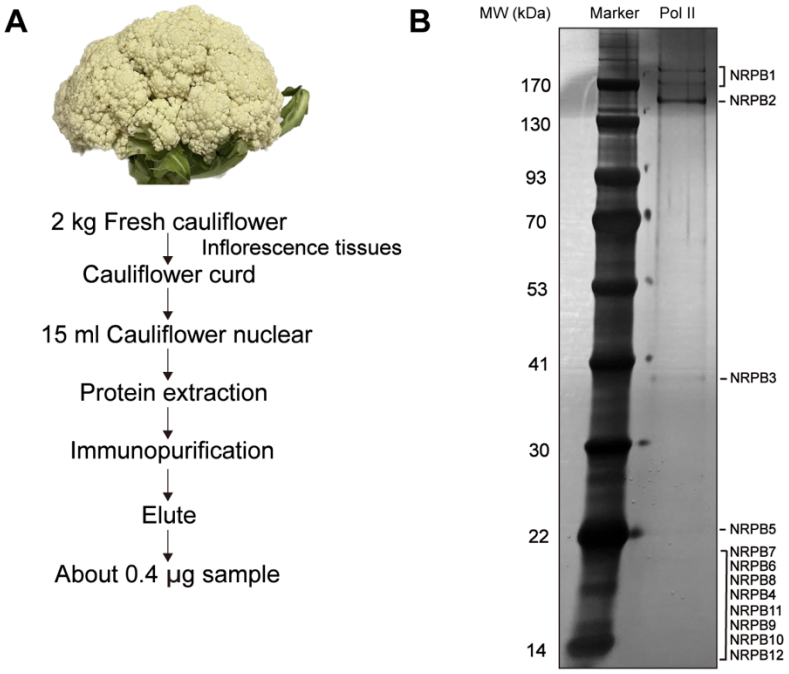


Fig. S9. The modeling of the unpaired region of the non-template DNA (DNA_{NT}-UPR) from our Pol V EC complex to Pols I-IV. (**A-C**) The superimposition of DNA_{NT}-UPR from our Pol V EC complex onto the structure of Pol I subunit RPA2 (A), Pol II subunit RPB2 (B), and Pol III subunit RPC2 (C) yielded steric conflicts as highlighted by red arrows (A, C) or less interaction (B). (**D**) The superimposition of DNA_{NT}-UPR from our Pol V EC complex onto the structure Pol IV NRPD2 subunit showed no conflict and good fitting.



C

The mass spectrometry results of affinity purified cauliflower Pol II

ACCESSION	Protein	MW (kDa)	Peptides	Unique peptides	Coverage (%)
XP_013627895	NRPB1	204.5	84	26	34.10
XP_013598194	NRPB2	134.9	49	9	29.37
XP_013609360	NRPB3	35.4	13	7	47.02
XP_013606250	NRPB4	15.9	2	4	21.74
XP_009135802	NRPB5a	24.2	7	7	39.81
XP_013634120	NRPB6a	16.6	2	2	33.33
XP_009126718	NRPB7	19.6	6	6	50.28
XP_009116628	NRPB8b	16.6	4	4	45.89
XP_009146283	NRPB9b	13.1	1	1	21.93
XP_009105127	NRPB10	8.2	2	2	28.17
XP_013602936	NRPB11	13.5	3	3	56.03
XP_048616672	NRPB12	6.0	1	1	13.73

Fig. S10. The purification and characterization of *Bo*Pol II. (**A**) The purification procedure. (**B**) The SDS-PAGE of the purified *Bo*Pol II stained by silver staining. (**C**) The mass-spectrum confirmation of the existence of all the 12 subunits of *Bo*Pol II.

