Proteomic analysis of haem-binding protein from *Arabidopsis thaliana* and *Cyanidioschyzon merolae*

Takayuki Shimizu¹,†, Rintaro Yasuda²,†, Yui Mukai², Ryo Tanoue³, Tomohiro Shimada³,†, Sousuke Imamura⁴, Kan Tanaka⁴, Satoru Watanabe² and Tatsuru Masuda¹

¹Graduate School of Arts and Sciences, The University of Tokyo, Komaba, Meguro-ku, Tokyo 153-8902, Japan
²Department of Bioscience, Tokyo University of Agriculture, Setagaya-ku, Tokyo 156-8502, Japan
³School of Agriculture, Meiji University, Kawasaki-shi, Kanagawa 214-8571, Japan
⁴Laboratory for Chemistry and Life Science, Institute of Innovative Research, Tokyo Institute of Technology, Yokohama-shi, Kanagawa 226-8503, Japan

†These authors contributed equally to this study.

Chloroplast biogenesis involves the coordinated expression of the plastid and nuclear genomes, requiring information to be sent from the nucleus to the developing chloroplasts and vice versa. Although it is well known how the nucleus controls chloroplast development, it is still poorly understood how the plastid communicates with the nucleus. Currently, haem is proposed as a plastid-to-nucleus (retrograde) signal that is involved in various physiological regulations, such as photosynthesis-associated nuclear genes expression and cell cycle in plants and algae. However, components that transduce haem-dependent signalling are still unidentified. In this study, by using haem-immobilized high-performance affinity beads, we performed proteomic analysis of haem-binding proteins from *Arabidopsis thaliana* and *Cyanidioschyzon merolae*. Most of the identified proteins were non-canonical haemoproteins localized in various organelles. Interestingly, half of the identified proteins were nucleus proteins, some of them have a similar function or localization in either or both organisms. Following biochemical analysis of selective proteins demonstrated haem binding. This study firstly demonstrates that nucleus proteins in plant and algae show haem-binding properties.

This article is part of the theme issue ‘Retrograde signalling from endosymbiotic organelles’.

### 1. Introduction

Haem serve as cofactors of haemoproteins in various organelles that function in mitochondria respiratory and chloroplast photosynthetic electron transport chains, and in the detoxification of reactive oxygen species and xenobiotics, as well as in oxygen storage and transport [1]. In addition, haem has been proposed to be a regulatory factor in control of transcription and intercellular signalling in yeast and animals [2,3].

The haem biosynthetic pathway begins with the synthesis of 5-aminolevulinic acid, the universal precursor of all tetrapyrroles. In photosynthetic organisms, the tetrapyrrole biosynthesis pathway branches into chlorophyll or haem synthesis, where the metabolite protoporphyrin IX (Proto) is the substrate of two structurally different metal chelatases. The Mg-chelatase converts Proto to Mg-protoporphyrin IX (MgProto) and ferrochelatase (FC) inserts Fe²⁺ into Proto to form haem (protohaem). All higher plants analysed so far possess two genes encoding FC (FC1 and FC2), which show differential tissue-specific and development-dependent expression profiles, such that FC2 is light-dependent and mainly...
expressed in photosynthetic tissues, whereas FCI is stress-responsive and ubiquitously expressed in all tissues [4,5]. Concerning the subcellular localization, the main FC activity is detected in chloroplasts and has very low activity in mitochondria [6,7], although the possibility of mitochondrial localization of FC cannot be excluded [8]. In the green algae Chlamydomonas reinhardtii, a single FC encodes a plastid-localized FC protein [9], while in the red algae Cyanidioschyzon merolae, FC is only found in mitochondrial extracts [10]. These results suggest that in Streptophyta and Chlorophyta, the dominant plastid FC activity supplies haem for the plastid as well as other organelle-localized haemoproteins, while distinct mitochondrial haem biosynthesis is employed in Rhodophyta. In these photosynthetic organisms, the function of haem is not limited to their roles as prosthetic groups, but they are also proposed to serve as signalling molecules [11,12].

Chloroplast biogenesis involves the coordinated expression of the plastid and nuclear genomes, requiring information to be sent from the nucleus to the developing chloroplasts and vice versa. The latter is achieved through plastid-to-nucleus (retrograde) signalling pathways in which plastids send a signal to regulate various physiological phenomena, such as photosynthesis-associated nuclear genes (PhANGs) expression [11], and cell cycle coordination [13], depending on their developmental and functional states. Genetic and biochemical analyses of this pathway suggest a major role for haem in retrograde signalling. In Arabidopsis thaliana, mutations affecting chloroplast function or treatments with inhibitors such as norflurazon (NF) or lincosmycin (Lin) result in the strong repression of many PhANGs. Characterization of genomes uncoupled (gun) mutants in which the expression of the nuclear gene Lhcb is maintained following chloroplast damage using NF treatment [14] suggests the involvement of tetrpyroles in retrograde signalling. Among the original five gun mutants described, gun2 and gun3 lack a functional haem oxygenase 1 and phytochromobilin synthase [15], and gun4 and gun5 are mutants of the regulator [16] and the H subunit of Mg-chelatase [15], respectively. More recently, the identification of a dominant gun6 mutant with increased FCI activity [17] restores PhANGs expression even when chloroplast development is blocked. These data suggest that increased flux through the FCI-producing haem may act as a signalling molecule that control PhANGs as a retrograde signal in A. thaliana.

Signalling function of haem is not limited in higher plants. In Ch. reinhardtii, haem along with MgProto has been proposed as a signalling molecule that may substitute for light [18]. Analysis of the transcriptome in Ch. reinhardtii showed that the expression of hundreds of genes was affected by exogenous haem treatment, but only a few of them were associated with photosynthesis [19]. In Cy. merolae, absicric acid (ABA) induced haem-scavenging tryptophan-rich sensory protein-related protein (TSPO), resulting in inhibition of the cell cycle G1/S transition [20]. Because the ABA-dependent inhibition of DNA replication was negated by addition of exogenous haem, it is proposed that ABA and haem have regulatory role in algal cell cycle initiation [20].

As described above, for assembly of holoproteins, haem is synthesized in plastids of A. thaliana and Ch. reinhardtii or in mitochondria of Cy. merolae should be transported to the appropriate cellular organelles, such as peroxisome, endoplasmic reticulum (ER) and nucleus. However, compared with bacteria, yeast and animals, the mechanism of haem trafficking from plastid or mitochondria to other organelles in photosynthetic organisms is still largely unknown. For membrane transport, involvement of the membrane-bound ABC (ATP-binding cassette) transporters and TSPO, was proposed in animal cells [11]. In fact, ABC transporters, such as ABCB6 and ABCG2/BCRP, are involved in tetrpyrole trafficking in mammalian cells [21,22] and Arabidopsis vacuolar ABC transporters AtMRP1–3 can transport chlorophyll catabolites to the vacuole during chlorophyll degradation [23]. In addition, homologues of TSPO in A. thaliana [24] and Cy. merolae [20] showed haem-binding properties and were induced by ABA treatment. However, the TSPO was localized to the secretory pathway [24]. In addition, because haem is poorly soluble in aqueous solutions under physiological conditions, involvement of haem carrier proteins was proposed [11]. The cytosolic p22HBP/SOUL protein which showed high affinity for haem was identified in animal cells [11]. A homologue of p22HBP/SOUL in A. thaliana was identified, which showed high affinity for haem, although its detailed function is unknown [25].

