Structural insights into the remodeled RNA polymerase II complex on an RNA template

Jie Hao,[1†] Zhuan Qin,[2†*] Junfei Ma,[1†] Jie Qu,[3†§] Yunhan Wang,[4†] Svetlana Y. Folimonova,[1,4] Bin Liu,[5,6] Wenwei Li,[2*] Ying Wang[1,4*]

[1] Department of Plant Pathology, University of Florida, Gainesville, FL 32611, USA
[2] Department of Microbial Pathogenesis, Yale University, New Haven, CT 06536, USA
[3] Department of Molecular Genetics, Ohio State University, Columbus, OH 43210, USA
[4] Plant Molecular and Cellular Biology Program, University of Florida, Gainesville, FL 32611, USA
[5] Food Science and Human Nutrition Department, University of Florida, FL 32611, USA
[6] Genetics Institute, University of Florida, FL 32611, USA
* To whom correspondence should be addressed. Email: zhuan.qin@yale.edu, wenwei.li@yale.edu, and ying.wang1@ufl.edu
† The first five authors should be regarded as Joint First Author.
§ Current address: Medpace Reference Laboratories, Cincinnati OH 45227, USA

Key words: viroid, RNA-templated transcription, RNA polymerase II (Pol II), Transcription Factor IIIA

Abstract

DNA-dependent RNA polymerases recognize not only DNA but also RNA templates. This RNA-dependent RNA polymerase activity was exploited by human hepatitis delta virus in mammalian cells and viroids in plants. Emerging evidence infers a re-organization of RNA polymerase II (Pol II) transcription complex on potato spindle tuber viroid (PSTVd) RNA template, an understudied layer of transcriptional regulation. Here, we provide evidence confirming that the reorganization of the 12-subunit Pol II to 7-subunit occurs in vivo for PSTVd transcription. Rpb4, Rpb5, Rpb6, Rpb7, and Rpb9 are not involved in PSTVd transcription. A splicing variant of the transcription factor IIIA with seven zinc finger domains (TFIIIA-7ZF) aids the remodeled Pol II in transcribing PSTVd. Based on AlphaFold3 prediction, we obtained the structure of the remodeled Pol II with PSTVd RNA and TFIIIA-7ZF. The predicted structure and experimental data both show that the N-terminus of TFIIIA-7ZF binds to the left terminal domain of PSTVd, while the C-terminus of TFIIIA-7ZF interacts with Rpb2. Altogether, the data illustrate an active form of heterogenous organization of the essential Pol II enzyme in vivo and provide structural insights into the remodeled Pol II and TFIIIA-7ZF transcription complex on RNA template.

Introduction

Transcription is a fundamental process that mediates the interplay between genetic information and phenotypes and, thus, is vital for organismal growth and responses to environmental cues (1-4). Transcription is catalyzed by DNA-dependent RNA polymerases (DdRPs) that generally polymerize RNA molecules based on DNA templates (5-7). Interestingly, it is well-known that some DdRPs can recognize both DNA and RNA templates for transcription (8). The RNA-dependent RNA polymerase (RdRP) activity of DdRPs regulates gene expression in diverse organisms across kingdoms. For instance, under nutrient-deficient conditions, the noncoding 6S RNA of *Escherichia coli* binds to the bacterial DdRP and prevents it from transcribing DNA templates. In response to nutrient availability, the DdRP transcribes a short RNA using the 6S RNA template that leads to the de-sequestration of the DdRP. This step, in turn, enables binding to DNA promoters and the synthesis of mRNAs (9). RNA polymerase II (Pol II) in mammalian cells binds to a hairpin within the noncoding B2 RNA and uses the longer strand of the hairpin sequence

63 as a template to extend the short strand by transcription. This extension subsequently
64 destabilizes the B2 RNA, representing a novel posttranscriptional mechanism to regulate
65 RNA stability (10). RNA-based pathogens, such as viroids and the human hepatitis delta
66 virus (HDV), also rely on this RdRP activity for replication in corresponding hosts (11).
67 Notably, both (+) and (-) strands of potato spindle tuber viroid (PSTVd, the type species
68 of the *Pospiviroidae* family) are transcribed by Pol II. Despite the importance, however,
69 the underlying regulatory mechanism of this RdRP activity of DdRPs remains elusive.

70 Taking Pol II on DNA templates as an example, this 12-subunit polymerase relies on
71 the coordination of multiple general transcription factors (i.e., TFIIA, TFIIB, TFIID, TFIIE,
72 TFIIF, and TFIIH) during transcription initiation around DNA promoter regions (12-15). By
73 and large, a minimal set of five general transcription factors (TFIIB, TFIID, TFIIE, TFIIF,
74 and TFIIH) is required for the promoter-driven transcription catalyzed by Pol II (16,17).
75 Transcription elongation, particularly in the cellular environment, requires additional
76 factors, including TFIIS, TFIIF, SPT6, etc (18-20). However, emerging evidence suggests
77 that Pol II complex on RNA templates may be organized in a distinct fashion.

78 In a protoplast cell-based replication system, we found that a TFIIS loss-of-function
79 mutant did not impair the propagation of potato spindle tuber viroid (PSTVd) (21), which
80 replicates in the nucleoplasm and relies on Pol II for de novo transcription (21-25). This
81 observation inspires us to reason that certain general transcription factors may not be
82 involved in the RdRP activity of Pol II. On the other hand, an alternative splicing product
83 of the transcription factor IIIA with seven zinc finger domains (TFIIIA-7ZF) is essential to
84 engage the remodeled Pol II to PSTVd RNA templates for transcription (21,22,25).
85 Previous studies showed that TFIIIA-7ZF interacts with Pol II as well as with (+) and (-)
86 PSTVd RNAs in vivo. Overexpression of TFIIIA-7ZF led to the elevated PSTVd titer, while
87 silencing of *TFIIIA-7ZF* reduced PSTVd accumulation (22). Moreover, TFIIIA-7ZF is
88 required in a Pol II-based in vitro transcription assay for generating full-length PSTVd
89 RNA products (21,22,25). This is akin to the replication of HDV that requires a virus-
90 encoded HDAg-S protein to ensure robust processivity in generating long RNA products
91 (26).

92 More recently, we found that a “remodeled” Pol II, which is composed of Rpb1, Rpb2,
93 Rpb3, Rpb8, Rpb10, Rpb11, and possibly Rpb12, actively catalyzes the RNA
94 polymerization using PSTVd RNA templates in the Pol II-based in vitro transcription
95 system (25). The heterogeneous organization of Pol II is supported by additional lines of
96 evidence. First of all, the absence of Rpb9 in the “remodeled” Pol II explains the higher
97 mutation rates of PSTVd during replication (27), when compared to the Pol II fidelity on
98 DNA templates (28). Furthermore, distinct Pol II subunits appear to control the
99 transcription of different sets of genes according to degron-based transiently degradation
100 of each subunit in mammalian cell lines (29). Taken together, evidence supports a model
101 of a distinctive organization of the transcription complex on viroid RNA templates. In spite
102 of the supporting evidence, whether such remodeled Pol II exists and functions in cells
103 remains to be directly determined. Moreover, it is unclear how TFIIIA-7ZF interacts with
104 the remodeled Pol II for RNA-templated transcription.

105 To gain further insights into the organization of the Pol II transcription complex on RNA
106 template, we tested whether Rpb4, Rpb5, Rpb6, Rpb7, Rpb9 and Rpb12 are involved in
107 binding PSTVd RNA templates in vivo. These subunits were considered absent in our
108 previously remodeled Pol II (25). Interestingly, we found that Rpb12 was involved in RNA
109 template binding, while the rest five subunits (Rpb4, Rpb5, Rpb6, Rpb7, and Rpb9) did
110 not interact with PSTVd RNA template in vivo. This observation is largely in agreement
111 with our recent nLC-MS/MS analysis of the active remodeled Pol II on RNA templates,
112 providing compelling evidence confirming the re-organization of Pol II for transcribing
113 viroid RNA templates in plants.

