Enzymatic components and physiological roles of mitochondrial fatty acid biosynthesis in plants

Xin Guan
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Biochemistry Commons

Recommended Citation
Guan, Xin, "Enzymatic components and physiological roles of mitochondrial fatty acid biosynthesis in plants" (2014). Graduate Theses and Dissertations. 13824.
https://lib.dr.iastate.edu/etd/13824

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Enzymatic components and physiological roles of mitochondrial fatty acid biosynthesis in plants

by

Xin Guan

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Biochemistry

Program of Study Committee:
Basil J. Nikolau, Major Professor
Thomas A. Bobik
Olga A. Zabotina
Eve S. Wurtele
Young Jin Lee
Dan Nettleton

Iowa State University
Ames, Iowa
2014

Copyright © Xin Guan, 2014. All rights reserved.
# TABLE OF CONTENTS

**Acknowledgements**

iv

**Abstract**

v

## Chapter I: General introduction

1

- Mitochondrial acyl carrier protein (mtACP) and its activation  2
- Mitochondrial acyl chain elongation cycle 2
- Mitochondrial malonyl-ACP, the donor of 2-carbon elongation units 4
- Lipoic acid and its functions in photorespiration 5
- Lipid A-like molecules 6
- Plant metabolomics methodology 7
- Dissertation organization 8
- References 8
- Figures 15

## Chapter II: Disruption of the mitochondrial \(\beta\)-ketoacyl-ACP synthase gene leads to defective lipid metabolism in Arabidopsis

22

- Abstract 22
- Introduction 23
- Results 26
- Discussion 40
- Methods 46
- References 54
- Figures 61

## Chapter III: Identification and reverse genetic characterization of the 3-hydroxyacyl-ACP dehydratase component of mitochondrial fatty acid synthase of Arabidopsis

75

- Abstract 75
- Introduction 77
- Results 79
- Discussion 90
- Methods 97
- References 102
- Figures 107

## Chapter IV: Biochemical identification and reverse genetic characterization of mitochondrial enoyl-ACP reductase, a redundant enzyme component of the mitochondrial Type II fatty acid synthase of Arabidopsis

123

- Abstract 123
Chapter V: Reverse genetic characterization of a phosphopantetheinyl transferase for activating acyl carrier proteins in Arabidopsis mitochondria

Chapter VI: Mitochondrial malonyl-CoA generation by malonyl-CoA synthetase in Arabidopsis

Chapter VII: General conclusions
ACKNOWLEDGEMENTS

I, Xin Guan (管鑫), would like to thank my major professor, Basil J Nikolau, and my committee members, Thomas A Bobik, Olga A Zabotina, Eve S Wurtele, Young Jin Lee, and Dan Nettleton, for their guidance and support throughout the course of this research.


I would also like to thank my collaborators at Iowa State University, Philip Dixon, Ling Li, Xuefeng Zhao, and Andrew Lithio. I would like to thank those who helped me in this research at Iowa State University, Tracey Pepper, Margie Carter, Jiqing Peng, and Siquan Luo.

I would like to thank my collaborators outside of Iowa State University, Kazuki Saito, Yozo Okazaki, Lloyd Sumner, Christian R Raetz, and Alex Abramson.

Finally, thanks to my father, Congsheng Guan (管从胜), my mother, Yang Zhao (赵阳), and my wife, Wei Zhang (张薇) for their love.
ABSTRACT

Plant cells appear to have at least two fatty acid forming systems, which occur in distinct subcellular compartments: plastids and mitochondria. Mitochondrial fatty acid synthase (mtFAS) is a Type II system that utilizes individual enzymes to catalyze the iterative reactions. Many of the mtFAS components have yet to be identified, and physiological functions of mtFAS are still poorly understood. In this dissertation, we characterized five Arabidopsis mtFAS components, namely mitochondrial β-ketoacyl-ACP synthase (mtKAS), mitochondrial 3-hydroxyacyl-ACP dehydratase (mtHD), mitochondrial enoyl-ACP reductase (mtER), mitochondrial phosphopantetheinyl transferase (mtPPT), and mitochondrial malonyl-CoA synthetase (mtMCS). MtKAS, mtHD, and mtER catalyze 3 of the 4 reactions of the acyl chain elongation cycle; mtPPT is responsible for activating mitochondrial acyl carrier protein (mtACP) by the addition of a phosphopantetheine cofactor; and mtMCS generates malonyl-CoA, the donor of 2-carbon elongation unit for the mtFAS system. Mitochondrial octanoyl-ACP is the precursor for the synthesis of lipoic acid, which is the coenzyme of H-subunit of glycine decarboxylase complex (GDC), a key enzyme of photorespiration. Accordingly, mutations in the mtFAS system lead to a typical deficiency in photorespiration due to the depletion of the lipoylation of H-subunit of GDC. Lipid profiles in these mutants are also altered as a secondary effect of the deficiency in photorespiration. In addition, mutant analyses demonstrate that mitochondrial 3-hydroxymyristyl-ACP contributes to the biosynthesis of lipid A-like molecules, suggesting a novel synthetic destination of the mtFAS intermediates. RNA-seq transcriptomic analyses of the mutants establish a regulatory network that is associated with the mtFAS functions. Taken together, our
results demonstrate that the mtFAS system has multiple functionalities and is crucial for plant metabolic homeostasis.
CHAPTER I: GENERAL INTRODUCTION

Plant cells appear to have at least three fatty acid forming systems, which take place in multiple subcellular compartments: plastids, the ER membrane, and mitochondria (1-3). These processes play distinct biological roles and do not reflect just metabolic redundancy (1,4,5). In contrast to the well-studied plastidial fatty acid synthase and ER-localized fatty acid elongase, the biochemical mechanism and physiological significance of mitochondrial fatty acid synthase (mtFAS) have yet to be elucidated.

The occurrence of a mtFAS system was first suggested by the discovery of a mitochondrial acyl carrier protein (mtACP) in Neurospora crassa (6). Since then, all of the mtFAS components have been biochemically characterized in yeast and humans (5). MtFAS is organized as a Type II system, which is comprised of dissociated, monofunctional enzymes. It mimics the Type II FAS in bacteria and plant plastids, but is distinct from the Type I FAS in the cytosol of yeast and humans, which is a single multifunctional protein containing all of the reaction centers that produce a fatty acid (1,4,7). In plants, most of the mtFAS components have yet to be discovered, with the exception of mtACP components and mitochondrial β-ketoacyl-ACP synthase (mtKAS) (8-10).

This introduction section summarizes the biochemical reactions in the mtFAS system in yeast and humans, providing a glimpse on the plant mtFAS system. We also introduce two metabolic destinations of Arabidopsis mitochondria-originated fatty acids, namely biosynthesis of lipoic acid and lipid A-like molecules. In addition, we discuss on metabolomics, a new methodology, which supports the investigation of the physiological significance of Arabidopsis mtFAS system.
Mitochondrial acyl carrier protein (mtACP) and its activation

MtACP is recruited as the carrier of acyl intermediates in the mtFAS system. It shuttles the acyl intermediates between the active sites of different FAS enzymes. Yeast and humans appear to have a single gene encoding the mtACP component (11,12). MtACP is produced as an inactive apo-protein, and is subsequently activated to the holo-form by the addition of a phosphopantetheine cofactor from CoA to a conserved serine residue, a reaction catalyzed by phosphopantetheinyl transferase (PPTase) (Fig 1). The holo-form of mtACP is used in the mtFAS system, in which the growing acyl chain is attached to the thiol group of phosphopantetheine cofactor during fatty acid biosynthesis.

In nature, two groups of PPTase appear to be responsible for the activation of acyl carrier protein. Group I PPTase is typified by E. coli AcpS, while Group II PPTase is exemplified by Bacillus subtilis Sfp (13,14). Yeast Ppt2 enzyme, which belongs to the Group I PPTase family, catalyzes the phosphopantetheinylation reaction on mtACP, a reaction occurs in mitochondria (15). By contrast, Human AASDHPPT enzyme, a member of the Group II PPTase family, supports the activation of both the mtACP component and the ACP domain of cytosolic Type I FAS in the cytosol (16).

Mitochondrial acyl chain elongation cycle

The iterative mitochondrial fatty acid elongation cycle is catalyzed by 4 enzymes, namely mitochondrial β-ketoacyl-ACP synthase (mtKAS), mitochondrial β-ketoacyl-ACP reductase (mtKR), mitochondrial 3-hydroxyacyl-ACP dehydratase (mtHD), and mitochondrial enoyl-ACP reductase (mtER) (Fig 2).

MtKAS supports the condensation reaction of the fatty acyl chain elongation cycle. In plants, mtKAS has been previously characterized in Arabidopsis (10,17,18).
Arabidopsis mtKAS is capable of initiating fatty acid biosynthesis by catalyzing the condensation reaction between two malonyl-ACP substrates to form acetoacetyl-ACP (18). In addition, this enzyme appears to be capable of catalyzing the elongation reaction between acyl-ACP and malonyl-ACP to generate β-ketoacyl-ACP, leading to the extension of acyl-ACP precursors containing acyl moieties of different chain lengths (18). In parallel, its ortholog in humans also displays the broad substrate specificity in vitro (19). The biochemical properties of the mtKAS enzyme are distinct from those of KAS isoforms in bacteria and plant plastids, in which KAS III is responsible for the initial condensation reaction between acetyl-CoA and malonyl-ACP, whereas KAS I and KAS II catalyze the remaining elongation reactions between acyl-ACP and malonyl-ACP (1,7).

MtKR is responsible for the second step of this pathway, a reaction that reduces β-ketoacyl-ACP to 3-hydroxyacyl-ACP. Yeast mtKR, which is encoded by the OAR1 gene, utilizes NADPH for the reduction of β-ketoacyl-ACP substrate (20). In comparison, human genome contains two homologous genes that encode mtKR isoforms (i.e., 17β-HSD8 and CBR4), which shares sequence similarity with yeast mtKR. The two human mtKR isoforms are organized in a heterotetramer mode, which support the reduction of β-ketoacyl-ACP in a NADH-dependent manner (21).

The third step of acyl chain elongation cycle is the iterative dehydration reaction that converts 3-hydroxyacyl-ACP to enoyl-ACP, which harbors a double bond at the trans-2 position. This reaction is catalyzed by mtHD. It is noteworthy that the mtHD enzyme in yeast does not share any sequence homology with the mtHD enzymes in humans and parasites (e.g., Trypanosoma brucei) (22-24).
MtER catalyzes the last reaction of fatty acyl chain elongation cycle, the reduction of enoyl-ACP to acyl-ACP. MtER in yeast and humans are structurally distinct from other known enoyl-ACP reductases in bacteria and plant plastids. Biochemical characterizations of these enzymes demonstrate a NADPH dependent mechanism (25-28).

**Mitochondrial malonyl-ACP, the donor of 2-carbon elongation units**

Malonyl-ACP, the donor of 2-carbon units for the acyl chain elongation cycle, is synthesized from malonyl-CoA and holo-ACP, a reaction catalyzed by mitochondrial malonyl-CoA ACP transacylase (mtMCAT). In this reaction, malonyl moiety is transferred from CoA to the thiol group of the phosphopantetheine cofactor of holo-ACP (Fig 3). MtMCAT have been characterized in yeast and humans, and they are homologous to each other, as well as all of the known bacterial counterparts (12,20).

Malonyl-CoA, the precursor substrate for the malonyl-ACP biosynthesis, is generated via two distinct biochemical reactions (Fig 3). The canonical mode of malonyl-CoA generation is via the ATP-dependent carboxylation of acetyl-CoA, a reaction catalyzed by acetyl-CoA carboxylase (ACCase). The ACCase in bacteria and plant plastids is an enzyme complex that is comprised of 3 protein subunits, including biotin-carboxyl carrier protein, biotin carboxylase, and carboxyl transferase, whereas the cytosolic ACCase in eukaryotes is a multifunctional enzyme that possesses all of the reaction centers required to catalyze the ACCase reactions (29). In yeast mitochondria, malonyl-CoA is also generated by an ACCase enzyme, which is encoded by *HFA1* and is homologous to cytosolic ACCase (30). In contrast to the ACCase reaction, human mitochondria appear to produce malonyl-CoA through an alternative route, in which a
mitochondrially localized malonyl-CoA synthetase (MCS) encoded by ACSF3 catalyzes the formation of malonyl-CoA from free malonate and CoA (31). Arabidopsis genome encodes a single MCS ortholog, which has been previously characterized to encode a cytosolically localized MCS enzyme (32).

**Lipoic acid and its functions in photorespiration**

Mitochondrial octanoyl-ACP is the precursor for the biosynthesis of lipoic acid (3,33) (Fig 4). Lipoic acid is an essential sulfur-containing cofactor, which is covalently bonded to the ε-amino group of a lysine residue of 4 proteins, namely H-subunit of glycine decarboxylase, E2-subunit of pyruvate dehydrogenase, E2-subunit of β-ketoglutarate dehydrogenase, and E2-subunit of branched chain α-ketoacid dehydrogenase (34-36). The octanoyl group of octanoyl-ACP is transferred to the lysine residue of these protein substrates by mitochondrial lipoyl (octanoyl) transferase (LT) and, then, the transferred octanoyl group is converted into lipoyl group by mitochondrial lipoyl synthase (LS) (37,38). In addition to the mitochondrial origin, lipoic acid biosynthesis appears to occur in plant plastids, a process catalyzed by plastidial homologs of LT and LS (39,40).

In plants, lipoic acid biosynthesis is associated with photorespiration, which recruits glycine decarboxylase complex (GDC) containing a lipoylated H-protein subunit. Photorespiration occurs in all oxygen-producing photosynthetic organisms, including plants, algae, and cyanobacteria (41,42). In Arabidopsis, this process involves biochemical reactions that take place in multiple subcellular compartments, namely chloroplasts, peroxisomes, and mitochondria (Fig 5). Photorespiration starts with O₂ substitution for CO₂ in the first reaction of photosynthetic CO₂ fixation, a reaction
catalyzed by ribulose 1,5-bisphosphate carboxylase-oxygenase (RuBisCO). The oxygenase reaction of RuBisCO generates the toxic compound phosphoglycolate, which is converted into glycolate in chloroplasts (43). The resulting glycolate is sequentially converted into glyoxylate and glycine in peroxisomes (41). In mitochondria, GDC, coordinated with serine hydroxymethyltransferase (SHMT), catalyzes the conversion of two molecules of glycine into one molecule each of serine, ammonium, and CO₂ (44-47). Serine is then sequentially converted into hydroxypyruvate, glycerate, and glycerate 3-phosphate, which re-enters Benson-Calvin cycle (41). GDC is comprised of 4 protein subunits, which are named P, T, L, and H protein (48). H protein does not carry any enzymatic activity, but it interacts with other 3 catalytic subunits via its lipoamide arm abd, thus, lipoylation of H protein is essential for GDC activity (49). Indeed, mutations in the mtFAS system result in the depletion of lipoylation of H protein, leading to a typical deficiency in photorespiration (10).

**Lipid A-like molecules**

Lipid A is the hydrophobic anchor of the outer monolayer of outer membrane of most Gram-negative bacteria (50). Arabidopsis mitochondria are characterized to contain orthologs (e.g., AtLpxA, AtLpxC, AtLpxD1, AtLpxD2, AtLpxB, AtLpxK, and AtKdtA) of *E. coli* enzymes of lipid A biosynthesis (51,52) (Fig 6). Lipid X, an intermediate of lipid A biosynthesis that contains two 3-hydroxymyristyl chains on the sugar molecule, has been detected in the wild-type Arabidopsis plants. The contents of lipid X are diminished when genes encoding either of the two acyltransferases (i.e., AtLpxA and AtLpxD1) are deleted, or when the expression of AtLpxC is suppressed. Mutant analysis also supports the continuation of this synthetic pathway beyond lipid X, namely AtLpxB-
catalyzed formation of tetraacyl disaccharide phosphate, AtLpxK-catalyzed addition of a second phosphate group, and AtKdtA-catalyzed addition of two Kdo-moieties. To date, the final structure and physiological functions of lipid A-like molecules have yet to be discovered in plants.

Plant metabolomics methodology

The total number of metabolites in the plant kingdom is estimated to range from 200,000 to 1,000,000 (53). A single accession of Arabidopsis thaliana is expected to produce about 5,000 metabolites, 2,632 of which are contained in AraCyc version 6.0 (54). Because metabolites are the final products of genes and proteins, plant metabolome (a complete set of metabolites) represents the ultimate phenotype in response to genetic or environmental alterations (55). Metabolites bear a large dynamic range (~10^6) in the relative concentration and the various biochemical properties and, thus, comprehensive coverage can be only achieved by using multiparallel complementary analytical platforms. Indeed, a combination of eight different mass spectrometry-based platforms successfully leads to the quantification of about 2,000 metabolites in Arabidopsis (56).

