Cell PRESS

Functional Divergence in the Growing Family of RNA Polymerases

Robert Landick[1,*]
[1]Departments of Biochemistry and Bacteriology, University of Wisconsin-Madison, Henry Mall, Madison, WI 53706, USA
*Correspondence: landick@bact.wisc.edu
DOI 10.1016/j.str.2009.02.006

In plants, unorthodox multisubunit RNA polymerases (RNAPs) play key roles in small interfering RNA (siRNA) genesis and function. In a recent issue of *Molecular Cell*, Ream et al. (2009) established a 12-subunit composition for *Arabidopsis* RNAPIV and RNAPV. Subunit and sequence divergence between RNAPIV-V and RNAPI-III suggests significant functional deviation of these intriguing RNAPs.

The long-held view that eukaryotes possess just three multi-subunit RNAPs (RNAPI for ribosomal RNA; RNAPII for messenger RNA and some small noncoding RNAs; and RNAPIII for transfer RNA, 5S RNA, and small noncoding RNAs) began to crumble in the plant world in 2001 with the discovery of *Arabidopsis* genes encoding additional large RNAP subunits. Multi-subunit RNAPs, which are responsible for genomic transcription in all cellular organisms, possess a highly conserved DNA- and RNA-binding cleft and catalytic center formed by two large, paralogous subunits called β′ and β in bacterial RNAP. RNAPI, RNAPII, and RNAPIII each possess unique orthologs of β′ and β, as well as 10 or more additional subunits, some shared and some unique. Starting from the clue that unexpected RNAP subunit genes exist in the *Arabidopsis* genome, Pikaard and coworkers, along with several other groups, established that land plants and some algae produce two alternative, nuclear RNAPs: RNAPIV and RNAPV, both of which are required for siRNA-mediated gene silencing in plants. RNAPIV appears to transcribe loci that encode siRNA precursors (at least a thousand such loci exist in *Arabidopsis*; Mosher et al., 2008). After being processed by Dicer to 24-bp RNA duplexes, the siRNAs target Argonaute-RISC complexes to cognate chromosomal regions, where they nucleate methylation and gene silencing. Independently, transcription of these heterochromatic regions by RNAPV, and possibly RNAPIV, appears necessary to maintain and spread the silenced state, likely by providing RNA scaffolds to which siRNA-Argonaute complexes bind in order to recruit the silencing machinery (Daxinger et al., 2009; Wierzbicki et al., 2008).

Although the involvement of alternative RNAPs in plant gene silencing has been evident since 2005, the exact subunit composition of the enzymes has not. Plants produce two alternative β′ orthologs (NRPD1 and NRPE1), but at least some plants produce only a single β ortholog (NRPD2), making it unclear whether two functionally distinct enzymes exist or if a single alternative RNAP possessing variant β′ orthologs might be responsible for all RNAP functions in siRNA-mediated gene silencing (thus, RNAPV was until recently called RNAPIVb). Ream et al. (2009) now resolve these questions unambiguously with a *tour de force* application of proteomics to alternative RNAPs that are epitope-tagged on their β′ orthologs and immunoprecipitated from cell lysates. By comparing the mapped peptides to the *Arabidopsis* genome sequence, Ream et al. established that RNAPIV and RNAPV are different enzymes that possess distinct orthologs of β′(RPB1) and RBP7, but that share single orthologs of both β(RPB2) and RPB4. RNAPIV and V possess all 12 of the canonical subunits present in RNAPI-III, but share RBP3, RBP9, and RBP11 with RNAPII and share RPB6, RBP8, RBP10, and RBP12 with RNAPI-III (Figure 1). RNAPV uses a unique RPB5 ortholog, whereas RNAPIV shares RPB5 with RNAPI-III. A subpopulation of RNAPV appears to use a unique RPB3 ortholog. Thus, RNAPIV and RNAPV possess the same basic 12-subunit structure as RNAPII, and like the better-known RNAPI-III, use a combination of shared and unique subunits. All 12 subunits may be required for activity, as RNAPV lacking subunits 4, 6, and 7 isolated from cauliflower was found to be inactive in vitro (Huang et al., 2009).

The apparent existence of two additional bona fide, multi-subunit RNAPs, possibly adapted for transcription of heterochromatin, raises the interesting question of the extent to which their clefts and catalytic centers may have diverged from RNAPI-III. Curiously, the RNAPIV-V orthologs of β′, which makes the most extensive nucleic acid and catalytic center contacts, are more divergent than the β orthologs, not only in being specific for RNAPIV or RNAPV but also in exhibiting amino acid substitution rates among plant orthologs 20 times that of the RNAPII subunit RPB1 and twice that of the RNAPIV-V β orthologs (Luo and Hall, 2007). Remarkably, for residues conserved among RNAPI-III, the corresponding positions in RNAPIV-V diverge more near the cleft and catalytic center than they do farther away (Haag et al., 2009). The divergence of the catalytic center is striking. Multi-subunit RNAPs catalyze nucleotidyl transfer using two Mg^{2+} ions bound to a highly conserved Asp triad in an NADFDGD loop that is invariant in all previously known multi-subunit RNAPs. Rapid nucleotide addition (i.e., at biologically relevant rates faster than $\sim$0.1 s^{-1}) requires participation of the trigger loop, a polymorphous sequence that alternates between a helical hairpin and a partially unfolded loop (Vassylyev et al., 2007; Wang et al., 2006). The helical hairpin form of the trigger loop is stabilized by interactions with the bridge helix and directly contacts the nucleoside triphosphate substrate to