114 Based on this information, we predicted the structure of the remodeled Pol II with
115 seven subunits (Rpb1, Rpb2, Rpb3, Rpb8, Rpb10, Rpb11, and Rpb12) as well as TFIIIA-
116 7ZF on the (+) PSTVd RNA template. The predicted model provides the opportunity to
117 understand the functional mechanism of an RNA-specific transcription factor (i.e., TFIIIA-
118 7ZF) in RNA-templated transcription. We experimentally tested the model to validate the
119 interaction and structural geometry of RNA templates, TFIIIA-7ZF, and the subunits in the
120 remodeled Pol II. Furthermore, we predicted the complex of TFIIIA-7ZF and (+) PSTVd
121 RNA that pinpoint zinc finger domains (ZF) 2, 3, 6 as possible RNA binding domains.

Experimental data clearly support the critical role of ZF3 and ZF6 are critical for RNA binding.

Altogether, our data support the in vivo existence of the remodeled Pol II and illustrate how TFIIIA-7ZF bridges the RNA template and the remodeled Pol II, which provides structural insights into the novel regulation of transcription by using an altered organization of polymerase complex.

Materials and Methods

Plant Materials

Nicotiana benthamiana plants were grown in a growth room with 28°C and a 14-h/10-h light/dark cycle. During agroinfiltration period, plants were moved into a growth room with 22°C and a 14-h/10-h light/dark cycle. PSTVd-infection and validation steps were the same as described previously (30,31).

Molecular Cloning

Following clones were ordered from Arabidopsis Biological Resource Center (ABRC) at the Ohio State University: Rpb4 (DKLAT5G09920) and Rpb7 (DKLAT5G59180). ORFs of Rpb5 (At3g22320), Rpb6 (At2g04630), Rpb9 (At4g16265), Rpb12 (At5g41010), TFIIF1 (At1g75510), and TFIIS (AT2G38560) were cloned via conventional RT-PCR and inserted into the pENTR-D-TOPO vector (ThermoFisherSci, Waltham, MA) and recombined to destination vectors. All clones were verified by Sanger Sequencing. Primers were listed in the Supp Table 1.

RNA-Immunoprecipitation (RIP)

RIP was performed following our established protocol (22). PSTVd-infected *N. benthamiana* plants were agro-infiltrated with destination vectors to express GFP-tagged Pol II subunits, TFIIF, and TFIIS, as well as HA-tagged TFIIIA-7ZF variants. Three days post infiltration, protein expressions were verified by microscopy observation of green fluorescence before RIP procedures. Collected leaf samples were lysed in RIP buffer (25 mM Tris pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 10% glycerol, 0.1% Triton X-100, 0.2% NP-40, protease inhibitor cocktail following manual) as described before (22). The lysates

152 were incubated with GFP-trap beads (ProteinTech, Rosemont, IL) anti-HA magnetic
153 beads (ThermoFisherSci) or anti-Myc magnetic beads (ThermoFisherSci) for 2 hr at 4°C
154 with rotation. The beads were washed twice in IP buffer and once in RNase-Free water,
155 before elution using acidic Glycine buffer following manual instructions (ThermoFisherSci).
156 The protein fractions were analyzed via immunoblotting as described before in details
157 (22). The precipitated RNAs were analyzed by semi-quantitative RT-PCR or RT-qPCR
158 (please see Supp Table 1 for primer sequences). For RT-qPCR, 5.8S rRNA was used as
159 negative control for normalizing PSTVd enrichment folds. Data from three biological
160 replicates were used for T-test analysis.
161
162 Bimolecular fluorescence complementation (BiFC) and Microscopy
163 For BiFC, PSTVd-infected *N. benthamiana* seedlings were used for agroinfiltration of
164 various combinations of constructs, all including 35S:RFP-Histone 2B (30) as the nucleus
165 marker. The N split (aa 1–174) YFP was fused in front of the N-terminus of RPB3 or
166 RPB12. The C split YFP (aa 175-end) was fused after the C-terminus of TFIIIA-7ZF and
167 its variant. We analyzed 10 randomly chosen regions of infiltrated leaves from at least
168 three plants for each treatment. ECHO Revolve microscope (Discover Echo Inc., San
169 Diego, CA) was used for observing the fluorescence signals.
170
171 Electrophoretic mobility shift assays (EMSA)
172 For EMSA, TFIIIA-7ZF and its variants were expressed using IMPACT system (New
173 England Biolabs, Ipswich, MA) in bacteria and purified using Chitin beads as previously
174 described (22). Protein-RNA binding reactions (20 mM Tris-HCl pH 7.5, 35 mM KCl, 5 μM
175 $ZnCl_2$, 3.5 mM $MgCl_2$, 10 nM yeast tRNA and 10% (v/v) glycerol) were performed at 25°C
176 using ^{32}P-labeled RNAs (20,000 dpm) with increasing concentrations of recombinant
177 proteins (described in Fig. 4). Electrophoresis was carried out on ice using 6% (w/v)
178 polyacrylamide (29:1) gels for 2 h at 120 V with Tris borate buffer (65 mM Tris, 22.5 mM
179 boric acid, pH 8). PhosphorImager cassette was used to visualize the signal in a Personal
180 Molecular Imager (Bio-Rad Laboratories, Hercules, CA). The intensity of each band on
181 the gel was quantified by using Quantity One software (Bio-Rad Laboratories). The

182 binding curves were obtained by plotting the fraction of RNA bound with proteins as
183 previously described (22).
184
185 **Results**

186 **The composition of in vivo Pol II complex on RNA templates**

187 Pol II has been well documented for catalyzing the replication of nuclear-replicating
188 viroids (32,33). Our recent findings demonstrate that Pol II can catalyze RNA-templated
189 transcription in vitro without a complete set of 12 subunits (25). We termed this distinct
190 organization the "remodeled" Pol II. To corroborate the remodeled Pol II organization in
191 vivo, we performed RNA immunoprecipitation (RIP) analyses to test the interaction
192 between PSTVd RNA and the Pol II subunits absent from the remodeled Pol II (i.e., Rpb4,
193 Rpb5, Rpb6, Rpb7, and Rpb9), using Rpb3 as a positive control. We could not detect
194 Rpb12 in our previous work (25), but it was suspected that the small size of Rpb12
195 rendered it being filtered out by size cut-off columns during sample preparation.
196 Therefore, we also included Rpb12 in this test. Since most of these subunits have multiple
197 gene copies in plants, we selected the gene copies that are expressed and incorporated
198 into Pol II based on published data (34). These subunits were expressed in the PSTVd-
199 infected *N. benthamiana* plants as fusions with a GFP-tag for RIP analyses. Plants
200 infiltrated with agrobacteria without the expression vector served as negative control.

201 As shown in Fig.1A, GFP-trap successfully precipitated all the target subunits. We
202 then used RT-qPCR to detect PSTVd RNA templates enriched in the immunoprecipitated
203 fraction. 5.8S rRNA was used as negative control for normalizing RT-qPCR data. As
204 shown in Fig. 1A, PSTVd RNA was significantly co-precipitated in samples expressing
205 subunits Rpb3 and Rpb12, confirming that Rpb3 and Rpb12 are part of the remodeled
206 Pol II in vivo. By contrast, PSTVd RNA was not significantly enriched in samples
207 expressing subunits Rpb4, Rpb5, Rpb6, Rpb7, or Rpb9 as compared with the sample
208 from plants infiltrated with an agrobacterium suspension without the expression vector.
209 Therefore, these subunits are unlikely part of the remodeled Pol II on PSTVd RNA
210 templates. The results are consistent with our previous in vitro analysis, except for Rpb12.

