



HAL
open science

The Plastid-Encoded RNA Polymerase-Associated Protein PAP9 Is a Superoxide Dismutase With Unusual Structural Features

Adrien Favier, Pierre Gans, Elisabetta Boeri Erba, Luca Signor, Soumiya Sankari Muthukumar, Thomas Pfannschmidt, Robert Blanvillain, David Cobessi

► To cite this version:

Adrien Favier, Pierre Gans, Elisabetta Boeri Erba, Luca Signor, Soumiya Sankari Muthukumar, et al.. The Plastid-Encoded RNA Polymerase-Associated Protein PAP9 Is a Superoxide Dismutase With Unusual Structural Features. *Frontiers in Plant Science*, 2021, 12, pp.668897. 10.3389/fpls.2021.668897. hal-03313484

HAL Id: hal-03313484

<https://hal.science/hal-03313484>

Submitted on 4 Aug 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



The Plastid-Encoded RNA Polymerase-Associated Protein PAP9 Is a Superoxide Dismutase With Unusual Structural Features

OPEN ACCESS

Edited by:

Ning Li,
Hong Kong University of Science
and Technology, China

Reviewed by:

Guang Zhu,
Hong Kong University of Science
and Technology, China
Takashi Shiina,
Kyoto Prefectural University, Japan

*Correspondence:

Robert Blanvillain
robert.blanvillain@cea.fr
David Cobessi
david.cobessi@ibs.fr

† Present address:

Thomas Pfannschmidt,
Pflanzenphysiologie, Institut für
Botanik, Leibniz-Universität Hannover,
Hanover, Germany

Specialty section:

This article was submitted to
Plant Proteomics and Protein
Structural Biology,
a section of the journal
Frontiers in Plant Science

Received: 17 February 2021

Accepted: 28 May 2021

Published: 30 June 2021

Citation:

Favier A, Gans P, Boeri Erba E,
Signor L, Muthukumar SS,
Pfannschmidt T, Blanvillain R and
Cobessi D (2021) The
Plastid-Encoded RNA
Polymerase-Associated Protein PAP9
Is a Superoxide Dismutase With
Unusual Structural Features.
Front. Plant Sci. 12:668897.
doi: 10.3389/fpls.2021.668897

Adrien Favier¹, Pierre Gans¹, Elisabetta Boeri Erba¹, Luca Signor¹,
Soumiya Sankari Muthukumar¹, Thomas Pfannschmidt^{2†}, Robert Blanvillain^{2*} and
David Cobessi^{1*}

¹ Université Grenoble Alpes, CEA, CNRS, IBS, Grenoble, France, ² Université Grenoble-Alpes, CNRS, CEA, INRA,
IRIG-LPCV, Grenoble, France

In Angiosperms, the plastid-encoded RNA polymerase (PEP) is a multimeric enzyme, essential for the proper expression of the plastid genome during chloroplast biogenesis. It is especially required for the light initiated expression of photosynthesis genes and the subsequent build-up of the photosynthetic apparatus. The PEP complex is composed of a prokaryotic-type core of four plastid-encoded subunits and 12 nuclear-encoded PEP-associated proteins (PAPs). Among them, there are two iron superoxide dismutases, FSD2/PAP9 and FSD3/PAP4. Superoxide dismutases usually are soluble enzymes not bound into larger protein complexes. To investigate this unusual feature, we characterized PAP9 using molecular genetics, fluorescence microscopy, mass spectrometry, X-ray diffraction, and solution-state NMR. Despite the presence of a predicted nuclear localization signal within the sequence of the predicted chloroplast transit peptide, PAP9 was mainly observed within plastids. Mass spectrometry experiments with the recombinant *Arabidopsis* PAP9 suggested that monomers and dimers of PAP9 could be associated to the PEP complex. In crystals, PAP9 occurred as a dimeric enzyme that displayed a similar fold to that of the FeSODs or manganese SOD (MnSODs). A zinc ion, instead of the expected iron, was found to be penta-coordinated with a trigonal-bipyramidal geometry in the catalytic center of the recombinant protein. The metal coordination involves a water molecule and highly conserved residues in FeSODs. Solution-state NMR and DOSY experiments revealed an unfolded C-terminal 34 amino-acid stretch in the stand-alone protein and few internal residues interacting with the rest of the protein. We hypothesize that this C-terminal extension had appeared during evolution as a distinct feature of the FSD2/PAP9 targeting it to the PEP complex. Close vicinity to the transcriptional apparatus may allow for the protection against the strongly oxidizing aerial environment during plant conquering of terrestrial habitats.

Keywords: plastid-encoded RNA polymerase, iron superoxide dismutase, chloroplast biogenesis, NMR, X-ray crystallography

INTRODUCTION

In the green lineage, the photosynthetic reactions in the chloroplast convert light energy into chemical energy with the release of di-oxygen. Other metabolic pathways take place in chloroplasts such as the biosynthesis of amino acids, fatty acids, vitamins, and hormones. Hence, the chloroplast functions sustain most life forms on Earth (Jarvis and López-Juez, 2013). According to the endosymbiosis theory of the origin of organelles, chloroplasts have evolved from a single ancient cyanobacterium engulfed around 1.5 billion years ago into a mitochondriate proto-eukaryote (Bobik and Burch-Smith, 2015). During evolution, a massive gene transfer occurred from the cyanobacterium into the nucleus of the host cell (Martin et al., 2002). Thus, the nuclear genome could encode from 1500 to 4500 chloroplast proteins whereas the plastid genome (plastome) encodes for about hundred proteins (Zybailov et al., 2008). The plastome (cpDNA) mainly encodes: (1) components of the plastid gene expression machinery (RNA polymerase, ribosomal proteins, tRNAs, and rRNAs), (2) subunits of each major functional photosynthesis-related complex (e.g., RuBisCO, Photosystem I and II, the cytochrome *b₆f* complex, NADPH dehydrogenase, and ATP synthase), and (3) a few proteins involved in other processes (e.g., ClpP1 and YCF3) (Sugiura, 1992; Majeran et al., 2012; Yu et al., 2014). Hence, the vast majority of chloroplast proteins are encoded by the nuclear genome. The pre-proteins are imported into the chloroplast from the cytosol mainly by the TOC-TIC machinery of the chloroplast envelope that recognizes and cleaves specific transit peptides (cTPs) at their N-terminal extremity (Jarvis, 2008). Once in the stroma, the proteins are properly folded. Since most of the protein complexes in the chloroplast contain nuclear and chloroplast-encoded proteins, coordination in expression of both genomes is essential (Liebers et al., 2017).

Two RNA polymerases are involved in plastid transcription: a nuclear-encoded RNA polymerase (NEP) and the plastid-encoded RNA polymerase (PEP). The NEP, a T3–T7 bacteriophage type RNA polymerase, transcribes the *rpo* genes (*rpoA*, *B*, *C1*, and *C2*), encoding the four subunits of the catalytic core of the PEP, and other housekeeping genes (Kremnev and Strand, 2014; Börner et al., 2015). During chloroplast biogenesis, the PEP core is reshaped in a multi-subunit RNA-polymerase of at least 16 different proteins (MW: ~1 MDa), which mainly transcribes photosynthesis related genes. The active PEP complex is composed of four *rpo* core subunits, and 12 nuclear-encoded PEP-associated proteins (PAPs) (Pfannschmidt et al., 2015). Mutations in most of the *pap* genes yield albino/ivory plants incapable of photosynthesis with a defect in the expression of PEP-dependent genes indicating that the PEP is not fully functional (Pfannschmidt et al., 2015). This shared phenotype triggered the idea of a PAPs-related developmental block corresponding to an epistasis effect. This effect occurs when all components are required for the stability of the entire complex ensuring that photosynthesis could be launched only if all the functions are present (Liebers et al., 2018).

The PAPs can be divided into four groups according to their hypothetical functions (Yu et al., 2014). PAP sequence analyses

and biochemical studies allowed to characterize four PAPs with potential known catalytic activities: PAP4, PAP7, PAP9, and PAP10. PAP7 belongs to methyltransferases (Gao et al., 2011), PAP10 is a thioredoxin (TrxZ) (Steiner et al., 2011) while PAP4 (FSD3) and PAP9 (FSD2) are both iron superoxide dismutases (FeSOD) (Myouga et al., 2008). Formation of superoxide radicals mainly occurs in electron transport chains of photosynthesis and respiration. Therefore, PAP4 and PAP9 may serve as protection against oxidative stresses generated during the first activities of the photosynthetic apparatus (Pfannschmidt et al., 2015). Indeed, superoxide radicals can damage sulfur containing amino acids, metals, and Fe-S clusters. SODs are cellular defenses against superoxide by catalyzing the dismutation of superoxide into hydrogen peroxide according to the overall reaction: $2\text{O}_2^- + 2\text{H}^+ \rightarrow \text{H}_2\text{O}_2 + \text{O}_2$ (Pfannschmidt, 2003; Abreu and Cabelli, 2010).

Besides the MnSODs and the copper-zinc SODs (Cu/ZnSODs, where Cu is the redox center), three iron superoxide dismutases (FeSODs) were characterized in plants. Dimeric MnSODs are found in the matrix of the mitochondria, with one Mn ion per monomer. Cu/ZnSODs are dimeric SODs found in the cytosol, peroxisomes, and plastids. Each monomer contains one Cu and one Zn ion. FeSODs are dimeric enzymes with one iron ion bound to each monomer. The fold of the FeSOD monomer is roughly similar to that of the MnSOD monomer and is completely different from the Cu/ZnSODs (Pilon et al., 2011). In plants, FSD1 is a cytoplasmic FeSOD, while PAP4 and PAP9 are FeSODs only observed in the chloroplast, both associated to the PEP (Myouga et al., 2008; Steiner et al., 2011). Surprisingly, the oligomeric assembly of PAP4 and PAP9 differ from that observed for FeSODs. PAP9 was reported as being a monomer in the PEP and PAP4 as a trimer (Steiner et al., 2011). In *Arabidopsis thaliana*, PAP4 and PAP9 could form a heterodimeric complex in the chloroplast nucleoids (Myouga et al., 2008). The *pap4-pap9* double mutant displayed an albino phenotype with no chloroplast development while the *pap4* or *pap9* single inactivation mutants showed pale green phenotypes and sensitivity to oxidative stress indicating some compensation effect but no full redundancy between the two proteins (Myouga et al., 2008). These observations strongly suggested that a heterodimeric complex PAP4/PAP9 could protect the transcriptionally active chromosome (TAC) during the early stages of chloroplast development from the superoxide radical produced during photosynthesis in the thylakoid membranes (Myouga et al., 2008). To better characterize PAP9 and understand how plastid-localized FeSODs were embedded in the PEP, we studied PAP9 using phylogenetic approaches, *in planta* experiments, mass spectrometry, X-ray diffraction, and solution-state NMR.