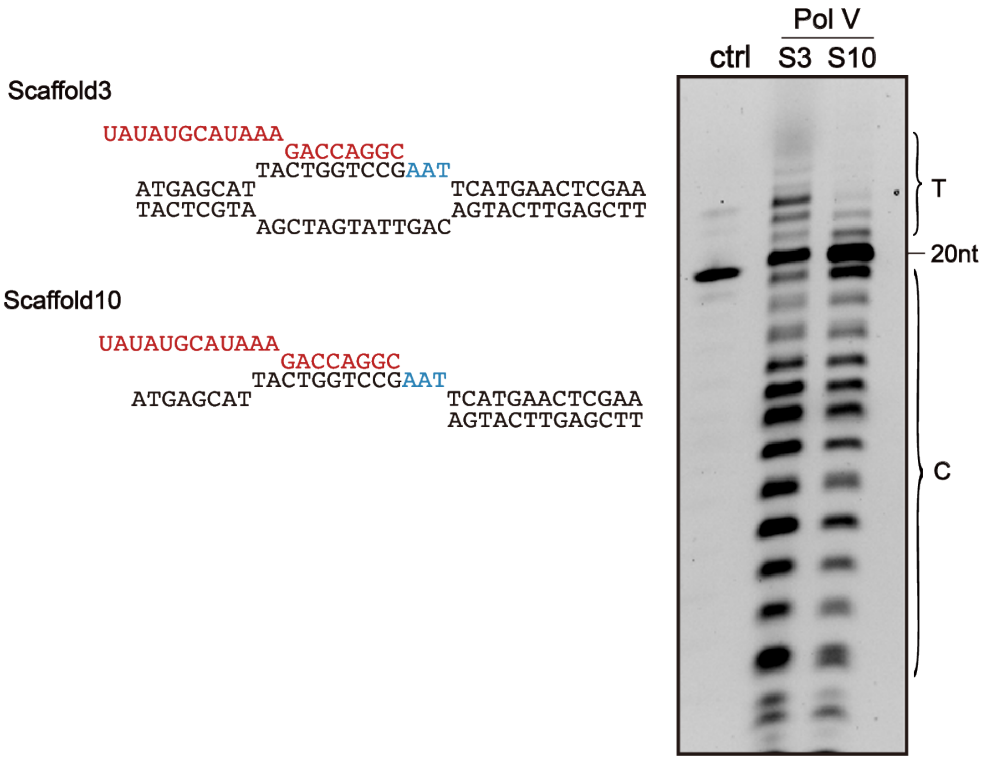


Fig. S11. The influence of the removal of DNA_{NT}-UPR and upstream paired region of DNA_{NT}. The transcription bubble designed to remove the DNA_{NT}-UPR and upstream paired region of DNA_{NT} (left panel) and the *in vitro* assay (right panel), which suggested that the removal of DNA_{NT}-UPR and upstream paired region of DNA_{NT} decreased both the elongation and cleavage activity of Pol V.

