

RESEARCH

Open Access



# The plant organellar primase-helicase directs template recognition and primosome assembly via its zinc finger domain

Antolin Peralta-Castro<sup>1</sup>, Francisco Cordoba-Andrade<sup>1</sup>, Corina Díaz-Quezada<sup>1</sup>, Rogério Sotelo-Mundo<sup>2</sup>, Robert Winkler<sup>1</sup> and Luis G. Brieba<sup>1\*</sup>

## Abstract

**Background** The mechanisms and regulation for DNA replication in plant organelles are largely unknown, as few proteins involved in replisome assembly have been biochemically studied. A primase-helicase dubbed Twinkle (T7 gp4-like protein with intramitochondrial nucleoid localization) unwinds double-stranded DNA in metazoan mitochondria and plant organelles. Twinkle in plants is a bifunctional enzyme with an active primase module. This contrast with animal Twinkle in which the primase module is inactive. The organellar primase-helicase of *Arabidopsis thaliana* (AtTwinkle) harbors a primase module (AtPrimase) that consists of an RNA polymerase domain (RPD) and a Zn++ finger domain (ZFD).

**Results** Herein, we investigate the mechanisms by which AtTwinkle recognizes its templating sequence and how primer synthesis and coupling to the organellar DNA polymerases occurs. Biochemical data show that the ZFD of the AtPrimase module is responsible for template recognition, and this recognition is achieved by residues N163, R166, and K168. The role of the ZFD in template recognition was also corroborated by swapping the RPDs of bacteriophage T7 primase and AtPrimase with their respective ZFDs. A chimeric primase harboring the ZFD of T7 primase and the RPD of AtPrimase synthesizes ribonucleotides from the T7 primase recognition sequence and conversely, a chimeric primase harboring the ZFD of AtPrimase and the RPD of T7 primase synthesizes ribonucleotides from the AtPrimase recognition sequence. A chimera harboring the RPDs of bacteriophage T7 and the ZBD of AtTwinkle efficiently synthesizes primers for the plant organellar DNA polymerase.

**Conclusions** We conclude that the ZFD is responsible for recognizing a single-stranded sequence and for primer hand-off into the organellar DNA polymerases active site. The primase activity of plant Twinkle is consistent with phylogeny-based reconstructions that concluded that Twinkle's last eukaryotic common ancestor (LECA) was an enzyme with primase and helicase activities. In plants, the primase domain is active, whereas the primase activity was lost in metazoans. Our data supports the notion that AtTwinkle synthesizes primers at the lagging-strand of the organellar replication fork.

**Keywords** RNA polymerase, Primase, Organelle, Flowering plants

\*Correspondence:

Luis G. Brieba

luis.brieba@cinvestav.mx

<sup>1</sup>Langebio-Cinvestav Sede Irapuato, Km. 9.6 Libramiento Norte Carretera.

Irapuato-León, 36821 Irapuato Guanajuato, Mexico

<sup>2</sup>Laboratorio de Estructura Biomolecular, Centro de Investigación en

Alimentación y Desarrollo, A.C. Carretera Gustavo Enrique Astiazarán

Rosas Núm. 46, Ejido a La Victoria, 83304 Hermosillo, Sonora, Mexico



## Background

A wealth of proteins involved in DNA transactions assemble as modular proteins with domains specialized in DNA binding, catalysis, or protein-protein interactions [1, 2]. Archetypical examples of modular proteins are the bifunctional primase-helicases. These proteins replicate plasmids, viral, and organellar genomes [3, 4]. Bifunctional primase-helicases are oligomeric proteins that assemble as hexamers or heptamers [5–8]. The primase-helicase from bacteriophage T7 is the representative member of this protein family [4]. In hexameric primase-helicases, the helicase domain unwinds double-stranded DNA using nucleotide hydrolysis while the primase domain recognizes a templated sequence to drive RNA synthesis. Primer synthesis by primase-helicases is achieved by the combination of their Zn<sup>++</sup> finger (ZFD) and RNA polymerase (RPD) domains of the primase module [9, 10]. DNA primases recognize a trinucleotide DNA sequence in which the 3' base is cryptic, meaning that this base is necessary for template recognition but is not instructional for the identity of the primer [6]. Primases recognize multiple template sequences. For instance, phylogenetically related primase-helicases from bacteriophages T7 and SP6 recognize the 5'-GTC-3' and 5'-GCA-3' sequences, respectively, where the underlined nucleotide is cryptic, as mentioned before meaning that this base is necessary for recognition but its complementary base is not copied into the RNA primer [6].

The mechanisms for DNA replication in plant mitochondria remains to be elucidated. Several groups debate over a mechanism involving the role of homologous recombination at DNA breaks as starting points for DNA replication and the formation of a T7-like replisome that generates primers for the plant mitochondrial DNAPs [11–14]. Plant mitochondria harbor enzymes capable to execute DNA replication by a T7-like replisome and by homologous recombination [11, 12, 14, 15]. In all eukaryotes, with the exception of metazoans, Twinkle (a T7 gp4-like protein with intramitochondrial nucleoid localization) is predicted to be active as a primase and helicase [16]. During lagging strand DNA synthesis, DNA primases perform multiple functions: (1) recognition of a ssDNA sequence, (2) *de novo* RNA synthesis, and (3) primer transfer to a DNAP. T7 primase-helicase, AtTwinkle, and *Dictyostelium discoideum* (DdTwinkle) are the only characterized enzymes from the primase-helicase family [17–19]. (Fig. 1A). The primase module consists of a ZFD and an RPD connected by a flexible amino acid linker. It is proposed that the ZFD fulfills roles in template recognition, primer stability, and primer transfer [5, 18, 20, 21]. The Twinkle ortholog of the plant model *Arabidopsis thaliana* (AtTwinkle) is active, both as a primase and as a helicase [13, 22, 23]. Contrary to all primases characterized to date, the primase module of AtTwinkle

recognizes two nucleotides (instead of one) as cryptic elements. The ssDNA recognition sequence of AtTwinkle is 3'-AGGG/C-5', in which the 3'-AG-5' pair is cryptic [13]. The ZFD of AtTwinkle displays unique amino acids that possibly interact with the ssDNA template sequence, suggesting that the recognition of two bases as cryptic elements is mediated by its ZFD [13]. Biochemical analysis indicates that template specificity is mediated by the RNAP and ZFD in RNA primases [20]. The “*cis*” model for *de novo* synthesis and ribonucleotide extension postulates that the Zn<sup>++</sup> finger recognizes the cryptic and the first two template nucleotides, directing the synthesis of the first two ribonucleotides [18]. Thus, in T7 Primase the ZFD is important for template recognition, transfer of the initial diribonucleotide from the RPD, and hand-off of the RNA primer to the T7 DNAP polymerase active site [15, 28–30]. In support of these biochemical studies, the structure of the T7 primosome illustrates that the ZFD of the T7 Primase domain intercalates between the fingers and thumb subdomains of T7 DNAP and that the RPD mainly interacts with the thumb subdomain [24].