211 As previously speculated, Rpb12 was likely lost during our sample preparation for nLC-
212 MS/MS analysis, which was reported in a similar study (20).

213

214 Structure overview of the TFIIIA-7ZF/Pol II complex on the (+) PSTVd RNA 215 templates

216 Previous structural studies have resolved multiple high-quality Pol II complexes with
217 general transcription factors at different transcription stages (19,35-41). The structure of
218 a canonical TFIIIA protein with nine zinc finger domains (ZFs) is also available (PDB 1tf6)
219 (42). However, the structure illustrating the complex of Pol II and RNA-specific
220 transcription factors (i.e., HDAg-S and TFIIIA-7ZF) is lacking. Nevertheless, these
221 resolved structures provided the basis for predicting a Pol II complex with other factors.
222 Taking the advantage of the emergence of AlphaFold3 (43), we predicted the structure of
223 the TFIIIA-7ZF/Pol II complex on the (+) PSTVd RNA template (359 nt). For comparison,
224 we also predicted the Pol II complex (with 12 subunits) on a similar length of a DNA
225 template (395 bp), using the sequence from an expression vector pRTL2 with the 35S
226 promoter. As shown in Fig. S1, the predicted models have very high confidence for the
227 Pol II and remodeled Pol II complexes, which reflects the presence of high-quality Pol II
228 data. The structures for the nucleic acid portions generally have low confidence, which is
229 expected since AlphaFold3 is generally thought to have less accuracy in predicting
230 nucleic acid structures.

231 The remodeled Pol II lacks the “Stalk” that consists of Rpb4/7 and its adjacent Rpb6
232 (Fig. 1B). This absence is possibly not caused by TFIIIA-7ZF occupancy directly because
233 TFIIIA-7ZF does not share any common binding sites with these three subunits (Fig. 1B).
234 The absence of Rpb5 and Rpb9 may be caused by TFIIIA-7ZF occupancy since their
235 binding sites crossover with the TFIIIA-7ZF binding site on the surface of Rpb1.

236

237 Testing the geometry of TFIIIA-7ZF in the Pol II/TFIIIA-7ZF/PSTVd complex

238 We noticed that TFIIIA-7ZF wraps the RNA template while attaches to the remodeled
239 Pol II complex (Fig. 1B). The N-terminal region of TFIIIA-7ZF is involved in binding with

240 the RNA template, and the TFIIIA-7ZF C-terminus appears to be critical for binding with
241 the Pol II core, particularly Rpb2. As shown in Fig. 2A, residues R243, R245, and K257
242 of TFIIIA-7ZF are predicted to be involved in binding with Rpb2 Q91 in the Protrusion
243 domain and E549/E550 in the External 2 domain. Residues R243, R245, and K257 are
244 outside of the seven zinc finger domains towards the C-terminus of the TFIIIA-7ZF. We
245 also noticed that the C-terminus end of TFIIIA-7ZF is relatively close to the N-terminus of
246 Rpb12 but away from the N-terminus of Rpb3 (Fig. 2B).

247 To empirically test the predicted structural geometry of the TFIIIA-7ZF/Pol II complex,
248 we used bimolecular fluorescence complementation assay (BiFC) to assess the relative
249 distance between TFIIIA-7ZF (with a C-terminal half of YFP fused to the end of the TFIIIA-
250 7ZF C-terminus) and Rpb12 or Rpb3 (with a N-terminal half of YFP fused to the subunit's
251 N-terminus). As shown in Fig. 2C, only the Rpb12/TFIIIA-7ZF pair, but not the
252 Rpb3/TFIIIA-7ZF pair, could complement the YFP fluorescence. We then designed a C-
253 terminus-lacking variant of TFIIIA-7ZF (TFIIIA-7ZF$^{C\text{-}del}$; aa 1-228) that still keeps all seven
254 zinc finger domains but lacks the critical residues for Rpb2 binding. This TFIIIA-7ZF$^{C\text{-}del}$
255 could not complement YFP fluorescence with Rpb12. This result confirms that the C-
256 terminus of TFIIIA-7ZF is critical for binding with the remodeled Pol II complex.
257 Importantly, the BiFC data corroborate the predicted structure of the complex in terms of
258 the relative distance between TFIIIA-7ZF and Pol II subunits (i.e., Rpb3 and Rpb12).

259

260 **Structure of TFIIIA-7ZF/viroid complex**

261 In our Pol II-based in vitro transcription assay, incubating RNA template with TFIIIA-
262 7ZF prior to the addition of Pol II is critical for optimized transcription efficiency (21), which
263 implies that TFIIIA-7ZF and PSTVd RNA-protein complex is formed first before recruiting
264 Pol II. However, it is unclear how TFIIIA-7ZF coordinates its seven ZFs for RNA template
265 binding. Therefore, we were interested in understanding the structure of TFIIIA-7ZF and
266 the PSTVd (+) genomic RNA by using AlphaFold3. Previously, it was speculated that the
267 left terminal region of PSTVd (+) RNA genome forms either a rod-shape structure or a
268 bifurcated structure during transcription (44). However, the model predicted by
269 AlphaFold3 provided an unexpected conformation (Fig. 3A). While PSTVd RNA bases

270 still pair similarly as in the rod-shape structure, the TFIIIA-7ZF protein bends the overall
271 rod-shape structure that leads to the formation of an enlarged left terminal region.

272 A close look at the predicted structure showed that ZF2 and ZF3 clamp the major
273 groove of the RNA template (Fig. 3B)., The C-terminus of TFIIIA-7ZF binds to the RNA
274 template, instead of binding with Rpb2 in the Pol II/TFIIIA-7ZF/PSTVd complex shown in
275 Fig. 2A. The TFIIIA-7ZF binding sites are estimated as G328 to G354 (Fig. 3A), akin to
276 our foot-printing assay that mapped the binding region from U331 to G346 (22). This
277 region is also proximal to the transcription initiation site between U359 and C1 (45). The
278 significant overlapping between the experimental data and the prediction model suggests
279 that the predicted model can indeed reflect the structure of the (+) PSTVd/TFIIIA-7ZF
280 complex. To further map the ZFs that are critical for RNA templates binding, we
281 introduced a point mutation changing the first histidine to asparagine to disrupt each
282 C2H2 ZF as previously reported (46,47) and performed RIP using these variants with an
283 HA tag. We could not detect the expression of the *zf5* variant in plants, potentially
284 attributable to poor stability of the protein. Nevertheless, the rest six variants were
285 successfully expressed and detected in the eluted RIP fractions. Since our primer sets
286 can detect both (+) and (-) PSTVd strands simultaneously, we only used semi-quantitative
287 RT-PCR to test for the enrichment of the (+) PSTVd RNA in the immunoprecipitated
288 fractions. As shown in Fig. 3C, the *zf2*, *zf3*, and *zf6* variants consistently failed to
289 precipitate PSTVd RNA, indicating that these domains are critical for RNA binding, which
290 provides a strong support for the predicted model of the RNP.