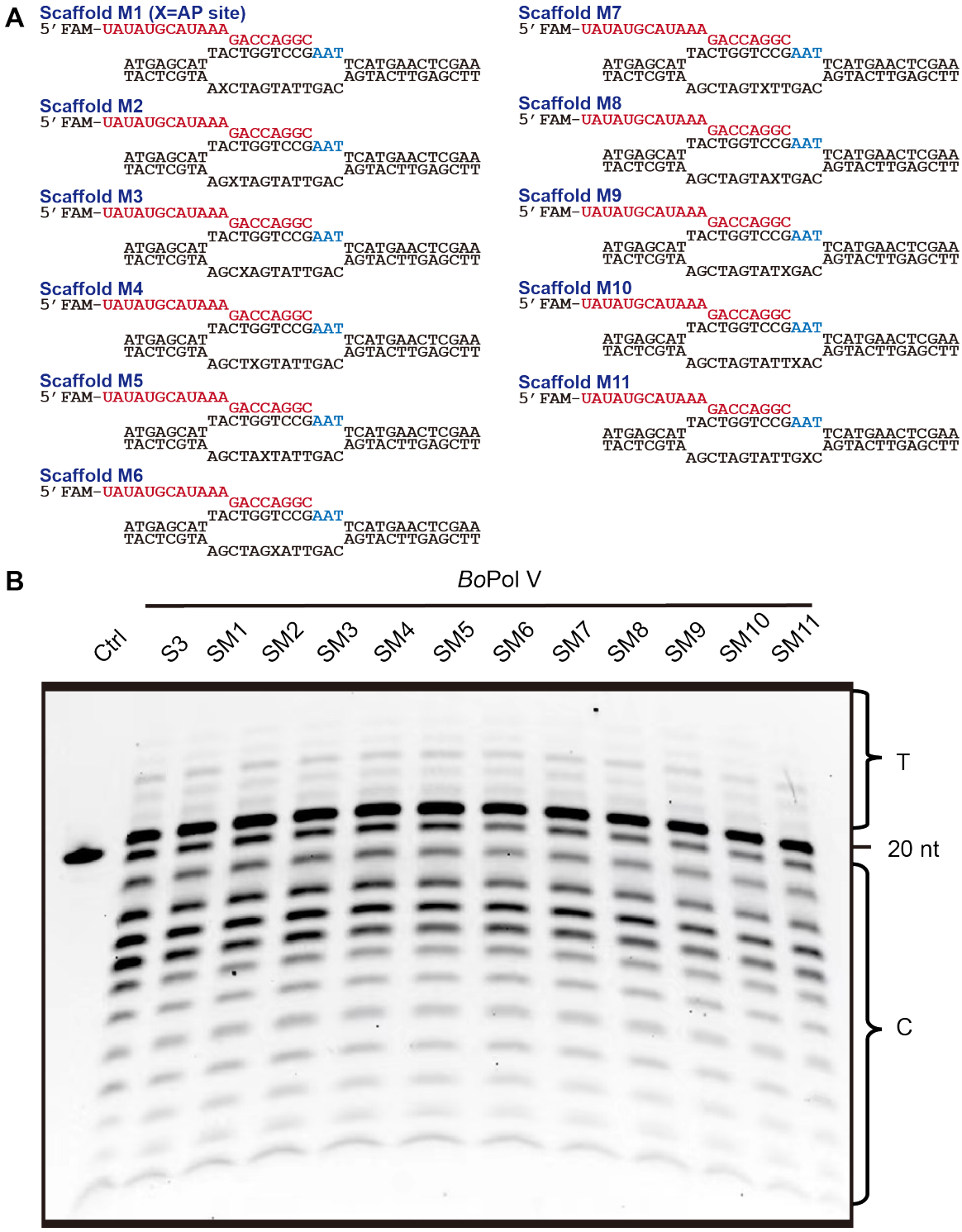


Fig. S12. The biochemical function of NRPE2-DNA$_{NT}$-UPR interaction. (**A**) The transcription bubble design to substitute DNA$_{NT}$-UPR nucleotide to AP (apurinic/apyrimidinic) site. (**B**) The *in vitro* assay suggested that the substitution of single DNA$_{NT}$-UPR base slightly decreased the cleavage activity but retaining the transcription activity.

Table S1. Oligos used in this research (X=AP site).