To understand the cellular alterations as a consequence of the deficiency in mtFAS, we analyze the mtfas mutants using 3 analytical platforms in this dissertation, namely HPLC, GC-MS, and LC-MS (Fig 7). HPLC method is sensitive in amino acids profiling when combined with OPA-based fluorescence derivatization for the amino group (57). GC-MS method is able to detect primary metabolites, such as hydrocarbons, organic acids, and sugars by employing chemical derivatization for the hydrophilic groups (58). LC-MS is efficient in the detection of small polar metabolites, but this
method can be also used for the profiling of lipids when utilizing different columns and mobile phases (52,59).

**Dissertation organization**

This dissertation is comprised of seven chapters.

Chapter I is the general introduction on the mtFAS system in eukaryotes, the metabolic destinations of Arabidopsis mitochondria-derived fatty acids, and the plant metabolomics methodology.

Chapter II to VI are research papers concerning five out of seven enzymes of the Arabidopsis mtFAS system. Specifically, Chapter II, III, and IV are focused on Arabidopsis mtKAS, mtHD, and mtER of the fatty acyl chain elongation cycle, respectively. Chapter V discusses on Arabidopsis mtPPT for its functions in the mtACP activation. Chapter VI presents on Arabidopsis mtMCS, an enzyme that generates the extension substrate for the mtFAS system. In these chapters, the lipidomics study was collaborated with Dr. Yozo Okazaki and Dr. Kazuki Saito (RIKEN Plant Science Center, Japan); the RNA-seq experiment was collaborated with Dr. Ling Li, Dr. Xuefeng Zhao, Andrew Lithio, and Dr. Dan Nettleton (Iowa State University).

Chapter VII summerizes the general conclusions on the biochemical and physiological significance of the Arabidopsis mtFAS system. Future directions and potential applications in the field of metabolic engineering are also discussed.

**REFERENCES**


FIGURES

Figure 1. Activation of mtACP.
Figure 2. Mitochondrial acyl chain elongation cycle.
Figure 3. Generation of mitochondrial malonyl-ACP.
Figure 4. Biosynthesis of lipoic acid.

Protein = H-subunit of glycine decarboxylase
E2-subunit of pyruvate dehydrogenase
E2-subunit of β-ketoglutarate dehydrogenase
E2-subunit of branched chain α-ketoacid dehydrogenase
Figure 5. Photorespiration process.

C1-THF: C1 Tetrahydrofolate; F6P: Fructose 6-Phosphate; GAP: Glyceraldehyde 3-Phosphate; GDC: Glycine Decarboxylase; 3PGA: 3-phosphoglycerate; RuBP: Ribulose 1,5-Bisphosphate; RuBisCO: Ribulose 1,5-Bisphosphate Carboxylase Oxygenase; SHMT: Serine Hydroxymethyltransferase
Figure 6. Biosynthesis of lipid A-like molecules in Arabidopsis.
Figure 7. Metabolomics pipelines in this dissertation.
CHAPTER II: DISRUPTION OF THE MITOCHONDRIAL $\beta$-KETOACYL-ACP SYNTHASE GENE LEADS TO DEFECTIVE LIPID METABOLISM IN ARABIDOPSIS

Manuscript in preparation and to be submitted to *The Plant Cell*

Xin Guan$^1$, Yozo Okazaki$^2$, Ling Li$^3$, Huanan Jin$^1$, Xuefeng Zhao$^4$, Andrew Lithio$^5$, Dan Nettleton$^5$, Kazuki Saito$^2$, and Basil J. Nikolau$^1$

1 Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, Iowa 50011

2 Metabolomic Function Research Group, RIKEN Plant Science Center, Yokohama 230-0045, Japan

3 Department of Genetics, Development, and Cellular Biology, Iowa State University, Ames, Iowa 50011

4 Plant Sciences Institute, Iowa State University, Ames, Iowa 50011

5 Department of Statistics, Iowa State University, Ames, Iowa 50011

ABSTRACT

Mitochondrial fatty acid synthase (mtFAS) is considered to be a multi-component Type II FAS system and previous genetic studies of one component ($\beta$-ketoacyl-ACP synthase; mtKAS; AT2G04540) indicates that mtFAS generates octanoyl-ACP for lipoic acid biosynthesis. The major manifestation of mutations in the mtKAS gene is a deficiency in photorespiration, caused by the near complete depletion of the lipoylation of the H-subunit of glycine decarboxylase complex, an enzyme essential for photorespiration. This study expanded on the earlier genetic characterizations of the mtKAS gene by characterizing the *in vitro* kinetic properties of the recombinant mtKAS
enzyme, which indicates that it displays comparable activities against all acyl-ACP substrates between 4 and 16 carbon acyl-chain lengths. We also show a novel metabolic destination of the mtFAS derived fatty acids, namely the 3-hydroxymyristic acid that is a hallmark of Arabidopsis lipid A-like molecules. Mutations in the \textit{mtKAS} gene present novel ultrastructural phenotypes, including impaired thylakoid structures and hyperaccumulation of plastoglobules. Consistent with these ultrastructural phenotypes, biochemical analyses show that these mutants present altered lipid profiles, namely reduced levels of the chloroplastic galactolipid and sulfolipid, and of surface lipid deposition, and hyperaccumulation of triacylglycerol. Feeding these plants intermediates of photorespiration illustrates that the hyperaccumulation of triacylglycerol is a result of the accumulated photorespiratory glycolate, but the reduced levels of galactolipids and sulfolipids are a consequence of sucrose depletion associated with the deficiency in photorespiration. The RNA-seq analysis of the transcriptome of the \textit{mtkas} mutant establishes that global changes in acyl-lipid metabolism are regulated at the level of gene transcripts. In addition, these data also indicate that the transcriptional homeostasis of signaling lipid metabolism is also disturbed. These results demonstrate the integral nature of the mtFAS system in regulating photorespiration and acyl-lipid metabolism.

**INTRODUCTION**

Plant cells utilize at least three fatty acid forming systems, which occur in multiple subcellular compartments: plastids, the ER membrane, and mitochondria (1-3). These processes have distinct biological roles and do not reflect just metabolic redundancy (1,4,5). Plastidial fatty acid synthase (ptFAS) generates the bulk of fatty acids, which serve as precursors for the biosynthesis of a variety of acyl-lipid molecules,
including phospholipids, galactolipids, sulfolipids, diacylglycerol, and triacylglycerol (TAG) (6,7). The ER-localized fatty acid elongase (FAE) utilizes preexisting acyl-CoA precursors derived from the ptFAS system to synthesize very-long-chain fatty acids, which are subsequently incorporated into a variety of lipids including surface lipids and the ceramide moiety of sphingolipids (2,8).

Mitochondrial fatty acid synthase (mtFAS) appears to primarily generate octanoic acid, which is required for the biosynthesis of lipoic acid (9,10). Lipoic acid is the cofactor that is essential for pyruvate dehydrogenase (PDH), β-ketoglutarate dehydrogenase (KGDH), and glycine decarboxylase complex (GDC) (11). To date, no other alternative metabolic destinations for the mtFAS-derived fatty acids have been demonstrated, although its role in generating acyl-ACP molecules for the remodeling of cardiolipins has been suggested (12,13).

Based on the biochemical characterization of the mtFAS system in fungi and mammals, mtFAS is organized as a Type II system, which recruits acyl carrier protein (ACP) as the carrier of acyl intermediates and utilizes dissociated, monofunctional enzymes to catalyze iterative reactions that produce fatty acids (5). MtFAS resembles the Type II FAS in bacteria and plant plastids (1,14), which contrasts with the Type I FAS that occurs in the cytosol of fungi and mammals that is a multifunctional protein that contains all of the catalytic centers for fatty acid production (4).

In plants, mitochondrial β-ketoacyl-ACP synthase (mtKAS) of the mtFAS system has been characterized in Arabidopsis (11,15,16). This enzyme is capable of initiating fatty acid biosynthesis by catalyzing the condensation reaction between two malonyl-ACP substrates to form acetoacetyl-ACP (16). In addition, as indicated by its ability to
complement an *E. coli kasI/kasII* double mutant strain, Arabidopsis mtKAS appears to be capable of catalyzing the condensation reactions between acyl-ACP and malonyl-ACP to generate β-ketoacyl-ACP, leading to the extension of the acyl-chain to different chain lengths (16). These biochemical properties are distinct from those of the β-ketoacyl-ACP synthase isoforms in bacteria and plant plastids, in which KASIII is responsible for the initial condensation reaction between acetyl-CoA and malonyl-ACP, whereas KASI and KASII catalyze the remaining condensation reactions between acyl-ACP and malonyl-ACP (1,14).

Initial genetic characterization has established that mutations in the *mtKAS* gene lead to diminished lipoylation of the H-subunit of GDC, an essential enzyme complex in the photorespiration process, however, the lipoylation of PHD and KGDH appears to be only slightly affected. The resulting loss of GDC activity manifests a deficiency in photorespiration (11).

To explore the significance of the mtFAS system in plant acyl-lipid metabolism, *mtkas* mutant alleles were further characterized. These plants are specifically deficient in the metabolism of a plethora of acyl-lipids, for example galactolipids, sulfolipids, triacylglycerol, and surface lipids. In addition, these plants display a deficiency in the biosynthesis of lipid-A like molecules, as a direct consequence of the deficiency in producing mitochondrial long-chain fatty acids. The global transcriptional alterations associated with the defective lipid metabolism are illustrated by RNA-seq analysis. Our findings reveal that the mtFAS system is essential for acyl-lipid metabolism and that the deficiencies in mtFAS cannot be compensated by other fatty acid forming systems.
RESULTS

1. Recombinant mtKAS enzyme exhibits broad acyl-chain specificity

Previous *in vivo* characterization have shown that when *E. coli* KAS I and KAS II are replaced by the Arabidopsis mtKAS gene, *E. coli* is still capable of producing acyl-ACP intermediates of up to 18-carbon chain length, suggesting a broad substrate specificity for the Arabidopsis mtKAS enzyme (16). To further determine the acyl chain specificity of this enzyme, *in vitro* enzymatic activities of the recombinant mtKAS enzyme was measured with eight different acyl-ACP substrates. Saturated acyl-ACP thioesters containing acyl-moieties of 4 to 16-carbon atoms are elongated by the recombinant mtKAS enzyme, whereas the elongation of a unsaturated substrate C\textsubscript{16:1}-ACP was near the detection limit of the assay system used (Fig 1A). The catalytic efficiency (\(k_{\text{cat}}/K_m\)) on all of the saturated acyl-ACP substrates tested is ranked in the following orders: C16 > C14 > C6 > C8 > C10 > C12 > C4 (Fig 1B). The \(k_{\text{cat}}/K_m\) values for the recombinant Arabidopsis mtKAS enzyme on short- and medium-chain (C4-C12) acyl-ACP substrates are comparable to those of its human ortholog (17). However, the \(k_{\text{cat}}/K_m\) values for the recombinant Arabidopsis mtKAS enzyme on long-chain (C14 and C16) acyl-ACP substrates is significantly higher than that of its human ortholog, which is dramatically reduced with myristyl-ACP and undetectable with palmityl-ACP substrate.

Taken together, our data confirm that the Arabidopsis mtKAS enzyme can support the condensation reactions with saturated acyl-ACP substrates of up to 16-carbon chain length.
2. The \textit{mtKAS} gene is expressed in various organs

Quantitative RT-PCR was performed with RNA templates to determine the spatial and temporal expression of the \textit{mtKAS} gene among aerial organs of different age plants, and roots, flowers, siliques and rosette leaves of Arabidopsis (Fig 2). These data indicate that the \textit{mtKAS} gene is expressed in all of the organs tested. The highest expression was found in aerial organs and flowers, followed by roots, siliques, and rosette leaves. The near ubiquitous expression pattern of the \textit{mtKAS} gene is consistent with the public transcriptome data visualized by Arabidopsis eFP-Browser (18,19).

3. The mutation in the \textit{mtKAS} gene exhibits a severe chloroplast phenotype

Two T-DNA insertion mutants of the \textit{mtKAS} gene (i.e., \textit{mtkas}-2 and \textit{mtkas}-3) were previously characterized, and both of them contain an insertion in the last exon of the gene (11). Characterizations of the \textit{mtkas} mutant alleles have indicated that they exhibit a growth phenotype that is typical of a deficiency in photorepiration. Namely, in the ambient air the mutants grow slower, exhibiting a yellowish leaf color (Fig 4A), but this phenotypic appearance can be reverted when the mutants are grown in an atmosphere enriched in CO$_2$ (11) (Fig 4B).

The leaf cells of \textit{mtkas}-2 and its wild-type siblings were examined by light microscopy (LM). These observations illustrate the distinct differences in the leaf cell morphology between the mutant and wild-type plants. Specifically, mesophyll cells were dramatically enlarged in \textit{mtkas}-2, making the leaves thicker than those of the wild-type plants (Fig 3A and 3B). In addition, the number of chloroplasts in mesophyll cells appears to be fewer and smaller in the \textit{mtkas}-2 mutant plants (Fig 3A and 3B).
Chloroplasts of the \textit{mtkas-2} and wild-type leaf mesophyll cells were compared by transmission electron microscopy (TEM) to explore the changes in morphogenesis and ultrastructure resulting from the mutation in the \textit{mtKAS} gene. In the \textit{mtkas-2} mutant allele the chloroplasts are smaller and shriveled in appearance (Fig 3C and 3D), which is consistent with the observations under LM. Thylakoid membrane assembly was also affected by the \textit{mtkas-2} allele, with less extensive granal stacking, and the thylakoid structures appear sparse and less extensive (Fig 3E and 3F). In addition, there are more large plastoglobules in chloroplasts of the mutants as compared to the wild-type siblings (Fig 3E and 3F).

4. Mutations in the \textit{mtKAS} gene alter soluble metabolites

Mutations in the \textit{mtKAS} gene diminish the lipoylation of the H protein subunit of glycine decarboxylase complex (GDC), leading to the inability to metabolize photorespiratory Gly (11). To evaluate the metabolic changes in response to the mutations in the \textit{mtKAS} gene, we quantified soluble metabolites, namely amino acids, glycolate, glyoxylate, and sucrose (Fig 5A).

The most dramatic aberrance is an about 130-fold increase in the Gly level in aerial organs of the mutants. Increased Gly levels were also detected in flowers and siliques of the \textit{mtkas} alleles (6-fold and 45-fold, respectively), but in roots Gly levels were near normal (data not shown). This asymmetry in Gly accumulation in different organs is consistent with the findings that enzymes of photorespiration pathway were mainly expressed in Arabidopsis leaves (20). In addition, several other amino acids exhibit significantly altered levels in aerial organs of the \textit{mtkas} mutant alleles, which might be a consequence of the defective nitrogen metabolism in the mutants. The levels
of glycolate and glyoxylate are increased to approximately 3-fold in aerial organs of the mutants. The level of sucrose, the transport form of sugar in Arabidopsis, drops to about 5% of the wild-type plants, which is consistent with the previous findings that sucrose levels were greatly reduced in other photorespiratory mutants, including 10-fdf and shmt1-2 (21).

These metabolic alterations of soluble metabolites are a consequence of the deficiency in photorespiration because they are reversed when the plants are grown in an elevated CO₂ atmosphere, a condition that inhibits photorespiration (Fig 5A). The only exception to this metabolic reversal is Gly, the levels of which are reduced considerably when the mutants are grown in an elevated CO₂ atmosphere, however, it is still 50-fold higher in aerial organs as compared to its wild-type siblings.

5. Mutations in the mtKAS gene alter acyl-lipid compositions

Fatty acid (FA) compositions are disturbed in the mtkas mutant alleles. The aerial organs of the mutants have significantly reduced levels of FA 16:1, 16:2, and 16:3, and elevated levels of FA 18:1 (Fig 6A). We also analyzed the fatty acid profiles in other organs (i.e., roots, flowers, siliques, and dry seeds), however, significant difference is not detected between the mtKAS mutants and the wild-type sibling plants (data not shown).