291 To further test the critical zinc fingers involved in (+) PSTVd RNA binding, we
292 performed electrophoretic mobility shift assays (EMSAs) using each *zf* variant. As shown
293 in Fig. 4, disruption of either ZF3 or ZF6 significantly impacts RNA binding, as these
294 variants lost the ability to achieve a complete RNA shifting in comparison to wildtype and
295 other *zf* variants. Interestingly, the rest of *zf* variants demonstrated similar or even higher
296 affinity to the PSTVd RNA as indicated by the dissociation constant (Kd). Notably, the *zf2*
297 variant could bind with (+) PSTVd RNA in EMSA tests, which is different from the RIP
298 outcomes presented in Fig. 3C. This discrepancy implies that the ZF2 domain may be
299 involved in other interactions in the cellular environment. On the other hand, when ZF2 is
300 impaired, ZF1 might contribute to stabilizing RNA binding in vitro through working together

301 with ZF3 for RNA binding. This is implied by the geometry between ZF1 and ZF3 on
302 PSTVd RNA in the predicted model (Fig. 3B). In spite of this discrepancy of ZF2, RIP,
303 EMSA, and the structural model all support the critical roles of ZF3 and ZF6 in (+) PSTVd
304 binding.
305
306 **Discussion**
307 The results presented in this study, together with the published data showing the
308 transcription-competent Pol II remodeled on PSTVd RNA template in vitro (25),
309 demonstrated a distinct organization of Pol II transcribing viroid RNA template in plants.
310 By using RNA-immunoprecipitation, we showed that only a subset of Pol II subunits is
311 associated with PSTVd RNA, in line with of our previous results from an in vitro setting
312 (25). Rpb1, Rpb2, Rpb3, Rpb8, Rpb10, Rpb11, and Rpb12 constitute the remodeled Pol
313 II for RNA-templated transcription. During the assembly of Pol II complex, there are
314 several intermediate complexes/subassemblies, including the Rpb1 subassembly
315 containing Rpb1/5/6/8, the Rpb2 subassembly containing Rpb2/9, the Rpb3 subassembly
316 containing Rpb3/10/11/12 and the Rpb4/7 subassembly (Stalk)(48). Interestingly, the
317 remodeled Pol II contains the intact Rpb3 subassembly, a modified Rpb1 subassembly
318 lacking Rpb5 and Rpb6, and a modified Rpb2 subassembly lacking Rpb9 (49,50). The
319 stalk portion of the Pol II is completely missing in the remodeled Pol II. It is known that
320 Rpb4, Rpb5, Rpb7, and Rpb9 tend to dissociate from the Pol II complex upon alpha-
321 Amanitin treatment (49), which is in line with our observation. Our findings also align with
322 a recent finding of possible heterogeneous organization of Pol II in mammalian cells (29),
323 suggesting that the transcriptional regulation at the polymerase organization level is more
324 complex than previously recognized.
325 Confirming the remodeled Pol II organization in vivo paves the way for further
326 investigations on the different regulations of DNA-dependent versus RNA-templated
327 transcription. For instance, the Pol II stalk is involved in binding with capping enzymes for
328 adding G-cap to nascent mRNAs (51,52). This explains the lack of G-cap in viroid RNAs
329 (53). Whether HDV RNA-templated transcription also uses a remodeled Pol II is not
330 immediately clear. Some HDV-derived mRNAs contain a G-cap, but Pol II-transcribed
331 genomic or antigenomic RNAs do not (54,55).

332 Although the detailed structural basis has largely been determined for almost every
333 step in RNA polymerases transcribing DNA templates, the molecular basis underlying the
334 widely existing RNA-templated transcription remains elusive. To date, the only known
335 structure, which contained 12-subunit Pol II, TFIIS, and a synthetic mosaic RNA template
336 with partial HDV sequence, was determined by using a crystallographic analysis (56).
337 This structure did not include the RNA-specific transcription factor (HDAg-S) for HDV
338 RNA templates. In addition, this complex does not possess strong processivity for yielding
339 long products (longer than 100 nt). Moreover, the TFIIS-based cleavage allows Pol II to
340 catalyze primed transcription on some RNA templates, but it is not required for
341 transcribing PSTVd as Pol II catalyzes de novo transcription using PSTVd RNA template
342 (21).
343 Our previous work established a critical role of TFIIIA-7ZF in enhancing Pol II
344 processivity on PSTVd RNA templates (21,22,25). In the simplest experimental setup,
345 products with hundreds of nucleotides were generated in vitro by simply mixing Pol II
346 purified from plants, purified TFIIIA-7ZF expressed in bacteria, and synthesized PSTVd
347 RNA templates (21,22,25). Here, we obtained a predicted structure of the transcription
348 complex and experimentally tested the structural model. Notably, TFIIIA-7ZF bends
349 around the left terminal part of PSTVd (+) RNA template to form an enlarged tertiary
350 structure (Fig. 3). This enlarged structure possibly contributes to the recruiting of Pol II.
351 This organization is different from previously reported hairpin-like RNA promoter used for
352 testing a chimeric HDV RNA template (56). The predicted model also suggests that the
353 viroid RNA promoters may not entirely rely on primary sequences. Instead, an overall
354 bended tertiary structure is critical for transcription. This model helps to explain that over
355 30 nuclear-replicating viroids with distinct sequences all use Pol II for replication (32).
356 When Pol II is recruited to the PSTVd/TFIIIA-7ZF complex, the C-terminus of TFIIIA-7ZF
357 somehow moves from RNA binding to the association with Rpb2. A C-terminus deletion
358 impairs TFIIIA-7ZF binding with the remodeled Pol II. The details of this structural
359 transition deserve future investigations.
360 In the Pol II-based in vitro transcription assay, factors that are needed for transcription
361 elongation in vivo (i.e., TFIIF, RTF1, PAF1, LEO1, CTR9, SPT6, SPT16, and SSRP1-B)
362 were all absent in the active remodeled Pol II on PSTVd RNA template (25). However,

363 whether any general transcription factors (in addition to TFIIIA-7ZF) or cofactors are
364 needed for PSTVd RNA-templated transcription in cells remains to be determined. We
365 also tested TFIIF and TFIIS in our RIP analyses. As shown in Fig. S2, there was no
366 evidence of a marked enrichment of PSTVd in the TFIIS immunoprecipitated fraction,
367 which is consistent with our previous data that TFIIS is dispensable for PSTVd replication
368 (21). Interestingly, we observed significant enrichment of PSTVd in TFIIF-
369 immunoprecipitated fraction (Fig. S2), suggesting that TFIIF could be involved in RNA-
370 templated transcription. It remains to be determined whether TFIIF is involved in the
371 initiation or elongation step in RNA-templated transcription. Nevertheless, this
372 observation also demonstrates that Pol II works with common and unique sets of general
373 transcription factors when transcribing DNA and RNA templates.
374

375 **Acknowledgments**
376 We are grateful for the technical support of Shachinthaka D. Dissanayaka
377 Mudiyanselage.
378

379 Author contributions: Jie Hao (Investigation, Methodology, Formal analysis, Validation,
380 Writing-review and editing), Zhuan Qin (Conceptualization, Visualization, Writing-review
381 and editing), Junfei Ma (Visualization, Writing-review and editing), Jie Qu (Investigation,
382 Methodology, Formal analysis, Validation), Yunhan Wang (Methodology, Writing-review
383 and editing), Svetlana Y. Folimonova (Resources, Writing-review and editing), Bin Liu
384 (Resources, Writing-review and editing), Wenwei Li (Conceptualization, Visualization,
385 Writing-review and editing), Ying Wang (Conceptualization, Funding acquisition, Project
386 administration, Resources, Supervision, Methodology, Formal analysis, Validation,
387 Writing-original draft, Writing-review and editing).
388

389 **Competing interests**
390 None declared.
391

392 **Funding**

This work was supported by US National Science Foundation MCB-2350392 (YW) and IOS-2410009 (YW), as well as US National Institutes of Health R01HL166169-01A1 (BL). Yunhan Wang is partially supported by CALS Deans' scholarship from University of Florida.

Data and materials availability

All data are available in the main text or the supplementary materials.