Name	Sequence	Purpose
DNA		
T-EM	AAGCTCAAGTATCGTAGCCTGGTCATTACTAGTA	Structure
NT-EM	TACTCGTAAGCTAGTATTGACGATACTTGAGCTT	Structure
T-S	AAGCTCAAGTACTTAAGCCTGGTCATTACGAGTA	Assay
NT-S2	TACTCGTAAGCTAGTATTGAAAGTACTTGAGCTT	Assay
NT-S3	TACTCGTAAGCTAGTATTGACAGTACTTGAGCTT	Assay
NT-S4	TACTCGTAAGCTAGTATTGACGGTACTTGAGCTT	Assay
NT-S5	TACTCGTAAGCTAGTATTGACGATACTTGAGCTT	Assay
NT-S6-1	TACTCGTAAGCTAGTATTGAC	Assay
NT-S6-2	AGTACTTGAGCTT	Assay
NT-S7-1	TACTCGTA	Assay
NT-S7-1	AGCTAGTATTGACAGTACTTGAGCTT	Assay
NT-S5X	TACTCGTAAGCTAXXXXXGACAGTACTTGAGCTT	Assay
NT-M1	TACTCGTAAXCTAGTATTGACAGTACTTGAGCTT	Assay
NT-M2	TACTCGTAAGXTAGTATTGACAGTACTTGAGCTT	Assay
NT-M3	TACTCGTAAGCXAGTATTGACAGTACTTGAGCTT	Assay
NT-M4	TACTCGTAAGCTXGTATTGACAGTACTTGAGCTT	Assay
NT-M5	TACTCGTAAGCTAXTATTGACAGTACTTGAGCTT	Assay
NT-M6	TACTCGTAAGCTAGXATTGACAGTACTTGAGCTT	Assay
NT-M7	TACTCGTAAGCTAGTXTTGACAGTACTTGAGCTT	Assay
NT-M8	TACTCGTAAGCTAGTAXTGACAGTACTTGAGCTT	Assay
NT-M9	TACTCGTAAGCTAGTATXGACAGTACTTGAGCTT	Assay
NT-M10	TACTCGTAAGCTAGTATTXACAGTACTTGAGCTT	Assay
NT-M11	TACTCGTAAGCTAGTATTGXCAGTACTTGAGCTT	Assay
RNA		
RNA-EM	UAUAUGCAUAAAGACCAGGC	Structure
RNA-20	FAM-UAUAUGCAUAAAGACCAGGC	Assay

Table S2. Cryo-EM data collection, refinement, and validation statistics.

	*Bo*Pol V apo	*Bo*Pol V EC
EMDB code	EMD-34820	EMD-34821
PDB code	8HIL	8HIM
Data collection and processing		
Microscopy	Titan Krios (#2, Guangming)	Titan Krios (#2, Guangming)
Voltage (kV)	300	300
Camera	Gatan K3 Summit	Gatan K3 Summit
Magnification	81,000	81,000
Pixel size (Å/pixel)	1.095	1.095
Total electron exposure ($e^-/Å^2$)	50	50
Exposure rate ($e^-/A^2/sec$)	20	20
Number of frames per movie (no.)	32	32
Defocus range (μm)	-1.0 to -2.5	-1.0 to -2.5
Automation software	EPU	EPU
Energy filter slit width (eV)	20	20
Micrographs collected (no.)	7,068	5,694
Micrographs used (no.)	6,805	5,694
Total extracted particles (no.)	1,609, 875	2,750,896
For each reconstruction		
Refined particles (no.)	63,603	153,968
Final particles (no.)	63,603	153,968
Data processing software	Relion3.1	Relion4.0
Point group	*C*1	*C*1
Resolution (global, Å)		
FSC 0.143 (unmasked/masked)	4.32/3.60	3.55/2.71
FSC 0.5 (unmasked/masked)	8.33/4.16	4.32/3.15
Local resolution range (Å)	3.0-6.0	2.6-5.0
Resolution range due to anisotropy (Å)	3.1-6.7	2.4-3.6
Map sharpening B factor ($Å^2$)	-85	-50
Map sharpening methods	Half-maps correlation	Half-maps correlation
Model composition		
No. atoms	23,134	23,946
Protein / RNA / DNA	23,127 / 0 / 0	22,632/239/1,068
Mg^{2+} / Zn^{2+}	1 / 6	1 / 6
Refinement		
Refinement package	Phenix	Phenix
- real or reciprocal space	Real space	Real space
- resolution cutoff (Å)	3.57	2.73
Model-Map scores		
- CC	0.79	0.84
- Average FSC (0.5 cutoff, Å)	3.57	2.89