 ©2009 Elsevier Ltd All rights reserved

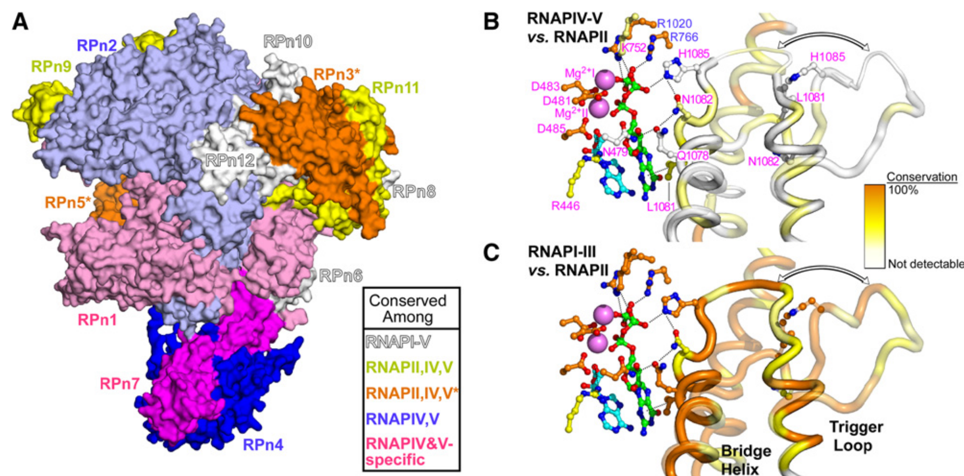

Figure 1. Subunit Sharing and Sequence Conservation in RNAPIV and RNAPV

(A) Shared subunits in *Arabidopsis* RNAPIV and RNAPV mapped onto the crystal structure of *S. cerevisiae* RNAPII (PDB ID 1WCM; Armache et al., 2005). White area: Subunits shared among all *Arabidopsis* RNAPs. Yellow and orange areas: Subunits shared among RNAPII, RNAPIV, and RNAPV (RNAPV contains a unique RPB5 ortholog; subpopulations of RNAPV appear to contain either RPB3 or an RNAPV-specific RPB3 ortholog). Dark and light blue areas: Subunits shared between RNAPIV and RNAPV. Dark and light magenta areas: subunits specific for RNAPIV and RNAPV.

(B) Conservation of the RNAPIV-V catalytic center. Shown are RPB1 residues (magenta) or RPB2 residues (blue) proposed to contact the catalytic Mg^{2+} ions (violet), the 3′ RNA nucleotide (cyan), or substrate GTP (green) in the crystal structure of an elongating complex of *S. cerevisiae* RNAPII with a helical hairpin form of the trigger loop (PBD ID 2E2H; Wang et al., 2006). The unfolded trigger loop conformation (PDB ID 2NVQ; Wang et al., 2006) is superimposed; the unfolded trigger loop and bridge helix are rendered semitransparent. Conservation of residues was calculated as the blosum62 information score in a comparison of six RPB1 sequences from plants and yeasts and six RPD1 and RPE1 sequences from plants, or of four RPB2 and four RPD2 sequences. Scores less than 0.625 were equated with undetectable conservation; scores of 1 were equated with 100% conservation. Scores were converted to a white-to-orange heat map of increasing conservation and mapped onto the RNAP catalytic center residues. Residues absent altogether in RNAPIV or V are depicted as thinner tubes. Details and alignments are available from the author.

(C) Conservation of the RNAPI-III catalytic center. The view is identical to that in (B), with conservation scores calculated using RPA1, RPB1, and RPC1 or RPA2, RPB2, and RPC2 as described for (B).

position it for catalysis (Figure 1). Interestingly, the Asp triad itself (DFDGD) is conserved in RNAPIV-V, and substitution of the Asp residues with Ala in RNAPIV or RNAPV disrupts both siRNA-mediated gene silencing and most transcription within silenced regions (Wierzbicki et al., 2008; Haag et al., 2009). However, other key residues in the trigger loop, bridge helix, and catalytic center that contact the reactive nucleotides and that are conserved among RNAPI-III reveal little if any detectable similarity in RNAPIV-V (Figure 1). Indeed the flexible tip of the trigger loop and a portion of the bridge helix appear missing altogether. The striking divergence of the RNAPIV-V cleft and catalytic center suggests that these enzymes may be adapted to an alternative form of DNA template and may either catalyze a reaction other than rapid nucleotide addition or use a different mechanism to achieve rapid nucleotide addition.