References

1. Spitz, F. and Furlong, E.E. (2012) Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet*, **13**, 613-626.
2. Levine, M. and Tjian, R. (2003) Transcription regulation and animal diversity. *Nature*, **424**, 147-151.
3. Crick, F.H. (1958) On protein synthesis. *Symp Soc Exp Biol*, **12**, 138-163.
4. Crick, F. (1970) Central dogma of molecular biology. *Nature*, **227**, 561-563.
5. Khatter, H., Vorlander, M.K. and Muller, C.W. (2017) RNA polymerase I and III: similar yet unique. *Curr Opin Struct Biol*, **47**, 88-94.
6. Thomas, M.C. and Chiang, C.M. (2006) The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol*, **41**, 105-178.
7. Ream, T.S., Haag, J.R., Wierzbicki, A.T., Nicora, C.D., Norbeck, A.D., Zhu, J.K., Hagen, G., Guilfoyle, T.J., Pasa-Tolic, L. and Pikaard, C.S. (2009) 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.
8. Dezelee, S., Sentenac, A. and Fromageot, P. (1974) Role of deoxyribonucleic acid-ribonucleic acid hybrids in eukaryotes. Synthetic ribo- and deoxyribopolynucleotides as template for yeast ribonucleic acid polymerase B (or II). *J Biol Chem*, **249**, 5978-5983.
9. Wassarman, K. and Saecker, R. (2006) Synthesis-mediated release of a small RNA inhibitor of RNA polymerase. *Science*, **314**, 1601-1603.
10. Wagner, S.D., Yakovchuk, P., Gilman, B., Ponicsan, S.L., Drullinger, L.F., Kugel, J.F. and Goodrich, J.A. (2013) RNA polymerase II acts as an RNA-dependent RNA polymerase to extend and destabilize a non-coding RNA. *EMBO J*, **32**, 781-790.
11. Flores, R., Owens, R.A. and Taylor, J. (2016) Pathogenesis by subviral agents: viroids and hepatitis delta virus. *Curr Opin Virol*, **17**, 87-94.
12. Roeder, R.G. (1998) Role of general and gene-specific cofactors in the regulation of eukaryotic transcription. *Cold Spring Harb Symp Quant Biol*, **63**, 201-218.
13. Lee, T.I. and Young, R.A. (2000) Transcription of eukaryotic protein-coding genes. *Annu Rev Genet*, **34**, 77-137.
14. Lemon, B. and Tjian, R. (2000) Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev*, **14**, 2551-2569.

15. Armache, K.J., Kettenberger, H. and Cramer, P. (2003) Architecture of initiation-competent 12-subunit RNA polymerase II. *Proc Natl Acad Sci U S A*, **100**, 6964-6968.
16. Liu, X., Bushnell, D.A. and Kornberg, R.D. (2013) RNA polymerase II transcription: structure and mechanism. *Biochim Biophys Acta*, **1829**, 2-8.
17. Bushnell, D.A., Bamdad, C. and Kornberg, R.D. (1996) A minimal set of RNA polymerase II transcription protein interactions. *J Biol Chem*, **271**, 20170-20174.
18. Schweikhard, V., Meng, C., Murakami, K., Kaplan, C.D., Kornberg, R.D. and Block, S.M. (2014) Transcription factors TFIIF and TFIIS promote transcript elongation by RNA polymerase II by synergistic and independent mechanisms. *Proc Natl Acad Sci U S A*, **111**, 6642-6647.
19. Vos, S.M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H. and Cramer, P. (2018) Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. *Nature*, **560**, 607-612.
20. Antosz, W., Pfab, A., Ehrnsberger, H.F., Holzinger, P., Kollen, K., Mortensen, S.A., Bruckmann, A., Schubert, T., Langst, G., Griesenbeck, J. *et al.* (2017) The Composition of the Arabidopsis RNA Polymerase II Transcript Elongation Complex Reveals the Interplay between Elongation and mRNA Processing Factors. *Plant Cell*, **29**, 854-870.
21. Dissanayaka Mudiyanselage, S.D. and Wang, Y. (2020) Evidence supporting that RNA polymerase II catalyzes de novo transcription using potato spindle tuber viroid circular RNA templates. *Viruses*, **12**, 371.
22. Wang, Y., Qu, J., Ji, S., Wallace, A.J., Wu, J., Li, Y., Gopalan, V. and Ding, B. (2016) A land plant-specific transcription factor directly enhances transcription of a pathogenic noncoding RNA template by DNA-dependent RNA polymerase II. *Plant Cell*, **28**, 1094-1107.
23. Rackwitz, H.R., Rohde, W. and Sanger, H.L. (1981) DNA-dependent RNA polymerase II of plant origin transcribes viroid RNA into full-length copies. *Nature*, **291**, 297-301.
24. Ma, J., Wang, Y., Folimonova, S.Y., Liu, B. and Wang, Y. (2024) Simultaneous observation of Pol II and potato spindle tuber viroid RNA in the nucleus by combining immunofluorescence and RNA in situ hybridization. *bioRxiv*, 2024.2009.2027.615539.
25. Dissanayaka Mudiyanselage, S.D., Ma, J., Pechan, T., Pechanova, O., Liu, B. and Wang, Y. (2022) A remodeled RNA polymerase II complex catalyzing viroid RNA-templated transcription. *PLoS Pathog.*, **18**, e1010850.
26. Yamaguchi, Y., Filipovska, J., Yano, K., Furuya, A., Inukai, N., Narita, T., Wada, T., Sugimoto, S., Konarska, M.M. and Handa, H. (2001) Stimulation of RNA polymerase II elongation by hepatitis delta antigen. *Science*, **293**, 124-127.
27. Lopez-Carrasco, A., Ballesteros, C., Sentandreu, V., Delgado, S., Gago-Zachert, S., Flores, R. and Sanjuan, R. (2017) Different rates of spontaneous mutation of chloroplastic and nuclear viroids as determined by high-fidelity ultra-deep sequencing. *PLoS Pathog*, **13**, e1006547.
28. Gout, J.F., Li, W., Fritsch, C., Li, A., Haroon, S., Singh, L., Hua, D., Fazelinia, H., Smith, Z., Seeholzer, S. *et al.* (2017) The landscape of transcription errors in eukaryotic cells. *Sci Adv*, **3**, e1701484.