Markedly reduced contents of galactolipids and sulfolipids are found in the mtkas mutants (Fig 6B). Two plastidial galactolipids, MGDG and DGDG, are reduced by 40% and 30%, respectively. For MGDG, although the levels of the two most abundant MGDG molecules, 34:6 and 36:6, are reduced only slightly, most other detected molecular forms (i.e., 34:1, 34:2, 34:3, 34:4, 34:5, and 36:5) are decreased by over 50%, while the reduction of DGDG is primarily associated with the same molecular species as the
MGDG reductions (i.e., 34:2, 34:3, 36:3, and 36:5). In addition, the plastidial anionic sulfolipid, SQDG, exhibits the most severe reduction in the \textit{mtKAS} mutants; it is decreased by over 50% relative to the wild-type plants. Specifically, the amounts of all detected SQDG species were significantly reduced in the mutants, but the reduction is most pronounced with the 34:2 and 36:5 SQDG species. Depleted galactolipids and sulfolipids are in keeping with the reduced levels of FA 16:1 and 16:3, because plastidial galactolipids and sulfolipids are considered to be the sole source of the two fatty acids (22). By contrast, the contents of phospholipids (i.e., PC, PE, PG, and PI) are only subtly altered by the \textit{mtkas} alleles (data not shown).

Dramatic increases in the triacylglycerol (TAG) levels are observed in the \textit{mtkas} mutant plants; more than 4-fold higher in the mutants relative to their wild-type siblings (Fig 6C). All TAG molecular species are affected and, thus, the affect appears to enhance the overall flux through the TAG assembly pathway, rather than differentially affecting a portion of the assembly pathway, which would be expected to change the distribution of TAG molecular species. In comparison to TAG, the contents of diacylglycerol (DAG), the precursor of TAG and other lipid species, are only slightly affected by the mutations in the \textit{mtKAS} gene, with reduced DAG 34:2 and elevated DAG 36:6 contents (Fig 6B).

Finally, the contents of surface lipids (i.e., very-long-chain fatty acid, alcohol, and alkane) are altered in the \textit{mtkas} mutant alleles (Fig 6D). Specifically, while the alcohol C30 and C32 are slightly accumulated, all the other surface lipids are significantly depleted in the mutants.
All of the alterations in acyl-lipids return to normal when the mutant plants are grown in the elevated CO₂ atmosphere (Fig 6A, 6B, 6C, and 6D), demonstrating that the defective acyl-lipid metabolism is a consequence of the deficiency in photorespiration.

6. Disturbed soluble metabolites lead to the morphological and acyl-lipid alterations

The deficiency in photorespiration leads to the alterations in morphological phenotype and metabolic changes in soluble metabolites and acyl-lipids, because they are reversed when photorespiration process is suppressed by CO₂ treatment. Considering that alterations in the soluble photorespiratory intermediates (e.g., glycolate, glyoxylate, glycine, and sucrose) are the direct effects in response to the deficiency in GDC function, we hypothesized that the morphological and acyl-lipid alterations are a result of the alterations in these photorespiratory intermediates. To test this hypothesis, we conducted a feeding experiment, in which the plants were treated with exogenous glycolate, glyoxylate, glycine, and sucrose, respectively.

The mtkas mutant alleles and their wild-type siblings were grown on media supplemented with 3% sucrose to test if sucrose complements the morphological and acyl-lipid alterations. The exogenous sucrose relieves the dwarf phenotype and yellowish leaf color of mtkas-2 and mtkas-3 as compared to the non-treated mtkas mutant alleles (Fig 4C). The presence of exogenous sucrose significantly alleviates the sucrose depletion that occurs in the mtkas mutant alleles, reaching about 25% of that in the sucrose-treated wild-type plants (Fig 5A). In addition, exogenous sucrose effectively complements the reductions in FA 16:1 and 16:3, and also relieved the alterations of 16:2 and 18:2 (Fig 6A). Furthermore, when these mutants were grown on the sucrose-containing medium, the affected MGDG, DGDG, SQDG, and DAG returned to near
normal levels, as compared to the sucrose-treated wild-type control (Fig 6B). In comparison, exogenous sucrose does not complement the hyperaccumulation of TAG in the mutants (data not shown). Taken together, these results demonstrate that sucrose depletion contributes to the morphological phenotype and acyl-lipid alterations in FA, galactolipids, and sulfolipids.

The wild-type plants were grown in the presence of 10mM glycolate, glyoxylate, and Gly, respectively, to evaluate the effects of these photorespiratory intermediates on the alterations in phenotypic appearance and acyl-lipid composition. Exogenous glycolate dramatically leads to a dwarf phenotype with the dark-green leaf color, however, exogenous glyoxylate and glycine had only subtle effects on the morphological appearance of the wild-type plants (Fig 4D). The presence of exogenous glycine only results in a 5-fold elevation in the glycine level (Fig 5B), and this elevation is insufficient to mimic the over 100-fold hyperaccumulation in the glycine level in the mutants. Glycolate-treated plants display a about 4-fold accumulation in the glycolate level, while glyoxylate-treated plants displayed a about 3-fold accumulation in the glyoxylate level (Fig 5B); the elevations in the two metabolites are similar to those in the non-treated mtkas mutant alleles, illustrating that the glycolate and glyoxylate treatments on the wild-type plants mimic the metabolic alterations in these two metabolites in the mtkas mutant alleles. Glycolate-treated wild-type plants exhibit slightly affected fatty acid contents (Fig 6A). Moreover, exogenous glycolate leads to the hyperaccumulation of TAG in the wild-type plants, by more than 6-fold higher, and the pattern of TAG species resembles that of the mtkas mutant plants (Fig 6C). By contrast, exogenous glyoxylate and glycine do not lead to any significant alteration in acyl-lipid composition (data not shown). In addition,
addition of glycolate, glyoxylate, and Gly does not lead to a *mtkas*-like changes in galactolipids, sulfolipids, and DAG. Collectively, these data demonstrate that the glycolate accumulation contributes to the FA alterations and hyperaccumulation of TAG.

7. **3-Hydroxymyristic acid is the fatty acyl build block of mitochondrial lipid A-like molecules**

   The recent discovery of Lipid A-like molecules in plants, and the characterization that the biosynthetic genes for these lipids are mitochondrially located (23), suggests that another metabolic destination of mitochondria-originated fatty acids may be these lipid A-like molecules. However, because the chemical structure of the final product in plants is still unclear (23), we took an alternative approach to determine the acyl chain specificity of the acyltransferases that catalyze the initial reactions of the biosynthesis of lipid A-like molecules, namely Arabidopsis *AtLpxA, AtLpxD1, and AtLpxD1*. These enzymes catalyze the acyltransferase reactions in which a 3-hydroxyacyl-chain is transferred from ACP to either a hydroxyl-group (*AtLpxA*) or amino-group (*AtLpxD1 and AtLpxD2*) of the UDP-glucosamine derivatives. These enzymes display very narrow acyl-chain length specificity, and in *E. coli* this results in the assembly of a Lipid A molecule that contains only 3-hydroxymyristoyl moiety (24).

   *AtLpxA* was recently characterized to prefer 3-hydroxymyristyl-ACP (25). We determined the acyl-chain length specificity of *AtLpxD1* and *AtLpxD2*, by expressing them in the *E. coli lpxD* mutant strains, respectively, and characterizing the consequential hybrid lipid A molecules that were produced by a hybrid pathway of plant and bacterial enzymes. Our first attempt to delete *E. coli lpxD* and functionally replace it with the plant homologs was unsuccessful when the full-length Arabidopsis homologs were used.
Considering that both AtLpxD1 and AtLpxD2 carry N-terminal mitochondrial targeting peptides, we considered that they may be interfering with functionality. Therefore, we removed the sequence coding for the N-terminal 60 and 71 residues for AtLpxD1 and AtLpxD2, respectively, and functionally replaced the E. coli lpxD with the two Arabidopsis homologs. Hybrid lipid A species were extracted from these E. coli lpxd mutant strains expressing AtlpxD1 or AtlpxD2, and analyzed by LC-MS/MS. The major [M-2H]^{2-} ion of lipid A species from lpxd::AtlpxD1 strain exhibits a mass-to-charge ratio (m/z) 897.7 (Fig 7A), and the [M-2H]^{2-} (m/z) 897.7 ion was confirmed to represent a lipid A molecule by MS/MS fragmentation (Fig 7B). This finding shows that the 4 primary hydroxyacyl chains in the hybrid lipid A molecule are 3-hydroxymyristyl chains and, thus AtLpxD1 prefers to use 3-hydroxymyristyl-ACP as the substrate for the transacylation reaction. In parallel, AtLpxD2 was confirmed to prefer 3-hydroxymyristyl-ACP as the optimal substrate (data not shown). The acyl chain specificities of AtLpxD1 and AtLpxD2 for 3-hydroxymyristyl-ACP, together with the same acyl chain specificity of AtLpxA (25), suggested that Arabidopsis lipid A-like molecules bear only the 3-hydroxymyristyl chain as the primary acyl moiety, which is consistent with the previously detected chemical structures of lipid A intermediates (e.g., lipid X and lipid IVₐ) in Arabidopsis lines with mutations in the LPX genes (23).

8. Mutations in the mtKAS gene lead to the reduced levels of 3-hydroxymyristic acid

The levels of 3-hydroxymyristic acid were characterized in different organs of the atlpxa-1 mutant allele and its wild-type siblings to determine if the total 3-hydroxymyristic acid is mainly hydrolyzed from lipid A-like molecules. A previous study demonstrated that atlpxa-1 allele reduces the amounts of lipid A intermediates in
Arabidopsis (23) and, thus, if lipid A-like molecules are the major sink of 3-hydroxymyristic acid, reduced levels of this hydroxyl-fatty acid would be expected to be observed in \textit{atlpxa-1} mutant. Although still detectable, the levels of 3-hydroxymyristic acid are greatly depleted in various organs of \textit{atlpxa-1} mutant, regardless of whether the plants were grown in ambient air or an elevated CO$_2$ atmosphere (Fig 8A to 8H). This finding is therefore consistent with the hypothesis that lipid A-like molecules are the major 3-hydroxymyristate sink in Arabidopsis, and the depletion in 3-hydroxymyristic acid would represent the reduction in lipid A-like molecules.

The profiling of 3-hydroxymyristic acid in various organs of the wild-type plants and \textit{mtkas} mutant alleles growing in both ambient air and the elevated CO$_2$ atmosphere was performed to test whether the mtFAS system contributes to the biosynthesis of lipid A-like molecules (Fig 8A to Fig 8H). The strongest reduction in 3-hydroxymyristic acid occurs in the aerial organs of the \textit{mtkas} mutant plants; in these plants levels are 20\% relative to the wild-type plants. Growing plants in an elevated CO$_2$ atmosphere did not alter the levels of 3-hydroxymyristic acid in the \textit{mtKAS} mutants. The levels of 3-hydroxymyristic acid are also decreased in roots, flowers, and siliques of the \textit{mtkas} alleles. The remaining 3-hydroxymyristic acid in the \textit{mtkas} mutant alleles is possibly due to other acyl-molecules, such as 3-hydroxymyristyl-ACP of fatty acid biosynthesis in plastids and mitochondria. Another model to explain the residual 3-hydroxymyristic acid is that the truncated mtKAS enzymes in the T-DNA-tagged mutants may display activities in the physiological conditions.

Confirmation that the reduced levels of 3-hydroxymyristic acid in the \textit{mtkas} mutant plants is not due to a secondary effect of the deficiency in photorespiration, we
analyzed shmt1-2 mutant allele as a positive control (Fig 8A to Fig 8H). SHMT1 gene encodes the mitochondrial hydroxymethyltransferase that coordinates with GDC to metabolize the photorespiratory Gly; the phenotype of shmt1-2 resembles that of the mtkas mutant plants (26). The levels of 3-Hydroxymyristic acid are not affected by the shmt1-2 allele whether the plants are grown in ambient air or an atmosphere enriched in CO2. As expected when the shmt1-2 mutants were grown in the elevated CO2 atmosphere, the aerial organs of these plants appeared near normal as compared to the wild-type plants. These results therefore demonstrate that the alteration in 3-hydroxymyristic acid in the mtkas mutant alleles is a direct consequence of the deficiency in mtFAS, rather than the indirect result of the deficiency in photorespiration.

9. Identification of differentially expressed genes by RNA-seq

To assess alterations in the gene expression associated with mutations in the mtkas gene, RNA-seq experiment was performed on two sibling genotypes, the wild-type plants and mtkas-2 mutant plants, and both were grown in the atmospheres at two CO2 concentrations (i.e., ambient CO2 and 1% CO2). Triplicate biological samples were isolated and RNA samples were subjected to sequencing with an Illumina HiSeq 2000. Gene expression levels were normalized for each gene model in each sample and determined as the number of fragments per kilobase of exon per million fragments mapped (FPKM).

A generalized linear model analysis was performed based on a negative binomial distribution to test the differentially expressed (DE) genes between the two genotypes that were grown in the same CO2 atmosphere. While controlling the false discovery rate
(FDR) at 0.5% using Benjamini and Hochberg’s method, we identified 3883 DE genes in ambient CO₂ atmosphere and 436 DE genes in an atmosphere of 1% CO₂.

10. Gene clusters and functional implications

We manually assigned these DE genes into 8 clusters (Fig 9A), each of which exhibits distinct patterns on the gene expression based on alterations associated with the mtkas-2 allele and on alterations in the growth condition (i.e., CO₂ levels in the atmosphere). In cluster 1 and 2, genes are affected by the mtFAS deficiency, due to their alterations in the elevated CO₂ atmosphere. However, we are not able to evaluate if the photorespiration deficiency contributes to the differential expression of these genes. In cluster 3, 4, 5, and 6, genes are likely to be affected by deficiencies in both mtFAS and photorespiration, which regulates the gene expressions in the opposite directions. Thus, in the elevated CO₂ atmosphere, DE genes are affected by the mtFAS deficiency, but when incorporating the photorespiration deficiency in ambient air, alterations in the gene expression either change to the opposite directions (cluster 3 and 4) or become unchanged (cluster 5 and 6). Finally, around 90% of the DE genes belongs to cluster 7 and 8, where the differential expression are only detected when the plants are grown in ambient air. These genes are only affected by the photorespiration deficiency. Therefore, the differential expressions of these genes are reversed when the photorespiration deficiency is supressed by the treatment of 1% CO₂.

Functional analysis of the genes within each cluster was conducted based on Gene Ontology (GO) annotations using TAIR GO Annotation Bulk Tool (www.arabidopsis.org). These genes are grouped into 14 broad functional categories (Fig 9B). Within each group, the proportion of genes that are involve in “response to stress”
and “response to abiotic or biotic stimulus” categories are higher than the proportion of genes in these two categories among all the genes in the genome. Furthermore, there are a higher proportions of genes that occur in cluster 6, 7, and 8 (i.e., genes regulated by both mtFAS and photorespiration) that belong to the functional categories of “transport”. These results illustrate that the deficiency in mtFAS and photorespiration lead to a stress response and a disturbed transportation system.

We also used Amigo GO Term Enrichment Tool (27) to explore the biological processes that are affected by the deficiency in mtFAS or photorespiration (Fig 9C). DE genes in cluster 1 to 6 (436 genes) are affected as a consequence of the deficiency in mtFAS; these genes mainly involve in the “response” processes to different endogenous or environmental factors (Fig 9D). In addition, mtFAS deficiency affects biosynthetic processes, including “myo-inositol hexakisphosphate biosynthesis” and “organic hydroxyl compound biosynthesis”, indicating that the mtFAS system may support other synthetic pathways, other than the biosynthesis of lipoic acid and lipid A-like molecules. DE genes in cluster 3 to 8 (3814 genes) are affected as a consequence of the deficiency in photorespiration. The photorespiration deficiency affects a broader set of pathways, including response process, metabolism, regulation, and transport (Fig 9E). Consistent with the alterations in lipid metabolism and thylakoid membrane that are discovered by chemical analysis and TEM microscopy, photorespiration deficiency significantly affects the processes of “response to lipid” and “thylakoid”.

11. Transcriptional regulation on the genes of acyl-lipid metabolism

The transcriptional impact on acyl-lipid metabolism was evaluated by analyzing the RNA-Seq data of the 798 genes of acyl-lipid metabolism, as categorized in ARALIP
One hundred and forty of these genes are DE in mtkas-2 mutant; 43 of them (31%) belong to cluster 7 (i.e., up-regulated by photorespiration deficiency), while 81 of them (58%) belong to cluster 8 (i.e., down-regulated by photorespiration deficiency). These results demonstrate that most of the DE genes of acyl-lipids metabolism are a consequence of the deficiency in photorespiration. These transcriptomic results are in agreement with the metabolic profiling data that most of the alterations in acyl-lipid profiles return to the wild-type level when the plants are grown in the CO2-enriched atmosphere, which suppresses the deficiency in photorespiration.