To elucidate the molecular mechanism of haem trafficking and signalling role, it is important to identify its molecular target(s). For this purpose, we have developed haem-immobilized high-performance affinity beads that allow single-step affinity purification of drug target proteins from crude cell extracts [26]. Here, we performed affinity purification of haem-binding proteins from A. thaliana and Cy. merolae cell extracts. Comparative analysis of these evolutionarily distant photosynthetic organisms will allow us to discuss shared features of the haem-binding proteins, as well as their diversity. Following proteomic analysis successfully identified possible candidate proteins that bind to haem. Our data suggest that haem is actually transferred into the nucleus and regulate not only transcription but also RNA metabolism and chromatin remodelling.

2. Material and methods

(a) Preparation of haemin-immobilized ferrite-glycidyl methacrylate bead

Magnetic ferrite-glycidyl methacrylate (FG) beads (5 mg) (Tama Seiki), were incubated with 10 mM 1-hydroxybenzo triazole, 10 mM 1-ethyl-3-(3-demethyl-aminopropyl)-carbodiimide HCl and 2 mM haemin in N,N-dimethylformamide for 4 h at room temperature. Unreacted residues were masked using 20% carbonic anhydride in N,N-dimethylformamide, and the resulting beads were stored at 4°C.

(b) Plant material and growth conditions

Arabidopsis thaliana wild-type (WT) was the Columbia-0 (Col-0) ecotype. Seeds were sown onto Murashige and Skoog medium supplemented with 1% (w/v) agar (pH 5.8) and incubated in white light (100 µmol m⁻² s⁻¹) for 2 h to induce germination. For protein extraction, seedlings were then grown for four weeks under continuous white light at 22°C. Cyanidioschyzon merolae 10D was grown at 40°C in MA2 medium under bubbling with 2% CO₂ and continuous illumination (50 µmol m⁻² s⁻¹) [10].

(c) Affinity purification of haem-binding proteins

Haemin-immobilized beads (0.5 mg) were equilibrated with 0.5% NP-40 lysis buffer (50 mM Tris–HCl (pH 8.0), 150 mM NaCl and 0.5% NP-40). Four-week-old A. thaliana seedlings (1.5 g) or Cy. merolae cells were harvested, ground into powder in liquid nitrogen, and then suspended in 5 ml of KCI lysis buffer (100 mM KCl, 12.5% glycerol, 20 mM HEPES-NaOH (pH 7.9), 1 mM MgCl₂, 0.2 mM CaCl₂, 0.2 mM EDTA, 0.1%
Bruker Daltonics, MI, USA) coupled with high-performance liquid chromatography (in-solution tryptic digestion and guanidination, Thermo Fisher Scientific, MI, USA) and purified with C18 column tip. Peptides were analysed with matrix-assisted laser desorption ionization-time of-flight mass spectrometer (MALDI-TOF-MS, Bruker Daltonics, MI, USA) coupled with high performance liquid chromatography. Data were analysed by the Mascot algorithm to identify proteins corresponding to the peaks.

(d) Identification of haem-binding proteins
The haem-bound proteins were subjected to trypsin digestion (in-solution tryptic digestion and guanidination, Thermo Fisher Scientific, MI, USA) and purified with C18 column tip. Peptides were analysed with matrix-assisted laser desorption ionization-time of-flight mass spectrometer (MALDI-TOF-MS, Bruker Daltonics, MI, USA) coupled with high performance liquid chromatography. Data were analysed by the Mascot algorithm to identify proteins corresponding to the peaks.

(e) Expression and purification of recombinant proteins
For *A. thaliana* candidate proteins (Atg09650 and Atg55760), DNA fragments were polymerase chain reaction (PCR)-amplified using respective RAFL clones [27] as templates with appropriate primer sets (electronic supplementary material, table S1). pET24 vector (Novagen) was also PCR amplified. After pre-culture in Luria-Bertani (LB) medium containing 50 μg ml\(^{-1}\) kanamycin, proteins were induced by adding 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 37°C for 3 h. For *Cy. merolae* candidate proteins (CM2023C and CML100C), pETNH or pColdTF vector and genes were PCR-amplified using appropriate primer sets (electronic supplementary material, table S1). The obtained gene fragments and vectors were cloned using an InFusion cloning kit (TaKaRa, Shiga, Japan). The resulting plasmids were introduced into *Escherichia coli* strain BL21(DE3) or Rosetta 2(DE3)pLysS competent cell (Merck Millipore, MI, USA). For expressing E. coli trigger factor (TF) as negative control, pColdTF vector (TaKaRa, Shiga, Japan) was also introduced into the Rosetta strain. The His-tagged fusion proteins were expressed as described previously [28].

*Escherichia coli* cells expressing recombinant proteins (500 ml LB medium) were suspended in 10 ml of Lysis buffer (20 mM Tris–HCl, pH 8.0, 500 mM NaCl, 10% glycerol, 5 mM imidazole) and disrupted by sonication. After centrifugation (10 000 g, 30 min, 4°C), the soluble fraction was passed through a 0.45 µm filter membrane and subjected to 1 ml of HisTrap column (GE Healthcare, IL, USA). After washing with 40 ml of Lysis buffer, the His-tag protein was eluted with a linear gradient of imidazole concentration (5–500 mM). Fractions containing purified proteins were collected and dialysed with buffer containing 20 mM Tris–HCl (pH 8.0), 500 mM NaCl, 6% glycerol.

(f) Haem-binding assay
*Escherichia coli* cell pellets expressing candidate proteins were suspended in KCl buffer and sonicated using a Branson Sonifier (Branson Instruments, CT, USA). After centrifugation, soluble extracts containing 1 mg of proteins were incubated with the 0.5 mg of beads for 2–4 h at 4°C. The beads were washed three times with 0.5% NP-40 lysis (or KCl) buffer, and bound proteins were eluted with the Laemmli SDS sample buffer. Eluted samples were separated by SDS–PAGE and detected by silver staining using a Pierce silver stain kit (Thermofisher Scientific, MI, USA).