29. Li, Y., Huang, J., Zhu, J., Bao, L., Wang, H., Jiang, Y., Tian, K., Wang, R., Zheng, H., Duan, W. *et al.* (2022) Targeted protein degradation reveals RNA Pol II heterogeneity and functional diversity. *Mol Cell*, **82**, 3943-3959 e3911.
30. Ma, J., Dissanayaka Mudiyanselage, S.D., Park, W.J., Wang, M., Takeda, R., Liu, B. and Wang, Y. (2022) A nuclear import pathway exploited by pathogenic noncoding RNAs. *Plant Cell*, **34**, 3543-3556.
31. Zhu, Y., Qi, Y., Xun, Y., Owens, R. and Ding, B. (2002) Movement of potato spindle tuber viroid reveals regulatory points of phloem-mediated RNA traffic. *PLANT PHYSIOLOGY*, **130**, 138-146.
32. Wang, Y. (2021) Current view and perspectives in viroid replication. *Curr. Opin. Virol.*, **47**, 32-37.
33. Ma, J., Dissanayaka Mudiyanselage, S.D., Hao, J. and Wang, Y. (2023) Cellular roadmaps of viroid infection. *Trends Microbiol*, **31**, 1179-1191.
34. Ream, T.S., Haag, J.R., Pontvianne, F., Nicora, C.D., Norbeck, A.D., Pasa-Tolic, L. and Pikaard, C.S. (2015) Subunit compositions of Arabidopsis RNA polymerases I and III reveal Pol I- and Pol III-specific forms of the AC40 subunit and alternative forms of the C53 subunit. *Nucleic Acids Res*, **43**, 4163-4178.
35. Plaschka, C., Lariviere, L., Wenzeck, L., Seizl, M., Hemann, M., Tegunov, D., Petrotchenko, E.V., Borchers, C.H., Baumeister, W., Herzog, F. *et al.* (2015) Architecture of the RNA polymerase II-Mediator core initiation complex. *Nature*, **518**, 376-380.
36. Robinson, P.J., Trnka, M.J., Bushnell, D.A., Davis, R.E., Mattei, P.J., Burlingame, A.L. and Kornberg, R.D. (2016) Structure of a Complete Mediator-RNA Polymerase II Pre-Initiation Complex. *Cell*, **166**, 1411-1422 e1416.
37. Ehara, H., Yokoyama, T., Shigematsu, H., Yokoyama, S., Shirouzu, M. and Sekine, S.I. (2017) Structure of the complete elongation complex of RNA polymerase II with basal factors. *Science*, **357**, 921-924.
38. Vos, S.M., Farnung, L., Urlaub, H. and Cramer, P. (2018) Structure of paused transcription complex Pol II-DSIF-NELF. *Nature*, **560**, 601-606.
39. Cramer, P., Bushnell, D.A. and Kornberg, R.D. (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science*, **292**, 1863-1876.
40. Dienemann, C., Schwalb, B., Schilbach, S. and Cramer, P. (2019) Promoter Distortion and Opening in the RNA Polymerase II Cleft. *Mol Cell*, **73**, 97-106 e104.
41. Kettenberger, H., Armache, K.J. and Cramer, P. (2004) Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol Cell*, **16**, 955-965.
42. Nolte, R.T., Conlin, R.M., Harrison, S.C. and Brown, R.S. (1998) Differing roles for zinc fingers in DNA recognition: structure of a six-finger transcription factor IIIA complex. *Proc Natl Acad Sci U S A*, **95**, 2938-2943.
43. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J. *et al.* (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, **630**, 493-500.
44. Dissanayaka Mudiyanselage, S.D., Qu, J., Tian, N., Jiang, J. and Wang, Y. (2018) Potato Spindle Tuber Viroid RNA-templated transcription: Factors and Regulation. *Viruses*, **10**.

45. Kolonko, N., Bannach, O., Aschermann, K., Hu, K.H., Moors, M., Schmitz, M., Steger, G. and Riesner, D. (2006) Transcription of potato spindle tuber viroid by RNA polymerase II starts in the left terminal loop. *Virology*, **347**, 392-404.
46. Lee, B.M., Xu, J., Clarkson, B.K., Martinez-Yamout, M.A., Dyson, H.J., Case, D.A., Gottesfeld, J.M. and Wright, P.E. (2006) Induced fit and "lock and key" recognition of 5S RNA by zinc fingers of transcription factor IIIA. *J Mol Biol*, **357**, 275-291.
47. Rothfels, K., Rowland, O. and Segall, J. (2007) Zinc fingers 1 and 7 of yeast TFIIIA are essential for assembly of a functional transcription complex on the 5 S RNA gene. *Nucleic Acids Res*, **35**, 4869-4881.
48. Wild, T. and Cramer, P. (2012) Biogenesis of multisubunit RNA polymerases. *Trends Biochem Sci*, **37**, 99-105.
49. Boulon, S., Pradet-Balade, B., Verheggen, C., Molle, D., Boireau, S., Georgieva, M., Azzag, K., Robert, M.C., Ahmad, Y., Neel, H. *et al.* (2010) HSP90 and its R2TP/Prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. *Mol Cell*, **39**, 912-924.
50. Gomez-Navarro, N., Peiro-Chova, L., Rodriguez-Navarro, S., Polaina, J. and Estruch, F. (2013) Rtp1p is a karyopherin-like protein required for RNA polymerase II biogenesis. *Mol Cell Biol*, **33**, 1756-1767.
51. Garg, G., Dienemann, C., Farnung, L., Schwarz, J., Linden, A., Urlaub, H. and Cramer, P. (2023) Structural insights into human co-transcriptional capping. *Mol Cell*, **83**, 2464-2477 e2465.
52. Li, Y., Wang, Q., Xu, Y. and Li, Z. (2024) Structures of co-transcriptional RNA capping enzymes on paused transcription complex. *Nat Commun*, **15**, 4622.
53. Ma, J., Mudiyanselage, S.D.D. and Wang, Y. (2022) Emerging value of the viroid model in molecular biology and beyond. *Virus Res.*, **313**, 198730.
54. Taylor, J.M. (2015) Hepatitis D Virus Replication. *Cold Spring Harb Perspect Med*, **5**.
55. Gudima, S., Wu, S.Y., Chiang, C.M., Moraleda, G. and Taylor, J. (2000) Origin of hepatitis delta virus mRNA. *J Virol*, **74**, 7204-7210.
56. Lehmann, E., Brueckner, F. and Cramer, P. (2007) Molecular basis of RNA-dependent RNA polymerase II activity. *Nature*, **450**, 445-449.