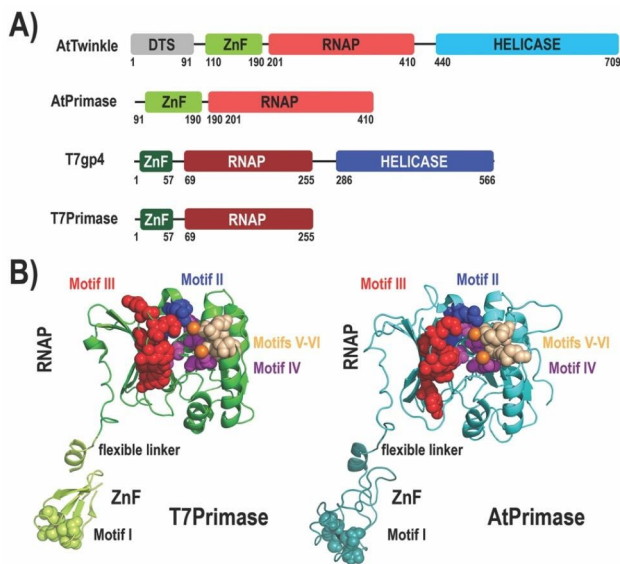
To study the possible role of the ZFD in template recognition, we performed a functional analysis of the amino acids that are possibly involved in template recognition and used swapping subdomain experiments between the Zn<sup>++</sup> fingers of T7 and Arabidopsis primases. Our results are congruent with the existence of a T7-like replisome in plant mitochondria that directs mitochondrial DNA replication and the formation of a T7 primosome to initiate mitochondrial DNA synthesis [12].

## Results

### AtPrimase is a modular protein poised to interact with ssDNA

The RNAP and ZFDs subdomains of T7 primase and AtTwinkle share 26% and 18% amino acid identity, respectively. A structural model of AtPrimase in comparison to T7 primase highlights the conservation of the active site residues in Motifs I to VI (Fig. 1B and Fig. 1S).

The main difference in primary sequence between the ZFDs of AtTwinkle and T7 primase is the insertion of eight amino acids between the second and third cysteines that coordinate a Zn<sup>++</sup> atom in the ZFD of AtTwinkle (Fig. 2A). According to several structural function studies, the T7 primase residues D31, S27, F29, H33, K41, and a group of aromatic amino acids (F29, F35, Y37, and W42) interact with the ssDNA template [21, 25–27]. AtPrimase does not retain this group of amino acids (Fig. 2A). Additionally, AtPrimase contains a number of residues close to the third and fourth Zn<sup>++</sup> finger cysteines that are not conserved with T7 primase, and because of their proximity to amino acids that interact with ssDNA, we hypothesize that these residues have the ability to interact with ssDNA and play a role in ssDNA



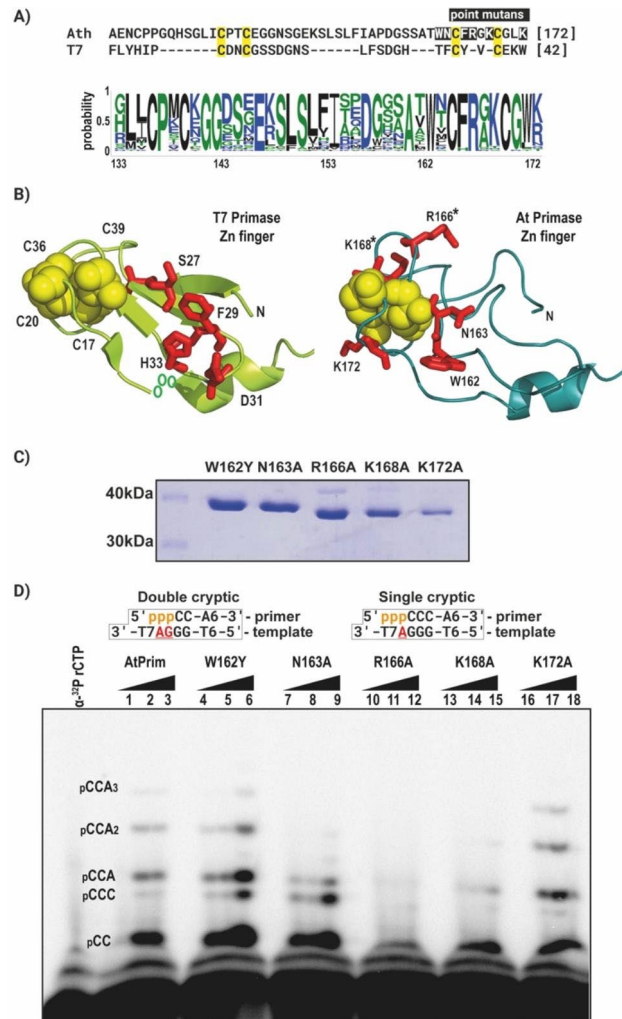
**Fig. 1** Domain and structural representation of T7 Primase and AtPrimase. **(A)** Domain organization of the hexameric primase-helicases of bacteriophage T7 and *A. thaliana*'s mitochondria. The boundaries of the helicase domains (blue and cyan blocks), primase (green blocks for the ZFD and red blocks for the RNAP domain, respectively) are indicated by numbers. AtTwinkle contains a signal peptide for mitochondrial and chloroplast targeting (grey block) at its N-terminus. **(B)** Crystal structure T7 Primase in comparison to a structural model of AtPrimase. Motifs II to VI that correspond to the RNAP domain are in ball-stick representation. The region corresponding to the Zn<sup>2+</sup> finger is the one that shows the greatest amino acid sequence divergence between primases

recognition. Those residues are W162, N163, R166, K168, and K172 (Fig. 2B and Fig. 1S).

To investigate this hypothesis, we mutated residue W162 for Try and the other residues for alanine (Fig. 2C) and evaluated their ribonucleotide synthesis on a canonical ssDNA template (Fig. 2D). AtPrimase-W162Y and AtPrimase-K172A synthesize RNA products that exhibited a similar pattern to the products synthesized by wild-type AtPrimase (Fig. 2D). Mutant AtPrimase-N163A, increases the synthesis of the pCCC ribonucleotide and AtPrimase-K168A strongly decreases ribonucleotide incorporation, whereas the AtPrimase-R166A mutant is largely inactive (Fig. 2D).