MATERIALS AND METHODS

Accessions

PAP9 At5g51100; accessions from the green lineage are given in **Supplementary Table 1**. Full-length coding sequences were retrieved from Blastp (**Supplementary Table 2**). The protein

sequences were aligned using Clustal Omega¹. The prediction of chloroplast pre-sequences (**Supplementary Table 3**) were established using ChloroP² (Emanuelsson et al., 1999). The predictions of the nuclear localization signals (NLS) were performed using NLS_Mapper³ and are given in **Supplementary Table 4** (Kosugi et al., 2009). Clustal Omega color-code as followed: [red (AVFPMILW): small + hydrophobic (includes aromatic Y); blue: (DE), acidic; magenta: (RHK), basic; green: (HSTYHCNGQ), hydroxyl + sulphydryl + amine + G].

Peptide Synthesis

The peptide ²²⁶QREQGTETEDDENPDDEVPEVYLDSDIDVSEVD²⁵⁹ corresponding to the last 34 residues of PAP9 was synthesized by Proteomic Solution with a purity (HPLC) of 98.29%. Its molecular mass (MW: 3925.85 Da) was checked using mass spectrometry.

Transient Transformation of Onion Cells

Gold Carrier Particles (Seashell technology) were coated with 1 µg of the expression vector and 1 µg of an internal control such as PAP10-RFP (Liebers et al., 2020). Gold particles were delivered into onion cells using a particle gun (BioRad). The transformed cells were allowed to express the construct for 16–24 h before fluorescence observation using proper filters. Signal profiles of the two fluorescence channels were acquired on pictures using ImageJ.

Cloning and Vector Construction

PAP9^{ΔcTP} (271 aa/31 kDa) in pBB408 corresponds to PAP9^{ΔcTP}-6His in the pEt21d backbone: RT-PCR fragment was obtained from seedling cDNA amplified with oP9^{ΔcTP}_FNco (5'-CCATGGGTGTTATCACAGCTGG)/oP9_RNot (5'-GCGGCCGCGTCAACCTCAGATACATCGATG), A-tailed and cloned in pGem-Teasy (pBB399a) then digested with *Nco*I, *Not*I and cloned in pET21d. PAP9-GFP in pAF04 (pEZX-NL backbone, Stanford): RT-PCR fragment was obtained from seedling cDNA amplified with oPAP9_FXho (5'-CTC GAGATGATGAATGTTGACAGTACAGCC) and oPAP9_RBH (5'-GGATCCCCGTC AACCTCAGATACATCGATGTCAC) cloned as above then digested with *Xho*I *Bam*HI and ligated in pEZX-NL. pBB301 (PA10-RFP) was used as internal control (Liebers et al., 2020).

Protein Expression and Purification

PAP9-6His (for ΔcTP-PAP9-6His) was overexpressed in *E. coli* Rosetta2 strain in LB with 100 µg/mL ampicillin and 34 µg/mL chloramphenicol. 6His-PAP9 (for ΔcTP-6His-PAP9) was overexpressed in *E. coli* Rosetta2 strain in LB with 100 µg/mL ampicillin and 50 µg/mL kanamycin. Cells were grown overnight in 50 mL of LB with antibiotics at 37°C. One liter of LB (with antibiotics) was then inoculated with the first culture to reach an initial OD₆₀₀ of 0.1. Growth was continued at 37°C. When the OD₆₀₀ reached 0.6, the temperature was decreased to 16°C and

isopropyl β-D-1-thiogalactopyranoside was added to give a final concentration of 0.5 mM. After an overnight induction, bacteria were harvested at 6619 g, for 25 min, at 4°C. The cell pellet was resuspended in 30 mL of lysis buffer (50 mM Tris-HCl pH 8.0, 0.5 M NaCl, 20 mM imidazole) containing a Complete Protease inhibitor Cocktail tablet (Roche). The lysate was centrifuged at 15,000 g, for 40 min, at 4°C. The purification was performed at room temperature. The supernatant was applied onto a NiNTA column in 50 mM Tris-HCl pH 8.0, 0.5 M NaCl, 20 mM imidazole. Proteins were eluted in one step in a buffer containing 50 mM Tris-HCl pH 8.0, 0.1 M NaCl, 300 mM imidazole. Then the eluate was diluted 2 times in 50 mM Tris-HCl pH 8.0 and loaded on a MonoQ column. Elution was performed using a linear NaCl gradient from 0 to 1 M in 50 mM Tris-HCl pH 8.0. The fractions containing PAP9-6His or 6His-PAP9 were pooled and concentrated with an Amicon Ultra 4 mL centrifugal filter and a 10 kDa membrane cut-off before loading on a HiLoad 16/60 Superdex 200 and then eluted with 10 mM Tris-HCl pH 8.0, 50 mM NaCl. The fractions containing the pure protein were pooled and concentrated for further experiments or stored at -20°C with 50% (v/v) glycerol.

¹⁵N,¹³C-6His-PAP9 was expressed in minimum media M9 supplemented with ¹⁵NH₄Cl, ¹³C-glucose and antibiotics. Briefly, 5 mL of LB were inoculated with *E. coli* Rosetta2 stock glycerol overexpressing 6His-PAP9. After 10 h of growing, 1 mL was added to 100 mL of minimum media supplemented as described above. After 1 night growing, when OD₆₀₀ was close to 2, the overnight culture was centrifuged to inoculate 1 L of minimum media M9 supplemented with ¹⁵NH₄Cl and ¹³C-glucose and antibiotics. Cell growth, overexpression and purification followed the procedure described above for 6His-PAP9 and PAP9-6His.

Enzymatic Assays

The superoxide dismutase activity of PAP9 was tested using pyrogallol. The pyrogallol auto-oxidation is characterized by increase of absorbance at 420 nm and superoxide dismutase inhibits the pyrogallol auto-oxidation. Briefly, 7 mM pyrogallol was dissolved in a *Tris*-succinate-EDTA buffer pH 8.2 and the pyrogallol auto-oxidation was followed by monitoring the absorbance increase at 420 nm. After 180 s, PAP9 at several concentrations (50, 100, 200, 500 µM, and 1 mM) or 5 µM Mn-SOD were added into the medium and the absorbance was monitored for further 3 min. Experiments were repeated three times for each concentration and the curves were plotted. Each curve correspond to the average of three enzymatic assays (**Supplementary Figure 1**).

LC/ESI and Native Mass Spectrometry

Liquid chromatography electrospray ionization mass spectrometry (LC/ESI-MS) was used to assess the masses of the intact PAP9-6His, and ¹⁵N,¹³C-6His-PAP9. All solvents were HPLC grade (Chromasolv, Sigma-Aldrich) and trifluoroacetic acid (TFA) was from Acros Organics (puriss, p.a.). Solvent A was 0.03% TFA in water, solvent B contained 95% acetonitrile, 5% water, and 0.03% TFA. A 6210 LC/ESI-TOF mass spectrometer interfaced with an HPLC binary pump system (Agilent Technologies) was used. The mass spectrometer was calibrated

¹<https://www.ebi.ac.uk/Tools/msa/clustalo/>

²<http://www.cbs.dtu.dk/services/ChloroP/>

³http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi

in the mass-to-charge (m/z) range 300–3000 using a standard calibrant (ESI-L, low concentration tuning mix, Agilent Technologies) before the measurements of protein samples. MS acquisition was carried out in positive ion mode and mass spectra were recorded in the 300–3200 m/z range. ESI source temperature was set at 573 K, nitrogen was used as drying gas (7 L/min) and as nebulizer gas (10 psi). The capillary needle voltage was set at 4000 V. Spectra acquisition rate was of 1.03 spectra/s. The MS spectra were acquired and the data processed with MassHunter workstation software (v. B.02.00, Agilent Technologies) and with GPMW software (v. 7.00b2, Lighthouse Data, Denmark). Immediately before the MS analysis, the protein samples were diluted to a final concentration of 8 μ M using solvent A. Samples were kept at 10°C in the autosampler and 8 μ L of each sample were injected into the system. They were first trapped and desalted on a reverse phase-C8 cartridge (Zorbax 300SB-C8, 5 μ m, 300 μ m ID \times 5 mm, Agilent Technologies) for 3 min at a flow rate of 50 μ L/min with 100% solvent A and then eluted and separated on a RP-HPLC column (Jupiter Proteo, 4 μ m, 90 Å, 1 mm ID \times 50 mm, Phenomenex) using a linear gradient from 5 to 95% solvent B in 15 min.

PAP9-6His was also analyzed by native MS (Boeri Erba and Petosa, 2015; Boeri Erba et al., 2020). Protein ions were generated using a nanoflow ESI (nano-ESI) source. Nanoflow platinum-coated borosilicate ESI capillaries were bought from Thermo Electron SAS (Courtaboeuf, France). MS analyses were carried out on a quadrupole time-of-flight mass spectrometer (Q-TOF Ultima, Waters Corporation, Manchester, United Kingdom). The instrument was modified for the detection of high masses (Sobott et al., 2002; van den Heuvel et al., 2006). The following instrumental parameters were used: capillary voltage = 1.2–1.3 kV, cone potential = 40 V, RF lens-1 potential = 40 V, RF lens-2 potential = 1 V, aperture-1 potential = 0 V, collision energy = 30–140 V, and microchannel plate (MCP) = 1900 V. All mass spectra were calibrated externally using a solution of cesium iodide (6 mg/mL in 50% isopropanol) and were processed with the Masslynx 4.0 software (Waters Corporation, Manchester, United Kingdom) and with Massign software package (Morgner and Robinson, 2012).

Solution-State NMR

One milligram of the 34 amino-acids C-terminal peptide of PAP9 was dissolved in 25 mM Na phosphate, pH 6.5 to a final concentration of 1 mM. For assignment of the peptide, homonuclear TOCSY, NOESY, and sensitivity-enhanced 13 C-HSQC experiments were recorded at 25°C on a Bruker ADVANCE III spectrometer operating at 1 H frequency of 600 MHz and equipped with a triple resonance pulsed field gradient cryoprobe.