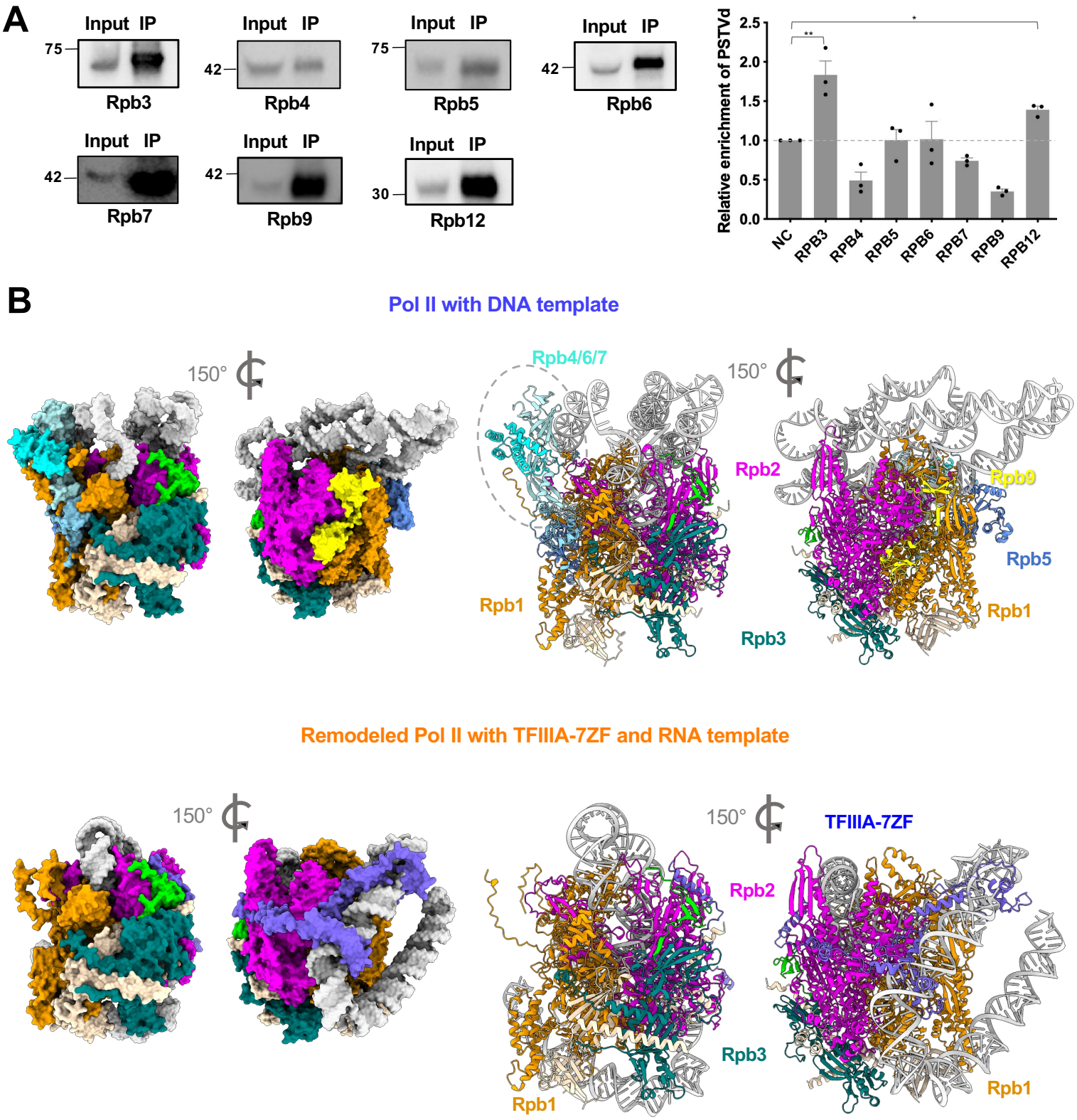

Figure 1. Remodeled Pol II in vivo and its predicted structure. A) RNA-immunoprecipitation using Pol II subunits. GFP-tagged Rpb3, Rpb4, Rpb5, Rpb6, Rpb7, Rpb9, and Rpb12 were subject to RNA immunoprecipitation to test their in vivo interaction with PSTVd RNA templates. Numbers indicate the molecular weight (Kda). RNAs from input and immunoprocipitated (IP) fraction were used for RT-qPCR analysis. 5.8S rRNA was used as negative control to normalize PSTVd enrichment. Data from there replicates were used for T-test analysis. B) Predicted structures of Pol II with DNA template and remodeled Pol II with TFIIIA-7ZF and RNA template. The catalytic subunits (Rpb1 and Rpb2) as well as the distinct subunits (Rpb4/5/6/7/9) and TFIIIA-7ZF are color coded. The stalk region is highlighted in a dash-line circle.

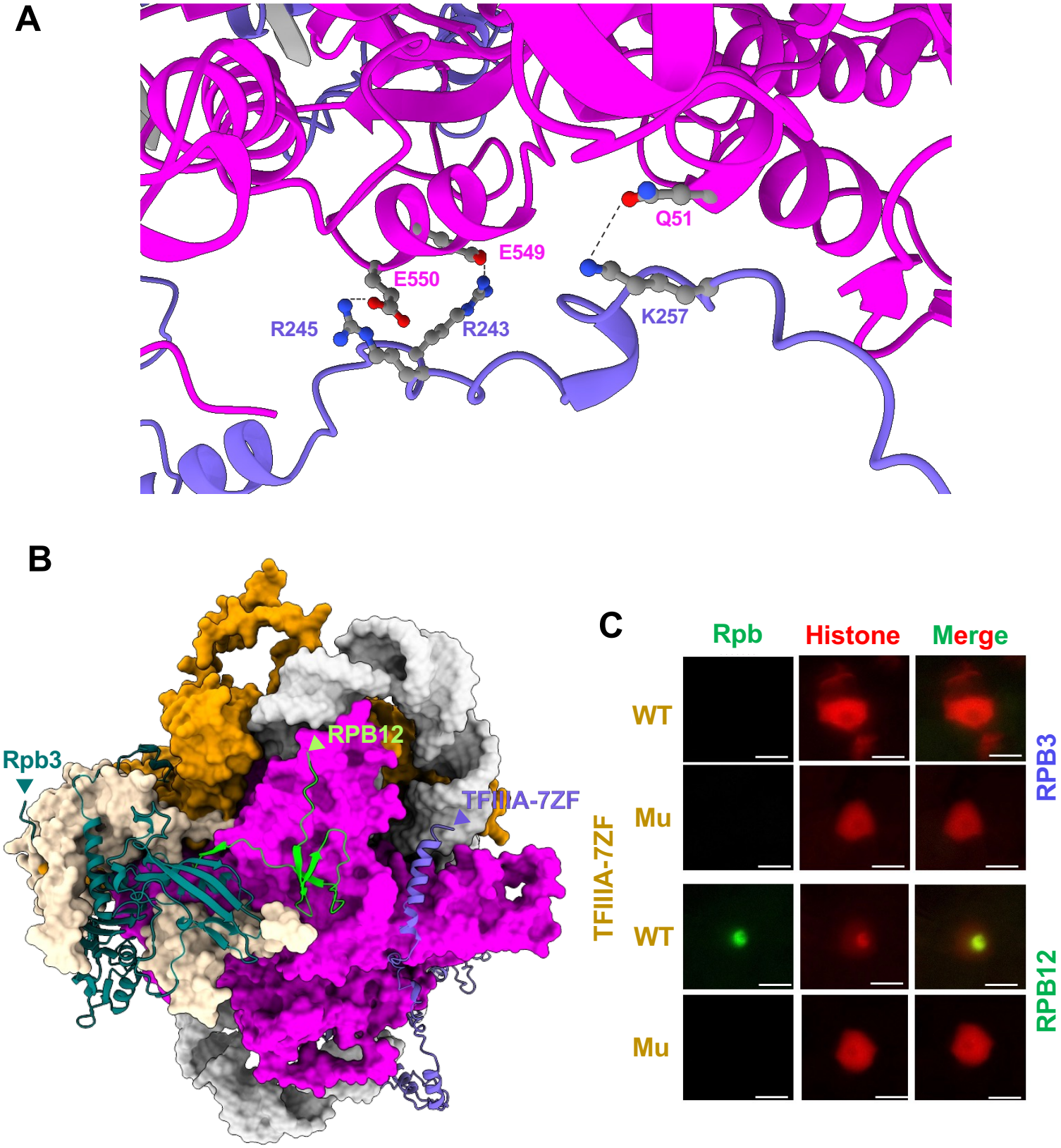

Figure 2. Relative position of TFIIIA-7ZF and Pol II subunits. A) Predicted interactions between Rpb2 and TFIIIA-7ZF. Rpb2 and TFIIIA-7ZF are depicted in magenta and blue, respectively. B) Predicted structure showing the relative position between the C-terminus of TFIIIA-7ZF (labeled by an arrow head) and the N-terminus of Rpb3 and Rpb12 (labeled by arrow heads). C) Bi-fluorescence complementation (BiFC) confirms that the C-terminus of TFIIIA-7ZF is close to the N-terminus of Rpb12 but not the N-terminus of Rpb3. Mu, TFIIIA-7ZF$^{C\text{-}del}$. Scale Bar, 10 μm.