### Residues R166 and K168 are directly involved in recognition of the cryptic pair

To determine the full extent of the interactions between the point mutants and the cryptic elements of the recognition site in AtPrimase, we tested the entire landscape of variations at the cryptic site. In this experiment, the cryptic 5'-GA-3' site was systematically changed for every nucleotide, while the G that templates for 5'rCTP was kept constant. AtPrimase mutants W162Y and K172A exhibited a similar pattern of RNA products as wild-type AtPrimase on all tested templates, although



**Fig. 2** The ZFD is involved in recognition of the cryptic pair within the single-stranded recognition sequence. **A)** Amino acid sequence alignment of the region corresponding to the ZFD of AtPrimase and T7 Primase. The cysteines that coordinate the zinc atom are highlighted in yellow and the residues that were mutated in AtPrimase are highlighted in black. The amino acid sequence logo corresponds to an alignment of the ZFD of plants. **B)** Structural comparison of the analyzed residues in the ZFD model of AtPrimase (red colored) and the structure of the ZFD of T7 primase (green colored). **C)** SDS-PAGE gel showing purified ZFD mutant proteins. **D)** Denaturing sequencing gel showing ribonucleotide synthesis by wild-type AtPrimase and the analyzed point mutations on a canonical 5'-T<sub>6</sub>GGGAT<sub>7</sub>-3' substrate, the underlined GA corresponds to the cryptic element of AtPrimase. This substrate drives the synthesis of a 5'-pppCC dimer than can be extended by the incorporation of six adenines to the template thymines. Wild-type AtPrimase (lanes 1 to 3) synthesizes the expected RNA products from a canonical template, while the K166A (lanes 10 to 12) and K168A (lanes 13 to 15) mutants display no activity on this substrate. The K172A mutant (lanes 16 to 18) shows a similar ribonucleotide synthesis with respect to the wild type. Mutant W162Y (lanes 4 to 6) and N163A (lanes 7 to 9) show an increase in the formation of the pCCC triribonucleotide. This product corresponding to the recognition of only one cryptic base

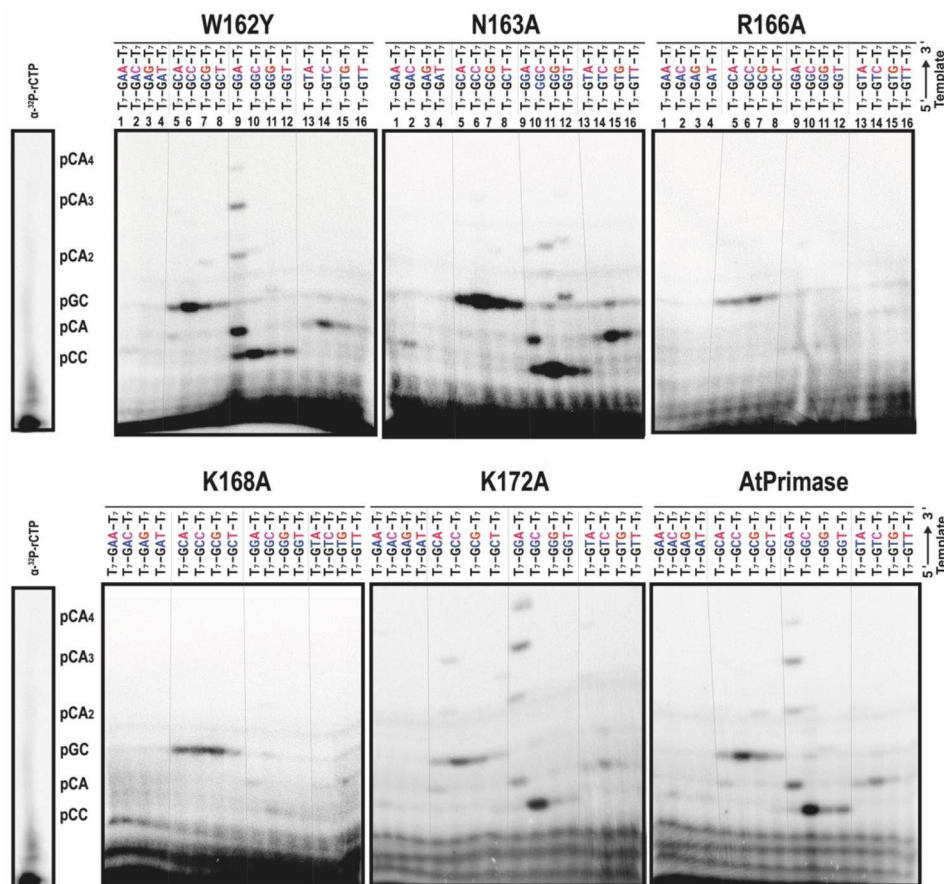
AtPrimase-W162Y exhibited a slight increase in formation of the pCC diribonucleotide on a canonical recognition sequence (Fig. 3).

In contrast, mutants R166A and K168A were unable to incorporate ribonucleotides on most of these templates. These mutants synthesize primers only in templates harboring the 5'-GCN-3' sequence, in which N represents any of the 4 bases. In this sequence these mutants synthesize a pGX diribonucleotide. However, these mutants were unable to extend the initial synthesized diribonucleotide (Fig. 3). Wild-type AtPrimase and the other point mutants synthesize a pGN diribonucleotide on the same templates, however they were able to extend them. Moreover, mutant N163A showed an increased diribonucleotide synthesis while ribonucleotide synthesis on the canonical template was severely hindered. The latter suggests that residues N163, R166, and K168 are involved in recognizing the cryptic element of the template recognition sequence 5'-(G/C)CGA-3'.

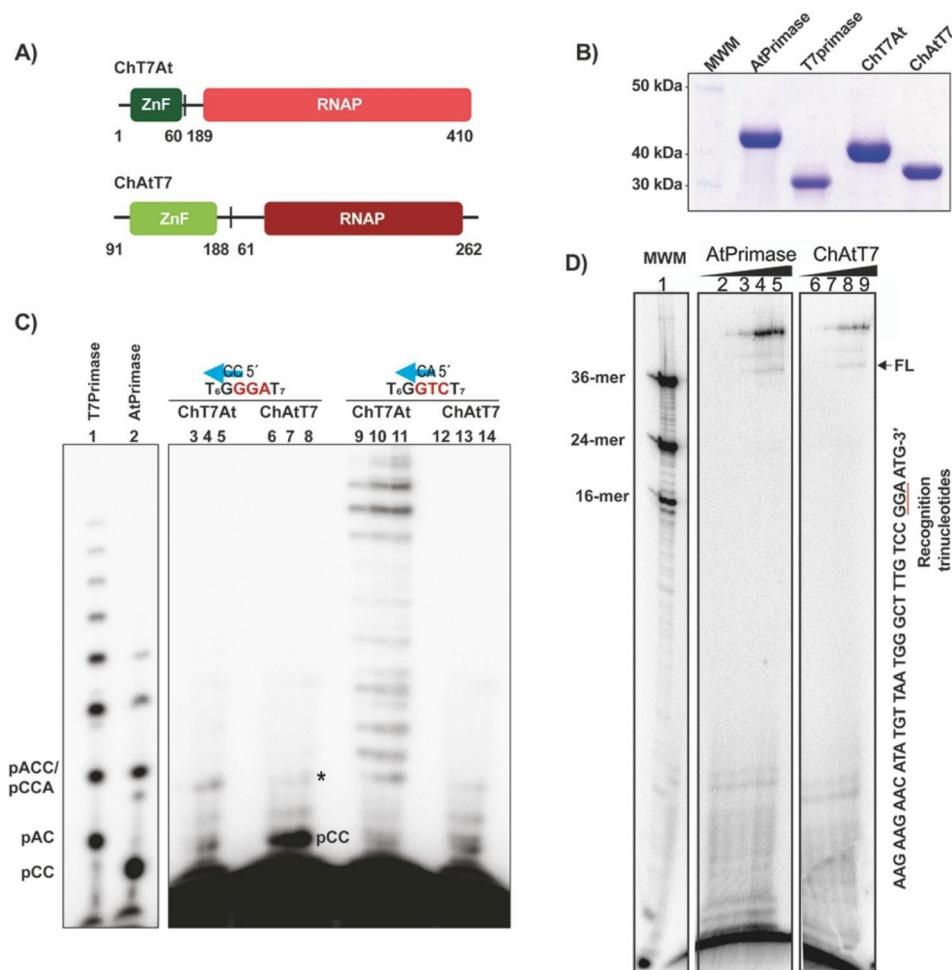
### A chimeric Primase harboring the ZFD of AtPrimase and the RNAPD of T7 Primase is active on the T7 primase recognition sequence