For assignment of 6His-PAP9, 100 μ M of 15 N, 13 C-6His-PAP9 in a 90:10 H₂O:D₂O, 10 mM Tris pH 8.0, 50 mM NaCl were used. Heteronuclear 3D Best-TROSY-HNCA, Best-TROSY-HNCACB, Best-TROSY-HNCOCANH (Favier and Brutscher, 2011; Solyom et al., 2013), sensitivity-enhanced 13 C-HSQC and 15 N-SOFAST experiments were recorded at 298 K on Bruker ADVANCE III HD spectrometers operating either at 1 H frequency of 600 or 700 MHz and equipped with a triple resonance pulsed field

gradient cryoprobe. [15 N, 1 H]-TRACT (to estimate the global correlation time) (Lee et al., 2006) and DOSY experiments (for measuring the translational diffusion) (Morris and Johnson, 1992) were recorded at 298 K on a Bruker ADVANCE III HD spectrometer operating at 1 H frequency of 700 MHz.

Crystallization, Data Collection, and Structure Resolution

6His-PAP9 and PAP9-6His at 5 mg/mL in 10 mM Tris-HCl, pH 8.0, 50 mM NaCl (+10% glycerol for 6His-PAP9) were subjected to crystallization using the sitting-drop vapor-diffusion technique and the high throughput crystallization facility at the EMBL, Grenoble, at 4°C. Crystallization hits were optimized using Limbro plates, at 293 K. Crystals of PAP9-6His were grown in PEG3350 from 15 to 19%, 0.1 M Bis-Tris pH 6.5, 0.2 M NaNO₃, for data collection. Crystals of 6His-PAP9 were grown in Bis-Tris pH 7.5, PEG3350 18%, 0.2 M NaNO₃.

Diffraction data for PAP9-6His were collected on ID23-1 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France, at 100 K, using a PILATUS detector and two crystals. Anomalous data at the peak and after the peak of the zinc K-edge for PAP9-6His and native data for 6His-PAP9 were collected on FIP-BM30A (Roth et al., 2002) at the ESRF, at 100 K, using an ADSC 315r detector. Diffraction data (Table 1) were processed and scaled using XDS (Kabsch, 2010).

Phasing was performed by molecular replacement using Phaser (McCoy et al., 2007) from CCP4 (Collaborative Computational Project, Number 4 (CCP4), 1994). To calculate the phases, the crystal structure of the eukaryotic FeSOD from *Vigna unguiculata* (PDB entry: 1UNF) (Muñoz et al., 2005) was used as a model after modifications based on sequence alignment with PAP9 from *A. thaliana* using CHAINSAW (Stein, 2008) from CCP4. The refinements and rebuilding were done using PHENIX (Adams et al., 2010) and COOT (Emsley et al., 2010), respectively. The model refinements were performed with the non-crystallographic symmetry and the water molecules were added using PHENIX in the last stages of the refinement. Refinement statistics are summarized in Table 2. Atomic coordinates and X-ray data for PAP9-6His were deposited in the PDB with the accession number 7BJK. Since 6His-PAP9 is similar to PAP9-6His, the diffraction data and the 3D-structure were not reported in the PDB.

RESULTS

Phylogeny of PAP9 in the Green Lineage

Significant sequence similarities with At-PAP9 were found as early as in clades representing the chlorophytes, indicating that salt-water algae acquired plastid-localized SODs early in evolution. However, sequence alignments (Figure 1) identified a critical domain, outside of the SOD catalytic domain (Figure 2A), at the C-terminal (C-ter) of the protein, which had strongly changed during evolution. Whereas absent in early separated clades (as represented by Chlamydomonas), a significant insertion after the last well-conserved arginine

TABLE 1 | Statistics of data collection.

	PAP9-6His	PAP9-6His	PAP9-6His
Wavelength (Å) and beamline	0.976250 (ID23-1)	1.280867 (FIP-BM30A)	1.284809 (FIP-BM30A)
Resolution range (Å)	48.20–2.25 (2.31–2.25)	107.0–2.59 (2.75–2.59)	48.69–3.14 (3.33–3.14)
Space group	C2	C2	C2
Unit cell parameters (Å, °)	a = 214.09, b = 83.01, c = 118.24, β = 115.759	a = 215.36, b = 83.39, c = 118.65, β = 115.57	a = 217.63, b = 83.86, c = 120.33, β = 116.13
Molecules in au	5	5	5
Number of total reflections	321,204 (13,098)	436,955 (66,107)	251,369 (38,638)
Unique reflections	83,998 (5642)	115,026 (17,963)	65,945 (10,382)
Average multiplicity	3.82 (2.32)	3.80 (3.68)	3.81 (3.72)
Data completeness (%)	94.5 (86.0)	99.0 (95.9)	99.2 (96.9)
R_{sym} (%)	10.8 (77.9)	13.5 (80.4)	15.1 (69.5)
$\langle I/\sigma(I) \rangle$	7.87 (1.03)	8.65 (1.82)	8.87 (2.00)
CC (1/2) (%)	99.5 (60.5)	99.2 (73.0)	99.1 (71.0)

$R_{\text{sym}} = \sum \sum |I_i - I_m| / \sum \sum I_i$, where I_i is the intensity of the measured reflection and I_m is the mean intensity of this reflection. Values indicated in parentheses correspond to the statistics in the highest resolution shell.

(Arg262) is found in *Selaginella* with a large proportion of acidic residues representing one third of the amino acids (Figure 2B). The C-terminal of PAP9 in its long form (i.e., 40 residues) is not essential in higher Angiosperms since different clades have a shorter domain of approximately 20 residues in *Physcomitrella*, basal clades of the ANA grade, *Apiales* from Eudicots, *Alismatales*, and *Asparagales* from Monocots. Interestingly, the PAP9 C-terminus is either totally absent in *Gyngko* and *Pinus* or present as the short sequence in *Picea*, suggesting that there is no *bona fide* PAP9 referring to the involvement of the protein to the PEP function. These observations corroborate the hypothesis according to which Gymnosperms had favored a different use of PEP complex canceling the use of some PAPs that are not found anymore in the clade. In most Eudicots, a largely acidic tail with a well-conserved tyrosine (Figure 2C) may be involved in the PEP function as it could also play

the role of electron donor with manganese clusters or as a signaling residue.

Subcellular Localization of PAP9-GFP Proteins

Some of the proteins associated to the PEP, like PAP9, possess a predicted NLS (Pfannschmidt et al., 2015). However, the putative NLS of PAP9 (Figure 2A and Supplementary Table 4) is nested within the cTP (Figure 2A and Supplementary Tables 3,4), which is conceptually cleaved off during plastid import through the TOC/TIC machinery. Hence the question arises whether the predicted sequence is actually a *bona fide* NLS. Since the NLS sequence at this position is not conserved in other species, it does not likely play an important role in PAP9 localization. This is experimentally supported by the transient localization of PAP9-GFP (Figures 3A,B), which appears to be mostly plastidial. However, the clear labeling of the stromules (Figure 3B), indicates that a part of the pool of fluorescent molecules is found in the stroma, released from the PEP/PAP complex. In some images, we could also detect some signals in the cytosol and nucleus (Supplementary Figure 2). The GFP fluorescent profile across plastids is more spread than that of the RFP, indicating that the PAP9-GFP signal is not as restricted as that of PAP10-RFP used here as specific marker of the PEP complex (Liebers et al., 2020). The translational fusion of GFP at the C-terminus may alter the function of the corresponding domain so that the localization may not reflect precisely that of PAP9. Such a perturbation has been observed for HMR/PAP5 (Chen et al., 2010) and pTAC6/PAP8 (Liebers et al., 2020) for which C-terminal GFP fusions alter the localization and/or the functionality of the protein.

Mass Spectrometry Analyzes

We utilized MS to assess the mass of PAP9-6His and ^{15}N , ^{13}C -6His-PAP9 under denaturing conditions. The experimental mass of PAP9-6His was 30,848 Da, matching the amino acidic

TABLE 2 | Refinement statistics.

	PAP9-6His
Resolution (Å)	48.20–2.25 (2.28–2.25)
R_{cryst} ($\sigma_F = 0$) (%)	17.94 (33.96)
R_{free} ($\sigma_F = 0$) (%)	22.10 (38.11)
Number of atoms	8997
Water molecules	399
B average (Å ²)	51.82
RMSD bonds (Å)	0.007
RMSD angle (°)	0.884
Ramachandran favored (%)	91.5
Ramachandran allowed (%)	7.4
Ramachandran disallowed (%)	0.5

Values indicated in parentheses correspond to the statistics in the highest resolution shell.

$R_{\text{cryst}} = \sum ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum |F_{\text{obs}}|$. R_{free} (Brünger, 1992) is the same as R_{cryst} but calculated for 5% data omitted from the refinement.

AtPAP9 **MMNVAVTATPSSLLYSPLLLPSQGP** **NRRMQWK**
 Poales2 MLGF PSLCPT SLPS.R
 Caryophyllales MSVAATSSISPSSSIFTHEGFKGGLNSIRFSPWK.
 Lamiales3 MLEGF F
 Vitis MLATPAPVSLASALLPAQGW KSSRSLLR.
 Magnoliids MVFQQLL L SRPCCQNPALLT.TTDICRPLR...IPNLH.
 Amborella MAMASQLLL LPNLRFHNSQSAPRVI.SNPICRPLK...LSKFQ.
 Selaginella MPPGCCA TAAAAIASIA PPRI SSN R
 Physcomitrella MGLRVSGIALRRVVVVVVVRF QLWENQAMATSSLA GSVFS.SNTRLNPIG...N
 Picea MAL.L.QHSLLLQHRPLH.TNIFRTKTK...T
 Marchantia MPRALHPIPTPHFPFSQRESIIPPSATVLSSERAAMAAVTGTRCLLPVPSLR.SSKAR...P...S
 Chlamydomonas MALAMKAQASSLVA..GQR..RAVR...P...A.....