![Figure 1](image)

**Figure 1.** Identification of haem-binding proteins from *A. thaliana* and *Cy. merolae*. Bound fractions of cell extract from *A. thaliana* (a) and *Cy. merolae* (b) were eluted from haem-immobilized FG beads with the SDS sample buffer and separated by SDS–PAGE. Proteins were visualized by silver staining.

3. Results

(a) Proteomic analysis of metal-tetrapyrrole binding proteins
To purify haemin-binding proteins, we performed affinity purification using FG beads [26]. Haem was covalently conjugated to the beads and incubated with extracts from *A. thaliana* and *Cy. merolae* cells. After extensive washing, bound proteins were eluted with the SDS sample buffer, and the eluate fractions were subjected to SDS gel electrophoresis and silver staining. As shown in figure 1, several bands were detected in fraction from haemin-immobilized beads, while almost no band was observed in beads without ligand showing negligible non-specific binding of proteins to the FG beads. These proteins were subjected to proteolytic digestion and MALDI-TOF-MS spectrometry. As a result, we identified 10 proteins from *A. thaliana* (table 1) and 10 proteins from *Cy. merolae* (table 2).

Subcellular localizations of *A. thaliana* candidate proteins (table 1) were predicted based on gene ontology cellular localization.
Table 1. Peptide identification for haem-binding proteins in *A. thaliana*.

| description                        | protein accession numbers | protein molecular size (kDa) | digest matches score (Mascot score) | peak no. | meas. m/z | ppm    | positions | peptide sequence a |
|------------------------------------|---------------------------|------------------------------|------------------------------------|----------|-----------|--------|-----------|-------------------|
| outer envelope pore protein 16–3  | At2g42210                 | 17.0                         | 41.6                               | peak 3   | 718.295   | −32.71 | 2–7       | DPAEMR            |
|                                   |                           |                              |                                    | peak 5   | 870.559   | 91.284 | 42–48     | DVPRVER           |
|                                   |                           |                              |                                    | peak 7   | 1930.938  | −1.204 | 140–154   | VDINGREYPYTVEKR   |
| PAUSED, a homologue of exportin-T | At1g72560                 | 111.5                        | 30.8                               | peak 2   | 613.332   | −56.41 | 704–708   | VEPLR             |
|                                   |                           |                              |                                    | peak 5   | 870.559   | 49.444 | 704–710   | VEPLRSK           |
|                                   |                           |                              |                                    | peak 6   | 1888.938  | 20.229 | 397–412   | NNLNSLDTKGLEEDR   |
|                                   |                           |                              |                                    | peak 7   | 1930.938  | −34.844 | 300–316   | VSALLTGAVEVLECK   |
| transcription factor bHLH110      | At1g27660                 | 49.6                         | 29.2                               | peak 2   | 613.332   | −33.685 | 388–392   | NRPCK             |
|                                   |                           |                              |                                    | peak 3   | 718.295   | −73.124 | 312–318   | AGENAIS           |
|                                   |                           |                              |                                    | peak 6   | 1888.938  | 26.17  | 118–134   | EELESSSTISDNQEGISK|
| protein kinase PINOID 2           | At2g26700                 | 59.3                         | 27.3                               | peak 3   | 718.295   | −22.414 | 400–405   | GDINKEK           |
|                                   |                           |                              |                                    | peak 4   | 842.52    | −30.461 | 406–412   | TNLVNLK           |
|                                   |                           |                              |                                    | peak 6   | 1888.938  | −40.444 | 475–489   | SKPPVVKPEETSHK    |
|                                   |                           |                              |                                    | peak 7   | 1930.938  | −20.503 | 496–510   | SVNYLYPPFMMSRK    |
| ABC transporter G family member 23| At5g19410                 | 70.0                         | 25.5                               | peak 3   | 718.295   | −71.102 | 163–167   | EREER             |
|                                   |                           |                              |                                    | peak 4   | 842.52    | 72.554  | 618–624   | KASKSTH           |
| NAD-dependent protein deacetylase SRT1| At5g55760        | 52.6                         | 24.6                               | peak 4   | 842.52    | 79.304  | 239–244   | TPDKSSK           |
|                                   |                           |                              |                                    | peak 5   | 870.559   | 91.284  | 76–82     | EGDLPK            |
|                                   |                           |                              |                                    | peak 6   | 1946.958  | −49.43  | 274–290   | IDLFQIILTQSISQQR  |
| pentatricopeptide repeat-containing protein | At3g09650        | 174.0                        | 24.5                               | peak 1   | 569.305   | 95.943  | 496–499   | GYCK              |
|                                   |                           |                              |                                    | peak 5   | 870.559   | 75.719  | 2907–303  | IDIKGSIK          |
|                                   |                           |                              |                                    | peak 6   | 1888.938  | 10.065  | 758–772   | FWLGLPMNSYGSEWK   |
| Asp—Glu—Ala—Asp (DEAD)-box ATP-dependent RNA helicase | At2g35920        | 111.0                        | 23.9                               | peak 2   | 613.332   | −56.41  | 370–374   | DLLPR             |
|                                   |                           |                              |                                    | peak 3   | 718.295   | −57.469 | 128–133   | AIDNER            |
|                                   |                           |                              |                                    | peak 4   | 842.52    | −43.806 | 150–156   | KLGSLKL           |
| calcium-dependent protein kinase 32| At3g57530                 | 60.9                         | 23                                 | peak 4   | 842.52    | −57.15  | 93–98     | SILKKK            |
|                                   |                           |                              |                                    | peak 5   | 870.559   | 49.444  | 199–205   | KETAPLKL          |
| transcription factor BOA           | At5g59570                 | 32.3                         | 20.7                               | peak 7   | 1930.938  | −32.749 | 101–116   | TAVIDADVREVEIMR   |
|                                   |                           |                              |                                    | peak 7   | 1930.938  | −94.079 | 138–152   | TSKRPRLVWTPQLHK   |
|                                   |                           |                              |                                    | peak 8   | 1946.958  | −93.288 | 153–169   | RFVDVIAHLGKNAVPK  |