The DE genes in cluster 7 and 8 were manually categorized based on their functions in acyl-lipid metabolism (Fig 10). In “fatty acid elongation and wax biosynthesis” pathway, 7 genes are up-regulated, whereas 25 genes are down-regulated in ambient air; such genes as CER1, CER3 (28), and KCS homologs (2) that encode key enzymes of the surface lipid biosynthesis are down-regulated. These data demonstrate that the transcriptional regulation of surface lipid metabolism is likely to be the reason that surface lipid contents are reduced in the mtkas mutants. In the “triacylglycerol & fatty acid degradation” pathway, 14 genes are up-regulated, while 4 genes are down-regulated. For example, the genes LACS7 (29), ACX (30), MFP (31), and KAT (32) that encode peroxisomal enzymes that convert free fatty acids to acyl-CoA are all up-regulated in response to the mtkas-2 allele. These results reveal an up-regulated degradation process of free fatty acids, which is likely to be hydrolyzed from lipids. In the “triacylglycerol biosynthesis” pathway, only DGAT1 is up-regulated, whereas 4 genes are down regulated. Considering that the enzyme encoded by DGAT1 is a plastidially localized enzyme that is responsible for TAG assembly (33), its induction in the mtkas
mutant is likely to contribute to the observed hyperaccumulation of plastoglobules and TAG in the mutants.

In addition, three biosynthetic pathways of signaling lipid are affected in the *mtkas-2* mutant. The biosynthetic genes of sphingolipids (6 gene being down-regulated) and signaling phospholipids (1 gene being up-regulated and 14 gene down-regulated) are suppressed in the mutant, whereas oxylipin metabolism genes are irregularly affected (7 genes being up-regulated and 5 genes being down-regulated). The differential expression of these genes illustrates a transcriptional regulation of lipid metabolism as a response to stress stimulus (8,34,35), which is consistent with the Gene Ontology (GO) annotation analysis that indicates the *mtkas* mutant are in a state of abiotic stress.

In contrast to those transcriptionally regulated pathways of acyl-lipid metabolism, most of the genes involved in the biosynthesis of galactolipids, sulfolipids, and lipid A-like molecules are not differentially expressed in the *mtkas-2* mutant allele relative to the wild-type plants. These data demonstrate that the defective metabolism of these lipids is unlikely to be regulated at the transcriptional levels.

**DISCUSSION**

Central to the process of fatty acid biosynthesis is the iterative Claissen condensation catalytic mechanism that drives the formation of new carbon-carbon bonds. These reactions are catalyzed by a series of homologous 3-ketoacyl synthase enzymes, which condense a malonyl-moiety with an acyl-moiety, resulting in the elongation of the acyl-chain. Both these acyl substrates are presented to the 3-ketoacyl synthase enzymes on thioester-bound carriers (acyl carrier protein (ACP) or Coenzyme A (CoA)), and the synthase homologs display different substrate specificities in terms of these carriers and
the nature of the acyl substrates (1,14). Plant cells appear to have at least three fatty acid forming systems, which occur in distinct subcellular compartments: the plastids, the ER membrane, and mitochondria. The plastidial fatty acid synthase (FAS) is a Type II system that utilizes ACP as the predominant carrier of the intermediates of the process. Moreover, initial characterizations of the distribution of ACP in plant cells lead to the concept that plastids are the sole location of de novo fatty acid biosynthesis from acetyl-CoA (36). In contrast to the plastidial system, the ER-located fatty acid biosynthesis system is analogous, but utilizes preexisting fatty acyl-CoAs as substrates, and elongates them using malonyl-CoA as the donor of the 2-carbon elongation unit.

The occurrence of a mitochondrial FAS (mtFAS) system was first suggested by the discovery of a mitochondrially located ACP in Neurospora crassa (37). Since then the mtFAS system has been characterized in yeast and humans (5). Although many of the enzyme components of mtFAS have yet to be identified in plant cells, the role of mtFAS in supplying octanoic acid, the precursor of lipoic acid biosynthesis has been established in plant cells (3). Apart from mtACP (38,39), the only other component of plant mtFAS that has been characterized is the mitochondrial 3-ketoacyl-ACP synthase (mtKAS), which has been genetically characterized via a genetic complementation experiment that demonstrates the ability of this enzyme to complement the E. coli kasI/kasII double mutant by contributing to the formation of acyl-ACP intermediates of up to 18-carbon chain length (16).
In vitro characterization of the acyl-chain length substrate specificity of the mtKAS enzyme

Our in vitro characterization of the recombinant mtKAS protein indicates that it displays broad substrate specificity with respect to acyl-ACP substrates, using acyl chains ranging from 4 to 16-carbons in length. These data are in agreement with the previous in vivo mtKAS gene complementation of the E. coli kasI/kasII double mutant, which contributes to the formation of fatty acids of up to 18-carbon chain length (16). It is generally believed that substrate specificity of KAS enzymes is one of the factors that determine the chain length distribution of fatty acids that different FAS systems can produce (16). In this respect, the mtFAS system is thought to preferentially produce medium chain fatty acids of less than 12-carbons, which fits with the need for octanoic acid needed to support lipoic acid biosynthesis. Indeed, in vitro assays using mitochondria crude extracts of pea suggest the enzymatic capacity prefers to produce such medium acyl-ACP products (40). Moreover, the activity of human mtKAS falls off significantly with acyl-ACP substrates longer than C12 (17). Therefore, our in vitro characterization of mtKAS suggests that this enzyme is not necessarily the determinant of the types of fatty acids that the mtFAS system can produce. Moreover, our own characterization of 3-hydroxymyristic acid as the acyl moiety of mitochondrialy located lipid A-like molecules indicate that the plant mtFAS system has the ability to not only produce octanoic acid for lipoic acid biosynthesis, but also longer chain-length fatty acids of at least 14-carbon and possibly 16-carbon atoms.
**Novel alterations of photorespiratory deficiency caused by the mtkas alleles**

T-DNA knockout alleles of the Arabidopsis *mtKAS* gene display a photorespiration-mutant phenotype, due to the loss of lipoylation of the H subunit of GDC, the essential photorespiration enzyme (11). Those previously identified visible and metabolic phenotypes include reduced plant size, pale-green leaf color, and the over-accumulation of Gly. The fact that these morphological and metabolic phenotypic differences are reversed by growth conditions that suppress photorespiration (e.g., growth in an elevated CO₂ atmosphere) indicated the role of this process in explaining the phenotypes of the *mtkas* mutations (11). Here, we expand on these prior characterizations and further dissect the connection between metabolic modulations associated with blocking mtFAS and the phenotypic consequence of this genetic lesion. Several novel morphological and ultrastructural phenotypes associated with the *mtkas* mutations include the increased size of leaf mesophyll cells, which appears to lead to considerably thicker leaves, even though overall leaf volume appears to be reduced. The ultrastructure of the chloroplasts is significantly altered, with a depletion of thylakoid membranes associated with the general “shriveled” appearance of these chloroplasts. Finally, the *mtkas* mutants hyper-accumulate plastoglobules, which normally are associated with leaf senescence (41). These findings clearly suggest that impaired chloroplast function is a consequence of the deficiency in photorespiration.

Some phenotypes are likely to be the results of the defective lipid metabolism. Our data show the severely reduced contents of galactolipids and sulfolipids in *mtkas* mutant alleles. These results suggest that the depleted thylakoid membrane in *mtkas* alleles is likely to be a result of the insufficient galactolipids and sulfolipids, because
these two classes of lipids compose thylakoid membrane systems and are mainly detected in plastidial membranes (7). In addition, TAG is accumulated in the mtKAS mutants, which is in agreement with the increased number of large plastoglobules observed in mtkas-2. It is possible that the exceeded TAG is synthesized in plastids, where it further facilitates the buildup of plastoglobules.

The defective lipid metabolism in the mtkas alleles is likely to be attributed to the poor sucrose and accumulated glycolate. Sucrose is the transport form of carbon in plants, and in photorespiratory mutants, impaired recycling of photorespiratory carbon results in low availability of carbon for sucrose synthesis (21,42). Indeed, our data show that sucrose levels are dramatically reduced to less than 5% in the mtkas alleles, implying an impaired sucrose metabolism due to the deficiency in photorespiration in the mtkas alleles. This deduction is also supported by the similar findings in other photorespiratory mutants, like 10-fdf and shmt1-2 (21). The effect of poor sucrose was evaluated by growing the mtkas mutant alleles on plates supplemented with sucrose. Interestingly, the reduced contents of galactolipids and sulfolipids are greatly alleviated by sucrose treatment. These data suggests that depleted sucrose leads to the defective galactolipid and sulfolipids metabolism, however, the exact biochemical mechanism is still to be investigated. Glycolate, an upstream metabolite of glycine decarboxylase, is accumulated in the mtkas alleles. The effect of the elevated glycolate was examined by growing the wild-type plants on plates containing glycolate; exogenous glycolate induced the TAG accumulate in a mtkas-like pattern. This result implies that the accumulated TAG may be the metabolic responses to the glycolate stress in the mtkas mutant alleles. However, we cannot exclude the possibility that accumulated TAG in mtKAS mutants may also
function as the transient intermediates to store those acyl chains de-esterified from galactolipids and sulfolipids, in order to prevent the detrimental levels of free fatty acids (43).

**Biochemical functions of long-chain acyl-ACP intermediates of mtFAS**

In eukaryotes, lipid A-like molecules are unique to plants, since only plants contain significant homologs encoding mitochondrial enzymes for the biosynthesis of lipid A-like molecules, namely AtLpxA, AtLpxC, AtLpxD, AtLpxB, AtLpxK, and AtKdtA (23,44). Although the structures of plant lipid A-like molecules have yet to be discovered, some of their intermediates have been characterized in the *lpx* and *kdta* mutant alleles (23). Our results of the substrate specificity of AtLpxD1 and AtLpxD2, in addition to that of AtLpxA (25), demonstrate that, regardless of the final structure, lipid A-like molecules bear 3-hydroxymyristyl chains as their primary acyl moiety. Therefore, we show that the levels of 3-hydroxymyristic acid can represent the contents of lipid A-like molecules, because they are depleted in the T-DNA knockout allele of the *AtLpxA* gene, which encodes the enzyme catalyzes the first reaction of biosynthesis of lipid A-like molecules. We found that the levels of 3-hydroxymyrystate are depleted in the *mtkas* mutant alleles, but remain normal levels in *shmt1-2* allele, a control plant with the deficiency in photorespiration. These data demonstrate that 3-hydroxymyristyl-ACP generated from the mtFAS system is donated for the biosynthesis of lipid A-like molecules in Arabidopsis, pointing to a novel biochemical destination of the mitochondrial long-chain acyl-ACP intermediates.

The physiological functions of plant lipid A-like molecules are still elusive. They are not essential for plant growth and development, since no morphological difference is
observed on null mutants of *AtLpxA*, *AtLpxD1*, *AtLpxD2*, *AtLpxB*, *AtLpxK*, nor *AtKdtA*, nor *AtLpxC-RNAi* knockdown plants (23,44). Metabolic changes are barely detected in the *AtLpxB*, *AtLpxK*, and *AtKdtA* mutants from NSF-funded Metabolomics 2010 Project data (45) visualized in Plant & Microbial Metabolomics Resource (PMR) website (http://www.metnetdb.org/PMR). Lipid A-like molecules could function similarly in plant mitochondria as bacteria for morphology maintenance by providing hydrophobic anchors of polysaccharide (23). This hypothesis receives supports from the observations that giant mitochondria are generated in Arabidopsis cells transiently overexpressing AtKdsB, an enzyme provides CMP-Kdo substrates for the modification of lipid A-like molecules (46). In addition, plant lipid A-like molecules have been suggested to transport to plastids after their synthesis in mitochondria and, thus, they are probably associated with certain functions in plastids (23). The physiological functions of Arabidopsis lipid A-like molecules deserve additional investigation in the future.

**METHODS**

**Plant materials and growth conditions**

The Arabidopsis genetic stocks carrying the *mtkas-2* (SALK_022295), *mtkas-3* (SALK_087186), and *shmt1-2* (SALK_083735) alleles are in Col-0 background, and were obtained from the Arabidopsis Biological Resource Center (Columbus, OH) (47). The *atlpxa-1* allele, which is in the WS background and isolated from the Arabidopsis Functional Genomics Consortium T-DNA mutant population (23,48), and was a kind gift from Dr. Christian R Raetz (Duke University).

Seeds were sterilized and sown on Murashige and Skoog agar medium as describe previously (49). In the biochemical feeding experiment, the medium was supplemented
with each biochemical (10mM glycolate, 10mM glyoxylate, 10mM glycine, or 3% sucrose), and the medium was adjusted to pH 5.7 before sterilization. After sowing seeds, plates were placed at 4°C for 4 days to break seed dormancy, and plants were grown in ambient air or the non-photorespiratory conditions (1% CO₂ in a growth chamber). The growth conditions were maintained at 23 ± 2°C with continuous illumination (100 µmol m⁻² s⁻¹). If needed, seedlings were transferred from plates to soil (LC1 Professional Growing Mix) at 7 to 10-day after germination.

**RNA extraction and quantitative RT-PCR**

RNA was extracted from 50 mg fresh tissues using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA). A 2µg aliquot of RNA was used for reverse transcription using the SuperScript III First-Strand Synthesis Super Mix (Life Technologies, Carlsbad, CA) with oligo (dT)₂₀ primer. Quantitative RT-PCR was performed on StepOnePlus Real-Time PCR System using SYBR Select Master Mix (Life Technologies, Carlsbad, CA). *Actin-2* (AT3G18780) was used as the reference RNA for relative quantification. Gene specific primers are K1 (GGATTCTCTAGATCAAGGGCTT) - K2 (CCTTCCCCTATCACAAAACCAT) for mtKAS and K3 (TCGTACAACCGGTATTGTGCTG) - K4 (AGGTCACGTCCAGCAAGGTCAA) for Actin-2.

**In vitro kinetic assays of the recombinant Arabidopsis mtKAS enzyme**

The ORF encoding Arabidopsis *mtKAS* and *Streptococcus pneumoniae* ACPS (*SpACPS*) were chemically synthesized by GenScript (Piscataway, NJ) after codon optimization for the expression in *E. coli*. *E. coli* genes encoding AcpP, *fabD*, and *fabG* were cloned from wild-type *E. coli* K12 strain (CGSC, Yale University). The PCR
product was cloned into the pET30b vector (EMD Millipore, Billerica, MA), and the resulting construct was named $mtKAS$-pET, $SpACPS$-ET, $AcpP$-pET, $fabD$-pET, and $fabG$-pET. All of these plasmids express recombinant enzymes with a His-tag located at the C-terminus. Protein expression was performed in E. coli BL21* strain (Life Technologies, Carlsbad, CA). His-tagged proteins were purified using Probond Nickel-Chelating Resin following the manufacturer’s instruction (Life Technologies, Carlsbad, CA).

To assay the mtKAS enzymatic activity, the holo-ACP and acyl-ACP substrates were prepared as previously described with modifications (17). Briefly, the recombinant ACP overexpressed by E. coli is a mixture of apo- and holo-forms. Apo-ACP in the mixture and acyl-CoA molecules were catalyzed by SpACPS to form acyl-ACP as described previously (50), resulting in the formation of a mixture of acyl-ACP and holo-ACP; the percentage of acyl-ACP is quantified by MALDI-QTOF analysis. Apo-ACP in the mixture and CoA were catalyzed by SpACPS to form holo-ACP. The mixtures of acyl-ACP and holo-ACP were used as the substrates for the kinetic assays as previously described (17) with modifications. The reaction mixture (total volume of 3mL) contains 0.2 M potassium phosphate buffer (pH 6.8), 1mM EDTA, 0.3mM dithiothreitol, 50µM holo-ACP, 50µM malonyl-CoA, 1µM recombinant fabD, 300µM NADPH, 1µM recombinant fabG, 1-50µM acyl-ACP, and 100nM recombinant mtKAS. Malonyl-ACP was first generated from malonyl-CoA and holo-ACP using recombinant fabD by incubating for 15min at 25°C. The reaction was started by addition of acyl-ACP, recombinant mtKAS, NADPH, and recombinant fabG, and incubated at 25°C. The
reduction in the absorbance at 340nm was measured every 3min for 5 times. Kinetic values are calculated using Prism Verion 5.0 (GraphPad Software, La Jolla CA).

**Fatty acid analysis**

Fatty acid analyses were conducted as described previously with some modifications (22). In brief, 50mg fresh tissue or 30 dry seeds were pulverized using Duall 20 Tissue grinder (Kontes Glass, Vineland, NJ), and were transmethylated at 80°C for 30min in 1mL 1N methanolic HCl containing 10µg/mL butylated hydroxytoluene and 5µg/mL pentadecanoic acid (15:0) as an internal standard. One milliliter of 0.9% NaCl and 1mL of heptane were added and the methylated mixture was vigorously vortexed. After centrifugation at 5000g for 10min, 1uL of the heptane phase was analyzed on GC-MS (Agilent 7890 series).