To test the hypothesis that the ZFD of AtPrimase drives template recognition, we constructed two chimeric Primases, one harboring the ZFD of the T7 Primase and the RNAP domain of AtPrimase (ChT7At) and another construct harboring the ZFD of AtPrimase and the RNAP domain of T7 Primase (ChAtT7) (Fig. 4A).

Both chimeric proteins were purified with similar yields and purity than AtPrimase or T7 Primase (Fig. 4B). The ChT7At chimera synthesized pAC and pCCA RNA products on the AtPrimase recognition sequence but failed to extend these products to longer RNA primers (Fig. 4C, lanes 3 to 5). In contrast, the ChT7At chimera was fully active on a T7 recognition sequence synthesizing ribonucleotides up to 11 nts (Fig. 4C, lanes 9 to 11). The ChAtT7 chimera was largely inactive on the T7 recognition sequence and only few pCC products were synthesized (Fig. 4C, lanes 12 to 14). In contrast, the ChAtT7 chimera synthesized an abundant amount of the pCC dimer on the AtPrimase recognition sequence and



**Fig. 3** Evaluation of the cryptic pair with point mutants of AtPrimase. Oligoribonucleotide synthesis reactions on substrates in which the cryptic pair was systematically modified to encompass all 16 possibilities. Reaction contained individual oligonucleotides harboring the 5'-T<sub>6</sub>GGNNT<sub>7</sub>-3' sequence. [ $\alpha$ -<sup>32</sup>P]-rCTP is used as radioactive label, because of the marked preference of AtPrimase for CTP as 5' ribonucleotide



**Fig. 4** Oligoribonucleotide synthesis by chimeric primases of *A. thaliana* mitochondria and bacteriophage T7. **A)** Representative diagram of the regions used to construct the chimeras between the ZFDs of T7 Primase and AtPrimase. The green blocks represent the ZFDs and the red blocks the RNAP domains. The numbers indicate the amino acids of the wild-type enzymes that were used to construct the chimeras by swapping the ZFD and RNAP domains. **B)** SDS-PAGE gel showing purified chimeric primases. **C)** Denaturing sequencing gel showing oligoribonucleotide synthesis by wild-type AtPrimase and T7 Primase on their canonical templates (lane 1 and 2, respectively), chT7At chimera (lanes 3 to 5 and 9 to 11) and chAtT7 chimera (lanes 6 to 8 and 12 to 14). The identity of the canonical substrates for AtPrimase and T7 Primase is 5'-T6GGGAT7-3' substrate and 5'-T6GGTCT7-3' substrate, respectively. The cryptic elements are underlined. The chT7At chimera efficiently synthesizes primers and harbors an increased substrate-independent oligoribonucleotide synthesis activity. This chimera was only able to synthesize primers on the substrate containing the bacteriophage T7-primase recognition trinucleotide. chAtT7 was extremely inefficient, synthesizing a product of maximum size of 4 nucleotides. Primase chimeras were tested at 100, 200 and 400 nM. The ssDNA substrate was used at 5 μM. **D)** Wild-type and chimeric AtPrimases prime organellar DNA polymerase (AtPollB). Coupled AtPrimase-DNA polymerase reactions on a canonical AtPrimase ssDNA recognition sequence labeled with [α-<sup>32</sup>P]-dATP. AtPollB, generates DNA:RNA products that correspond to products longer than 36 nts in the presence of increasing concentrations of AtPrimase and ChAtT7 chimera (lanes 2 to 5 and 6 to 9, respectively). The AtPrimase recognition sequence is underlined. In this experiment AtPrimase was present at concentration of 150, 300, and 600 nM whereas AtPollB was maintained at a fixed concentration of 150 nM

few of those products were extended to an pCCA RNA product and as in the case of the ChT7At these products failed to be extended to tetra-, penta- or longer ribonucleotides (Fig. 4C, lanes 6 to 8). The pCC product on the AtPrimase recognition sequence suggest that the ChAtT7 chimera specifically recognizes the AtPrimase sequence.

#### ChT7At and ChAtT7 chimeras prime plant Organellar DNA polymerases

We have shown that AtPrimase selectively generate primers that are used by AtPolls, but those primers are not used by the Klenow Fragment of *E. coli* and T7 DNAP [13]. We were curious to test if the tri ribonucleotide primer (pCCA) synthesized by ChAtT7 (Fig. 4C, lanes 6 to 8 and labeled with an asterisk) could be used as a primer by AtPollB (Fig. 4D). In this experiment, a single-stranded DNA template containing the 5'-CCGGA-3'

sequence was incubated with increased concentrations of AtPrimase and ChAtT7 in the presence of radioactively labeled dNTPs. A reaction incubated with wild-type AtPrimase and AtPolIB makes for an initial amplification product of 36 nts that is extended to longer products (Fig. 4D, lanes 2 to 5), possibly by the terminal transferase activity of AtPolIB [28]. Reactions incubated in the presence of ChAtT7 also resulted in an amplification product of 36nts that are extended to longer primer extension products (Fig. 4D, lanes 6 to 9), although the efficiency of the amplification reaction was significantly lower than the one incubated with the wild-type AtPrimase (Fig. 4D).

### Individually ZFD and RPD domains of AtPrimase functionally assemble

We individually expressed the ZFD and the RPD of AtPrimase and evaluated if these individually expressed domains can functionally complement each other (Fig. 5).

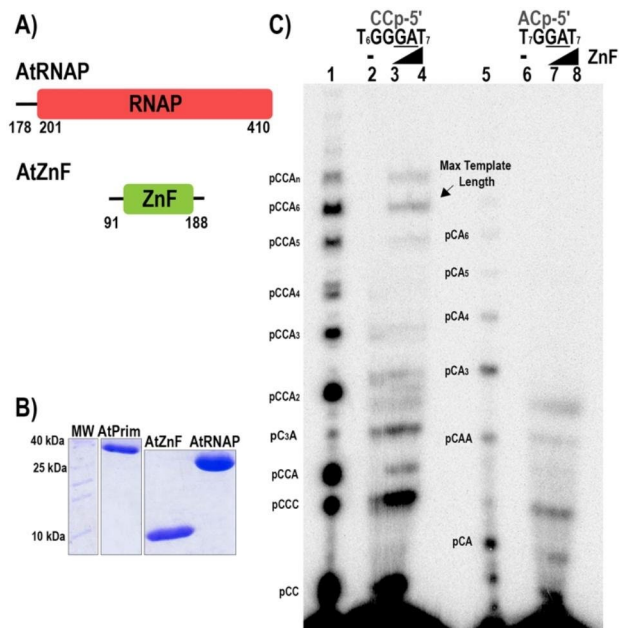
In this experiment, we use a canonical and a modified AtPrimase recognition sequence in the absence or presence of increasing concentrations of ZFD. The RPD

domain by itself lacks the ability to initiate oligoribonucleotide synthesis. However, the addition of increasing concentrations of ZFD result in an increase of synthesized oligoribonucleotides (Fig. 5C, lanes 1 to 8). The stimulation in oligoribonucleotide synthesis in the presence of ZFD is stronger in reactions containing the wild-type recognition sequence (Fig. 5C, lanes 1 to 4) in comparison to reactions assayed on a modified recognition sequence (Fig. 5C, lanes 5 to 8).