B-factors (Å^2)	67.8	76.3
Protein / RNA / DNA	67.8 / - / -	70.6/134.6 /184.5
Mg^{2+} / Zn^{2+}	78.9 / 130.4	34.7 / 118.5
R.m.s. deviations		
Bond lengths (Å)	0.008	0.008
Bond angles (°)	1.332	1.217
Validation		
MolProbity score	2.00	1.74
Clashscore	6.01	6.32
Poor rotamers (%)	1.12	0.56
C-beta outlier (%)	0.00	0.04
CaBLAM outliers	7.03	3.21
EMRinger score	2.26	3.18
Ramachandran plot[b]		
Favored (%)	86.4	94.3
Allowed (%)	13.3	5.6
Outlier (%)	0.3	0.1

References and Notes

1. R. G. Roeder, W. J. Rutter, Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* **224**, 234–237 (1969). doi:10.1038/224234a0 Medline

2. J. R. Haag, C. S. Pikaard, Multisubunit RNA polymerases IV and V: Purveyors of non-coding RNA for plant gene silencing. *Nat. Rev. Mol. Cell Biol.* **12**, 483–492 (2011). doi:10.1038/nrm3152 Medline

3. A. J. Herr, M. B. Jensen, T. Dalmay, D. C. Baulcombe, RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**, 118–120 (2005). doi:10.1126/science.1106910 Medline

4. T. Kanno, B. Huettel, M. F. Mette, W. Aufsatz, E. Jaligot, L. Daxinger, D. P. Kreil, M. Matzke, A. J. M. Matzke, Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat. Genet.* **37**, 761–765 (2005). doi:10.1038/ng1580 Medline

5. M. A. Matzke, R. A. Mosher, RNA-directed DNA methylation: An epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* **15**, 394–408 (2014). doi:10.1038/nrg3683 Medline

6. Y. Onodera, J. R. Haag, T. Ream, P. Costa Nunes, O. Pontes, C. S. Pikaard, Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120**, 613–622 (2005). doi:10.1016/j.cell.2005.02.007 Medline

7. D. Pontier, G. Yahubyan, D. Vega, A. Bulski, J. Saez-Vasquez, M.-A. Hakimi, S. Lerbs-Mache, V. Colot, T. Lagrange, Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis*. *Genes Dev.* **19**, 2030–2040 (2005). doi:10.1101/gad.348405 Medline

8. T. Blevins, R. Podicheti, V. Mishra, M. Marasco, J. Wang, D. Rusch, H. Tang, C. S. Pikaard, Identification of Pol IV and RDR2-dependent precursors of 24 nt siRNAs guiding de novo DNA methylation in *Arabidopsis*. *eLife* **4**, e09591 (2015). doi:10.7554/eLife.09591 Medline

9. X. Du, Z. Yang, A. J. F. Ariza, Q. Wang, G. Xie, S. Li, J. Du, Structure of plant RNA-DEPENDENT RNA POLYMERASE 2, an enzyme involved in small interfering RNA production. *Plant Cell* **34**, 2140–2149 (2022). doi:10.1093/plcell/koac067 Medline

10. A. Fukudome, J. Singh, V. Mishra, E. Reddem, F. Martinez-Marquez, S. Wenzel, R. Yan, M. Shiozaki, Z. Yu, J. C.-Y. Wang, Y. Takagi, C. S. Pikaard, Structure and RNA template requirements of *Arabidopsis* RNA-DEPENDENT RNA POLYMERASE 2. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2115899118 (2021). doi:10.1073/pnas.2115899118 Medline

11. V. Gasciolli, A. C. Mallory, D. P. Bartel, H. Vaucheret, Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. *Curr. Biol.* **15**, 1494–1500 (2005). doi:10.1016/j.cub.2005.07.024 Medline

12. J. R. Haag, T. S. Ream, M. Marasco, C. D. Nicora, A. D. Norbeck, L. Pasa-Tolic, C. S. Pikaard, In vitro transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. *Mol. Cell* **48**, 811–818 (2012). doi:10.1016/j.molcel.2012.09.027 Medline