AtPAP9
 1 10 20 30 40 50 60
 α1 α2
 AtPAP9 **NGKRRLLGTVAVS**GVITAGFELKPP**PY**PLD**ALEP**HM**SRET**LDY**HWGKH**HKTYVENL**N**KQILG.TDLDALSLEEIVLLSYN
 Poales2 .QSRRRNNVRRRAPVAAALLELKPP**PY**PLD**ALEP**HL**SREAVEH****HWGRH**QRSHVDC**L**NAAIAG.TQLEGMLEMD.....
 Caryophyllales .LKEQTRVRRAGAGVVAKFDLKP**PY**FD**ALEP**HM**SQTFEY****HWGKH**HRAYVDN**L**KQIVG.TELDGLPLEEVVRYITN
 Lamiales3 .LQKLNIRRAAGGVKAIIEELKPP**PY**PLD**ALEP**HM**SRET**F**Y****HWGKH**HRAYVDN**L**KQIVN.TELDGKRLLEEIIIDTNY
 Vitis .GKRRTCSSKGNSSLIIAKFELKPP**PY**PL**S****ALEP**HM**SRET**E**Y****HWGKH**HRRGYVDN**L**RQIVG.TELDMGTLEDIIITITN
 Magnoliids .LRKRQSCVSRSSKIFSFGLTTP**PY**KL**D****ALEP**YM**SREALEH****HWGVH**HRDYVDS**L**KQLENNDPIYGYTMEEMVKFTYN
 Amborella .KRYKQS...LRDSRVLAFYGRKTP**PY**KL**D****ALEP**YM**SQRT**L**V****HWGKH**HQEVYDGL**N**KQLEN.SPIYGTLEELVKVTYN
 Selaginella .GGSKRSQCNRGIARAKFALPAL**PY**ESS**ALEP**YM**S**KKT**L****V****HWGKH**QRGYIEN**L**RQIAN.TPLEGYLLEDIVKISYN
 Physcomitrella .VATMRKRLVRMRAPRAEFLRPL**PY**EL**D****ALEP**HM**S**K**T**L**F****HWGKH**HKAYVDN**L**KQVEG.TEAERQNLSELVLSAYN
 Picea .RGAMLR.SRMAPLTVLFLGKEP**PY**KL**D****ALEP**FM**S**RRT**L****F****HWGKH**HRRGYVDN**L**KQIEG.NELEGFTLEELIKVTYN
 Marchantia .NATSLKVKLPLGAPVAKFDLAD**PY**PL**D****ALEP**HM**S**K**T**L**F****HWGKH**HRAYVNN**L**NGQIEN.SDLERNTLEEIVQSSYN
 Chlamydomonas .SGR.....RAVITRAALELKSP**PY**AL**D****ALEP**HM**S**Q**T**L**F****HWGKH**HRAYVDN**M**KQVAG.TPLDGSLEEVIVLASWN

AtPAP9
 70 80 90 100 110 120 130 140
 α3 α4 α5 β1
 AtPAP9 KGNMLPAFNAAQ**WNH**EF**F**WESIQ**GG**GKPT**G**ELLRL**I**ERDFG**S**FEEFLERFKSAASN**FGSGW**W**L**AYKANRLDVAN
 Poales2I**WNH**DF**F**WQSMK**PGGG**GEPS**G**RLMEL**I**ERDFG**S**YERMIKELKHAAL**TQ****FGSGW**AW**L**AYKANRLVEGN
 Caryophyllales KGDMLPSFNAAQ**WNH**EF**F**WESMK**PGGG**GKPS**G**ELLAQ**I**EKDFG**S**FEAFVNEFKAAAT**Q****FGSGW**AW**L**VYKANKLDVGN
 Lamiales3 NGDFLPPFNAAQ**WNH**EF**F**WGSMK**PSGG**GEPS**G**DLDEL**I**NRDFG**S**FEAFVNEFKAAAT**Q****FGSGW**AW**L**VYKANKLDVGN
 Vitis KGDLLPAFNAAQ**WNH**TS**F**WESMK**PGGG**GEPS**G**DLLE**L**I**K**DRFG**S**FERFVEEFKAAST**Q****FGSGW**AW**L**AYKANRLDVGN
 Magnoliids SGNPLPEFNVAE**WNH**DF**F**WESMQ**PEGG**KTP**S**GGVLQ**I**EKDFG**S**FEKFRDEFVEAALS**L****FGSGW**W**L**VLKT.....
 Amborella NGNPMPFNAAQ**WNH**DF**F**WESMQ**PEGG**KSP**W**GVLQ**I**EKDFG**S**LEKFRDEFVEAALS**L****FGSGW**W**L**VLKT.....
 Selaginella RGSPQAFNAAQ**WNH**DF**F**FTSL**K**PG**GG**KSP**T**GEIA**L**IDK**Q**L**S**YDNFVKEFKLAAAS**Q****FGAGW**AW**L**AL.....
 Physcomitrella DGNPKAYFNAAQ**WNH**EF**F**WSSMS**PSGG**GE**P**GE**L**LE**L**ISSNFV**S**YKFAEQFKRAAGS**Q****FGSGW**W**L**VV.....
 Picea NGNPMPAFNAAQ**WNH**DL**F**WECME**PEGG**GDCT**G**EMIQ**E**I**R**DFG**S**YNNFVEEFKQAA**TQ****FGSGW**W**L**VV.....
 Marchantia NGNP TPAFNAAQ**WNH**DF**F**WQSMK**PGGG**GP**T**GEIA**L**INRDFG**S**YDNFVKEFKAGGAT**Q****FGSGW**W**L**VL.....
 Chlamydomonas NGQP TPFVFNAAQ**WNH**T**F**WESMK**PN****GG**GAP**T**GALAE**A**I**T**ERDFG**S**LDFKKEEFKQAGM**TQ****FGSGW**AW**L**NAD.....

AtPAP9
 150 160 170 180 190 200 210 220
 β2 β3 α6 α7
 AtPAP9 AVNPLPKEEDKLVIV**KTP****N**AVN**P**LVWDYS**PL**LTID**T**WEHAY**Y**LD**F**FENR**R**AE**Y**INT**F**MEK**L**VSW**E**T**V**ST**R**LES**A**IARAVQ
 Poales2 AVNCPSEKDNKLVV**KTP****N**AVN**P**LLWDYS**PL**LAID**V**WEHAY**Y**LD**F**YENR**R**AD**Y**VST**F**LDK**L**VSW**E**VV**S**SR**L**EN**A**INRAT**E**
 Caryophyllales AVNRPSEDDKLVIV**KSP****N**AVN**P**LVWDY**Y**PL**L**TID**V**WEHAY**Y**LD**F**QNR**R**PD**Y**ISI**F**EMEN**L**VSW**D**AVN**A**R**Y**EAA**K**SLATE
 Lamiales3 AANPLPSDDDKLVV**KSP****N**AVN**P**LVWDY**S**PL**L**TLD**V**WEHAY**Y**LD**F**QNR**R**PD**Y**ISI**F**LEK**L**VSW**E**AV**S**SR**L**ES**A**KARAE
 Vitis AVNRPSEEDKLVV**KSP****N**AIN**P**LLWDC**N**PL**L**TID**V**WEHAY**Y**LD**F**QNR**R**PD**Y**ISI**F**MEK**L**VSW**E**AV**S**SR**L**ES**A**KAQVAN
 MagnoliidsQEKRLAVV**KTS****N**AVN**P**LCVGD**I**PIIN**L**DM**W**EHAY**Y**LD**F**YKDN**R**KV**V**NI**F**MNH**L**VSW**N**AA**M**GR**M**AR**A**ES**F**VNL
 AmborellaKEK**K**F**V**M**KTS****N**AVN**P**LVG**E**I**P**IG**L**DM**W**EHAY**Y**LD**F**YKDD**R**KV**M**T**L**NH**L**VSW**H**T**V**AL**R**M**I**R**A**E**F**VNL
 SelaginellaKEAK**L**V**KTS****N**AIN**P**LIWDH**I**PL**L**VD**V**WEHAY**Y**LD**F**QNR**R**PD**Y**VST**F**LEH**L**VSW**D**AVN**A**R**L**DR**A**KAFM**N**W
 PhyscomitrellaKDG**K**LE**L****KTP****N**AL**T**PIV**W**GH**I**PL**L**V**D**WEHAY**Y**LD**F**YHN**R**RP**Y**LS**K**LD**N**VSW**Q**AV**G**R**L**AQ**A**KAFV**N**Y
 PiceaKDS**K**L**A**VE**KTS****N**AL**N**PI**L**W**H**I**P**LL**T**ID**V**WEHAY**Y**LD**F**QNR**R**AD**Y**IS**A**MD**N**L**V**SW**D**AVN**A**R**F**GR**A**QAFAN**L**
 MarchantiaKNG**K**L**A**IE**KTP****N**AV**T**PI**L**W**G**S**I**PL**L**V**D**WEHAY**Y**LD**F**QNR**R**PD**Y**LS**T**V**N**E**L**IS**W**DAV**N**R**L**AR**A**KASV**N**L
 ChlamydomonasKTG**K**L**S**IS**KSP****N**AV**N**P**V**VE**G**KT**P**IL**T**LV**D**WEHAY**Y**LD**V**QNR**R**PD**Y**IT**T**F**M**E**K**L**N**W**D**AV**A**Q**R**Y**A**AA**T**K.....

AtPAP9
 230 240 250
 AtPAP9 REQEGTE...TEDEENPDDEVPVYLDSDIDVSEVD..
 Poales2 RAKVDEKR.RQDDVEATSRKPVEMYLDSDNDSETE..
 Caryophyllales RGKEAVQN..
 Lamiales3 REKEELE.KREYEE**T**DRP**I**SEATEVY**L**ESDADIAEAE..
 Vitis REREERRKRAEEEEQMPYSEAVKMYLESDDGGDSEAE
 Magnoliids GEPRI**P**VA..
 Amborella GEPK**I**PVA..
 Selaginella GEPDWLFDKDA.....DILTYEEAGIDVVEDSS.
 Physcomitrella GEPV**T**SEL..
 Picea REPI**E**PEL..
 Marchantia GTPS**I**PEQ..
 Chlamydomonas

FIGURE 1 | PAP9 secondary structures mapping on a sequence alignment including orthologous proteins from different clades of the green lineage. The PAP9 secondary structure from *Arabidopsis thaliana* is drawn as followed: the α -helices are displayed as squiggles and β -strands as arrows. The conserved residues are highlighted in red. The residues involved in the metal binding, Zn^{2+} in the crystal structure of *A. thaliana* PAP9, are indicated with a blue triangle. The cTP and NLS of the *A. thaliana* PAP9 are highlighted in green and magenta, respectively. The drawing was prepared using ESPript (Robert and Gouët, 2014).