### Discussion

To date, the only characterized primase activity in bifunctional primase-helicases from eukaryotes are those from *Dictyostelium discoideum* (DdTwinkle) and *A. thaliana* (AtTwinkle) [19, 31]. In contrast to the animal primase-helicases, DdTwinkle and AtTwinkle efficiently execute oligoribonucleotide synthesis. AtTwinkle preferentially uses the 5'-(G/C)CGA-3' sequence to start ribonucleotide synthesis. In contrast to all primases characterized to date, AtTwinkle uses two nucleotides, 5'-GA-3', as cryptic elements [19]. Given the role of the ZFD of orthologous DNA primases in template recognition, we hypothesized that the ZFD of AtPrimase determines the ability of this enzyme to use two nucleotides as cryptic elements.

We hypothesized that a set of divergent amino acids between T7 and Arabidopsis primases are involved in template recognition (Fig. 2). Residues W162, N163, R166, K168, and K172 are in an optimal position to interact with ssDNA (Fig. 2). The primase activity of the K172A mutant resulted in the synthesis of RNA products that are similar to those synthesized by wild-type AtPrimase (Fig. 2). Mutants W162Y and N163A increase pCCC ribonucleotide synthesis, which is the expected RNA product of a primase that only recognizes the last base of the cryptic pair, while R166A and K168A mutants strongly decrease ribonucleotide incorporation. These results suggest that residues W162 and N163 are involved in the recognition of two nucleotides as the cryptic element, while residues R166 and K168 abolished the ability of the ZFD to bind to ssDNA. To further determine the full extent of the interactions between the mutants and the cryptic element of the recognition site, we tested variations of the cryptic site. In this experiment, the 5'-GA-3' cryptic site was consistently changed for each of its sixteen possible sequences. The initial G that templates for C was kept constant using substrates with the 5'-T<sub>7</sub>-GNN-T<sub>7</sub>-3' sequence (Fig. 3). AtPrimase-K171A exhibited almost identical ribonucleotide synthesis than wild-type AtPrimase on all tested substrates. Interestingly, the R166A and K168A mutants were unable to incorporate ribonucleotides in most of the assayed templates. Only in templates with the 5'-GCN-3' sequence these mutants synthesize a pGN dinucleotide but were



**Fig. 5** Complementation of AtPrimase subunits results in oligoribonucleotide synthesis **A**) Diagram of the deletion mutants of the primase domain. The numbers below the blocks indicate the position of the residues included in each construct **B**) SDS-PAGE gel showing the purified enzymes used in the primase activity assay. **C**) Complementation assay of activity of deletion mutants of the primase domain. Reaction using AtPrimase on both substrates serve as molecular weight marker. In a substrate with the canonical sequence 5'-T<sub>7</sub>-GGGA-T<sub>7</sub>-3', the individual AtPrim domains were able to restore their primase activity (lanes 2 to 4). The primase activity of the domains acting individually is severely affected in the substrate with a modified sequence 5'-T<sub>7</sub>-GGA-T<sub>7</sub>-3' (lanes 6 to 8). AtPrim and AtRPD were used at 2000 nM, AtZnF was used at 2000 and 4000 nM

unable to extend it. The N163A mutant exhibited a strong decrease in RNA synthesis on the canonical template sequence, suggesting that this residue participates in the recognition of the cryptic base pair (Fig. 3). Finally, the W162Y mutant exhibited a similar ribonucleotide synthesis with respect to the wild type, but also show an increase in the formation of a product corresponding to the recognition of only one cryptic base. This suggests that this residue is also involved in the use of two bases as cryptic elements. This experiment suggests that residues W162 and N163, are involved in the recognition of the purine pair (GA) as a cryptic element (Fig. 3). Accordingly to this suggestion, Zn<sup>++</sup> fingers that recognize sequences rich in purines (such as GAG or GAGA) harbor an asparagine residue for the recognition of N<sub>7</sub> atom of adenosine and the recognition of guanosines is determined by residues like tryptophan of phenylalanine [29, 30] (Fig. 3).

To investigate whether the RNA polymerase domain, adjacent to the ZFD is involved in template recognition, we constructed chimeric enzymes (Fig. 4). These chimeras harbor the ZFD of bacteriophage T7 primase with the RNAP domain of the *A. thaliana* primase, and a chimera assembled by the ZFD of the primase of *A. thaliana* and the RNA polymerase domain of the bacteriophage T7 primase. We dubbed those chimeras as ChT7At and ChAtT7, respectively (Fig. 4). The ChT7At and ChAtT7 chimeras show contrasting results. In one hand, the ChT7At chimera is fully active on a T7 primase recognition sequence, whereas the ChAtT7 chimera only synthesized a pCC dimer on the AtPrimase recognition sequence. A minimal amount of those products was extended to a pCCA trimer, and this chimera was inefficient in extending the initial diribonucleotide and tri ribonucleotide products to longer elongation products (Fig. 4). It is proposed that the ZFD positions the initial diribonucleotide from the active site of the RPD to its N-terminal domain to allow diribonucleotide to 3mers or 4mers [26]. Thus, the high amount of pCC dimer synthesized by the ChAtT7 chimera and the complete activity of the ChT7At chimera on a T7 primase recognition indicates the role of the ZFD in recognizing a cryptic sequence during diribonucleotide synthesis.

As the ChAtT7 chimera synthesizes triribonucleotide primers we were curious to know if those RNA primers could be used by the plant mitochondrial DNAPs, in this case AtPolIB. We found that the primers synthesized by the ChAtT7 chimera are extended by AtPolIB, even when those primers were synthesized by the catalytic domain of bacteriophage T7 primase. During the evolution of plant organelles, the ancestral bacteriophage T7-like DNA polymerase was lost. Plant organelles harbor a DNA polymerase that is non-related to T7 DNA polymerase [31–33]. In contrast, plant Twinkle and T7

helicase-primase are phylogenetically related and share biochemical similarities [13, 16]. We think that AtTwinkle is able to efficiently synthesize primers for a DNA polymerase that is distant to T7 DNA polymerase, because the ZFD of AtPrimase as an independent module evolved to interact with the Plant Organellar DNA polymerases and execute primer transfer into their active site.