a Carbamidomethyl modification.
| Description | Protein accession numbers | Closest homologue of *A. thaliana* | Putative localization | Protein molecular size (kDa) | Digest matches score (Mascot score) | Peak no. | Mass (m/z) | PPM | Positions | Peptide sequence |
|-------------|--------------------------|-----------------------------------|----------------------|-----------------------------|-----------------------------------|---------|-----------|------|-----------|-----------------|
| starch-associated protein R1 | CMT547C | At1g10760 | cytosol | 176.9 | 31.6 | peak 2 | 569.331 | 80.12 | 80.15 | ALIPR |
| - | | | | | | peak 3 | 842.528 | 34.41 | 1532–1538 | ILSIGK |
| - | | | | | | peak 5 | 944.552 | 6.692 | 85–91 | KARVCR |
| - | | | | | | peak 7 | 1808.882 | 19.997 | 1460–1474 | GTQVRDYKYPSTKK |
| - | | | | | | peak 9 | 2251.172 | 13.135 | 904–924 | AELASPGALEEFLMGAR |
| - | | | | | | peak 2 | 569.331 | 80.178 | 238–242 | TPLAK |
| - | | | | | | peak 8 | 1851.882 | 75.502 | 275–289 | KIREILQPDPAR |
| - | | | | | | peak 9 | 2251.172 | 7.024 | 427–446 | LALFSLESRESSKSR |
| - | | | | | | peak 10 | 2254.161 | 30.957 | 16–35 | OSUDEDGEREELMWIR |
| mutS family DNA mismatch repair protein MSH5 | CMN192C | At3g20475 | nucleus | 98.5 | 31.2 | peak 2 | 569.331 | 80.12 | 80.15 | ALIPR |
| - | | | | | | peak 3 | 842.528 | 34.41 | 1532–1538 | ILSIGK |
| - | | | | | | peak 5 | 944.552 | 6.692 | 85–91 | KARVCR |
| - | | | | | | peak 7 | 1808.882 | 19.997 | 1460–1474 | GTQVRDYKYPSTKK |
| - | | | | | | peak 9 | 2251.172 | 13.135 | 904–924 | AELASPGALEEFLMGAR |
| - | | | | | | peak 2 | 569.331 | 80.178 | 238–242 | TPLAK |
| - | | | | | | peak 8 | 1851.882 | 75.502 | 275–289 | KIREILQPDPAR |
| - | | | | | | peak 9 | 2251.172 | 7.024 | 427–446 | LALFSLESRESSKSR |
| - | | | | | | peak 10 | 2254.161 | 30.957 | 16–35 | OSUDEDGEREELMWIR |
| hypothetical protein | CMS174C | None | extracellular? | 35.7 | 29.8 | peak 2 | 569.331 | 80.12 | 80.15 | ALIPR |
| cystathionine beta-synthase | CMS037C | At2g20430 | chloroplast | 56.4 | 26.9 | peak 4 | 870.56 | 21.87 | 125–132 | LLGAEIVR |
| - | | | | | | peak 7 | 1808.882 | 75.502 | 275–289 | KIREILQPDPAR |
| - | | | | | | peak 9 | 2251.172 | 7.024 | 427–446 | LALFSLESRESSKSR |
| - | | | | | | peak 10 | 2254.161 | 30.957 | 16–35 | OSUDEDGEREELMWIR |
| nuclear receptor co-repressor/ HDAC3 complex subunit | CML100C | At5g67320 | nucleus | 60.7 | 25.7 | peak 4 | 870.56 | 21.87 | 125–132 | LLGAEIVR |
| - | | | | | | peak 7 | 1808.882 | 75.502 | 275–289 | KIREILQPDPAR |
| - | | | | | | peak 9 | 2251.172 | 7.024 | 427–446 | LALFSLESRESSKSR |
| - | | | | | | peak 10 | 2254.161 | 30.957 | 16–35 | OSUDEDGEREELMWIR |
| probable leucine aminopeptidase | CMH135C | At3g59760 | cytosol | 68.9 | 24.5 | peak 4 | 870.56 | 21.87 | 125–132 | LLGAEIVR |
| Asp–Glu–Ala–Asp (DEAD)-box ATP-dependent RNA helicase | CML137C | At2g24200 | nucleus | 66.4 | 24.4 | peak 4 | 870.56 | 21.87 | 125–132 | LLGAEIVR |
| - | | | | | | peak 7 | 1808.882 | 75.502 | 275–289 | KIREILQPDPAR |
| - | | | | | | peak 9 | 2251.172 | 7.024 | 427–446 | LALFSLESRESSKSR |
| - | | | | | | peak 10 | 2254.161 | 30.957 | 16–35 | OSUDEDGEREELMWIR |
| similar to GTPase-activating protein | CML230C | At4g15850 | vesicle | 55.0 | 23.3 | peak 3 | 842.528 | 34.39 | 14–20 | ALTRER |
| - | | | | | | peak 8 | 1851.882 | 38.367 | 338–353 | LVEHHPLPLPGRSPSR |
| - | | | | | | peak 10 | 2251.172 | 15.379 | 317–337 | TANPQYQAARAAASHPAR |
| - | | | | | | peak 4 | 870.56 | 19.981 | 165–172 | WSLLR |
| - | | | | | | peak 5 | 944.552 | 11.224 | 391–398 | LLAEEISK |
| - | | | | | | peak 8 | 1851.882 | 53.312 | 57–73 | IQQEVRPGQPILK |
| - | | | | | | peak 3 | 842.528 | 34.39 | 14–20 | ALTRER |
| - | | | | | | peak 8 | 1851.882 | 38.367 | 338–353 | LVEHHPLPLPGRSPSR |
| - | | | | | | peak 10 | 2251.172 | 15.379 | 317–337 | TANPQYQAARAAASHPAR |
| - | | | | | | peak 4 | 870.56 | 19.981 | 165–172 | WSLLR |
| - | | | | | | peak 5 | 944.552 | 11.224 | 391–398 | LLAEEISK |
| - | | | | | | peak 8 | 1851.882 | 53.312 | 57–73 | IQQEVRPGQPILK |
| - | | | | | | peak 3 | 842.528 | 34.39 | 14–20 | ALTRER |
| - | | | | | | peak 8 | 1851.882 | 38.367 | 338–353 | LVEHHPLPLPGRSPSR |
| - | | | | | | peak 10 | 2251.172 | 15.379 | 317–337 | TANPQYQAARAAASHPAR |
| - | | | | | | peak 4 | 870.56 | 19.981 | 165–172 | WSLLR |
| - | | | | | | peak 5 | 944.552 | 11.224 | 391–398 | LLAEEISK |
| - | | | | | | peak 8 | 1851.882 | 53.312 | 57–73 | IQQEVRPGQPILK |
| - | | | | | | peak 3 | 842.528 | 34.39 | 14–20 | ALTRER |
| - | | | | | | peak 8 | 1851.882 | 38.367 | 338–353 | LVEHHPLPLPGRSPSR |
| - | | | | | | peak 10 | 2251.172 | 15.379 | 317–337 | TANPQYQAARAAASHPAR |
| - | | | | | | peak 4 | 870.56 | 19.981 | 165–172 | WSLLR |
| - | | | | | | peak 5 | 944.552 | 11.224 | 391–398 | LLAEEISK |
| - | | | | | | peak 8 | 1851.882 | 53.312 | 57–73 | IQQEVRPGQPILK |

* c, carbamidomethyl modification.
localization in the TAIR database (https://www.arabidopsis.org/). For \textit{Cy. merolae} candidate proteins (table 2), subcellular localizations were predicted by the TargetP program [29], as well as from those of closest homologues of \textit{A. thaliana}. Among candidate proteins, four proteins from \textit{A. thaliana} and three proteins from \textit{Cy. merolae} were predicted to be plastid-localized. In both organisms, half of the candidates were occupied by putative nucleus-localized proteins (five proteins in \textit{A. thaliana} and five proteins in \textit{Cy. merolae}). It is interesting to note that nuclear proteins with similar functions were obtained from either or both organisms (see below). Others were hypothetical, cytosolic, mitochondrial or vesicle proteins.