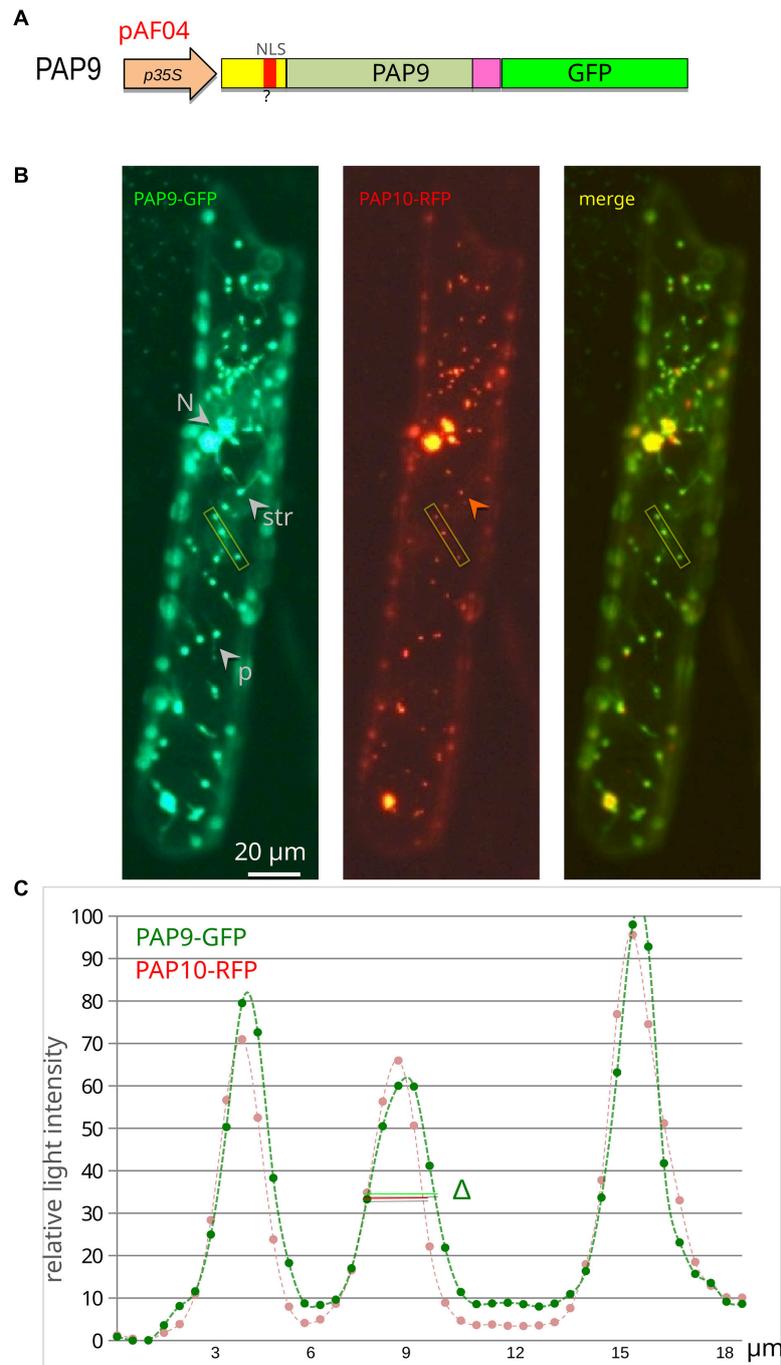
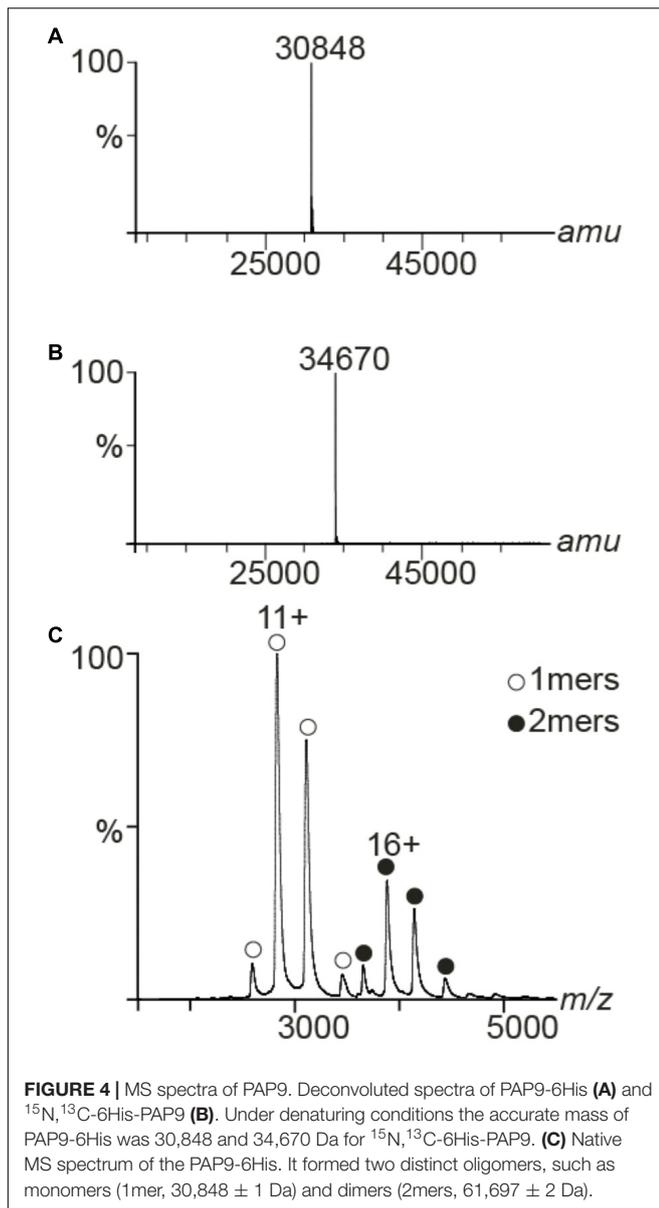


FIGURE 3 | PAP9 is localized in plastids. **(A)** Schematic illustration of the PAP9-GFP construction in pAF04. p35S, CaMV35S promoter region; cTP, chloroplast transit peptide in yellow; ?, predicted NLS (nuclear localization signal) in red; C-terminal domain in magenta; GFP in green. **(B)** Transiently expressed PAP9-GFP in onion epidermal cells. N, nucleus; str, stromule; p, plastid. The red arrowhead points to the absence of red fluorescence in stromules. The yellow rectangle represents the analyzed segment in panel **(C)**. **(C)** Fluorescent signal quantitative profile on an 18- μm -long segment of the image across three plastids. Δ represents the difference in width of the GFP signal compared to the red signal of PAP10-RFP.

X-Ray Structure Analyzes

Five molecules of PAP9 are in the asymmetric unit. Four of them form two dimers. The fifth interacts with a molecule from another asymmetric unit to form also a dimer. Both monomers

in the dimer are related by a non-crystallographic twofold axis. The monomers are very similar with a value of root mean square deviation (RMSD) ranging from 0.14 to 0.21 \AA between monomers when calculated between the $\text{C}\alpha$ atoms. The buried



area calculated using PISA (Krissinel and Henrick, 2007) in the dimer interface is 1785 \AA^2 . PAP9 is folded in two domains similar to those observed in FeSODs or MnSODs. The N-terminal domain extends from Gly1 to Gly93 and contains three α -helices. The C-terminal domain (Gly94–Gln229) displays an α/β fold with a three anti-parallel β -strands sandwiched by four α -helices and the N-terminal domain (Figure 5). No electron density is observed for residues from Arg141 to Glu155 and for the last 29 residues from Gly231 to Asp259 suggesting flexibility. Crystallographic analysis of the 6His-PAP9, produced to decrease the C-terminal flexibility, did not allow to better observe the electron density of the C-terminal part and the structures of 6His-PAP9 and PAP9-6His were similar. The catalytic center is at the interface of the N- and C-terminal domains. Surprisingly a zinc ion, instead of the expected iron ion, is penta-coordinated in

the catalytic center. Anomalous difference electron density map calculated at the zinc K-edge showed a strong peak of anomalous density (Figure 6) while the map computed with diffraction data collected just after the zinc K-edge does not show any strong peak. The zinc ion is penta-coordinated by the His31, His83 side chains of the N-terminal domain, the Asp182, and His186 side chains from the α/β fold domain, and a water molecule supposed to mimic the position of the hydroxide ion (Figure 6). The arrangement of the five coordinating ligands forms a trigonal bipyramid with His31 and the water molecule as the axial ligands. The side chains of His35, Tyr39, Gln79, and Trp184 close the catalytic site (Figure 6). Since PAP9 mainly binds Zn^{2+} in our expression/purification steps no catalytic activity could be observed excepted at very high PAP9 concentrations (Supplementary Figure 1).

Structure Comparisons and the PAP9 Family

Rms deviations calculated using PDBfold (Krissinel and Henrick, 2004) between the monomer of PAP9 and more than 200 monomers of SODs from the PDB range from 0.71 \AA (PAP9 vs. FeSOD from *V. unguiculata*, PDB entry: 1UNF) to 1.6 \AA with the FeSOD from *Aquifex pyrophilus* (PDB entry: 1COJ) (Lim et al., 1997). The structure comparisons revealed that the fold of PAP9, the ligands involved in the metal coordination and residues closing the catalytic site are conserved. Dimer interface comparison with FeSOD from *V. unguiculata* revealed also a conservation of residues involved in interactions by hydrogen bonds between the subunits. The Glu185 carboxylate group from one monomer interacts with the Ser130 hydroxyl group involving a water molecule and also with the His186 imidazole ring of the catalytic center from the other monomer. Additionally, the hydroxyl group of Ser130 interacts with the hydroxyl group of Ser130 from the other monomer (Figure 7). The main difference originates from the metal center occupied by a zinc ion in AtPAP9 instead of an iron ion. The conserved interaction described in FeSOD from *V. unguiculata* between His35 of one monomer and Tyr188 of the other monomer is not observed in PAP9. The residues Gly156 to Ser164 of the cytosolic FeSOD from *V. unguiculata* corresponding to Val144 to Pro152 of the flexible loop Arg141–Glu155 in PAP9 are not observed in the electron density.