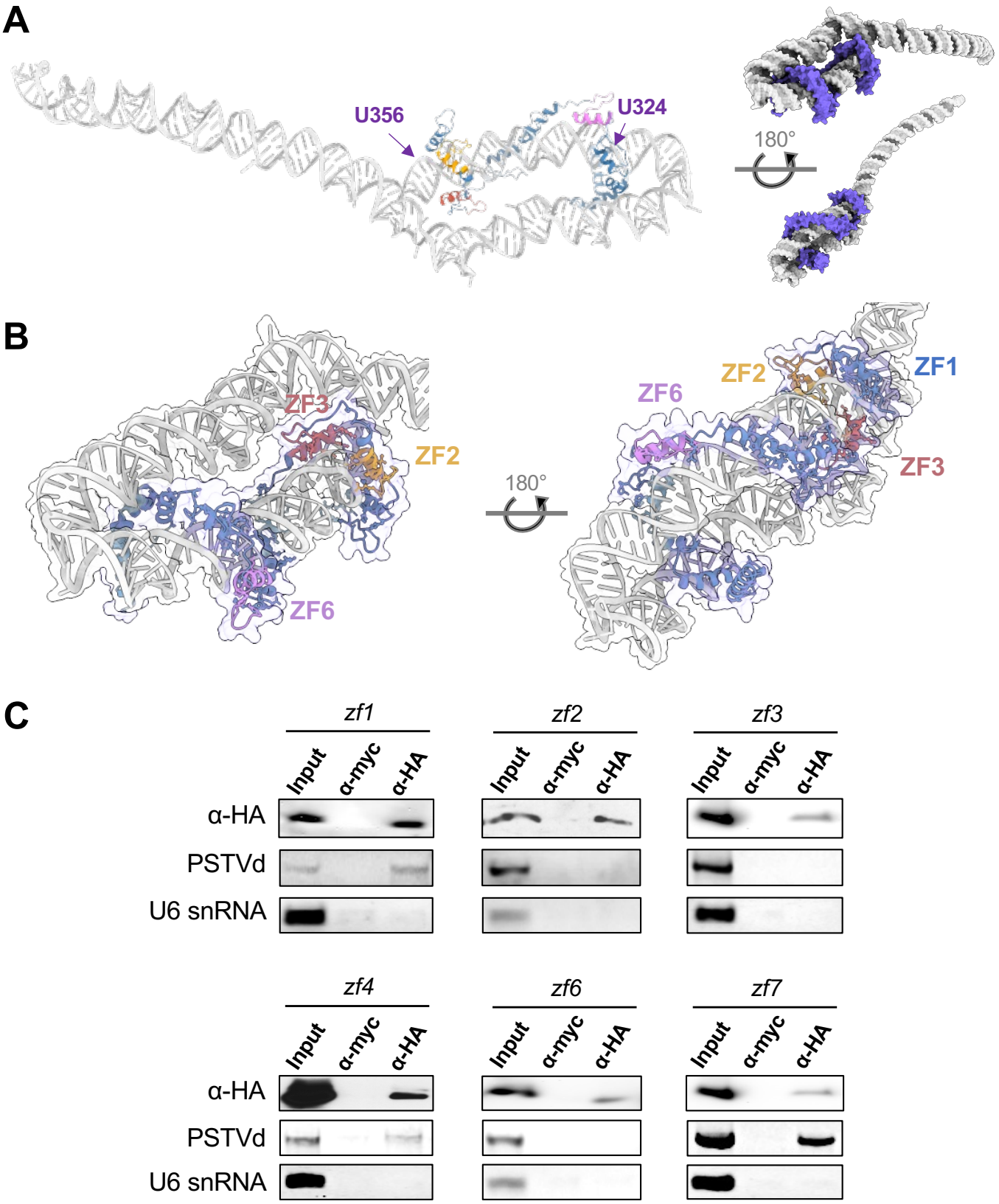

Figure 3. Structure of TFIIIA-7ZF and (+) PSTVd RNA. A) Predicted structure of TFIIIA-7ZF and PSTVd RNA. Different zinc finger domains are color highlighted. Regions between U356 and U324 are predicted to be covered by TFIIIA-7ZF. B) A zoom-in view of the critical zinc fingers for RNA binding. C) RNA-immunoprecipitation demonstrating the function of the zinc fingers. HA-tagged proteins were expressed in PSTVd-infected plants. α-HA and α-MYC were used for immunoprecipitation (IP) and control treatments, respectively. Semi quantitative RT-PCR were used for examining PSTVd enrichment in IP fractions, with U6 snRNA served as negative control. All experiments were repeated at least twice.

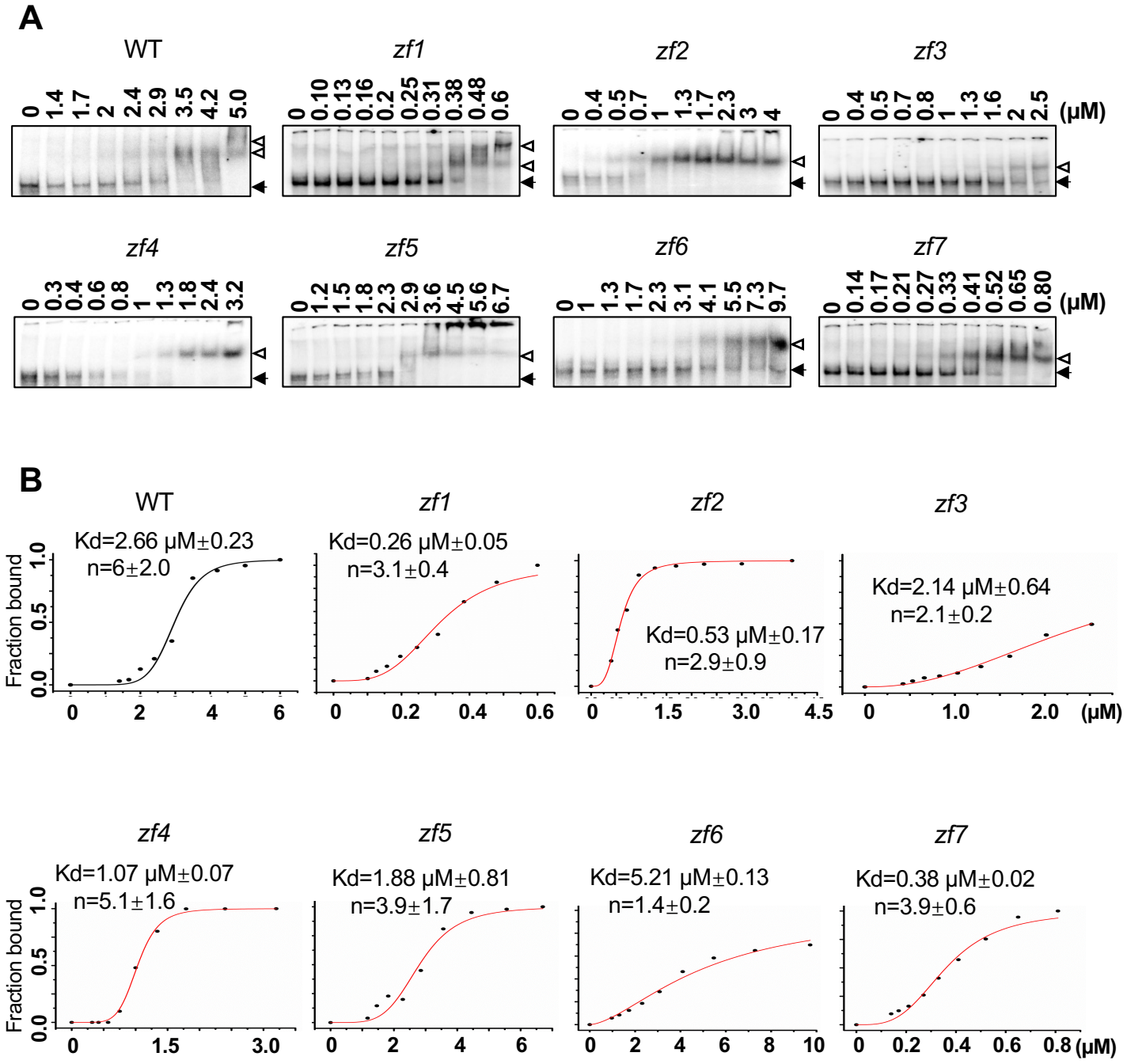

Figure 4. Interaction of TFIIIA-7ZF variants with PSTVd in EMSAs. A) Representative EMSA result for the binding between TFIIIA-7ZF zinc finger variants and PSTVd. Protein concentrations are listed for each lane. **B)** Representative binding curves for each TFIIIA-7ZF variants. Each data point was calculated based on three replicates. Protein concentrations are listed on X axes.