13. K. Huang, X.-X. Wu, C.-L. Fang, Z.-G. Xu, H.-W. Zhang, J. Gao, C.-M. Zhou, L.-L. You, Z.-X. Gu, W.-H. Mu, Y. Feng, J.-W. Wang, Y. Zhang, Pol IV and RDR2: A two-RNA-polymerase machine that produces double-stranded RNA. *Science* **374**, 1579–1586 (2021). doi:10.1126/science.abj9184 Medline

14. J. Singh, V. Mishra, F. Wang, H. Y. Huang, C. S. Pikaard, Reaction mechanisms of Pol IV, RDR2, and DCL3 drive RNA channeling in the siRNA-directed DNA methylation pathway. *Mol. Cell* **75**, 576–589.e5 (2019). doi:10.1016/j.molcel.2019.07.008 Medline

15. Q. Wang, Y. Xue, L. Zhang, Z. Zhong, S. Feng, C. Wang, L. Xiao, Z. Yang, C. J. Harris, Z. Wu, J. Zhai, M. Yang, S. Li, S. E. Jacobsen, J. Du, Mechanism of siRNA production by a plant Dicer-RNA complex in dicing-competent conformation. *Science* **374**, 1152–1157 (2021). doi:10.1126/science.abl4546 Medline

16. Z. Xie, E. Allen, A. Wilken, J. C. Carrington, DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12984–12989 (2005). doi:10.1073/pnas.0506426102 Medline

17. J. Zhai, S. Bischof, H. Wang, S. Feng, T. F. Lee, C. Teng, X. Chen, S. Y. Park, L. Liu, J. Gallego-Bartolome, W. Liu, I. R. Henderson, B. C. Meyers, I. Ausin, S. E. Jacobsen, A One precursor one siRNA model for Pol IV-dependent siRNA biogenesis. *Cell* **163**, 445–455 (2015). doi:10.1016/j.cell.2015.09.032 Medline

18. D. Zilberman, X. Cao, S. E. Jacobsen, ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* **299**, 716–719 (2003). doi:10.1126/science.1079695 Medline

19. D. Zilberman, X. Cao, L. K. Johansen, Z. Xie, J. C. Carrington, S. E. Jacobsen, Role of *Arabidopsis* ARGONAUTE4 in RNA-directed DNA methylation triggered by inverted repeats. *Curr. Biol.* **14**, 1214–1220 (2004). doi:10.1016/j.cub.2004.06.055 Medline

20. X. Cao, S. E. Jacobsen, Role of the *Arabidopsis* DRM methyltransferases in de novo DNA methylation and gene silencing. *Curr. Biol.* **12**, 1138–1144 (2002). doi:10.1016/S0960-9822(02)00925-9 Medline

21. X. Cao, S. E. Jacobsen, Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 16491–16498 (2002). doi:10.1073/pnas.162371599 Medline

22. G. Moissiard, S. J. Cokus, J. Cary, S. Feng, A. C. Billi, H. Stroud, D. Husmann, Y. Zhan, B. R. Lajoie, R. P. McCord, C. J. Hale, W. Feng, S. D. Michaels, A. R. Frand, M. Pellegrini, J. Dekker, J. K. Kim, S. E. Jacobsen, MORC family ATPases required for heterochromatin condensation and gene silencing. *Science* **336**, 1448–1451 (2012). doi:10.1126/science.1221472 Medline

23. A. T. Wierzbicki, T. S. Ream, J. R. Haag, C. S. Pikaard, RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nat. Genet.* **41**, 630–634 (2009). doi:10.1038/ng.365 Medline

24. X. Zhong, J. Du, C. J. Hale, J. Gallego-Bartolome, S. Feng, A. A. Vashisht, J. Chory, J. A. Wohlschlegel, D. J. Patel, S. E. Jacobsen, Molecular mechanism of action of plant DRM de novo DNA methyltransferases. *Cell* **157**, 1050–1060 (2014). doi:10.1016/j.cell.2014.03.056 Medline