Sequence comparisons between PAP9 and SODs of the PDB showed that the flexible C-terminal part (Gly231 to Asp259) of PAP9 is not observed in the sequences of SODs of the PDB. The longest C-terminal extension is observed in FeSOD of *Helicobacter pylori* (PDB entry: 3CEI) (Esposito et al., 2008). However, it is 19 residues shorter than in PAP9 and is folded as a kinked α -helix that interacts with the N-terminal domain. The 29 last residues unobserved in the electron density map of PAP9 are found in several sequences reported as plastid SODs. Indeed, the PAP9 C-terminal part alone, used in alignment searches of the UniProtKB database restricted to plants, matches FeSODs; some of which being not annotated as plastid-localized, despite individual detection of a chloroplast transit peptide using the ChloroP prediction tool. Most of hits are *bona fide*

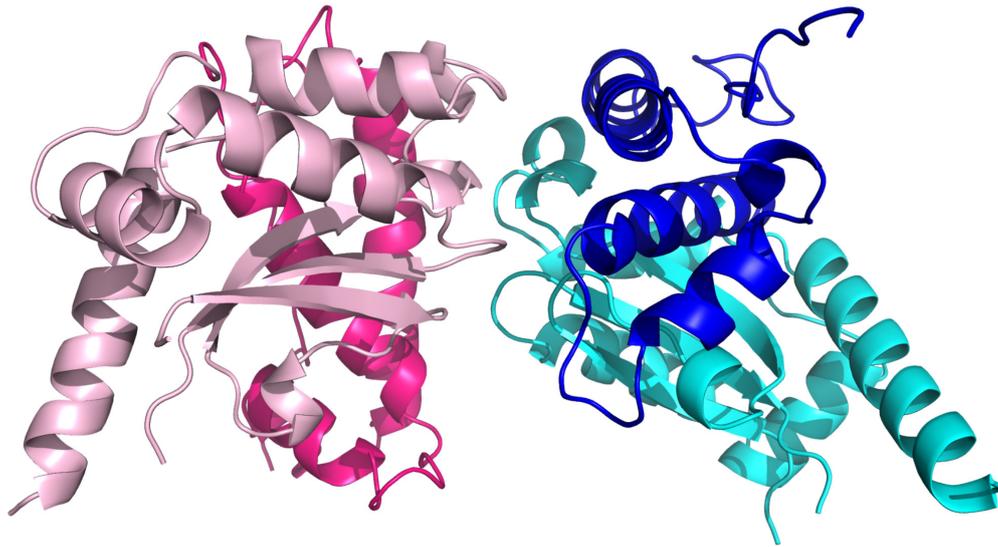


FIGURE 5 | View of the PAP9 dimer. The β -strands are drawn in arrows and the α -helices are represented in ribbons. The N-terminal domains are colored in dark pink and dark blue. The C-terminal domains are in cyan and light pink.

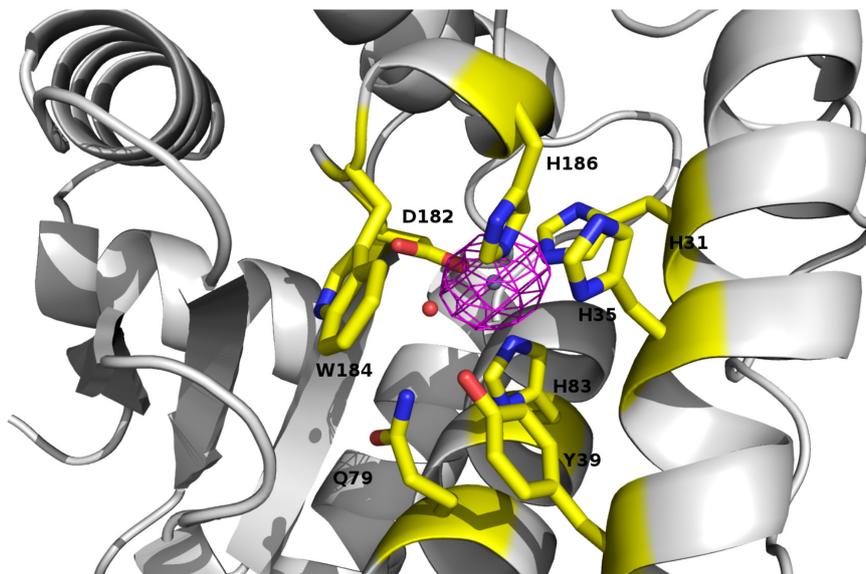


FIGURE 6 | View of the catalytic site of PAP9 superimposed with the anomalous electron density map calculated at the zinc K-edge. Residues of the catalytic site and closing the catalytic site are drawn in sticks. The zinc ion is drawn as gray sphere. The water molecule corresponding to the hydroxide ion is represented as a red sphere.

PAP9 orthologous SODs, and the C-terminal sequence represents a signature of this protein family. In addition, the sequence homology between PAP9 and PAP4/FSD3 (MW: 25657.94 Da) from *A. thaliana* is very high, suggesting that both FeSODs have a similar fold. However, PAP4 does not have the C-terminal extension found in PAP9. PAP9 and PAP4 should be functionally distinct and partially redundant as suggested by comparison of individual light-green phenotypes to the more severe albino phenotype of the double mutant (Myouga et al., 2008).

Solution-State NMR Analyses

Two segments, suggesting a dynamic structure, are not observed in the crystal structure of PAP9, i.e., the loop Arg141–Glu155 and the C-terminal part Gly231–Asp259 and are supposed to behave a fast dynamic. In order to further investigate the structural and dynamic properties of these unseen parts in the PAP9 crystal structure, we produced ^{15}N , ^{13}C -6His-PAP9. In our conditions (see section “Materials and Methods”), only about forty peaks can be observed above the background in the ^{15}N -SOFAS spectrum

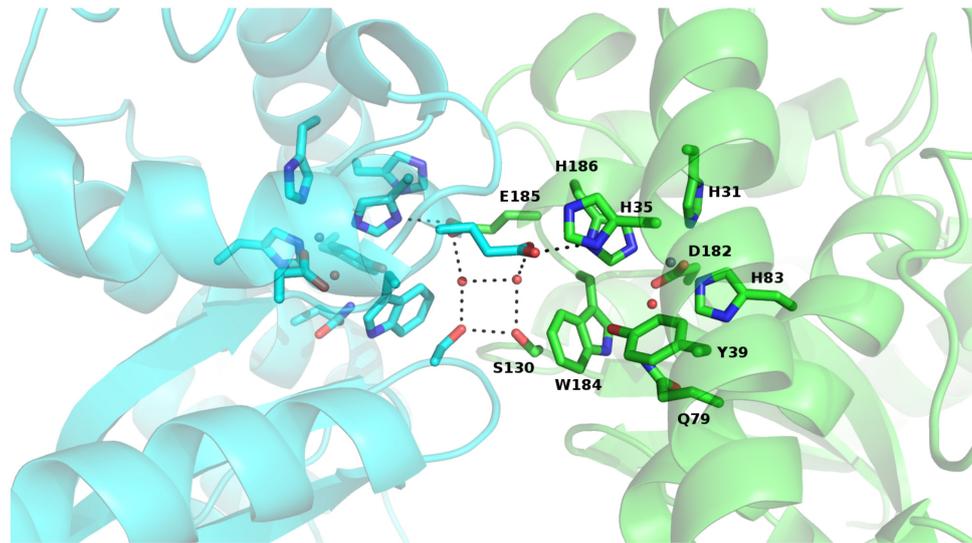


FIGURE 7 | View of the conserved interactions between both monomers of PAP9 *and* observed in FeSODs, with *each monomer* of PAP9 in a different color. The residues involved are drawn in sticks, the hydrogen bonds are represented in dark dashed lines and the water molecules are shown as spheres. The zinc ion is drawn as gray sphere. The β -strands are drawn in arrows and the α -helices are represented in ribbons.

in agreement with the presence of some dynamic residues. The most intense residues have an apparent rotational correlation time of 3 ns measured using $[^{15}\text{N},^1\text{H}]$ -TRACT technique (Lee et al., 2006), a value near those expected for free peptides or small proteins such as ubiquitin. In the other hand, the translational diffusion coefficient measured using DOSY experiment at 293 K is of $7 \times 10^{-7} \text{ cm}^2/\text{s}$, indicating that PAP9, from the point of view of translational diffusion, behaves like an object of 80 kDa. For such molecular weight, the residues located in the structured regions of the protein are expected to be line broadened supporting the fact that only the flexible residues can be observed in the NMR spectra. These results indicate that the observed residues have a fast movement while being included in a much larger species. We performed a set of 3D-experiments to assign these residues: HNCA, HNCACB, and HNCOCANH. Of these residues, only fifteen present detectable correlations in HNCACB experiments. A first analysis allows characterizing unambiguously a GTxTx sequence that corresponds only to the GTETE sequence located in the C-terminal tail of PAP9. In order to help to identify other residues within this part and characterize secondary structures, we studied a peptide composed of the 34 last residues of PAP9. We have entirely assigned the protons and carbons of the peptide using homonuclear TOCSY, NOESY, and ^{13}C -HSQC experiments at natural abundance. SSP program (Marsh et al., 2006) using $\text{C}\alpha$, $\text{C}\beta$, $\text{H}\alpha$ chemical shift data sets show that the peptide does not present any secondary structure propensity at all (Supplementary Figure 3). In the same way, the ^{13}C -HSQC experiment of the integer $^{15}\text{N},^{13}\text{C}$ -6His-PAP9 presents the very similar correlations than those observed for the peptide (Supplementary Figure 4), strongly suggesting that the C-terminal tail in 6His-PAP9 is also dynamic. Analysis of the observable $\text{C}\alpha$ and $\text{C}\beta$ chemical shift values in the protein together with comparison of those of the peptide allowed us

to assign the Gly231–Glu238 and the Ser251–Asp259 stretches. Assignments of Asn239, Val247–Leu249 can be proposed on basis of the HNCA experiment. The assigned ^1H - ^{15}N correlation spectrum of 6His-PAP9 is shown in Figure 8. No residue of the Gln226–Glu230 stretches were identified in agreement with their position in the last helix of the protein. Interestingly, the correlations of the residues, when observable, located in the middle of the tail: Asn239–Asp250 showed weaker intensities than those in the Gly231–Glu238 and Ser251–Asp259 stretches.