(b) Plastid-localized candidate proteins

Among the seven identified plastid-localized proteins, we are interested in the \textit{A. thaliana} ABC transporter (At5g19410), which corresponds to ABCG23, as an energy-dependent transport mechanism is required for moving hydrophilic haem through or out of the lipid bilayer [11]. In \textit{A. thaliana}, there are 129 genes encoding the ABC transporter superfamily and ABCG23 is one of 29 members of the WBC subfamily [30] and is a plastid-envelope localized half-molecule type of ABC transporter [31].

Another interesting protein family are the pentatricopeptide repeat (PPR) proteins. In \textit{A. thaliana}, one plastid-localized PPR protein (At3g09650) is identified. In plants, most PPR proteins are supposed to bind RNA with sequence-specific manner and functions in post-transcriptional processes, including RNA editing, RNA splicing, RNA cleavage and translation [32]. At3g09650, corresponds to HCF152, and is involved in the processing of the chloroplast psbB-psbT-psbH-petB-petD transcript unit [33]. Concerning haem regulation, we recently found that GUN1, which is a PPR protein with a small MutS-related (SMR) domain and is the central integrator of retrograde signalling [34], binds to haem and modulates tetrahydrobiopterin biosynthesis [35].

(c) Nuclear-localized candidate proteins

Identification of haem-binding nuclear-localized proteins suggests that produced haem is actually transferred and functions in the nucleus in \textit{A. thaliana} and \textit{Cy. merolae}. In this study, we are focused on four functional groups that were identified from either or both organisms.

Haem is known to bind to transcription factors in yeast and mammalian cells [11]. Among identified nuclear proteins, two proteins from \textit{A. thaliana} were transcription factors. At1g27660 belongs to basic/helix-loophelix (bHLH) superfamily proteins [36]. In the \textit{A. thaliana} genome, 147 bHLH encoding genes have been identified and At1g27660 is assigned as bHLH110, although its function is unknown. Another transcription factor is At5g59570, which corresponds to BROTHER OF LUX ARRHYTHMO (BOA), a component of the circadian clock [37].

The second group is Asp–Glu–Ala–Asp (DEAD)-box ATP-dependent RNA helicases (DDBRHs): one protein from \textit{A. thaliana} (At2g59920) and one protein from \textit{Cy. merolae} (CML137C). The DDBRH family participates in broad aspects of RNA metabolism, such as transcription, translation, RNA decay and miRNA processing. This is also involved in cell cycle regulation, tumorigenesis, apoptosis, cancer development and viral infection [38], although its physiological function in plants and algae are poorly known.

The third group contains components of nuclear pore proteins: PAUSED (PSD) (At1g72560) from \textit{A. thaliana} and NUP107 homologue (CMC129C) from \textit{Cy. merolae}. In \textit{A. thaliana}, PSD encodes an orthologue of exportin-T, which mediates the nuclear transport of rRNA in yeast and mammals [39]. A null psd mutant of \textit{A. thaliana} showed defect in various developmental events [39]. NUP107 is localized to the nuclear rim and is an essential component of the nuclear pore complex.

The fourth group contains components of histone deacetylase (HDA): SRT1 (At5g55760) and HDA3 complex subunit (CML100C). HDA removes an acetyl group from Lys residues of histone, resulting in the histone wrapping DNA more tightly that represses the gene expression from the removed chromatin region. SRT1 is involved in the siruin family. In humans, SRT1 functions in ageing and metabolism [40]. In \textit{A. thaliana}, it is reported that SRT1 negatively regulates stress tolerance and glycolysis but stimulates mitochondrial respiration through interaction with cMyc-binding protein 1 (AtMBP-1) [41]. The closest homologue of CML100C in \textit{A. thaliana} is At5g67320 corresponding to a WD40 protein HOS15, which interacts with HDA9 to repress transcription of the GIGAENTIA-mediated photoperiodic flowering pathway [42].

(d) Characterization of haem-binding proteins

To verify whether candidate proteins actually bind to haem, we produced several recombinant proteins for the haem-binding assay. We chose soluble globular proteins for \textit{in vivo} expression in \textit{E. coli}: HCF152 (At3g09650), SRT1 (At5g55760), GTPase-activating protein (CMJ230C) and HDA3 complex subunit (CML100C).

For proteins from \textit{A. thaliana}, full-length cDNA fragments of HCF152 and SRT1 were cloned into pET24 in BL21(DE3). After induction, cell lysates were separated into soluble and precipitated fractions by centrifugation. HCF152 protein was expressed as an inclusion body (figure 2a) and refolding of the recombinant protein was not successful. Meanwhile, a certain portion of recombinant SRT1 was expressed in soluble fraction as a 52 kDa protein (figure 2a). We tested the ability of SRT1 to bind haem using haemin-agarose beads. As shown in figure 2b, SRT1 demonstrated haem-binding activity (figure 2b). It is noted that when we tested non-haemoprotein (lysozyme) and haemoproteins (catalase, myoglobin and apohorseradish peroxidase (HRP)), no binding to haemin-agarose was observed (electronic supplementary material, figure S1). To further characterize the haem-binding property, SRT1 was purified to homogeneity by using His-tag for affinity purification (electronic supplementary material, figure S2). Then, we monitored the haemin binding by absorbance, following the evolution of the Soret peak at 415 nm which appears in the presence of SRT1 (figure 3e). The interaction led to an increase in absorbance at this wavelength when
increasing the haemin concentration. The absorbance values, plotted in the inset of figure 2c, gave a saturation curve from which a $K_d$ of 0.68 ± 0.40 µM was estimated, assuming one haemin bound per domain. To further analyse the effect of haem on SRT1, we measured the siruin activity of SRT1. The obtained SRT1 exhibited the siruin activity, but exogenous haemin had no effect on the activity (figure 2c).

For proteins from *Cy. merolae*, cDNA of CMJ230C and CML100C were cloned into pETNH in Rosseta II. In this experiment, *A. thaliana* p22HBP protein was used as a positive control and *E. coli* TF protein was as a negative control. Although most proteins were detected in precipitated fractions in both cases, certain portions of recombinant proteins of CMJ230C and CML100C were detected in soluble fractions at 56 and 62 kDa bands, respectively (electronic supplementary material, figure 3). When cell lysates were subjected to haemin-agarose beads, both proteins were detected in eluted fractions (figure 3a). Binding profiles of positive (p22HBP) and negative (TF) confirmed the specificity of haem binding. A faint band of CMJ230C may be caused by poor expression in the soluble fraction. In addition, the binding of CMJ230C and CML100C to haemin-agarose was competitively decreased by the incubation with soluble haemin, indicating the specific interaction between CMJ230C/CML100C and haemin (figure 3a).