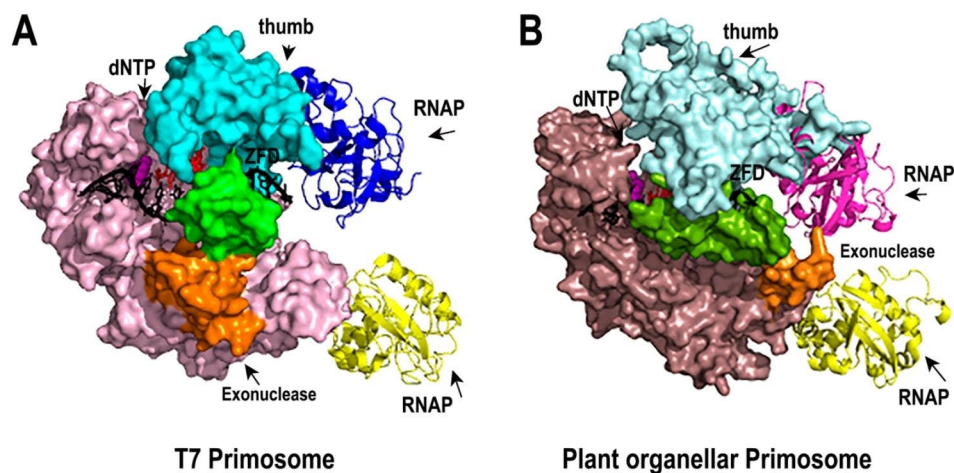
In T7 primase-helicase, the ZFD acts in collaboration with an adjacent RNAP domain rather than with its RNAP domain to execute primer synthesis. To evaluate if the ZFD of AtPrimase could restore the activity of a RNAP domain of AtPrimase we used independently purified ZFD and RNAP domains to evaluate their ability to synthesize ribonucleotides. We found that the RNAP domain and the ZFD function as an active primase, even when they are physically separated (Fig. 5). The latter suggests that AtTwinkle is able to execute priming *in trans* as is the case for T7 primase [18].

The electron microscopy structure of T7 DNA Primase and T7 DNAP after priming delivery shows a direct set of interactions between the ZFD of the T7 primase and the exonuclease and thumb subdomain of T7 DNAP, those interactions strongly suggest that the ZFD and T7DNAP have co-evolved to perform specific protein-protein interactions and that an exogenous ZFD could not execute primer delivery (PDB ID: 6n9u) (Fig. 6A) [24, 34]. A structural model of AtPolIB with AtPrimase during primer delivery, shows that AtPolIB is predicted to interact with the ZFD of AtPrimase via its insertion 1 and an exonuclease domain a helix, two structural elements that are unique to plant mitochondrial DNAPs [13, 35] (Fig. 6B).

## Conclusions

D-loop, theta model, rolling circle, and recombination-dependent replication processes are proposed mechanisms for the replication of plant organellar genomes [36–40]. The recombinogenic nature of plant mitochondrial DNA and the presence of enzymes involved in homologous recombination in plant organelles suggests that plant organellar DNA replication may be accomplished through recombination-dependent replication (RDR) [11, 14]. As an alternative, Morley and coworkers proposed that a mechanism similar to the coordinated leader and lagging-strand replication model of bacteriophage T7 is used to replicate plant mitochondrial DNA [12, 41]. According to this scenario, short ribonucleotides are synthesized by the primase module of AtTwinkle and those ribonucleotides are transferred into the active site of the lagging strand DNA polymerase [13].

Interestingly, plants harbor another primase dubbed PrimPol [42, 43]. In Arabidopsis, AtPrimPol is found in the nucleus, mitochondria, and chloroplast. Human PrimPol reinitiates mitochondrial and nuclear DNA



**Fig. 6** Structural model of the plant organellar primosome. **(A)** Cryo electron structure of the T7 primosome [24]. T7DNAP is drawn in a surface representation. The ZFD is green colored and locates at a cleft between the thumb subdomain (cyan) and a unique helix of the exonuclease domain (orange). The RNAP domain of T7 primase (*cis* in blue and *trans* in yellow) interact with the thumb subdomain (blue) and the exonuclease domain, respectively. **(B)** Structural model of the proposed plant organellar primosome. The structural model of AtPollB is in surface representation and colored in light brown. A unique loop of the exonuclease domain is colored in orange and the thumb subdomain is colored in aquamarine. The ZFD of plant mitochondrial AtPrimase is colored in olive and is positioned between a unique insertion (Insertion 1) of the thumb subdomain and an exposed loop of the exonuclease domain (orange). The RNAP domain of AtPrimase (*cis* in magenta and *trans* in yellow) are poised to interact with the thumb subdomain and the exonuclease domain respectively

replication at DNA lesions, chain terminating nucleotides or complex structures, like G-quadruplexes, to assure the continuous progression of the replication fork to maintain DNA replication [44–47]. AtPrimPol synthesizes primers that are elongated by organellar DNA polymerases, suggesting that in plant organelles AtPrimPol may participate in the assembling of replication forks as an alternative mechanism to avoid replication fork collapse.

## Methods

### Site directed mutagenesis and chimeric proteins construction

The open reading frame of AtPrimase (residues 92–410) was PCR amplified from a codon optimized AtTwinkle vector using AtPrimase\_Nter (5′-aggaaaCATATGacccc-ggttgatacc) and AtPrimase\_Cter (5′-agaagGGATCCt-taatagcgtttcgg) oligonucleotides. The PCR product was subcloned between the *Nde* I and *Bam* HI restriction sites (capital letters) of a modified pET19 vector to create the pET19-AtPrimase plasmid. Site-directed mutagenesis was carried out using a modified version of the Q5 Site-Directed Mutagenesis Kit (New England Biolabs). For each point mutants, the pET19-AtPrimase vector was subject to PCR amplification using the oligonucleotides depicted in Table S1. Each PCR product was incubated with T4 kinase and T4 ligase for 30 min and subject to Dpn I treatment for 2 h. 1 μl of the reaction was used to transform chemically competent *E. coli*. The expression plasmid for the primase domain of bacteriophage T7 primase-helicase (residues 1-262) was a gift from Dr. Charles C. Richardson laboratory [21].

AtZnF was constructed using a PCR reaction using the AtPrimase\_Nter oligonucleotide in combination with the AtZnF oligonucleotide (5′-ctcttccGGATCCt-tacggatctgctgaggccaggcc). AtRPD was constructed via PCR amplification using the NtRPD oligonucleotide (5′-aaagaaaCATATGaccgttgaaggattgaactg) and the AtPrimase\_Cter oligonucleotide. ChT7AT and ChAtT7 chimeric proteins were constructed using the Gene Splicing by Overlap Extension (SOEing) protocol [48]. Detailed explanation of this methodology is presented in supplementary materials. AtZnF, AtRPD, and chimeras were subcloned into a modified pET19 vector between *Nde* I and *Bam* HI restriction sites. The amino acid boundaries for each subdomain of the chimeras are defined in the figure's legends.