For the analysis of 3-hydroxy fatty acid, the methylation reaction was extended to 16h at 80°C in order to facilitate the transmethylation of the amide bonds in the lipid A-like molecules (51). The heptane phase was collected, dried under a stream of N₂ gas, and the hydroxyl groups were silylated in 100µL BSTFA at 65°C for 20min. The samples were dried under a stream of N₂ gas and dissolved in heptane for GC-MS analysis.

**Amino acid analysis**

Amino acids were analyzed by HPLC in the form of o-phthalaldehyde (OPA)-derivatized fluorophore. In brief, free amino acids were extracted at 85°C for 10min from 5 mg lyophilized tissue powder with 1mL water containing 10uM butylamine as an internal standard. After brief centrifugation, the samples were filtrated through the 0.45µm Iso-Disc Filter (Sigma-Aldrich, St. Louis, MO). OPA Buffer was prepared by mixing 12.25mg OPA, 312.5uL methanol, 6mL 0.4M potassium tetraborate tetrahydrate
(pH 9.5, adjusted with 10% H₂SO₄), and 19.22μL β-mercaptoethanol. The amino groups of metabolites were derivatized with OPA in a 3min in-loop reaction between 6μL samples and 6μL OPA Buffer as described previously (52). Chromatographic analysis was performed using a Hypersil ODS column (Thermo, 250mm X 4mm, 5μm) on an HPLC system (Agilent 1200 series). The two mobile phases were (A) 10mM sodium phosphate (pH=7.3)/ methanol (90/10, v/v) and (B) 10mM sodium phosphate (pH=7.3)/ methanol (20/80, v/v). The elusion program was as follow: it started with 100% mobile phase A; the mobile phase B was increased stepwise to 15%, 20%, 30%, 40%, 80%, and 100% over 5min, 8min, 8min, 4min, 20min, 1min intervals, respectively; the solvent was held at 100% mobile phase B for 5min; the LC gradient was then returned to 100% mobile phase A over 1min and held at 100% mobile phase A for 8min to equilibrate the column. The flow rate was 1mL/min. The fluorescence signal was detected at λ_{ex}/λ_{em} of 337nm/454nm.

**Glycolate, glyoxylate, and sucrose analysis**

Soluble metabolites were extracted from the lyophilized tissue powder as described previously (53), and divided to two portions. One portion was subjected to GC-MS analysis for glycolate and glyoxylate (Agilent 7890 series). For these analyses, the samples were methoximated as previously described (53), followed by the silylation by adding 50μL of 20% (v/v) MTBSTFA in acetonitrile and incubating at 65°C for 30min. 2-Ethylbutyric acid was used as the internal standard for the quantification of glycolate and glyoxylate. The other portion of the extract was subjected to sucrose analysis by GC-MS as described previously (21). Ribitol was used as internal standard for quantification.
Lipidomics analysis

Lipids were extracted from 16-day-old Arabidopsis aerial organs as previously reported (54) with modifications. In brief, 50mg fresh tissue were collected in 2mL Eppendorf tubes, quickly immersed in liquid N₂, and pulverized using 2 stainless steel beads in a Mixer Mill MM301 (Retsch, Haan, Germany) at 15Hz for 2min. The lipids were extracted with 0.8mL of tert-butyl methyl ether/methanol (75/25, v/v) containing 1,2-didecanoyl-sn-glycero-3-phosphocholine as an internal standard. The mixture was mixed for 5min at 25°C, and then placed on ice for 15min. After centrifugation at 1000g at 4°C for 10min, 100µL of the upper layer was transferred to a new Eppendorf tube and dried under a stream of N₂ gas. After removal of the solvents using a centrifugal concentrator, the residue was dissolved in 125µL of ethanol, and centrifuged at 10,000g at 4°C for 15 min.

Supernatant was subjected to LC-MS analysis using an LC-MS system equipped with an electrospray ionization (ESI) interface (HPLC, Waters Acquity UPLC system; MS, Waters Xevo G2 QTOF) as reported previously (55). The lipidomics analysis detected MGDG, DGDG, SQDG, PC, PE, PI, PG, PA, DAG, and TAG. Within each lipid species defined by the head group, mass data specifically identify the total number of carbons and the total number of double bonds present in the acyl groups.

Genetic complementation of an E. coli lpxD knockout strain

The ORFs for AtlpxD1 (AT4G05210) and AtlpxD2 (AT4G21220) were cloned from Arabidopsis cDNA generated from the 16-day-old aerial organs. PCR was conducted, removing 5’ sequences that encode the putative mitochondrial transit peptides. The PCR products were then cloned into the pBE522 vector (56), resulting in
the constructions of \textit{AtlpxD1-pBE} and \textit{AtlpxD2-pBE}. The bacterial \textit{lpxD} gene was deleted by using a PCR-based method (57). The entire \textit{lpxD} coding sequence was removed and replaced with the gentamycin resistant gene. The gene replacement allele was verified by PCR and sequencing.

**Isolation and characterization of hybrid lipid A species**

Hybrid lipid A molecules were extracted from the \textit{E. coli lpxd} mutant strain as described previously (25). The lipid A extracts were dissolved in chloroform/methanol (2/1, v/v), and analyzed using a LC-MS system equipped with an electrospray ionization (ESI) interface (HPLC, Agilent 1100 Series; Column, Ascentis Si Supelguard cartridge; MS, Agilent MSD Trap). Two mobile phases were consisted of (A) chloroform/methanol/ aqueous ammonium hydroxide (800/195/5, v/v/v) and (B) chloroform/methanol/ water/ aqueous ammonium hydroxide (600/340/5/5, v/v/v/v). The elution program was as follow: 100% mobile phase A was held for 2min and then linearly increased to 100% mobile phase B over 2min and held at 100% mobile phase B for 5min; the LC gradient was then returned to 100% mobile phase A over a 1 min interval and held at 100% mobile phase A for 2min. The flow rate was 1mL/min. MS data were collected in the negative mode. The major MS peak \([M-2H]^{2-} (m/e 898) was selected for MS/MS fragmentation.

**LM and TEM of leaves**

For light microscopy (LM) and transmission electron microscopy (TEM), samples were collected and fixed with 2% glutaraldehyde (w/v) and 2% paraformaldehyde (w/v) in 0.1M cacodylate buffer, pH 7.2 for 48 hours at 4°C. Samples were rinsed 3 times in 0.1M cacodylate buffer and then post-fixed in 1% osmium tetroxide in 0.1 M cacodylate
buffer for 2 hours at room temperature. The samples were rinsed in deionized distilled water and stained with 2% aqueous uranyl acetate for 4 hours, dehydrated in a graded acetone series, cleared with ultra-pure acetone, infiltrated and embedded using Spurr’s epoxy resin (Electron Microscopy Sciences, Ft. Washington, PA). Resin blocks were polymerized for 48 hours at 65°C. Thick and ultrathin sections were made using a Reichert Ultracut S ultramicrotome (Leeds Precision Instruments, Minneapolis, MN). Thick 1µm sections were stained with 1% toluidine blue and imaged with an Olympus BX-40 light microscope (Olympus America Inc, Center Valley, PA). Ultrathin sections were collected onto carbon coated copper grids. Images were captured using a JEOL 2100 scanning and transmission electron microscope (Japan Electron Optic Laboratories, Peabody, MA).

**RNA-seq**

Total RNA was extracted from pooled leaf samples according to the TRIzol RNA isolation protocol. It was purified using the QIAGEN RNasy Mini Kit (QIAGEN, Valencia, CA) with DNAse I treatment to remove any DNA contamination.

The 200 bp short-insert library was constructed. Transcriptome sequencing was performed on Illumina HiSeq2000 using V3 Reagent (91 air end sequencing). The low quality reads were filtered out by removing reads with adaptors, reads with unknown nucleotides larger than 5%, and low quality reads (The bases which quality ≤ 10 is more than 20% of the reads). The cleaned reads were aligned to the reference genome of *Arabidopsis thaliana* in Phytozome v 8.0 (http://www.phytozome.net) respectively using TopHat (58). The mapped reads were counted by htseq-count (http://www-huber.embl.de/users/anders/HTSeq/doc/count.html).
The count data were analyzed as a completely randomized design with four treatments, one for each combination of genotype and growth condition. Each treatment was replicated three times. Genes with an average of at least one uniquely mapped read across samples and a number of nonzero read counts at least as large as the number of treatments were tested for differential expression between genotypes within each growth condition, as well as tested for differential expression between growth condition within each genotype, using the R package QuasiSeq (http://cran.r-project.org/web/packages/QuasiSeq). The negative binomial QLShrink method implemented in the QuasiSeq package as previously described (59) was used to compute a p-value for each gene and each genotype comparison. The log of each count mean was modeled as the sum of an intercept term, a genotype effect, a growth condition effect, an interaction between genotype and growth condition, and an offset normalization factor, determined for each sample by the log of the TMM normalization factor (60). Estimates of the fold change between control conditions and the elevated CO₂ atmosphere, and between genotypes, were computed by evaluating the exponential function at estimates of effect differences. Using the p-values for each comparison, the approach as previously described (61) was used to estimate the number of genes with true null hypotheses among all genes tested, and this estimate was used to convert the p-values to q-values (62). To obtain approximate control of the false discovery rate at 5%, genes with q-values no larger than 0.05 were declared to be differentially expressed.

REFERENCES


Proceedings of the National Academy of Sciences of the United States of America **108**, 11387-11392


FIGURES

Figure 1. Substrate specificity of the recombinant mtKAS enzyme. (A) Graph of the substrate concentration dependence. (B) Kinetic parameters.

<table>
<thead>
<tr>
<th>Acyl-ACP</th>
<th>Km (µM)</th>
<th>Vmax (nmol/min/mg)</th>
<th>$k_{cat}/K_m$ (nM$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C4</td>
<td>19.1 ± 3.0</td>
<td>144.7 ± 8.1</td>
<td>9.4</td>
</tr>
<tr>
<td>C6</td>
<td>8.0 ± 1.2</td>
<td>193.1 ± 2.0</td>
<td>29.8</td>
</tr>
<tr>
<td>C8</td>
<td>22.9 ± 2.4</td>
<td>444.1 ± 17.0</td>
<td>24.0</td>
</tr>
<tr>
<td>C10</td>
<td>14.1 ± 1.7</td>
<td>237.4 ± 7.5</td>
<td>20.9</td>
</tr>
<tr>
<td>C12</td>
<td>25.6 ± 6.4</td>
<td>317.6 ± 30.2</td>
<td>15.4</td>
</tr>
<tr>
<td>C14</td>
<td>14.0 ± 1.2</td>
<td>340.8 ± 9.3</td>
<td>30.2</td>
</tr>
<tr>
<td>C16</td>
<td>8.2 ± 1.2</td>
<td>359.9 ± 14.3</td>
<td>51.2</td>
</tr>
<tr>
<td>C16:1</td>
<td></td>
<td>Activity too low for kinetic assay</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2. Expression profiles of the *mtKAS* gene in various organs.
Figure 3. LM and TEM of the leaf sections of the wild-type and \textit{mtkas}-2 mutant plants. (A) and (B) LM of the leaf sections of the wild-type and \textit{mtkas}-2 mutant plants. (C) and (E) TEM of chloroplasts and thylakoid membranes in the wild-type plants. (D) and (F) TEM chloroplasts and thylakoid membranes in \textit{mtkas}-2 mutant plants.
Figure 4. Morphological appearance of the wild-type, *mtkas*-2, and *mtkas*-3 plants as affected by elevated CO₂, glycolate, glyoxylate, Gly, and sucrose. (A) The wild-type, *mtkas*-2, and *mtkas*-3 plants grown on the normal plates in ambient air. (B) The wild-type, *mtkas*-2, and *mtkas*-3 plants grown on normal plates in 1% CO₂ atmosphere. (C) The wild-type, *mtkas*-2, and *mtkas*-3 plants grown with exogenous 3% sucrose in ambient air. (D) The wild-type plants grown with exogenous 10mM glycolate, 10mM glyoxylate, and 10mM Gly in ambient air.
Figure 5. Alterations in soluble metabolites. (A) Log2 transformation of the ratio of alterations between mtkas-2 and the wild-type plants in ambient air, CO2 treatment condition, and sucrose treatment condition. (B) Log2 transformation of the ratio of alterations between the wild-type plants with exogenous chemicals (i.e., 10mM glycolate, glyoxylate, and Gly) and the non-treated wild-type plants.
Figure 6. Alterations in acyl-lipids. (A) Log2 transformation of the ratio of alterations in FA. (B) Alterations in Galactolipids, sulfolipids, and DAG. (C) Alterations in TAG. (D) Alterations in Surface lipids.
Figure 6 continued.
Figure 6 continued.
Figure 7. LC-MS/MS analysis of the hybrid lipid A species. (A) Negative ion m/z 897.7 of lipid A species isolated from *E. coli lpxd::AtlpxD1* allele. (B) The daughter ions m/z 784.3 and 679.4.
Figure 8. Levels of 3-hydroxymyristic acid in various organs. (A) to (D) Levels of 3-hydroxymyristic acid in aerial organs, roots, flowers, and siliques of plants grown in ambient air. (E) to (H) Levels of 3-hydroxymyristic acid in aerial organs, roots, flowers, and siliques of plants grown in a 1% CO₂ atmosphere.
Figure 9. RNA-seq. (A) Gene clusters. (B) Functional analysis. (C) Genes of mtFAS-effect and Photorespiration-effect. (D) GO term enrichment analysis of genes affected by the deficiency in mtFAS. (E) GO term enrichment analysis of genes affected by the deficiency in photorespiration.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Alterations in ambient CO₂</th>
<th>Alterations in elevated CO₂</th>
<th># of GE genes</th>
<th>Biological description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>Up</td>
<td>Up</td>
<td>60</td>
<td>mtFAS deficiency (Up)</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>Down</td>
<td>Down</td>
<td>162</td>
<td>mtFAS deficiency (Down)</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>Up</td>
<td>Down</td>
<td>14</td>
<td>mtFAS deficiency (Down) Photorespiration deficiency (Up)</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>Down</td>
<td>Up</td>
<td>47</td>
<td>mtFAS deficiency (Up) Photorespiration deficiency (Down)</td>
</tr>
<tr>
<td>Cluster 5</td>
<td>--</td>
<td>Up</td>
<td>90</td>
<td>mtFAS deficiency (Up) Photorespiration deficiency (Down)</td>
</tr>
<tr>
<td>Cluster 6</td>
<td>--</td>
<td>Down</td>
<td>63</td>
<td>mtFAS deficiency (Down) Photorespiration deficiency (Up)</td>
</tr>
<tr>
<td>Cluster 7</td>
<td>Up</td>
<td>--</td>
<td>1574</td>
<td>Photorespiration deficiency (Up)</td>
</tr>
<tr>
<td>Cluster 8</td>
<td>Down</td>
<td>--</td>
<td>2026</td>
<td>Photorespiration deficiency (Down)</td>
</tr>
</tbody>
</table>

B

- Electron transport or energy pathways
- DNA or RNA metabolism
- Transcription,DNA-dependent
- Signal transduction
- Other biological processes
- Cell organization and biogenesis
- Transport
- Protein metabolism
- Response to abiotic or biotic stimulus
- Developmental processes
- Unknown biological processes
- Response to stress
- Other metabolic processes
- Other cellular processes
Figure 9 continued.