DISCUSSION

In Angiosperms, the developmental program following germination in the dark is skotomorphogenesis. Inside the cell, chloroplast biogenesis is blocked, allowing for the formation of yellow etiolated seedlings without the chlorophylls. After light perception etiolated seedlings start the photomorphogenesis program leading to chloroplast biogenesis (Liebers et al., 2018). This essential step toward photo-autotrophy involves the rapid assembly of the photosynthetic apparatus within the thylakoid membranes. Jointly, chlorophylls are quickly synthesized from the stored precursors, protochlorophyllides, by the light-activated protochlorophyllide oxidoreductase (POR). Chlorophylls are then inserted in the light harvesting antenna proteins. Transcription of photosynthesis associated plastid genes is ensured by PEP and is rapidly promoted after light perception owing to the PAP assembly into the active PEP complex. Two of the PAPs are FeSODs (Myouga et al., 2008; Steiner et al., 2011). FeSODs catalyze the dismutation of superoxide radicals into peroxides and may protect the transcriptional machinery from the newly acquired photosynthetic capacity (Pfannschmidt et al., 2015). Once

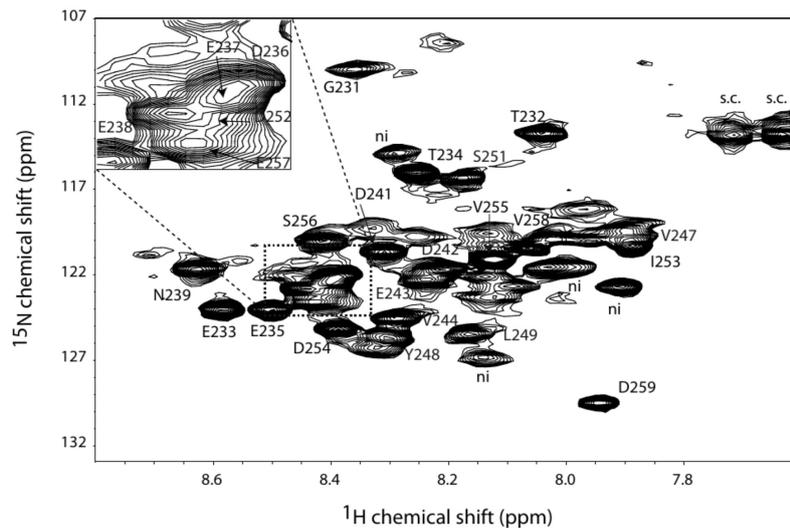


FIGURE 8 | ^1H - ^{15}N correlation spectrum of PAP9 with the assigned amino acid residue labels annotated "ni" standing for not identified.

the chloroplast is formed and fully photosynthetically active, the PEP activity substantially decreases.

Transmembrane translocation of PAP9 into the chloroplast results from the recognition of its N-terminal plastid transit peptide by the transmembrane TOC/TIC machinery. Fluorescence microscopy experiments showed that PAP9 is mainly located in the chloroplast stroma (Figure 3); the stroma localization may result from the lack of developed thylakoids in onion epidermal cells. Therefore, the predicted nuclear localization sequence observed within the cTP (Figure 2A and Supplementary Table 4) may not serve a localization purpose. It is cleaved off instead during the chloroplast import leading to a mature protein of 30,848 Da as observed using mass spectrometry analysis in denaturing conditions (Figure 4A). The native MS data indicated that PAP9 assembles as dimers. Monomers were also detected, suggesting protein dynamics during assembly. The ionization efficiency of the different oligomeric states affects the relative abundance of the different species in the MS spectra. Therefore, it is not possible to judge whether the monomers are more abundant than the dimers. Moreover, the native MS experiments were performed at 5 μM concentration and in ammonium acetate, which is a different buffer used for purification, NMR, and crystallographic experiments. The buffer conditions may affect the relative abundance of the species.

In the crystals, PAP9 is a symmetric dimer (Figure 5) as revealed by the low RMSD values between both monomers. The buried surface of the dimer interface suggests that the dimer is the biological form of PAP9. The FeSODs and MnSODs are active as dimeric or tetrameric (dimer of dimers) enzymes (Perry et al., 2010). In the PEP, PAP9 has been observed as a monomer (Steiner et al., 2011); a form of the protein also observed in our mass spectrometry analyses. The main difference between PAP9 analyzed here, and the FeSODs or MnSODs, is the metal ion bound to the catalytic site. In our crystal structure a zinc ion, instead of an iron ion, is penta-coordinated by a water

molecule, supposed to mimic the position of the hydroxide ion (Muñoz et al., 2005), the His31, His83, Asp182, and His186 side chains (Figure 6) as observed in the cytoplasmic FeSODs and MnSODs. The zinc ion cannot be the catalytic ion to perform the dismutation of superoxide since it has only the redox state II, in opposition to Fe and Mn that both have several redox states from II to VI and II to VIII, respectively. Since PAP9 is an active FeSOD even when overexpressed in *E. coli* (Myouga et al., 2008), the replacement of Fe by Zn occurred during either overexpression or purification of the protein although the metal center is hidden from the solvent by the conserved side chains of His35, Tyr39, Gln79, and Trp184 (Figure 6). As observed in FeSODs and MnSODs, His186 from the catalytic site of one monomer interacts also with the Glu185 carboxylate group from the other monomer. All these observations suggest a conserved catalytic mechanism in cytoplasmic and plastid FeSODs.

The main difference between PAP9 and the other FeSODs, and even MnSODs, is the additional residues of the C-terminal part. In the crystal structures of PAP9-6His and 6His-PAP9, no electron density was observed for the 29 last residues of the C-terminal part resulting from flexibility. Proteolysis can be excluded since the correct molecular weight of the 6His-tagged PAP9 was observed using mass spectrometry (Figure 4A). The flexibility does also not result from the construction of the over-expressed recombinant protein since the electron density of the C-terminal part is not observed for 6His-PAP9. The only observable residues of ^{13}C , ^{15}N -6His-PAP9 using NMR correspond essentially to the C-terminal residues whose dynamic is identical to that of the free peptide (Supplementary Figure 4). This result clearly shows that the C-terminal part is flexible with its central part (weaker intensities of correlations) not as free as the two other parts, probably due to some interactions of this part with residues at the protein surface. As in FeSOD from *V. unguiculata* (Muñoz et al., 2005), no electron density is observed for residues Val144 to Pro152 of the loop

Arg141–Glu155 suggesting flexibility in cytosolic FeSODs (FSD1) from plants and PAP9. The C-terminal extension observed in PAP9 could then allow distinguishing between PAP9, as a component of the PEP, and other plant FeSODs. We hypothesize that the C-terminal tail anchors PAP9 to the PEP complex and its observed flexibility arises from the isolation of a subunit that normally belongs to a larger multisubunit complex.

The C-terminal part of the protein had strongly changed during evolution (Figures 1, 2). It is absent in early clades of the green lineage. A first significant C-terminal modification is found in Charales and Physcomitrella while a second longer fragment appears in *Selaginella*. Such events are dating back to the conquest of fresh waters and terrestrial life. It is then possible that the C-terminal part could have appeared along with a complete set of new features for controlling chloroplast transcription; namely the assembly of PEP-PAP complex. The acquisition of these features, including SOD activities in a stoichiometry of four units per complex (three PAP4 and one PAP9), may provide sufficient protection of the organelle while the photosynthetic cells are exposed to a more oxidizing environment. This C-terminal part is totally absent in Gymnosperms, which seem to have evolved a completely different strategy of photo-autotrophy acquisition with, for example, no light regulation of chloroplast biogenesis since seedlings can green in darkness.

The PEP is composed of at least 16 subunits of unknown structures. Interactions between some of them were only reported by using non-direct observations, using yeast-two-hybrid assays (Yu et al., 2013) and fluorescent microscopy (Myouga et al., 2008). We have recently shown by NMR that PAP5 interacts with PAP8 (Liebers et al., 2020). PAP9 was proposed to interact with PAP4 therefore forming a hetero-complex of FeSODs (Myouga et al., 2008), and we show here that PAP9 can have several oligomeric states. Surprisingly, neither this heterocomplex nor the PAP9 dimer have been described (Steiner et al., 2011) suggesting that the PEP is probably a dynamic complex, still poorly characterized at the level of its structure and composition.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <http://www.wwpdb.org/>, 7BJK.

AUTHOR CONTRIBUTIONS

RB and DC designed the research. AF, PG, EBE, LS, SSM, RB, and DC performed the research. EBE and LS contributed mass

spectrometry data. AF and PG contributed NMR data. RB and DC wrote the manuscript with contributions from AF, PG, EBE, LS, and TP. All authors approved the manuscript.

FUNDING

This work used the platforms of the Grenoble Instruct-ERIC center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-0005-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). This work was supported by the Agence National de la Recherche (ANR-17-CE11-0031).

ACKNOWLEDGMENTS

The diffraction experiments were conducted on beamline FIP-BM30A and ID23-1 at the ESRF (Grenoble, France). We thank the beamline staff for technical help, Auriane Bron and Florence Prunier-Bossion for their technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.668897/full#supplementary-material>

Supplementary Figure 1 | Enzymatic assay of PAP9. The superoxide dismutase activity of PAP9 was tested using pyrogallol. The pyrogallol auto-oxidation was followed by monitoring the absorbance increase at 420 nm. After 180 s, PAP9 at several concentrations [50 (orange), 100 (gray), 200 (yellow), 500 μ M (light blue), and 1 mM (green)] or 5 μ M Mn-SOD (dark blue) were added into the medium and the absorbance was monitored for further 3 min.

Supplementary Figure 2 | Transiently expressed PAP9-GFP in onion epidermal cells. N, nucleus; str, stromule; p, plastid. The red arrowhead points to the absence of red fluorescence in stromules.

Supplementary Figure 3 | Secondary structure propensity (SSP) scores for the Cter-PAP9 peptide (circles) and the C-terminal tail of integer PAP9 (squares) using $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$, and $\text{H}\alpha$ chemical shifts. Between residues 245 and 252, the SSP score (star) was obtained from the $^{13}\text{C}\alpha$ and $\text{H}\alpha$ chemical shifts only. Positive values represent α -structure propensity and negative values represent β -structure propensity. The SSP is near zero along the sequence indicating the absence of any secondary structure in the peptide. The numbering of the residues corresponds to the whole protein.

Supplementary Figure 4 | Overlay of ^1H - ^{13}C correlation spectra (sensitivity-enhanced HSQC) of PAP9 and the Cter-PAP9 peptide. The peptide signals and the PAP9 peaks are shown in red and in black, respectively. All resonances of the peptide are observable in the PAP9 spectrum indicating the presence of the same mobility in the C-terminal tail of PAP9.