### Protein purification

Protein purification of AtPrimase mutants and chimeric proteins was carried out in an *E. coli* BL21(DE3) strain in LB medium supplemented with 0.1 mM ZnCl<sub>2</sub> as previously described [13]. Wild-type T7 Primase was expressed and purified as previously described (33).

### Molecular modeling

The homology model of AtPrimase (residues 92 to 410) was built using the molecular operating environment (MOE) platform using the homology model algorithm. The amino acid sequence of AtPrimase was aligned with the amino acid sequence used to solve the crystal structure of bacteriophage T7 primase (PDB ID: 1NUI) [21] using the default settings of the align panel. The



MOE platform generated ten homology intermediate models and a final model that is minimized using the CHARMM27 force field.

### Primase reactions

Primase reactions were executed in a buffer containing 40 mM Tris–HCl pH 7.5, 50 mM potassium glutamate, 10 mM MgCl<sub>2</sub> and 10 mM DTT. Primase reactions contained 100 μM of the indicated NTPs and 10 μCi of [ $\alpha$ -<sup>32</sup>P]-NTP (3,000 Ci/mM). Reactions were terminated by adding an equal volume of stop buffer (98% formamide, 0.1% xylene cyanol, 20 mM of EDTA). Reaction products were separated on a 27% denaturing polyacrylamide gel containing 3 M urea. Each primase reaction contained varying amounts of recombinant protein or single-stranded DNA template as indicated in the figure legends. Primase reactions were analyzed by phosphorimaging. Primase reactions were carried out using 5 μM of the synthetic template and 0.5 μM of each recombinant AtPrimase-helicase.

### Polymerase-primase coupled reactions

Coupled primase–DNA polymerase reactions were carried out under the same conditions as the primase reaction, adding 100 μM of dNTPs and labeled with [ $\alpha$ -<sup>32</sup>P]-dCTP. AtPrimase and chimeric proteins were present at 150, 300, and 600 nM whereas DNA polymerases were present at 150 nM. The template sequence used in the coupled primase–polymerase assay contained both the AtPrimase and T7 primase recognition sites.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-023-04477-4>.

Supplementary Material 1  
Supplementary Material 2  
Supplementary Material 3  
Supplementary Material 4

### Acknowledgements

APC and FCA thank CONACYT for the graduate studies fellowship.

### Authors' contributions

APC; wrote the manuscript, experimental design, and execution, data interpretation, FCA and CDQ, mutant construction and protein expression-purification, RSM and RW revised the manuscript, experimental design data interpretation, and funding. LGB, wrote and revised the manuscript, original idea, experimental design and data interpretation, funding.

### Funding

This work was supported by grant Ciencia de Fronteras-CONACYT # 70713 and the Organización de Estados Iberoamericanos-programa FORCYT.

### Data Availability

All raw data generated or analyzed during this study are included in this published article.

### Declarations

#### Ethics approval and consent to participate

Not Applicable.

#### Consent to publish

Not Applicable.

#### Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

#### Competing interests

The authors declare no competing interests.

Received: 22 May 2023 / Accepted: 19 September 2023

Published online: 06 October 2023

### References

- Sousa R. Structural and mechanistic relationships between nucleic acid polymerases. *Trends Biochem Sci.* 1996;21(5):186–90.
- Yao NY, O'Donnell ME. Evolution of replication machines. *Crit Rev Biochem Mol Biol.* 2016;51(3):135–49.
- Peter B, Falkenberg M. TWINKLE and other human mitochondrial DNA helicases: structure, function and disease. *Genes (Basel)* 2020, 11(4).
- Kulczyk AW, Richardson CC. The replication system of bacteriophage T7. *Enzymes.* 2016;39:89–136.
- Kato M, Frick DN, Lee J, Tabor S, Richardson CC, Ellenberger T. A complex of the bacteriophage T7 primase-helicase and DNA polymerase directs primer utilization. *J Biol Chem.* 2001;276(24):21809–20.
- Frick DN, Richardson CC. DNA primases. *Annu Rev Biochem.* 2001;70:39–80.
- Ilyina TV, Gorbalenya AE, Koonin EV. Organization and evolution of bacterial and bacteriophage primase-helicase systems. *J Mol Evol.* 1992;34(4):351–7.
- Fernandez-Millan P, Lazaro M, Cansiz-Arda S, Gerhold JM, Rajala N, Schmitz CA, Silva-Espina C, Gil D, Bernado P, Valle M, et al. The hexameric structure of the human mitochondrial replicative helicase twinkle. *Nucleic Acids Res.* 2015;43(8):4284–95.
- Donmez I, Patel SS. Mechanisms of a ring shaped helicase. *Nucleic Acids Res.* 2006;34(15):4216–24.
- Hamdan SM, Richardson CC. Motors, switches, and contacts in the replisome. *Annu Rev Biochem.* 2009;78:205–43.
- Cupp JD, Nielsen BL. Minireview: DNA replication in plant mitochondria. *Mitochondrion.* 2014;19(Pt B):231–7.
- Morley SA, Peralta-Castro A, Briebe LG, Miller J, Ong KL, Ridge PG, Oliphant A, Aldous S, Nielsen BL. Arabidopsis thaliana organelles mimic the T7 phage DNA replisome with specific interactions between twinkle protein and DNA polymerases Pol1A and Pol1B. *BMC Plant Biol.* 2019;19(1):241.
- Peralta-Castro A, Baruch-Torres N, Briebe LG. Plant organellar DNA primase-helicase synthesizes RNA primers for organellar DNA polymerases using a unique recognition sequence. *Nucleic Acids Res.* 2017;45(18):10764–74.
- Gualberto JM, Newton KJ. Plant mitochondrial genomes: Dynamics and Mechanisms of Mutation. *Annu Rev Plant Biol.* 2017;68:225–52.
- Briebe LG. Structure-function analysis reveals the singularity of Plant mitochondrial DNA replication components: a Mosaic and Redundant System. *Plants (Basel)* 2019, 8(12).
- Shutt TE, Gray MW. Twinkle, the mitochondrial replicative DNA helicase, is widespread in the eukaryotic radiation and may also be the mitochondrial DNA primase in most eukaryotes. *J Mol Evol.* 2006;62(5):588–99.
- Gupta A, Patil S, Vijayakumar R, Kondabagil K. The Polyphyletic Origins of Primase-Helicase Bifunctional Proteins. *J Mol Evol.* 2017;85(5–6):188–204.
- Qimron U, Lee SJ, Hamdan SM, Richardson CC. Primer initiation and extension by T7 DNA primase. *EMBO J.* 2006;25(10):2199–208.
- Harman A, Barth C. The Dictyostelium discoideum homologue of Twinkle, Twm1, is a mitochondrial DNA helicase, an active primase and promotes mitochondrial DNA replication. *BMC Mol Biol.* 2018;19(1):12.