At least three mechanistic requirements could explain the preferential divergence of the cleft and catalytic center of RNAPIV-V. First, it is possible that transcription of heterochromatic DNA, which is both heavily methylated and tightly packed in nucleosomal arrays, places atypical demands on the coupling of the nucleotidyl transfer reaction to the movement of DNA strands through the enzyme. Even mononucleosomes pose significant barriers to RNAPII transcription. It seems probable that the more tightly packed, higher-order chromatin structures in heterochromatin could create special challenges for an RNAP. This hypothesis is attractive because it could resolve the paradox that active transcription of heterochromatin is required for maintenance of its silenced state. By providing enzymes specific for a DNA template with altered structural characteristics, RNAPIV-V would allow this structure to become refractory to RNAPI-III transcription yet maintain an ability to be transcribed. A second possibility is that RNAPIV-V diverged to accommodate use of either RNA or ssDNA as templates in addition to the natural double-stranded template. The complete pathway of siRNA-mediated gene silencing remains incompletely defined, and use of such alternate templates could be important either in the contemporary pathway or at some point in the evolution of the mechanism in land plants. Finally, it remains at least theoretically possible that RNAPIV-V are not polymerases at all, but rather play some role in the processing of RNA as unconventional nucleases. Although the requirement for RNAPIV in the synthesis of siRNA precursor transcripts and of RNAPV in the generation of transcripts from silenced loci are unambiguous (Wierzbicki et al., 2008; Haag et al., 2009), RNA synthesis by purified RNAPIV-V using duplex DNA templates in vitro remains to be demonstrated. RNAPs are known to possess hydrolytic endonuclease and exonuclease activities that utilize the same catalytic center as the nucleotidyl transfer reaction (Sosunov et al., 2003). Thus, an alternative role—possibly in RNA processing—cannot yet be excluded.

Distinguishing among the possibilities and defining the roles of RNAPIV-V in siRNA-mediated silencing will require their purification in active form and their study using in vitro assays. It is a testament to the power of genetics that the roles of these interesting enzymes could be brought so far along without the complement of in vitro transcription. Such biochemical approaches have

 ©2009 Elsevier Ltd All rights reserved

Cell PRESS

been essential to defining the mechanisms of RNAPI-III. In vitro studies of RNAPIV-V should provide fundamental information on their rates of transcription and on the mechanism by which they deal with modified DNA structures. This biochemical information will be invaluable in piecing together the puzzle of their role in siRNA-mediated gene silencing.

Even then, the mystery of how RNAPII apparently substitutes for the function of RNAPIV and RNAPV in metazoans may remain. Abundant evidence suggests noncoding transcription of heterochromatin by RNAPII plays a role in establishment and maintenance of siRNA-mediated silencing in metazoans, but how RNAPII meets the mechanistic requirements evident in the divergence of the RNAPIV-V cleft and catalytic center remains unclear. RNAPII is capable of remarkable plasticity; it can be ubiquitinated within its main cleft even when bound to DNA, and it can use RNA as a template when bound by hepatitis delta antigen. Perhaps other modifications alter the properties of RNAPII in ways that mimic the diverged sequences of RNAPIV and RNAPV in plants and allow alternative RNAPII subpopulations to substitute for their functions in other eukaryotes.

REFERENCES

Armache, K.J., Mitterweger, S., Meinhart, A., and Cramer, P. (2005). J. Biol. Chem. *280*, 7131–7134.

Daxinger, L., Kanno, T., Bucher, E., van der Winden, J., Naumann, U., Matzke, A.J., and Matzke, M. (2009). EMBO J. *28*, 48–57.

Haag, J.R., Pontes, O., and Pikaard, C.S. (2009). PLoS ONE *4*, e4110.

Huang, L., Jones, A.M., Searle, I., Patel, K., Vogler, H., Hubner, N.C., and Baulcombe, D.C. (2009). Nat. Struct. Mol. Biol. *16*, 91–93.

Luo, J., and Hall, B.D. (2007). J. Mol. Evol. *64*, 101–112.

Mosher, R.A., Schwach, F., Studholme, D., and Baulcombe, D.C. (2008). Proc. Natl. Acad. Sci. USA *105*, 3145–3150.

Ream, T.S., Haag, J.R., Wierzbicki, A.T., Nicora, C.D., Norbeck, A., Zhu, J.K., Hagen, G., Guilfoyle, T.J., Pasa-Tolic, L., and Pikaard, C.S. (2009). Mol. Cell *33*, 192–203.

Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V., and Goldfarb, A. (2003). EMBO J. *22*, 2234–2244.

Vassylyev, D.G., Vassylyeva, M., Zhang, J., Palangat, M., Artsimovitch, I., and Landick, R. (2007). Nature *448*, 163–168.

Wang, D., Bushnell, D., Westover, K., Kaplan, C., and Kornberg, R. (2006). Cell *127*, 941–954.

Wierzbicki, A.T., Haag, J.R., and Pikaard, C.S. (2008). Cell *135*, 635–648.