25. J. R. Haag, O. Pontes, C. S. Pikaard, Metal A and metal B sites of nuclear RNA polymerases Pol IV and Pol V are required for siRNA-dependent DNA methylation and gene silencing. *PLOS ONE* **4**, e4110 (2009). doi:10.1371/journal.pone.0004110 Medline

26. R. Landick, Functional divergence in the growing family of RNA polymerases. *Structure* **17**, 323–325 (2009). doi:10.1016/j.str.2009.02.006 Medline

27. T. S. Ream, J. R. Haag, A. T. Wierzbicki, C. D. Nicora, A. D. Norbeck, J.-K. Zhu, G. Hagen, T. J. Guilfoyle, L. Pasa-Tolić, C. S. Pikaard, Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. *Mol. Cell* **33**, 192–203 (2009). doi:10.1016/j.molcel.2008.12.015 Medline

28. M. Marasco, W. Li, M. Lynch, C. S. Pikaard, Catalytic properties of RNA polymerases IV and V: Accuracy, nucleotide incorporation and rNTP/dNTP discrimination. *Nucleic Acids Res.* **45**, 11315–11326 (2017). doi:10.1093/nar/gkx794 Medline

29. P. Cramer, D. A. Bushnell, J. Fu, A. L. Gnatt, B. Maier-Davis, N. E. Thompson, R. R. Burgess, A. M. Edwards, P. R. David, R. D. Kornberg, Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288**, 640–649 (2000). doi:10.1126/science.288.5466.640 Medline

30. P. Cramer, D. A. Bushnell, R. D. Kornberg, Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**, 1863–1876 (2001). doi:10.1126/science.1059493 Medline

31. N. A. Hoffmann, A. J. Jakobi, M. Moreno-Morcillo, S. Glatt, J. Kosinski, W. J. H. Hagen, C. Sachse, C. W. Müller, Molecular structures of unbound and transcribing RNA polymerase III. *Nature* **528**, 231–236 (2015). doi:10.1038/nature16143 Medline

32. C. Engel, S. Sainsbury, A. C. Cheung, D. Kostrewa, P. Cramer, RNA polymerase I structure and transcription regulation. *Nature* **502**, 650–655 (2013). doi:10.1038/nature12712 Medline

33. C. Fernández-Tornero, M. Moreno-Morcillo, U. J. Rashid, N. M. I. Taylor, F. M. Ruiz, T. Gruene, P. Legrand, U. Steuerwald, C. W. Müller, Crystal structure of the 14-subunit RNA polymerase I. *Nature* **502**, 644–649 (2013). doi:10.1038/nature12636 Medline

34. M. Girbig, A. D. Misiaszek, C. W. Müller, Structural insights into nuclear transcription by eukaryotic DNA-dependent RNA polymerases. *Nat. Rev. Mol. Cell Biol.* **23**, 603–622 (2022). doi:10.1038/s41580-022-00476-9 Medline

35. T. J. Guilfoyle, Purification, subunit structure, and immunological properties of chromatin-bound ribonucleic acid polymerase I from cauliflower inflorescence. *Biochemistry* **19**, 5966–5972 (1980). doi:10.1021/bi00567a004 Medline

36. L. Huang, A. M. E. Jones, I. Searle, K. Patel, H. Vogler, N. C. Hubner, D. C. Baulcombe, An atypical RNA polymerase involved in RNA silencing shares small subunits with RNA polymerase II. *Nat. Struct. Mol. Biol.* **16**, 91–93 (2009). doi:10.1038/nsmb.1539 Medline

37. K. J. Armache, S. Mitterweger, A. Meinhart, P. Cramer, Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *J. Biol. Chem.* **280**, 7131–7134 (2005). doi:10.1074/jbc.M413038200 Medline

38. S. L. Tucker, J. Reece, T. S. Ream, C. S. Pikaard, Evolutionary history of plant multisubunit RNA polymerases IV and V: Subunit origins via genome-wide and segmental gene

duplications, retrotransposition, and lineage-specific subfunctionalization. *Cold Spring Harb. Symp. Quant. Biol.* **75**, 285–297 (2010). doi:10.1101/sqb.2010.75.037 Medline