To verify spectral effects of proteins upon haem binding, we purified CMJ230C and CML100C by using His-tag for affinity purification (electronic supplementary material, figure S4). In the case of CML100C, required amounts of soluble purified protein for spectral analysis were only obtained when expressed with the pColdTF system. Thus, CML100C was expressed as a fusion protein of *E. coli* TF (CML100C–TF) with this system and purified, while purified TF was used as a negative control. Similar to *Arabidopsis* SRT1, mixing with equal molar concentration of haem solution with CMJ230C (6 µM) caused the red-shift of the haemin peak to 412 nm (figure 3c). In the case of CML100C–TF, a low concentration of purified protein (3.5 µM) was mixed with threefold molar concentration of haem solution (10.5 µM), which resulted in a slight red-shift of the haemin peak to 415 nm (figure 3c). When equal molar of TF and haemin was mixed, no spectral shift was observed.
Interestingly, canonical haemoproteins were not involved in haemin binding by absorbance, following the evolution of the Soret peak at 415 nm which appears in the presence of CML100C. The absorbance values, plotted in the inset of figure 3c, gave a saturation curve from which a $K_d$ of 1.33 ± 0.25 µM was estimated, assuming one haem bound per domain. These results demonstrated that, as well as Arabidopsis SRT1, Cy. merolae and CML100C bind to haem with high specificity.

4. Discussion

In this study, we performed proteomic analysis of haem-binding proteins in A. thaliana and Cy. merolae by using haemin-immobilized high-performance magnetic FG beads. As designed [26], FG beads showed extremely low non-specific binding of proteins (figure 1) and we could identify several candidates of haem-binding proteins from both organisms. Interestingly, canonical haemoproteins were not involved in haemin-binding proteins, probably covalently or non-covalently attached haem prevented the binding to proteins. In fact, when we analysed haemoproteins (catalase, myoglobin, apo-HRP) for haemin-agarose assay, no binding was observed (electronic supplementary material, figure S4). It is interesting to note that apo-HRP, which spontaneously binds to haem to form an active holo-enzyme, did not bind to haem-agrose. Thus, it is likely that only proteins which can bind to haem at the surface with substantial specificity can bind to haemin-ligated beads. In addition, already reported haem-binding proteins such as Fbx3, TSPO and p22HBP/SOUL were not included in this list, probably protein binding was dependent on their expression, solubility and affinity to the beads. It should be noted that because many identified proteins had not been annotated as haem-binding proteins, we should be careful to check whether each listed protein is actually binding to haem.

Some candidate proteins were possibly involved in haem transfer. Plastid-envelope localized A. thaliana ABC transporter ABCC23 (At5g19410) is a potential candidate protein for haem transfer. As designed [26], FG beads showed extremely low non-specific binding of proteins (figure 1) and we could identify several candidates of haem-binding proteins from both organisms. Interestingly, canonical haemoproteins were not involved in haemin-binding proteins, probably covalently or non-covalently attached haem prevented the binding to proteins. In fact, when we analysed haemoproteins (catalase, myoglobin, apo-HRP) for haemin-agarose assay, no binding was observed (electronic supplementary material, figure S4). It is interesting to note that apo-HRP, which spontaneously binds to haem to form an active holo-enzyme, did not bind to haem-agrose. Thus, it is likely that only proteins which can bind to haem at the surface with substantial specificity can bind to haemin-ligated beads. In addition, already reported haem-binding proteins such as Fbx3, TSPO and p22HBP/SOUL were not included in this list, probably protein binding was dependent on their expression, solubility and affinity to the beads. It should be noted that because many identified proteins had not been annotated as haem-binding proteins, we should be careful to check whether each listed protein is actually binding to haem.
located in the ER-to-Golgi membrane protein, is involved in haem scavenging. In *Cy. merolae*, ABA-inducible TSPO may decrease in the level of unbound haem that inhibits DNA replication [20]. It is therefore important to characterize how haem trafficking in the ER-to-Golgi membrane system, which is mediated by vesicle transport, occurs. In this sense, further analysis of CMJ230C, together with *Cy. merolae* TSPO (CMS231C), is necessary.

Surprisingly, half of the candidate proteins were nucleus proteins in both organisms, supporting the hypothesis that haem is actually transferred to the nucleus for regulatory and/or signalling purposes in these organisms like animal and yeast cells. Furthermore, identification of nuclear haem-binding proteins with similar functional or localization from either or both organisms may indicate the fundamental function of haem in these organisms. Because the detailed function of candidate proteins has not been elucidated in these organisms, further analysis is still needed for understanding. However, considering the general function of candidate proteins, it is possible that haem is involved in transcription through transcription factors, RNA metabolism through RBDHs and nucleoporins, and epigenetic histone modification through HDAs in *A. thaliana* and *Cy. merolae*.

For transcriptional regulation, haem is known to bind the transcription factor HAP1 in yeast to mediate oxidative stress [44]. In mammals, haem also binds to the basic leucine zipper protein Bach1, which represses genes such as haem oxygenase 1 [45]. Haem also coordinates regulation of metabolism with the circadian clock via the Rev-erb haem sensors [46]. For haem binding, a haem-regulatory motif (HRM) is found in bacteria and eukaryotic systems [47]. In fact, HAP1 and Bach1 contain 7 and 6 HRMs, respectively, while Rev-erb binds to haem with non-classical HRM. In bHLH110 and BOA, we could not detect any classic HRM, so haem may bind to distinct domains if they really bind to haem. It is interesting to note that in mammalian cells, haem biosynthesis is circadian-regulated and several components including Rev-erb bind haem [46]. In tobacco, the FC activity is inversely regulated with that of Mg-chelatase during cyclic photoperiods [48], but the involvement of haem in circadian regulation is totally unknown in plants and algae.

For RNA metabolism, haem is known to bind the haem-binding protein DGCR8 (DIGeorge critical region-8), which is a key miRNA processing enzyme in human cells and requires bound haem for its activity [49]. At present, the functions of candidate DBRHs on RNA metabolism are totally unknown. However, considering the effects of DBRHs [38] and miRNA [50] on cell cycle regulation, testing of the involvement of these components on the haem-dependent cell cycle regulation in *Cy. merolae* [13,20,51] is attractive. Currently, there is no report about haem-dependent regulation on nuclear pore transport. Considering haem is imported into the nucleus through the nuclear pore, it is possible that PSD and NUP107 are involved in haem transport into the nucleus in *A. thaliana* and *Cy. merolae*, respectively.