C

Cluster 1
mtFAS-effect: +

Cluster 4
Cluster 5

Photorespiration-effect: -
Cluster 8

Cluster 2
mtFAS-effect: -

Cluster 3
Cluster 6

Photorespiration-effect: +
Cluster 7

D

Response to monosaccharide
Response to alcohol
Response to hexose
Response to inorganic substance
Response to carbohydrate
Response to temperature stimulus
Response to karrikin
Response to fructose
Response to salt stress
Circadian rhythm rhythmic process
Response to osmotic stress
Response to acid
Polyol metabolic process
Inositol phosphate metabolic process
Response to sucrose
Response to disaccharide
Response to hormone
Response to endogenous stimulus
Response to stress
Polyol biosynthetic process
Inositol phosphate biosynthetic process
Myo-inositol hexakisphosphate biosynthetic process
Myo-inositol hexakisphosphate metabolic process
Response to radiation
Response to light stimulus
Response to organic substance
Response to chemical
Response to oxygen-containing compound
Response to abiotic stimulus
Response to stimulus

-log(_10)(P-value)

0 5 10 15 20 25 30 35 40
Figure 9 continued.
Figure 10. Alterations in the genes of acyl-lipid metabolism.
CHAPTER III: IDENTIFICATION AND REVERSE GENETIC CHARACTERIZATION OF THE 3-HYDROXYACYL-ACP DEHYDRATASE COMPONENT OF MITOCHONDRIAL FATTY ACID SYNTHASE OF ARABIDOPSIS

Manuscript in preparation and to be submitted to The Plant Cell

Xin Guan¹, Ling Li², Yozo Okazaki³, Xuefeng Zhao⁴, Andrew Lithio⁵, Dan Nettleton⁵, Kazuki Saito³, and Basil J. Nikolau¹

¹ Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, Iowa 50011
² Department of Genetics, Development, and Cellular Biology, Iowa State University, Ames, Iowa 50011
³ Metabolomic Function Research Group, RIKEN Plant Science Center, Yokohama 230-0045, Japan
⁴ Plant Sciences Institute, Iowa State University, Ames, Iowa 50011
⁵ Department of Statistics, Iowa State University, Ames, Iowa 50011

ABSTRACT

Arabidopsis mitochondrial fatty acid synthase (mtFAS) is considered to be a multi-component Type II FAS system, which generates fatty acid precursors for the biosynthesis of lipoic acid and lipid A-like molecules. Three of the four enzyme components that catalyze the iterative reactions of this process are unknown. In this study, we report the identification and characterization of the Arabidopsis gene locus (AT5G60335) as the mitochondrial 3-hydroxyacyl-ACP dehydratase (mtHD), which catalyzes the 3rd reaction in the 4-reaction cycle. Evidence for this conclusion includes: 1)
the targeting of a AT5G60335-GFP chimeric protein to mitochondria; 2) expression of the AT5G60335-encoding protein can functionally replace the yeast mtHD counterpart; and 3) the bacterially expressed recombinant AT5G60335-encoding protein can support the \textit{in vitro} catalysis of the hydration of enoyl-CoA, the reverse reaction of mtHD. Using an RNAi transgene, we generated a series of Arabidopsis lines with a range of down-regulated \textit{mtHD} expression levels. Plants with the decreased levels of the \textit{mtHD} expression display the same morphological and metabolic phenotypes as mutations in the mitochondrial β-ketoacyl-ACP sy nthase (mtKAS), the only known enzyme of the mtFAS system. These morphological and metabolic phenotypes include miniaturized aerial organs, yellowish leaf color, alterations in the pool sizes of photorespiratory intermediates, and defective metabolism of a series of lipids and amino acids. In addition, altered levels of chlorophylls and their degradation intermediates were detected in these mutants. Most of these alterations are a consequence of the deficiency in photorespiration due to the block in lipoic acid biosynthesis that requires mtFAS to generate the octanoic acid precursor. As such, these photorespiration effects can be reversed when the plants are grown in a non-photorespiratory condition (i.e., an atmosphere of elevated CO$_2$ concentrations). In parallel, knockdown of the \textit{mtHD} gene expression in a mutant lacking the mtKAS activity led to the more severe morphological and metabolic phenotypes as compared to the single mutation in the \textit{mtKAS} gene. RNA-seq analysis of the transcriptome of \textit{mtHD} mutant establishes the global changes in carbon and nitrogen metabolism. These results demonstrate that the mtFAS system is an essential metabolic process needed for plant growth and development.
INTRODUCTION

In eukaryotic cells, de novo fatty acid biosynthesis occurs in multiple subcellular compartments: the cytosol, mitochondria, and plant plastids. These processes play distinct biological roles and do not reflect just metabolic redundancy (1-3).

Mitochondrial FAS (mtFAS) is considered to be a multi-component Type II system that recruits discrete, monofunctional enzymes to produce fatty acids (2). The mtFAS system resembles the canonical Type II FAS that occurs in bacteria and plant plastids (3,4). These systems contrast with the Type I FAS that occur in the cytosol of mammals and fungi; these are multifunctional proteins that contain all of the reaction centers that are needed to produce fatty acids (1). In yeast and humans, all of the mtFAS components have been characterized (2). In yeast, knockout of any gene encoding the enzyme of the mtFAS system leads to loss of cytochromes, inability to grow on the non-fermentable carbon source, and deficiency in mitochondrial RNA processing (2).

Although the physiological function of mtFAS is still somewhat elusive, the contribution of fatty acid precursor for lipoic acid biosynthesis has been elucidated by suppressing the expression of mitochondrial acyl carrier protein (mtACP) in the human cell culture (5).

In plants, however, most mtFAS components have yet to be biochemically identified, with the exception being mtACP and β-ketoacyl-ACP synthase (mtKAS) (6-9) (Chapter 2). Previous study on the mtKAS gene reveals that the Arabidopsis mtFAS system contributes to the biosynthesis of lipoic acid, the cofactor of pyruvate dehydrogenase (PDH), β-ketoglutarate dehydrogenase (KGDH), and glycine decarboxylase complex (GDC) (9) (Chapter 2). Knockout of the mtKAS gene leads to the diminished but still detectable lipoylation of the H-subunit of GDC, an essential enzyme
complex in the photorespiration process, resulting in a typical deficiency in photorespiration (9). However, the lipoylation of PHD and KGDH is only slightly affected by the mutations in the \( mtKAS \) gene (9); this may indicate another lipoic acid generation system in plant plastids, other than mitochondria (10,11). Furthermore, mutations in the \( mtKAS \) gene lead to defective metabolism of a series of lipids, namely surface lipids, chloroplastidic galactolipids and sulfolipids, and triacylglycerol, as a secondary effect of the deficiency in photorespiration (Chapter 2). In addition to contributing to lipoic acid biosynthesis, the Arabidopsis mtFAS system also supports the biosynthesis of lipid A-like molecules, whose physiological functions are currently unknown (Chapter 2). In addition, fatty acids generated by the Arabidopsis mtFAS system have been proposed to be required for the remodeling of cardiolipins in plants (12,13).

Here, we identify and characterize the Arabidopsis \( mtHD \) enzyme, which is responsible for the 3-hydroxyacyl-ACP dehydratase reaction. Knockdown mutants of the \( mtHD \) gene exhibit morphological and metabolic phenotypes that are similar to the mutations in the \( mtKAS \) gene, while double mutants of the \( mtHD \) and \( mtKAS \) genes display a more severe alteration relative to the \( mtKAS \) mutant. The global transcriptional alterations associated with the defective carbon and nitrogen metabolism are illustrated by RNA-seq analysis. These findings conclude that the Arabidopsis mtFAS system is essential for plant metabolic homeostasis.
RESULTS

1. Identification of the putative Arabidopsis mtHD gene

Using the human HsHTD sequence as the query, BLAST searches of protein database reveal a single mtHD candidate (AT5G60335) in the Arabidopsis genome (total score 76, query cover 76%, and E value 8e-17). When yeast HTD2 is used as the query, BLAST searches do not suggest any significant homolog in the Arabidopsis genome. Human HsHTD and its Arabidopsis ortholog share 40.1% similarity and 22.0% identity, however, they do not exhibit sequence homology at N-termini, which are rich in basic amino acids and lack acidic residues (Fig 1). These features are the characteristic of mitochondrial targeting sequence elements, and indeed, the human HsHTD N-terminal sequence functions as a mitochondria transit peptide (14). Arabidopsis AT5G60335 gene product is predicted to be a mitochondrially localized protein by MitoProt II (24 N-terminal residues with a score of 0.9957) (15), PSORT (with a score of 0.751) (16), and Target P (17 N-terminal residues with a score of 0.680) (17,18).

2. Subcellular localization of the protein encoded by AT5G60335

To determine the subcellular localization of Arabidopsis AT5G60335 gene product, a C-terminal GFP fusion was constructed by cloning the 5’ terminus of the AT5G60335 gene encoding 40 amino acid residues in-frame with the GFP-coding sequence, and stably expressed in Arabidopsis under the control of 35S promoter. The transgenic plants were examined by confocal microscopy. In mesophyll cells, GFP signals localize to small but distinct areas, and they do not overlap with chlorophyll auto-fluorescence (Fig 2A), revealing that this fusion protein does not reside in plastids. Root cells of these transgenic plants were further investigated to avoid interference from
chlorophyll auto-fluorescence. MitoTracker Orange was applied in parallel as the mitochondrial marker, and its signal was recorded simultaneously with the GFP signal. The wild-type control does not exhibit any GFP fluorescence (Fig 2B). The P35S::GFP control localizes to cytosol and nuclear, which are distinct from mitochondria (Fig 2C). GFP fluorescence of the AT5G60335-GFP fusion protein overlaps with MitoTracker Orange fluorescence (Fig 2D), illustrating that the N-terminal sequence of the protein encoded by AT5G60335 directs GFP to Arabidopsis mitochondria. We therefore conclude that the Arabidopsis AT5G60335 gene encodes a protein that is mitochondrially localized.

3. **Biochemical identification of the protein encoded by AT5G60335**

To determine whether the Arabidopsis AT5G60335 gene encodes a functional equivalent of the yeast Htd2p protein, the yeast yhr067w mutant strain that carries a mutation in the HTD2 gene was transformed with the YEp351 empty vector, HTD2-YEp351, AT5G60335-YEp351, and AT5G60335-YEp351M (Fig 3). The negative control, YEp351 empty vector, does not rescue the growth deficiency of the mutant strain, whereas the positive control, HTD2-YEp351, reversed the normal growth of the mutant on the glycerol containing media. In parallel, the mutant strains can also grow on the media containing glycerol as the sole carbon source when expressing AT5G60335-YEp351 (full length protein) or AT5G60335-YEp351M (N-terminally truncated protein with a yeast COQ3 mitochondrial targeting peptide), demonstrating that the Arabidopsis AT5G60335 gene encodes a mitochondrial 3-hydroxyacyl-ACP dehydratase (mtHD) component of the mtFAS system.
4. Kinetic properties of the recombinant mtHD protein

Based on the sequence alignment information, a truncated mtHD protein was expressed in *E. coli* without the N-terminus 24 amino acid residues that act as a mitochondrial targeting sequence. The His-tagged protein was purified by the nickel affinity column chromatography. *In vitro* enzymatic assays of the recombinant mtHD enzymes in humans and its *Trypanosoma brucei* ortholog suggested that kinetic values in the (3R)-specific hydratase activity of these enzymes can be measured to determine the acyl chain specificity of the mtHD enzyme (14,19). Therefore, the $V_{\text{max}}$ and $k_{\text{cat}}$ values in the (3R)-specific hydratase activity of the recombinant Arabidopsis mtHD protein were measured using trans-2-decenoyl-CoA and trans-2-hexadecenoyl-CoA substrates (Fig 4A). The $K_{\text{m}}$, $V_{\text{max}}$, and catalytic efficiency ($k_{\text{cat}}/K_{\text{m}}$) on the two enoyl-CoA substrates are comparable (Fig 4B), and are also comparable to those of its orthologs from humans and *T. brucei* (14,19). Collectively, the Arabidopsis mtHD enzyme exhibits broad acyl chain specificity for medium-chain (decenoyl-CoA) and long-chain (hexadecenoyl-CoA) acyl-CoAs. This is in agreement with the acyl chain specificities of the Arabidopsis mtKAS enzyme, which is responsible for the first reactions of mitochondrial fatty acid biosynthesis (Chapter 2), establishing that the Arabidopsis mtFAS system may support the synthesis of both medium- and long-chain fatty acids.

5. Expression pattern of the *mtHD* gene

Quantitative RT-PCR was performed with RNA templates to determine the spatial and temporal expression of the *mtHD* gene among aerial organs of different age plants, roots, flowers, siliques, and rosette leaves of wild-type plants (Fig 5). These data shows that the *mtHD* gene is expressed in all of the organs tested. The highest expression was
found in flowers. The near ubiquitous expression pattern of the \textit{mtHD} gene is consistent with the public microarray data visualized by Arabidopsis eFP-Browser (20,21). The near ubiquitous expression pattern of the \textit{mtHD} gene is also in keeping with that of the \textit{mtKAS} gene that involves in the same pathway (Chapter 2).

6. Compromised growth and development of the \textit{mthd-rnai} mutant alleles

To understand the physiological significance of the \textit{mtHD} gene, we identified a T-DNA-tagged mutant allele stock (CS856112) by searching SALK Insertion Sequence Database (http://signal.salk.edu). Plant genomic DNA sequences flanking the T-DNA border were amplified and sequenced, and the molecular structure of the mutant allele was deduced by aligning these flanking DNA sequences with the Arabidopsis genome sequence. This allele carries a chimeric T-DNA sequence that deletes the stop codon of the \textit{mtHD} gene and introduces a 3’-terminus encoding 7 extra amino acid residues. Quantitative RT-PCR shows that the expression level of the mutated \textit{mtHD} gene in this \textit{mthd} allele is comparable to that of the \textit{mtHD} gene in the wild-type plants (data not shown). In addition, the kinetic properties of the recombinant mutant \textit{mtHD} enzyme on trans-2-decenoyl-CoA and trans-2-hexadecenoyl-CoA substrates are comparable to the wild-type enzyme (data not shown). Our results suggest that the T-DNA-tagged allele in the CS856112 stock does not impair the \textit{mtHD} gene functions. Indeed, the mutant allele exhibits a wild-type-like appearance (data not shown).

As an alternative strategy, RNAi transgenic Arabidopsis plants expressing a double-strand \textit{mtHD} RNA fragment under the control of 35S promoter were generated and characterized (Fig. 6A). As indicated by PCR amplification of the P35S::\textit{mtHD} RNAi transgene, we recovered 36 independent lines that carry the transgene (data not
shown). About one fifth of these transgenic lines exhibit significantly reduced size of aerial organs in the 16-day-old plants relative to the wild-type controls. Two transgenic lines with the most severe phenotypic differences are designated as \textit{mthd-rnai-1} and \textit{mthd-rnai-2} (Fig. 6B). These two alleles express 18±4% and 22±5% of the \textit{mtHD} transcript levels, respectively, as compared to the wild-type plants. The morphological differences associated with the expression of RNAi transgene could be reversed when the mutants are grown in an atmosphere enriched in CO$_2$ (Fig. 6C), illustrating that the compromised growth is a result of the deficiency in photorespiration.

\section*{7. Metabolic alterations in soluble metabolites and acyl-lipids}

Metabolic profiling was performed on the aerial organs of the \textit{mthd-rnai-1} and \textit{mthd-rnai-2} mutant alleles. The most dramatic alteration is an ~70-fold increase in Gly levels in the mutants (Fig 7A); it is in agreement with the previous finding that mutations in the mtFAS system diminish the lipoylation of the H-subunit of GDC, leading to the inability to metabolize photorespiratory Gly (9). Several other amino acids exhibit significant changes (Fig 7A), which illustrate the defective nitrogen metabolism in the mutants. Glycolate is the upstream metabolite of Gly in the photorespiration process. Its level increases to approximately 2.5-fold in the \textit{mthd-rnai} mutant alleles relative to the wild-type plants (Fig 7C), which is likely to be a result of the Gly accumulation. Sucrose is the transport form of sugar in Arabidopsis, and is depleted in the photorespiratory mutants due to the inefficient carbon recovery (22,23). In the \textit{mthd-rnai} mutant alleles, the sucrose levels drop to about 10% of the normal level (Fig 7D), which is consistent with the sucrose depletion in other mutants deficient in photorespiration, including \textit{10-fdf}, \textit{shmt1-2}, and \textit{mtkas-2} and \textit{mtkas-3} (23) (Chapter 2).
Fatty acid (FA) profiling indicates that the *mtHD* mutants have significantly reduced levels of FA 16:2 and FA 16:3 and elevated level of FA 18:1 (Fig 7E). In addition, in the mutants the level of 3-hydroxymyristic acid drops to about 35% of the wild-type level (Fig 7G). Furthermore, impaired lipid content is found in the *mthd-rnai* mutant alleles (Fig 7H). Chloroplastic galactolipids (MGDG and DGDG) are reduced to 70% and 80% relative to the normal level, respectively, and levels of the sulfolipids, SQDG, are also decreased, to about 60% of the wild-type levels. The reduction in galactolipids and sulfolipids suggest the defective thylakoid membrane systems, which are consistent with the reduced levels of the fatty acid 16:3, which is exclusively found in thylakoid membranes of Arabidopsis (24). Hyper-accumulation of TAG is observed to be 4-fold higher in the mutants relative to the wild-type control (Fig 7H). Finally, the contents of surface lipids (i.e., very-long-chain fatty acid (VLCFA), alcohol, and alkane) are reduced in the *mthd-rnai* mutant alleles (Fig 7J).