REFERENCES

Abreu, I. A., and Cabelli, D. E. (2010). Superoxide dismutases – a review of the metal-associated mechanistic variations. *Biochim. Biophys. Acta.* 1804, 263–274. doi: 10.1016/j.bbapap.2009.11.005

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst. D* 66, 213–221. doi: 10.1107/S0907444909052925

Bobik, K., and Burch-Smith, T. M. (2015). Chloroplast signaling within, between and beyond cells. *Front. Plant Sci.* 6:781. doi: 10.3389/fpls.2015.00781

- Boeri Erba, E., and Petosa, C. (2015). The emerging role of native mass spectrometry in characterizing the structure and dynamics of macromolecular complexes. *Protein Sci.* 24, 1176–1192. doi: 10.1002/pro.2661
- Boeri Erba, E., Signor, L., and Petosa, C. (2020). Exploring the structure and dynamics of macromolecular complexes by native mass spectrometry. *J. Proteomics* 222:103799. doi: 10.1016/j.jpro.2020.103799
- Börner, T., Aleynikova, A. Y., Zubo, Y. O., and Kusnetsov, V. V. (2015). Chloroplast RNA polymerases: role in chloroplast biogenesis. *Biochim. Biophys. Acta* 1847, 761–769. doi: 10.1016/j.bbabi.2015.02.004
- Brünger, A. T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355, 472–475. doi: 10.1038/355472a0
- Chen, M., Galvao, R. M., Li, M., Burger, B., Bugea, J., Bolado, J., et al. (2010). *Arabidopsis* HEMERA/pTAC12 initiates photomorphogenesis by phytochromes. *Cell* 141, 1230–1240. doi: 10.1016/j.cell.2010.05.007
- Collaborative Computational Project, Number 4 (CCP4) (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* 50, 760–763. doi: 10.1107/S0907444994003112
- Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004
- Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* 8, 978–984. doi: 10.1110/ps.8.5.978
- Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and development of Coot. *Acta Cryst. D* 66, 486–501. doi: 10.1107/S0907444910007493
- Esposito, L., Seydel, A., Aiello, R., Sorrentino, G., Cendron, L., Zanotti, G., et al. (2008). The crystal structure of the superoxide dismutase from *Helicobacter pylori* reveals a structured C-terminal extension. *Biochim. Biophys. Acta* 1784, 1601–1606. doi: 10.1016/j.bbapap.2008.04.024
- Favier, A., and Brutscher, B. (2011). Recovering lost magnetization: polarization enhancement in biomolecular NMR. *J. Biomol. NMR* 49, 9–15. doi: 10.1007/s10858-010-9461-5
- Gao, Z.-P., Yu, Q.-B., Zhao, T.-T., Ma, Q., Chen, G.-X., and Yang, Z.-N. (2011). A functional component of the transcriptionally active chromosome complex, *Arabidopsis* pTAC14, interacts with pTAC12/HEMERA and regulates plastid gene expression. *Plant Physiol.* 157, 1733–1745. doi: 10.1104/pp.111.184762
- Hirel, P. H., Schmitter, M. J., Dessen, P., Fayat, G., and Blanquet, S. (1989). Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. *Proc. Natl. Acad. Sci. U.S.A.* 86, 8247–8251. doi: 10.1073/pnas.86.21.8247
- Jarvis, P. (2008). Targeting of nucleus-encoded proteins to chloroplasts in plants. *New Phytol.* 179, 257–285. doi: 10.1111/j.1469-8137.2008.02452.x
- Jarvis, P., and López-Juez, E. (2013). Biogenesis and homeostasis of chloroplasts and other plastids (2013). *Nat. Rev. Mol. Cell. Biol.* 14, 787–802. doi: 10.1038/nrm3702
- Kabsch, W. (2010). XDS. *Acta Cryst. D* 66, 125–132. doi: 10.1107/S0907444909047337
- Kosugi, S., Hasebe, M., Tomita, M., and Yanagawa, H. (2009). Systematic identification of yeast cell cycle-dependent nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10171–10176. doi: 10.1073/pnas.0900604106
- Kremnev, D., and Strand, A. (2014). Plastid encoded RNA polymerase activity and expression of photosynthesis genes required for embryo and seed development in *Arabidopsis*. *Front Plant Sci.* 5:385. doi: 10.3389/fpls.2014.00385
- Krissinel, E., and Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Cryst. D* 60, 2256–2268. doi: 10.1107/S0907444904026460
- Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774–797. doi: 10.1016/j.jmb.2007.05.022
- Lee, D., Hilty, C., Wider, G., and Wüthrich, K. (2006). Effective rotational correlation times of proteins from NMR relaxation interference. *J. Magn. Reson.* 178, 72–76. doi: 10.1016/j.jmr.2005.08.014
- Liebers, M., Chevalier, F., Blanvillain, R., and Pfannschmidt, T. (2018). PAP genes are tissue- and cell-specific markers of chloroplast development. *Planta* 248, 629–646. doi: 10.1007/s00425-018-2924-8
- Liebers, M., Gillet, F. X., Israel, A., Pounot, K., Chambon, L., Chieb, M., et al. (2020). Nucleo-plastidic PAP8/pTAC6 couples chloroplast formation with photomorphogenesis. *EMBO J.* 39:e104941. doi: 10.15252/embj.2020104941
- Liebers, M., Grübler, B., Chevalier, F., Lerbs-Mache, S., Merendino, L., Blanvillain, R., et al. (2017). Regulatory shifts in plastid transcription play a key role in morphological conversions of plastids during plant development. *Front. Plant Sci.* 19:23. doi: 10.3389/fpls.2017.00023
- Lim, J. H., Yu, Y. G., Han, Y. S., Cho, S., Ahn, B. Y., Kim, S. H., et al. (1997). The crystal structure of an Fe-superoxide dismutase from the hyperthermophile *Aquifex pyrophilus* at 1.9 Å resolution: structural basis for thermostability. *J. Mol. Biol.* 270, 259–274. doi: 10.1006/jmbi.1997.1105
- Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., et al. (2012). Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. *Plant Physiol.* 158, 156–189. doi: 10.1104/pp.111.188474
- Marsh, J. A., Singh, V. K., Jia, Z., and Forman-Kay, J. D. (2006). Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Protein Sci.* 15, 2795–2804. doi: 10.1110/ps.062465306
- Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., et al. (2002). Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl. Acad. Sci. U.S.A.* 99, 12246–12251. doi: 10.1073/pnas.182432999
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Cryst.* 40, 658–674. doi: 10.1107/S0021889807021206
- Morgner, N., and Robinson, C. V. (2012). Massign: an assignment strategy for maximizing information from the mass spectra of heterogeneous protein assemblies. *Anal. Chem.* 84, 2939–2948. doi: 10.1021/ac300056a
- Morris, K. F., and Johnson, C. S. (1992). Diffusion-ordered two-dimensional nuclear magnetic resonance spectroscopy. *J. Am. Chem. Soc.* 114, 3139–3141. doi: 10.1021/ja00034a071
- Muñoz, I. G., Moran, J. F., Becana, M., and Montoya, G. (2005). The crystal structure of an eukaryotic iron superoxide dismutase suggests intersubunit cooperation during catalysis. *Protein Sci.* 14, 387–394. doi: 10.1110/ps.04979505
- Myouga, F., Hosoda, C., Umezawa, T., Izumi, H., Kuromori, T., Motohashi, R., et al. (2008). A heterocomplex of iron superoxide dismutases defends chloroplast nucleoids against oxidative stress and is essential for chloroplast development in *Arabidopsis*. *Plant Cell* 20, 3148–3162. doi: 10.1105/tpc.108.061341
- Perry, J. J., Shin, D. S., Getzoff, E. D., and Tainer, J. A. (2010). The structural biochemistry of the superoxide dismutases. *Biochim. Biophys. Acta* 1804, 245–262. doi: 10.1016/j.bbapap.2009.11.004
- Pfannschmidt, T. (2003). Chloroplast redox signals: how photosynthesis controls its own genes. *Trends Plant Sci.* 8, 33–41. doi: 10.1016/s1360-1385(02)00005-5
- Pfannschmidt, T., Blanvillain, R., Merendino, L., Courtois, F., Chevalier, F., Liebers, M., et al. (2015). Plastid RNA polymerases: orchestration of enzymes with different evolutionary origins controls chloroplast biogenesis during the plant life cycle. *J. Exp. Bot.* 66, 6957–6973. doi: 10.1093/jxb/erv415
- Pilon, M., Ravet, K., and Tapken, W. (2011). The biogenesis and physiological function of chloroplast superoxide dismutases. *Biochim. Biophys. Acta* 1807, 989–998. doi: 10.1016/j.bbabi.2010.11.002
- Robert, X., and Gouët, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucl. Acids Res.* 42, W320–W324. doi: 10.1093/nar/gku316
- Roth, M., Carpentier, P., Kaikati, O., Joly, J., Charrault, P., Pirocchi, M., et al. (2002). FIP: a highly automated beamline for multiwavelength anomalous diffraction experiments. *Acta Cryst. D* 58, 805–814. doi: 10.1107/s0907444902003943
- Sobott, F., Hernandez, H., McCammon, M. G., Tito, M. A., and Robinson, C. V. (2002). A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Anal. Chem.* 74, 1402–1407. doi: 10.1021/ac0110552
- Solyom, Z., Schwarten, M., Geist, L., Konrat, R., Willbold, D., and Brutscher, B. (2013). BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. *J. Biomol. NMR* 55, 311–321. doi: 10.1007/s10858-013-9715-0

- Stein, N. (2008). CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *J. Appl. Cryst.* 41, 641–643. doi: 10.1107/S0021889808006985
- Steiner, S., Schröter, Y., Pfalz, J., and Pfannschmidt, T. (2011). Identification of essential subunits in the plastid-encoded RNA polymerase complex reveals building blocks for proper plastid development. *Plant Physiol.* 157, 1043–1055. doi: 10.1104/pp.111.184515
- Sugiura, M. (1992). The chloroplast genome. *Plant Mol. Biol.* 19, 149–168. doi: 10.1007/BF00015612
- van den Heuvel, R. H., van Duijn, E., Mazon, H., Synowsky, S. A., Lorenzen, K., Versluis, C., et al. (2006). Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry. *Anal. Chem.* 78, 7473–7483. doi: 10.1021/ac061039a
- Yu, Q. B., Huang, C., and Yang, Z. N. (2014). Nuclear-encoded factors associated with the chloroplast transcription machinery of higher plants. *Front Plant Sci.* 5:316. doi: 10.3389/fpls.2014.00316
- Yu, Q. B., Lu, Y., Ma, Q., Zhao, T. T., Huang, C., Zhao, H. F., et al. (2013). TAC7, an essential component of the plastid transcriptionally active chromosome complex, interacts with FLN1, TAC10, TAC12 and TAC14 to regulate chloroplast gene expression in *Arabidopsis thaliana*. *Physiol. Plant.* 148, 408–421. doi: 10.1111/j.1399-3054.2012.01718.x
- Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q., et al. (2008). Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS One* 3:e1994. doi: 10.1371/journal.pone.0001994

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.

Copyright © 2021 Favier, Gans, Boeri Erba, Signor, Muthukumar, Pfannschmidt, Blanvillain and Cobessi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.