For epigenetic regulation, histone modification-dependent gene repression is suggested in *A. thaliana* retrograde signalling [52]. In this paper, a chloroplast envelope-bound plant homeodomain transcription factor (PTM) is identified and the proteolytic cleavage of PTM occurs in response to retrograde signals and amino-terminal PTM accumulates in the nucleus, where it activates ABI4 transcription by histone modifications. However, because recent careful analysis showed no significant involvement of PTM [53] and ABI4 [54] in the retrograde signalling, involvement of haem on such epigenetic regulation needs to be elucidated. In this study, we demonstrated that both *A. thaliana* SRT1 (figure 2) and *Cy. merolae* HDA3 complex subunit (figure 3) have haem-binding activity. Because the sirtuin activity of SRT1 was not affected by haemin, it is possible that haem affects the complex formation, stability and/or localization of SRT1 rather than the SRT1 activity. It is noted that the HDA3 complex subunit itself had no significant HDA activity in our assay. Because exogenously treated haem affected cell cycle regulation in *Cy. merolae* [13,20,51] and global gene expression in *Ch. reinhardtii* [18], analysis of haem-dependent histone modification and transcriptome should be investigated in the future.

By using haemin-immobilized high-performance beads, we have succeeded in identification of novel haem-binding candidate proteins from *A. thaliana* and *Cy. merolae*. As half of the candidates were occupied with nucleus proteins, it is likely that haem functions as an actual signal molecule in these organisms. The identification of nucleus proteins with similar function or localization suggests the fundamental but unknown function of haem, which may lead significant studies in retrograde signalling.

In summary, it has been considered that haem acts as a retrograde signalling molecule in *A. thaliana* and *Cy. merolae*. We recently reported that the major retrograde signalling protein GUN1 can bind haem, activate the FC1 activity and regulate the flow through the tetrapyrrole biosynthesis pathway [35], that supports a role for haem in mediating retrograde signalling and opens up the opportunity to develop a unifying hypothesis for this pathway. Therefore, our comprehensive analysis of haem-binding proteins will significantly contribute for the elucidation of this pathway in the future.

**Data accessibility.** This article has no additional data.

**Authors’ contributions.** S.W. analysed and interpreted data and co-wrote the article. Ta.S. expressed recombinant proteins, and analysed and interpreted data. Y.M. and R.Y. performed haem-binding experiments. R.T. measured the sirtuin activity. To.S. and K.T. performed proteomic analysis of haem-binding proteins. S.I. analysed and interpreted data. T.M. designed the study, performed experiments, analysed and interpreted data and co-wrote the article.

**Competing interests.** We declare we have no competing interests.

**Funding.** We are grateful for funding from JSPS KAKENHI (grant nos JP16H02217, JP16K07675, JP16K07393, JP17K07444, JP18H03941, JP18K05386 and JP18K14650). This work was performed under the Cooperative Research Program of ‘NJRC Mater. & Dev’.

**Acknowledgements.** We thank Biomaterial Analysis Center, Technical Department of Tokyo Institute of Technology for technical support.
diverse regulators ranging from transcription factors to MAP kinases. Cell Res. 16, 681–692. (doi:10.1038/sj.cr.710086)

3. Tsigfouglou A, Tsimadou A, Papadopoulou L. 2006 Heme as key regulator of major mammalian cellular functions: molecular, cellular, and pharmacological aspects. Pharmacol. Ther. 111, 327–345. (doi:10.1016/j.pharmthera.2005.10.017)

4. Chow K-S, Singh DP, Walker AR, Smith AG. 1998 Two different genes encode ferrochelatase in Arabidopsis: mapping, expression and subcellular targeting of the precursor proteins. Plant J. 15, 531–541. (doi:10.1046/j.1365-313X.1998.00235.x)

5. Nagai S, Koide M, Takashashi S, Kikuta A, Aono M, Sasaki-Sekimoto Y, Ohta H, Takamika Y, Masuda T. 2007 Induction of isoforms of tetrapyrrole biosynthetic enzymes, AthEMA2 and AthFC1, under stress conditions and their physiological functions in Arabidopsis. Plant Physiol. 144, 1039–1051. (doi:10.1104/pp.107.100065)

6. Comah J, Roger J, Pal Singh D, Smith A. 2002 Measurement of ferrochelatase activity using a novel assay suggests that plastids are the major site of haem biosynthesis in both photosynthetic and non-photosynthetic cells of pea (Pisum sativum L.). Biochem. J. 362, 423–432. (doi:10.1042/00062042)

7. Masuda T, Suzuki T, Shimada H, Ohta H, Takamika K. 2003 Subcellular localization of two types of ferrochelatase in cucurbit. Plant Cell 217, 602–609. (doi:10.1007/s00245-003-1019-2)

8. Hey D, Ortega-Rodes P, Fan T, Schurrer F, Brings L, Hedtke B, Grimm B. 2016 Transgenic tobacco lines expressing sense or antisense ferrochelatase 1 RNA show modified ferrochelatase activity in roots and provide experimental evidence for dual localization of ferrochelatase 1. Plant Cell Physiol. 57, 171–2585. (doi:10.1093/pcp/pctw171)

9. van Lis R, Attia A, Nogaj LA, Beale SI. 2005 The cell cycle coordinator from organelle to nuclear DNA replication in plant cells. Proc. Natl Acad. Sci. USA 102, 803–807. (doi:10.1073/pnas.0804270105)

10. Susek R, Ausubel F, Chory J. 1993 Signal transduction mutants of Arabidopsis uncouple nuclear CBG and RBCS gene expression from chloroplast development. Cell 74, 787–799. (doi:10.1016/0092-8674(93)90459-4)

11. Mochizuki N, Brusslan J, Larkin R, Nagatani A, Chory J. 2001 Arabidopsis genomies uncoupled S (GUNS) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. Proc. Natl Acad. Sci. USA 98, 2053–2058. (doi:10.1073/pnas.98.4.2053)

12. Larkin R, Alonso J, Ecker J, Chory J. 2003 GUN4, a regulator of chlorophyll synthesis and intracellular signaling. Science 299, 902–906. (doi:10.1126/science.1079978)

13. Woodson JD, Perez-Ruiz JM, Chory J. 2011 Heme synthesis by plastid ferrochelatase I regulates nuclear gene expression in plants. Curr. Biol. 21, 897–903. (doi:10.1016/j.cub.2011.04.004)

14. von Gromoff E, Akawady A, Meinecke L, Grimm B, Beck C. 2008 Heme, a plastid-derived regulator of nuclear gene expression in Chlamydomonas. Plant Cell 20, 552–567. (doi:10.1105/tpc.105.045650)