The metabolic alterations in all of these metabolites, with the only exception of Gly and 3-hydroxymyristic acid, are a consequence of the deficiency in photorespiration, because they are reversed when the plants are grown in an elevated CO2 atmosphere, which inhibits photorespiration (Fig 7B, C, D, F, I, and K). 3-Hydroxymyristic acid levels are markedly reduced in the *mthd-rnai* mutants, regardless of whether the plants are grown in ambient air or an elevated CO2 atmosphere. Therefore, the depletion of 3-hydroxymyristate is a direct result of the deficiency in the mtFAS system, rather than a secondary effect of the deficiency in photorespiration. Thus, mtFAS contributes 3-hydroxymyristyl-ACP substrate for the synthesis of Lipid A-like molecules, a unique sink of mitochondrially derived fatty acids in Arabidopsis.
8. Metabolic alterations in chlorophylls and their degradation intermediates

Chlorophyll-like molecules (i.e., chlorophyll a, chlorophyll b, and three breakdown intermediates (i.e., pheophytin a, pheophytin b, and pheophorbide a)) were quantified by lipidomics analysis. In the mthd-rnai mutant alleles the total amount of these molecules is decreased to about 75% of wild-type levels (Fig 8A). Chlorophyll a maintains normal levels in the mthd-rnai mutant alleles, whereas chlorophyll b levels decreases to about 80% of the wild-type levels (Fig 8A). The de-Mg forms, pheophytin a and pheophytin b, were depleted, however, the dephytelated form, pheophorbide a, was slightly accumulated as compared to the wild-type plants (Fig 8A). Similar alterations in chlorophyll-like molecules are also found in the mtKAS mutants (data not shown), confirming that the reduced levels of chlorophyll-like molecules are a result of the deficiency in the mtFAS system. These observations are consistent with an increased rate of chlorophyll degradation, which would lead to increased accumulation of chlorophyll degradation intermediates (i.e., pheophorbide a) and pale-green leaf color of the mtHD and mtKAS mutants. Similar to the disturbance in lipid metabolism, the alterations in chlorophylls and their degradation intermediates are a consequence of the deficiency in photorespiration, because they are reversed when photorespiration is suppressed by growing plants in an elevated CO₂ atmosphere (Fig 8B).

9. Morphological and metabolic phenotypes of the mtkas-2/mthd-rnai double mutants

The RNAi mutants of the mtHD gene exhibit the similar morphological and metabolic phenotypes as T-DNA-tagged mutants of the mtKAS gene, concluding that the mtHD gene and mtKAS gene encoding enzymes that function in the same pathway.
Further evidence to this effect was obtained by the analysis of the \textit{mtkas-2/mthd-rnaI} double mutant lines. These lines were generated by the direct transformation of the \textit{mtkas-2} mutant line with the \textsc{p35S::mtHD RNAi} transgene. This transformation and selection process was conducted in an atmosphere of \% CO\textsubscript{2}, which renders the \textit{mtkas-2} allele plants a normal appearance (Chapter 2). Two independent T2-generation \textit{mtHD} RNAI lines in the \textit{mtkas-2} background were selected for phenotypic evaluation. qRT-PCR analysis of these two lines indicated that they express the \textit{mtHD} transcript at 27±7% and 30±6% of the wild-type levels. These two double mutant lines do not exhibit any morphological difference as compared to \textit{mtkas-2} (or wild-type) when the plants are grown in an elevated CO\textsubscript{2} atmosphere (Fig 9A). However, when grown in an ambient atmosphere, these double mutant plants display a more severe phenotype than the \textit{mtkas-2} mutant line. Specifically, when grown in ambient air, these double mutants appear to be non-viable at constant illumination; their growth is arrested at the cotyledon stage (Fig 9A), and they do not flower even after growth for 4-weeks, by which stage both the \textit{mtkas2} and \textit{mtHD} RNAI lines have already flowered (data not shown).

Profiling of amino acid pools and 3-hydroxymyristic acid was used to examine if knocking down of the \textit{mtHD} gene expression in the \textit{mtkas-2} mutant background further impairs the mtFAS system. The two double mutant lines have significantly elevated Gly levels as compared to the \textit{mtkas-2} allele, regardless of whether they are grown in ambient air or an elevated CO\textsubscript{2} atmosphere (Fig 9C and D). In addition, the two double mutants show significant reduced levels of 3-hydroxymyristic acid, to about 40\% of that in \textit{mtkas-2}, when grown in an atmosphere of elevated CO\textsubscript{2} concentration (Fig 9E). These data demonstrate that suppression of the \textit{mtHD} expression in an \textit{mtkas-2} mutant background
further impairs the mtFAS function, leading to the decreased ability to metabolize photorespiratory Gly and to synthesize 3-hydroxymyristyl-ACP precursor needed for the assembly of lipid A-like molecules. These metabolic alterations also illustrate that mutations in the mtHD gene and mtKAS gene may result in an inability in the morphological development in ambient air. However, we can not exclude the possibility that mtHD could be involved in other metabolic processes other than mtFAS and photorespiration, but these affections may not be dominated, because most phenotypic changes of the mthd-rnai mutant allele and mtkas-2/mthd-rnai double mutant alleles are reversed by CO₂ treatment (Fig 9B). Taken together, knockdown of the mtHD gene expression in the mtkas-2 allele introduces a more severe impairment on the mtFAS system, which further results in an inability to survive in ambient air, demonstrating that the Arabidopsis mtFAS system is essential for growth and development.

10. Identification of differentially expressed genes by RNA-seq

RNA-seq experiments were performed to access the effect on the global transcriptome of the mutations in the mtHD gene. Triplicate biological samples were isolated and RNA samples were sequenced from two sibling genotypes: wild-type plants and those carrying the mthd-rnai-1 mutant allele, each of which were grown in the atmosphere of ambient CO₂ and 1% CO₂. Transcript abundance levels were normalized for each gene model in each sample and determined as the number of fragments sequenced per kilobase of exon per million fragments mapped (FPKM).

A generalized linear model analysis was performed based on a negative binomial distribution to test for differentially expressed (DE) genes between the two genotypes that were grown in the same atmosphere. While controlling the false discovery rate (FDR) at
1% using Benjamini and Hochberg’s method, we identified 3352 DE genes in the ambient CO2 atmosphere and 66 DE genes in an atmosphere enriched in 1% CO2.

11. Transcriptional regulation on the genes of carbon and nitrogen metabolism

Functional analysis on the DE genes between the wild-type plants and *mthd-rnai-1* grown in ambient air were carried out using MetNet Online (25), which is based on the biochemical pathways summarized in AraCyc (26). Totally 452 metabolic pathways are overrepresented by these DE genes.

Alterations in metabolites (e.g., amino acids, sucrose, and acyl-lipids) suggest that the suppressed expression of the *mtHD* gene disturbs the metabolism in carbon (C) and nitrogen (N). To evaluate the transcriptional impact on C and N metabolism, we analyzed those DE genes associated with metabolic processes of C and N metabolism, namely photorespiration, photosynthesis, TCA cycle, sucrose metabolism, nitrogen assimilation, amino acid metabolism, and chlorophyll metabolism (Fig 10). In the “photorespiration” process, 7 genes are DE, including the genes encoding the H subunit and P subunit of GDC (27), which is likely to be a consequence of the dysfunction of GDC due to the disruption in lipoic acid biosynthesis. By contrast, other genes encoding the essential components of photorespiration exhibit normal transcript levels, suggesting that the deficiency in photorespiration is not regulated at the transcriptional levels. Thirty-one genes in the “photosynthesis” process are DE in the *mthd-rnai-1* mutant. Twenty-two of these “photosynthesis”-related DE genes encode subunits of the photosynthetic reaction center, and all of them are down-regulated, demonstrating that there is a defective photosynthesis process as a consequence of the deficiency in photorespiration. In the “TCA cycle”, 18 genes are DE, most of which are up-regulated, for example the E1
subunit of β-ketoglutarate dehydrogenase and flavoprotein subunit 1 of succinate dehydrogenase. Considering that the lipoylation of the E2 subunit of pyruvate dehydrogenase and E2 subunit of β-ketoglutarate dehydrogenase are only slightly reduced as a result of the mutations in mtFAS (9), the increased abundance of the transcripts of genes in “TCA cycle” reveals that respiration is likely to be induced in the \textit{mthd-rnai-1} mutant. In “sucrose metabolism”, 29 genes are DE in the mutant, and most of these genes are up-regulated (e.g., genes encoding phosphoglucoisomerase, phosphoglucomutase, and sucrose phosphate synthase) illustrating a transcriptionally up-regulated sucrose biosynthetic pathway. This implies that the sucrose depletion in the mutant is unlikely to be a result of the transcriptional regulation on the sucrose biosynthesis process, but is likely to be metabolically regulated by the precursors, due to the insufficient photorespiratory C recovery and reduced photosynthetic CO$_2$ fixation.

In N metabolism, namely “ammonia assimilation” and “Asp/n, Glu/n” metabolism, 26 genes are DE in the mutant. Specifically, the up-regulation on the genes encoding nitrilase isoforms (i.e., the enzymes that transform nitrile into ammonia) and glutamine synthetase isoforms (i.e., the enzymes that assimilates ammonia in the form of Gln) demonstrate an up-regulated N assimilation, which is likely to be in response to the inadequate photorespiratory N recovery in the mutant. In addition, the DE genes encoding asparagine synthetase and glutamine synthetase establish a transcriptional explanation on the altered levels of Asp, Asn, Glu, and Gln in the mutants.

Finally, 13 genes of “chlorophyll metabolism” are DE. These data show that chlorophyll biosynthesis is down-regulated, because genes encoding the key enzymes of this pathway (e.g., glutamyl-tRNA reductase, magnesium-protoporphyrin IX
monomethyl ester cyclase, magnesium-protoporphyrin IX methyltransferase, magnesium chelatase, and protochlorophyllide reductase A) exhibit reduced expression levels. These data also illustrate an up-regulation of the chlorophyll degradation pathway, due to the increased abundance of the transcripts encoding enzymes of this process (e.g., pheophorbide A oxygenase, pheophytinase, and chlorophyllase). Collectively, the RNA-seq data of chlorophyll metabolism are in agreement with the reduced levels of the chlorophylls and their degradation intermediates.

All of these DE genes, with the exception of AT5G54190, are a consequence of the deficiency in photorespiration, because they are DE in ambient air and exhibit wild-type expression levels in an elevated CO₂ atmosphere. AT5G54190, which encodes the protochlorophyllide reductase A of chlorophyll degradation pathway, is down-regulated in both ambient air and an atmosphere enriched in CO₂, and it displays a significantly higher expression level in ambient air relative to that in higher CO₂ atmosphere, suggesting that the expression of this gene is regulated by both the deficiency in photorespiration and the mtFAS system. Taken together, the RNA-seq results reveal the defective metabolism of C and N in the mthd-rnai-1 mutant, demonstrating the integral functionalities of photorespiration as well as mtFAS in the central C and N metabolism.

**DISCUSSION**

**Characterization of Arabidopsis mitochondrial 3-hydroxyacyl-ACP dehydratase**

The third reaction in the iterative process of fatty acyl chain elongation cycle is the dehydration of 3-hydroxyacyl-ACP to form trans-2-enoyl-ACP. This reaction is catalyzed by 3-hydroxyacyl-ACP dehydratase. Plant mitochondria are known to have the 3-hydroxyacyl-ACP dehydratase activity (28), however, the enzyme that catalyzes this
reaction has yet to be identified. This study establishes that the Arabidopsis mtHD enzyme is encoded by the gene locus AT5G60335, which is responsible for catalyzing this reaction in mitochondria. This fact was experimentally established by the heterologous expression of the enzyme encoded by AT5G60335 in the mitochondria of the yeast mutant strain yhr067w (carries knockout allele of HTD2), which restores its deficiency in the HTD2 function, leading to the complementation of the growth on media containing glycerol as the sole carbon source. Furthermore, the E. coli produced recombinant Arabidopsis mtHD enzyme is enzymatically active, and can catalyze the hydratase reaction, which converts trans-2-decenoyl-CoA and trans-2-hexadecenoyl-CoA to the relevant 3-hydroxyacy-CoA molecules. This is a reaction that is also catalyzed by its orthologs isolated from humans and T. brucei (14,19). Finally, the N-terminal extension sequence of the Arabidopsis protein is characterized to be a mitochondrial transit peptide that directs the mtHD-GFP fusion protein to Arabidopsis mitochondria.

It is noteworthy that the Arabidopsis mtHD amino acid sequence is homologous to the mtHD enzymes from humans and T. brucei, and it also shares similarity with Mycobacterium tuberculosis 3-hydroxyacyl-ACP dehydratase (29). By comparison, the Arabidopsis mtHD enzyme does not exhibit any homology as compared to the yeast mtHD enzyme (Htd2p), as well as other known Type II HD enzymes of fatty acid biosynthesis, including bacterial fabA/fabZ and plant plastidial HD isoforms (3,4), which belong to the thioesterase/thioester dehydratase/isomerase (TED1) superfamily. This variance in sequence homology of these HD enzymes demonstrates a divergent evolution of the Arabidopsis, human and T. brucei mtHD family of enzymes.
Linkage between mtHD and photorespiration

The role of the Arabidopsis mtFAS system in providing octanoyl-ACP, the precursor of lipoic acid biosynthesis, has been previously established (30). Accordingly, mutations in an Arabidopsis mtFAS component (mitochondrial β-ketoacyl-ACP synthase; mtKAS; AT2G04540) display a typical photorespiration-deficient appearance and hyperaccumulation of glycine, due to the reduction in the lipoylation of the H-subunit of glycine decarboxylase complex (GDC), an essential enzyme complex in photorespiration (9) (Chapter 2). Acyl lipid metabolism is also altered in the mtKAS mutants as a secondary effect of the deficiency in photorespiration (Chapter 2).

Furthermore, a novel metabolic destination of the fatty acids generated by the mtFAS system has been discovered; that is mitochondrial 3-hydroxymyristyl-ACP, which is used for the biosynthesis of lipid A-like molecules (Chapter 2). In this study, we show that the mthd-rnai mutant alleles exhibit similar deficiency in photorespiration, acyl-lipid metabolism, and biosynthesis of lipid A-like molecules as compared to the mtkas mutant alleles. We therefore conclude that mtHD is indeed a component of the Type II mtFAS system.

The mtFAS system is the only source of octanoyl-ACP used for lipoic acid biosynthesis in yeast (2). By contrast, in Arabidopsis lipoic acid biosynthesis occurs in two distinct subcellular localizations, mitochondria and plastids, and both processes utilize octanoyl-ACP substrates derived from mtFAS and ptFAS systems, respectively(10,11,31,32). The lipoylated E2 subunits of pyruvate dehydrogenase and β-ketoglutarate dehydrogenase occur in both plastids and mitochondria, but these proteins’ lipoylation status is only partially reduced by mutations in the mtFAS component genes.
(e.g., \textit{mtKAS} mutants) \cite{9}. In contrast however, the lipoylation of the mitochondrially-located H-subunit of GDC is almost completely eliminated, with only a small amount of lipoylation being maintained independent of the \textit{mtKAS}-dependent mtFAS system \cite{9}. The \textit{mtKAS}-independent lipoylation of the H protein is considered to utilize the lipoic acid (or octanoic acid) that that is sourced from the chloroplast FAS and lipoic acid biosynthesis pathways \cite{9}. However, its not clear how the chloroplast-generated lipoic acid (or octanoic acid) is transported into mitochondria. For this reason, despite the apparent loss of mtFAS activity in the \textit{mtKAS} mutants, these mutants may still provide sufficient octanoic acid to maintain the small amount of H-protein lipoylation to enable the development of stunted but viable seedlings in ambient air atmosphere. To test this hypothesis, we generated and characterized plants that carried mutations in two mtFAS genes, the double mutant of the \textit{mthD} gene and the \textit{mtKAS} gene. As compared to the \textit{mtkas-2} single mutant, the growth, development, and reproduction of the plants carrying both \textit{mtkas-2} and \textit{mthd-rnai} alleles are more severely affected at many stages of the plant’s life cycle. Under all of the growth conditions tested, defective growth are observed on the double mutant. Particularly, in a condition where photorespiration is highly active (i.e., constant illumination and ambient atmosphere), the \textit{mtkas-2/mthd-rnai} double mutants are not viable; their growth is arrested at the cotyledon stage, and they do not flower. In addition, glycine is hyper-accumulated in the double mutant relative to the \textit{mtkas-2} mutant. These results suggest that in the double mutant there is a more dramatic impairment of protein lipoylation than in the \textit{mtkas-2} mutant, implying that mtFAS is more dramatically blocked in the \textit{mtkas-2/mthd-rnai} double mutant. We therefore conclude that lipoylation of the H protein is exclusively dependent on the mtFAS system,
and that the import of lipoic acid or octanoic acid derived from ptFAS is a very minor or non-contributor to this lipoylation (9).