39. A. C. Cheung, P. Cramer, Structural basis of RNA polymerase II backtracking, arrest and reactivation. *Nature* **471**, 249–253 (2011). doi:10.1038/nature09785 Medline

40. S. Sainsbury, J. Niesser, P. Cramer, Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. *Nature* **493**, 437–440 (2013). doi:10.1038/nature11715 Medline

41. D. Wang, D. A. Bushnell, K. D. Westover, C. D. Kaplan, R. D. Kornberg, Structural basis of transcription: Role of the trigger loop in substrate specificity and catalysis. *Cell* **127**, 941–954 (2006). doi:10.1016/j.cell.2006.11.023 Medline

42. S. Neyer, M. Kunz, C. Geiss, M. Hantsche, V.-V. Hodirnau, A. Seybert, C. Engel, M. P. Scheffer, P. Cramer, A. S. Frangakis, Structure of RNA polymerase I transcribing ribosomal DNA genes. *Nature* **540**, 607–610 (2016). doi:10.1038/nature20561 Medline

43. C. Bernecky, F. Herzog, W. Baumeister, J. M. Plitzko, P. Cramer, Structure of transcribing mammalian RNA polymerase II. *Nature* **529**, 551–554 (2016). doi:10.1038/nature16482 Medline

44. H. Kettenberger, K. J. Armache, P. Cramer, Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol. Cell* **16**, 955–965 (2004). doi:10.1016/j.molcel.2004.11.040 Medline

45. E. Nudler, RNA polymerase backtracking in gene regulation and genome instability. *Cell* **149**, 1438–1445 (2012). doi:10.1016/j.cell.2012.06.003 Medline

46. N. E. Thompson, D. B. Aronson, R. R. Burgess, Purification of eukaryotic RNA polymerase II by immunoaffinity chromatography. Elution of active enzyme with protein stabilizing agents from a polyol-responsive monoclonal antibody. *J. Biol. Chem.* **265**, 7069–7077 (1990). doi:10.1016/S0021-9258(19)39260-9 Medline

47. J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. H. Hagen, E. Lindahl, S. H. W. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018). doi:10.7554/eLife.42166 Medline

48. S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017). doi:10.1038/nmeth.4193 Medline

49. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015). doi:10.1016/j.jsb.2015.08.008 Medline

50. J. Zivanov, T. Nakane, S. H. W. Scheres, Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in *RELION*-3.1. *IUCrJ* **7**, 253–267 (2020). doi:10.1107/S2052252520000081 Medline

51. D. Kimanius, L. Dong, G. Sharov, T. Nakane, S. H. W. Scheres, New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem. J.* **478**, 4169–4185 (2021). doi:10.1042/BCJ20210708 Medline

52. T. Bepler, A. Morin, M. Rapp, J. Brasch, L. Shapiro, A. J. Noble, B. Berger, Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* **16**, 1153–1160 (2019). doi:10.1038/s41592-019-0575-8 Medline

53. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004). doi:10.1107/S0907444904019158 Medline

54. P. V. Afonine, B. K. Poon, R. J. Read, O. V. Sobolev, T. C. Terwilliger, A. Urzhumtsev, P. D. Adams, Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.* **74**, 531–544 (2018). doi:10.1107/S2059798318006551 Medline

55. Y. Z. Tan, P. R. Baldwin, J. H. Davis, J. R. Williamson, C. S. Potter, B. Carragher, D. Lyumkis, Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796 (2017). doi:10.1038/nmeth.4347 Medline

56. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004). doi:10.1002/jcc.20084 Medline

57. P. Di Tommaso, S. Moretti, I. Xenarios, M. Orobitg, A. Montanyola, J.-M. Chang, J.-F. Taly, C. Notredame, T-Coffee: A web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res.* **39**, W13–W17 (2011). doi:10.1093/nar/gkr245 Medline

58. X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014). doi:10.1093/nar/gku316 Medline