15. Voss B, Meinecke L, Kurtz T, Al-Babili S, Beck CF, Hess WR. 2011 Hemin and magnesium-protoporphyrin IX induce global changes in gene expression in Chlamydomonas reinhardtii. Plant Physiol. 155, 892–905. (doi:10.1104/pp.110.158683)

16. Kobayashi Y, Ando H, Hanaoka M, Tanaka K. 2016 Abscistic acid participates in the control of cell cycle initiation through heme homeostasis in the unicellular red alga Cyanidioschyzon merolae. Plant Cell Physiol. 57, 953–960. (doi:10.1093/pcp/ppw054)

17. Krishnamurthy PC et al. 2006 Identification of a mammalian mitochondrial porphyrin transporter. Nature 443, 586–589. (doi:10.1038/nature05125)

18. Latunde-Dada GO, Simpson RJ, McKie AT. 2006 Recent advances in mammalian haem transport. Trends Biochem. Sci. 31, 182–188. (doi:10.1016/j.tibs.2006.01.005)

19. Lu Y, Li Z, Drozdowicz Y, Hortensteiner S, Martinoia E, Rea P. 1998 AtMRP2, an ABC protein GUN1 regulates tetrapyrrole biosynthesis in Arabidopsis. Plant Physiol. 116, 892–897. (doi:10.1109/tpc.1013997)

20. Koussevitzky S, Nott A, Mockler TC, Hong F, Sachetto-Martins G, Surpin M, Linn J, Mittler R, Chory J. 2007 Signals from chloroplasts converge to regulate nuclear gene expression. Science 316, 715–719.

21. Shimizu T et al. 2019 The retrograde signaling protein GUN1 regulates tetrapyrrole biosynthesis. Proc. Natl Acad. Sci. USA 116, 24 900–24 906. (doi:10.1073/pnas.1911251116)

22. Meierhoff K, Felder S, Nakamura T, Bechtold N, Schuster G. 2003 HCF152, an Arabidopsis RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast psbD-psbB-psbD-psbB-pret-pretD RNAs. Plant Cell 15, 1480–1495. (doi:10.1105/tpc.010397)

23. Schmitzlinneweber C, Small I. 2008 Pentatricopeptide repeat proteins: a socket set for organelle gene expression. Trends Plant Sci. 13, 663–670. (doi:10.1016/j.tplants.2008.10.001)

24. Corbesier L, Felder S, Nakamura T, Bechtold N, Schuster G. 2003 HCF152, an Arabidopsis RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast psbD-psbB-psbD-psbB-pret-pretD RNAs. Plant Cell 15, 1480–1495. (doi:10.1105/tpc.010397)

25. Houtkooper RH, Pirinen E, Auwerx J. 2012 Sirtuins and metabolic disease: from bench to bedside. Nat. Rev. Endocrinol. 8, 553–564. (doi:10.1038/nrendo.2012.102)
41. Liu X, Wei W, Zhu W, Su L, Xiong Z, Zhou M, Zheng Y, Zhou D-X. 2017 Histone deacetylase AtSRT1 links metabolic flux and stress response in Arabidopsis. Mol. Plant 10, 1510–1522. (doi:10.1016/j.molp.2017.10.010)

42. Park HJ, Baek D, Cha JY, Liao X, Kang SH, Mcclung CR, Lee SY, Yun D-J, Kim W-Y. 2019 HOS15 interacts with the histone deactetylase HDA9 and the evening complex to epigenetically regulate the floral activator GIGANTEA. Plant Cell 31, 37–51. (doi:10.1105/tpc.18.00721)

43. Min MK, Jang M, Lee M, Lee J, Song K, Lee Y, Choi KY, Robinson DG, Hwang I. 2013 Recruitment of Arf1-GDP to golgi by Glo3p-type ArfGAPs is crucial for golgi maintenance and plant growth. Plant Physiol. 161, 676–691. (doi:10.1104/pp.112.209148)

44. Zhang L, Hach A, Wang C. 1998 Molecular mechanism governing heme signalling in yeast: a higher-order complex mediates heme regulation of the transcriptional activator HAP1. Mol. Cell. Biol. 18, 3819–3828. (doi:10.1128/MCB.18.7.3819)

45. Zenke-Kawasaki Y, Dohi Y, Katoh Y, Ikura T, Ikura M, Asahara T, Tokunaga F, Iwai K, Igarashi K. 2007 Heme induces ubiquitination and degradation of the transcription factor Bach1. Mol. Cell. Biol. 27, 6962–6971. (doi:10.1128/MCB.02415-06)

46. Yin L et al. 2007 Rev-erbalpha, a heme sensor that coordinates metabolic and circadian pathways. Science 318, 1786–1789. (doi:10.1126/science.1150179)

47. Yang J, Ishimori K, O’Brian MR. 2005 Two heme binding sites are involved in the regulated degradation of the bacterial iron response regulator (Irr) protein. J. Biol. Chem. 280, 7671–7676. (doi:10.1074/jbc.M411664200)

48. Papenbrock J, Mock H-P, Kruse E, Grimm B. 1999 Expression studies in tetrapyrrole biosynthesis: inverse maxima of magnesium chelatase and ferrochelatase activity during cyclic photoperiods. Planta 208, 264. (doi:10.1007/s004250050558)

49. Faller M, Matsunaga M, Yin S, Loo JA, Guo F. 2007 Heme is involved in microRNA processing. Nat. Struct. Mol. Biol. 14, 23–29. (doi:10.1038/nsmb1182)

50. Bueno MJ, Malumbres M. 2011 MicroRNAs and the cell cycle. Biochim. Biophys. Acta 1812, 592–601. (doi:10.1016/j.bbadis.2011.02.002)

51. Kobayashi Y, Tanaka K. 2016 Transcriptional regulation of tetrapyrrole biosynthetic genes explains abscisic acid-induced heme accumulation in the unicellular red alga Cyanidioschyzon merolae. Front. Plant Sci. 7, 1300. (doi:10.3389/fpls.2016.01300)

52. Sun X, Feng P, Xu X, Guo H, Ma J, Chi W, Lin R, Lu C, Zhang L. 2011 A chloroplast envelope-bound PHD transcription factor mediates chloroplast signals to the nucleus. Nat. Commun. 2, 477. (doi:10.1038/ncomms1486)

53. Page MT, Kacprzak SM, Mochizuki N, Okamoto H, Smith AG, Terry MJ. 2017 Seedlings lacking the PTM protein do not show a genomes uncoupled (gun) mutant phenotype. Plant Physiol. 174, 21–26. (doi:10.1104/pp.16.01930)

54. Kacprzak SM, Mochizuki N, Nanjao B, Xu D, Leister D, Kleine T, Okamoto H, Terry MJ. 2019 Plastid-to-nucleus retrograde signalling during chloroplast biogenesis does not require AB4. Plant Physiol. 179, 18–23. (doi:10.1104/pp.18.01047)