20. Kusakabe T, Hine AV, Hyberts SG, Richardson CC. The Cys4 zinc finger of bacteriophage T7 primase in sequence-specific single-stranded DNA recognition. *Proc Natl Acad Sci U S A*. 1999;96(8):4295–300.
21. Kato M, Ito T, Wagner G, Richardson CC, Ellenberger T. Modular architecture of the bacteriophage T7 primase couples RNA primer synthesis to DNA synthesis. *Mol Cell*. 2003;11(5):1349–60.
22. Diray-Arce J, Liu B, Cupp JD, Hunt T, Nielsen BL. The Arabidopsis At1g30680 gene encodes a homologue to the phage T7 gp4 protein that has both DNA primase and DNA helicase activities. *BMC Plant Biol*. 2013;13:36.
23. Towle-Weicksel JB, Cao Y, Crislip LJ, Thurlow DL, Crampton DJ. Chimeric proteins constructed from bacteriophage T7 gp4 and a putative primase-helicase from Arabidopsis thaliana. *Mol Biol Rep*. 2014;41(12):7783–95.
24. Gao Y, Cui Y, Fox T, Lin S, Wang H, de Val N, Zhou ZH, Yang W. Structures and operating principles of the replisome. *Science* 2019, 363(6429).
25. Akabayov B, Lee SJ, Akabayov SR, Rekhi S, Zhu B, Richardson CC. DNA recognition by the DNA primase of bacteriophage T7: a structure-function study of the zinc-binding domain. *Biochemistry*. 2009;48(8):1763–73.
26. Zhang H, Lee SJ, Richardson CC. The roles of tryptophans in primer synthesis by the DNA primase of bacteriophage T7. *J Biol Chem*. 2012;287(28):23644–56.
27. Lee SJ, Zhu B, Akabayov B, Richardson CC. Zinc-binding domain of the bacteriophage T7 DNA primase modulates binding to the DNA template. *J Biol Chem*. 2012;287(46):39030–40.
28. Garcia-Medel PL, Baruch-Torres N, Peralta-Castro A, Trasvina-Arenas CH, Torres-Larios A, Briebe LG. Plant organellar DNA polymerases repair double-stranded breaks by microhomology-mediated end-joining. *Nucleic Acids Res*. 2019;47(6):3028–44.
29. Clore MG, Omichinski JG. GAGA: structural basis for single Cys2His2 Zinc Finger-DNA Interaction. *Madame Curie Bioscience Database [Internet] (Landes Bioscience)* 2013.
30. Anzellotti AI, Liu Q, Bloemink MJ, Scarsdale JN, Farrell N. Targeting retroviral zinc finger-DNA interactions: a small-molecule approach using the electrophilic nature of trans-platinum-nucleobase compounds. *Chem Biol*. 2006;13(5):539–48.
31. Moriyama T, Terasawa K, Sato N. Conservation of POPs, the plant organellar DNA polymerases, in eukaryotes. *Protist*. 2011;162(1):177–87.
32. Moriyama T, Sato N. Enzymes involved in organellar DNA replication in photosynthetic eukaryotes. *Front Plant Sci*. 2014;5:480.
33. Czernecki D, Nourisson A, Legrand P, Delarue M. Reclassification of family A DNA polymerases reveals novel functional subfamilies and distinctive structural features. *Nucleic Acids Res* 2023.
34. Juarez-Quintero V, Peralta-Castro A, Benitez Cardoza CG, Ellenberger T, Briebe LG. Structure of an open conformation of T7 DNA polymerase reveals novel structural features regulating primer-template stabilization at the polymerization active site. *Biochem J*. 2021;478(13):2665–79.
35. Baruch-Torres N, Briebe LG. Plant organellar DNA polymerases are replicative and translesion DNA synthesis polymerases. *Nucleic Acids Res*. 2017;45(18):10751–63.
36. de Haas JM, Hille J, Kors F, van der Meer B, Kool AJ, Folkerts O, Nijkamp HJ. Two potential *Petunia hybrida* mitochondrial DNA replication origins show structural and in vitro functional homology with the animal mitochondrial DNA heavy and light strand replication origins. *Curr Genet*. 1991;20(6):503–13.
37. Backert S, Borner T. Phage T4-like intermediates of DNA replication and recombination in the mitochondria of the higher plant *Chenopodium album* (L). *Curr Genet*. 2000;37(5):304–14.
38. Oldenburg DJ, Bendich AJ. DNA maintenance in plastids and mitochondria of plants. *Front Plant Sci*. 2015;6:883.
39. Oldenburg DJ, Bendich AJ. The structure of mitochondrial DNA from the liverwort, *Marchantia polymorpha*. *J Mol Biol*. 1998;276(4):745–58.
40. Cheng N, Lo YS, Ansari MI, Ho KC, Jeng ST, Lin NS, Dai H. Correlation between mtDNA complexity and mtDNA replication mode in developing cotyledon mitochondria during mung bean seed germination. *New Phytol*. 2017;213(2):751–63.
41. Lee SJ, Richardson CC. Choreography of bacteriophage T7 DNA replication. *Curr Opin Chem Biol*. 2011;15(5):580–6.
42. Garcia-Medel PL, Peralta-Castro A, Baruch-Torres N, Fuentes-Pascacio A, Pedroza-Garcia JA, Cruz-Ramirez A, Briebe LG. Arabidopsis thaliana PrimPol is a primase and lesion bypass DNA polymerase with the biochemical characteristics to cope with DNA damage in the nucleus, mitochondria, and chloroplast. *Sci Rep*. 2021;11(1):20582.
43. Garcia-Gomez S, Reyes A, Martinez-Jimenez MI, Chocron ES, Mouron S, Terrados G, Powell C, Salido E, Mendez J, Holt IJ, et al. PrimPol, an archaic primase/polymerase operating in human cells. *Mol Cell*. 2013;52(4):541–53.
44. Torregrosa-Munumer R, Forslund JME, Goffart S, Pfeiffer A, Stojkovic G, Carvalho G, Al-Furoukh N, Blanco L, Wanrooij S, Pohjoismaki JLO. PrimPol is required for replication reinitiation after mtDNA damage. *Proc Natl Acad Sci U S A*. 2017;114(43):11398–403.
45. Gonzalez-Acosta D, Blanco-Romero E, Ubieta-Capella P, Mutreja K, Miguez S, Llanos S, Garcia F, Munoz J, Blanco L, Lopes M, et al. PrimPol-mediated repriming facilitates replication traverse of DNA interstrand crosslinks. *EMBO J*. 2021;40(14):e106355.
46. Jacobs K, Doerdelmann C, Krietsch J, Gonzalez-Acosta D, Mathis N, Kushinsky S, Guarino E, Gomez-Escolar C, Martinez D, Schmid JA, et al. Stress-triggered hematopoietic stem cell proliferation relies on PrimPol-mediated repriming. *Mol Cell*. 2022;82(21):4176–4188e4178.
47. Kobayashi K, Guillian TA, Tsuda M, Yamamoto J, Bailey LJ, Iwai S, Takeda S, Doherty AJ, Hirota K. Repriming by PrimPol is critical for DNA replication restart downstream of lesions and chain-terminating nucleosides. *Cell Cycle*. 2016;15(15):1997–2008.
48. Horton RM, Cai Z, Ho SM, Pease LR. Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *BioTechniques* 8(5):528–535 (November 1990). *BioTechniques* 2013, 54(3):129–133.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.