This conclusion is further supported from the quantification of 3-hydroxymyristic acid levels, the hallmark metabolite of lipid A-like molecules. The 3-hydroxymyristate level is significantly reduced in the mtkas-2 single mutant, but is even more dramatically reduced in the plants that carry both the mtkas-2 mutant allele and the mthd-rnai allele. The simplest model to explain these results is that the mtkas-2 allele is not a null allele, despite the fact that a relatively sensitive qRT-PCR protocol could not detect the mtKAS mRNA in these mutant plants. The residual mtKAS activity in the mtkas mutants can therefore support sufficient mtFAS activity to allow a small amount of the H protein lipoylation. But mtFAS activity is further reduced in the mtkas-2/mthd-rnai double mutants to enhance the metabolic and morphological phenotype of these plants. Although this is the simplest explanation for these observations, there are also more complex models possible, such as a yet unknown, but inefficient condensation enzyme in Arabidopsis mitochondria that could support the mtFAS activity in the mtkas mutants. This could be the dual localization of the plastidial β-ketoacyl-ACP synthases to both plastids and mitochondria, as occurs with DNA topoisomerase (AT4G31210) (33) or alternatively a mitochondrially targeted KAS that is weakly expressed.

**Defective carbon and nitrogen metabolism**

The depletion of sucrose and alterations in amino acid levels in plants defective in mtFAS (i.e., the mtkas and mthd-rnai mutants) illustrate the importance of photorespiration and carbon (C) and nitrogen (N) metabolism. The connectivity between mtFAS and C and N metabolism is via the lipoylation state of the H protein of GDC,
which is dependent of mtFAS. C and N metabolism is considered to be a complex metabolic network that is coordinated by an intricate regulatory machinery, which integrates processes that occur in different subcellular organelles, including chloroplasts, mitochondria and peroxisomes (34). Photorespiration is a key process in coordinating C and N metabolism, and is also similarly distributed among several subcellular compartments.

Photorespiration is a consequence of the inability of ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO) to distinguish between CO₂ and O₂ (35). 2-Phosphoglycolate, the product of the oxygenase reaction of RuBisCO, is recycled to 3-phosphoglycerate via the process of photorespiration that requires at least 16 enzymes distributed across four compartments (e.g., chloroplasts, mitochondria, peroxisomes, and the cytosol). Flux through the photorespiration cycle is maximal during high-light intensity and low atmospheric CO₂ concentrations; in these conditions it is estimated that in C₃ plants about every four catalytic cycles of the RuBisCO reaction leads to the formation of a 2-phosphoglycolate molecule (36). In the stoichiometry of photorespiration one of the carbon atoms of phosphoglycolate is lost as CO₂ and the process also leads to the deamination of glycine resulting in the formation of one molecules of ammonia. Thus, blocking mtFAS, which leads to reduced GDC activity due to the under-lipooylation of H protein affects C metabolism because recovery of the photorespiratory C in the form of phosphoglycolate metabolism cannot be correctly processed. The defective C metabolism appears to contribute to inefficient photosynthesis, indicated by the transcriptional down-regulation of the 22 genes encoding
subunits of the photosynthetic reaction center in the *mthd-rnai-1* mutant, which leads to the reduced levels of sucrose in this mutant.

In addition, the under-lipooylation of H protein also affects N metabolism, altering amino acid pool sizes. The most dramatic change is in glycine levels, which is the substrate for GDC. Smaller but significant alterations occur in other amino acids, which reflect the block in the release of ammonia by GDC (35). Normally this released ammonia is re-assimilated by glutamine synthetase (GS) and glutamate synthase (GOGAT) into glutamine and glutamate, the latter of which is used in the conversion of glyoxylate to glycine that is required in the recovery of photorespiratory carbon (37). Therefore, mutations that block mtFAS function manifest alterations in C and N metabolism via the inability to support photorespiration due to the under-lipooylation of the H protein of GDC (34).

Consistent with this metabolic model, characterization of the transcriptome of the *mthd-rnai-1* mutant indicates that genes encoding GS isoforms are up-regulated, as are the genes encoding nitrilase isoforms, which generate ammonia from nitrile as a photorespiration-independent N source. In addition, the expression of the genes encoding enzymes that interconvert aspartate and asparagine are DE in the mutant, revealing that the transcriptional regulation of the amino acid metabolism is beyond the GS-GOGAT pathway, which explains the alteration in amino acid homeostasis.

Finally, in addition to amino acids, N is a constituent of chlorophylls, which are synthesized from glutamate (38). The yellowish leaf color of the *mtFAS* mutants (i.e., *mthd-rnai* and *mtkas* mutants) indicates that both chlorophyll biosynthesis and/or catabolism may be affected. This is unlikely a consequence of metabolic regulation of
chlorophyll metabolism, because glutamate, the precursor of chlorophyll biosynthesis, exhibits elevated levels in the mthd-rnai and mtkas mutants. Indeed, they are likely to be regulated at the level of gene expression as indicated by the RNA-seq experiment.

METHODS

Plant materials and growth conditions

Arabidopsis thaliana ecotype Columbia (Col-0) and a T-DNA insertion mutant (mthd, CS856112) were obtained from the Arabidopsis Biological Resource Center (Columbus, OH). Seeds were sterilized and sown on Murashige and Skoog agar medium as describe in Chapter 2. After seeds were sown, plates were placed at 4°C for 4 days to break seed dormancy, and plants were grown in ambient air or the non-photorespiratory condition (1% CO2 in a growth chamber). The temperature was maintained at 23 ± 2°C with continuous illumination (100 µmol m⁻² s⁻¹). If needed, seedlings were transferred from plates to soil (LC1 Professional Growing Mix) at 7 to 10-day after germination.

RNA extraction and quantitative RT-PCR

RNA was extracted from 50 mg fresh tissues using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA). A 2µg aliquot of RNA was used for reverse transcription using the SuperScript III First-Strand Synthesis Super Mix (Life Technologies, Carlsbad, CA) with oligo (dT)₂₀ primer. Quantitative RT-PCR was performed on StepOnePlus Real-Time PCR System using SYBR Select Master Mix (Life Technologies, Carlsbad, CA). Actin-2 (AT3G18780) was used as the reference RNA for relative quantification. The gene specific primers are H1 (CTGAGGTGAGCCATGATTGGAA) - H2 (GTATACTGCTCCAGGGAAATGG) for mtHD and H3
Recombinant DNA construction

Arabidopsis cDNA encoding the *mtHD* gene was cloned from reverse transcribed RNA preparations from 16-day-old aerial organs of the wild-type Arabidopsis plants. The 5’-terminal coding sequence, which encodes 40 amino acid residues, was amplified and subcloned into pEarleyGate103 (39) for subcellular localization study. GFP-coding sequence was cloned from pEarleyGate103 and subcloned into pEarleyGate100 (39) to confirm the localization of the *P35S::GFP* product in Arabidopsis. A *mtHD*-specific sequence was amplified using H5 (CACCATGCTGAGGTGAGCCATGATTG) - H6 (TCATCACCTATGTAAACCGGTG) and subcloned into pB7GWIWG2(II) (40) for the RNAi experiment.

For yeast molecular genetic assays, the yeast *HTD2* was amplified from yeast genome DNA extracts, while the full-length Arabidopsis *mtHD* was amplified from the Arabidopsis mRNA preparation. These two genes were cloned into the yeast expression vector YEp351 (41) under the control of PGK promoter; the two constructs were named *HTD2*-YEp351 and *mtHD*-Yep351, respectively. In case the N-terminal mitochondrial transit peptide of Arabidopsis *mtHD* is not recognized by yeast as a targeting peptide to mitochondria, a yeast mitochondrial expression vector, YEp351M, was constructed by amplifying the 5’-terminal 105bp coding sequence of yeast *COQ3* (a yeast mitochondrial transit peptide-encoding sequence (42)) and cloned into YEp351. Truncated Arabidopsis *mtHD* was amplified, removing the 5’-terminal sequence that encodes the putative 31
amino acid transit peptide. The PCR product was cloned into YEp351M, resulting in a construct named \textit{mtHD-YEp351M}.

For protein expression, the ORF, which encodes Arabidopsis \textit{mtHD}, was chemically synthesized by GenScript (Piscataway, NJ) after codon optimization for expression in \textit{E. coli}. The \textit{mtHD} ORF was amplified, removing the organelle targeting peptide sequence. The PCR product was cloned into pET30b vector (EMD Millipore, Billerica, MA), and the resulting construct was named \textit{mtHD-pET}. The \textit{mtHD-pET} construct expresses a recombinant mtHD protein containing a His-tag located at the C-terminus.

**Plant transformation and selection**

Arabidopsis transformation was performed using the \textit{Agrobacterium tumefaciens}-mediated infiltration method (43) to generate transgenic plants that express GFP fusion transgenes and the \textit{mtHD} RNAi transgene. T1 seeds were collected and subsequently selected for Bar-resistance by spraying soil-grown seedlings with 1:1500 Finale Weed&Grass Killer. For transformation of the \textit{mtHD} RNAi transgene into the \textit{mtkas-2} mutant allele, plants were grown in the elevated CO\textsubscript{2} atmosphere for both transformation and selection.

**Confocal microscopy**

T2 generation of transgenic plants, which express \textit{35S::mtHD-GFP} and \textit{35::GFP}, were used for the subcellular localization study by confocal microscopy at Confocal Microscope Facility, Iowa State University. Seven-day-old seedling grown on MS plates were harvested and stained for 15min in the MS media containing 200nM MitoTracker Orange CMTMRos (Life Technologies, Carlsbad, CA). After a 15min wash in the MS
media, the seedlings were examined using a Leica TCS NT confocal microscope system. The 489nm/500-535nm (excitation/emission wavelength) were used to detect GFP fluorescence in both roots and leaves; 543nm/540-600nm were used to detect the fluorescence signals of MitoTracker Orange in roots; 543nm/600-790nm was used to detect fluorescence of chlorophyll in leaves.

Yeast molecular genetic assays

Yeast htd2 knockout strain (YHR067W, BY4741 background, mating-type A) was obtained from Thermo Scientific (Waltham, MA). The deletion of the HTD2 gene in the mutant was confirmed by PCR and sequencing. The yeast mutant strain was transformed with yeast expression constructs (YEp351, HTD2-YEp351, mtHD-YEp351, and mtHD-YEp351M) using MicroPulser Electroporator (BioRAD) following manufacturer’s instruction. Transgenic yeast cells were screened in Synthetic Complete (SC) medium without leucine, a selection marker of YEp351 vector. Three individual cell lines were selected from each transgenic event for the following genetic complementation assays of the htd2 mutant strain. Cultures were first grown in SC medium containing 2% glucose until OD$_{600}$ reach 1.5. Cells were collected by centrifugation, washed with SC medium, and plated on solid SC medium supplemented with 3% glycerol as sole carbon source as described previously (14,19,44).

Hydratase 2 activity assays

The mtHD-pET construct was transformed into E. coli BL21* strain (Life Technologies, Carlsbad, CA) for recombinant protein production. His-tagged mtHD was purified using Probond Nickel-Chelating Resin (Life Technologies, Carlsbad, CA) following the manufacturer’s instruction. Hydratase 2 activity of the recombinant mtHD
was measured using trans-2-decenoyl-CoA and trans-2-hexadecenoyl-CoA as substrates following a previous procedure (14). These enoyl-CoA molecules were chemically synthesized from free fatty acids and CoA, and were purified using HPLC as previously described (45). The reaction mixtures were lyophilized and, then, the acyl moieties on acyl-CoA molecules were transmethylated at 80°C for 30 min in 1 mL 1N methanolic HCl containing 1 µg/mL pentadecanoic acid (15:0) as an internal standard. After adding 1 mL of 0.9% NaCl and 1 mL of heptane, the methylated mixture was vigorously vortexed, and centrifuged at 5000 g for 10 min. The heptane phase was separated in a new vial and dried under a stream of N₂ gas. The hydroxyl groups of 3-hydroxydecenoic acid and 3-hydroxyhexadecenoic acid, which were produced by the recombinant mtHD enzyme, were silylated in 100 µL BSTFA at 65°C for 20 min. The samples were dried under a stream of N₂ gas, dissolved in heptane, and quantified by GC-MS. Kinetic parameters were calculated using Prism Verion 5.0 (GraphPad Software, La Jolla CA).

**Metabolomics analysis**

Metabolites (i.e., amino acids, fatty acids, hydroxyfatty acids, surface lipids, sucrose, glycolate, and membrane lipids) were extracted from 16-day-old Arabidopsis aerial organs and analyzed using HPLC, GC-MS, or LC-MS as described in Chapter 2. In addition chlorophylls and their degradation intermediates were quantified using LC-MS (46).

**Transcriptomics analysis by RNA-seq**

RNA extraction, sequencing, sequence assembly, and statistical analysis were conducted as described in Chapter 2.
REFERENCES


FIGURES

Figure 1. Sequence alignment of Human HsHtd2 and its Arabidopsis ortholog.
Figure 2. Subcellular localization of the protein encoded by AT5G60335.
Figure 3. Biochemical identification of the enzyme encoded by AT5G60335.

\[
\begin{align*}
\text{yhr067w + YE}\text{p351} \\
\text{yhr067w + HTD2-YE}\text{p351} \\
\text{yhr067w + AT5G60335-YE}\text{p351} \\
\text{Yhr067w + AT5G60335-YE}\text{p351M}
\end{align*}
\]

Serial dilution
Figure 4. Kinetic properties of the recombinant mtHD enzyme. (A) Graph of the substrate concentration dependence. (B) Kinetic parameters.

<table>
<thead>
<tr>
<th>Enoyl-CoA</th>
<th>Km (μM)</th>
<th>Vmax (μM/min/mg)</th>
<th>$k_{cat}/K_m$ (μM$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C10</td>
<td>27.82 ± 4.45</td>
<td>7.85 ± 0.41</td>
<td>0.97</td>
</tr>
<tr>
<td>C16</td>
<td>32.75 ± 5.67</td>
<td>11.44 ± 0.68</td>
<td>1.20</td>
</tr>
</tbody>
</table>
Figure 5. Expression profiles of Arabidopsis *mtER* in various organs.
Figure 6. Characterization of the *mthd-rna1* mutant alleles. (A) RNAi transgene of the *mtHD* gene. (B) Morphological appearances in ambient air. (C) Morphological appearances in the elevated CO$_2$ atmosphere.
Figure 7. Metabolic alterations in the *mthd-rna1* mutant alleles. (A) Alterations in amino acids in ambient air. (B) Alterations in amino acids in the elevated CO$_2$ atmosphere. (C) Alterations in glycolate. (D) Alterations in sucrose. (E) Alterations in fatty acids in ambient air. (F) Profiles of fatty acids in the elevated CO$_2$ atmosphere. (G) Alterations in 3-hydroxymyristic acid in the elevated CO$_2$ atmosphere. (H) Alterations in galactolipids and sulfolipids in ambient air. (I) Profiles of galactolipids and sulfolipids in the elevated CO$_2$ atmosphere. (J) Alterations in surface lipids in ambient air. (K) Profiles of surface lipids in the elevated CO$_2$ atmosphere.
Figure 7 continued.
Figure 7 continued.
Figure 7 continued.
Figure 7 continued.
Figure 8. Alterations in chlorophylls and their degradation intermediates. (A) Alterations in ambient air. (B) Alterations in the elevated CO₂ atmosphere.
Figure 9. Double mutations in the *mtHD* and *mtKAS* genes. (A) Morphological appearances in ambient air. (B) Morphological appearances in the elevated CO$_2$ atmosphere. (C) Alterations in amino acids in ambient air. (D) Alterations in amino acids in the elevated CO$_2$ atmosphere. (E) Alterations in 3-hydroxymyristic acid in the elevated CO$_2$ atmosphere.
Figure 9 continued.
Figure 9 continued.
Figure 10. DE genes in carbon and nitrogen metabolism.
CHAPTER IV: BIOCHEMICAL IDENTIFICATION AND REVERSE GENETIC CHARACTERIZATION OF MITOCHONDRIAL ENOYL-ACP REDUCTASE, A REDUNDANT ENZYME COMPONENT OF THE MITOCHONDRIAL TYPE II FATTY ACID SYNTHASE OF ARABIDOPSIS

Manuscript in preparation and to be submitted to *Plant Physiology*

Xin Guan¹, Yozo Okazaki², Kazuki Saito², and Basil J. Nikolau¹

¹ Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, Iowa 50011

² Metabolomic Function Research Group, RIKEN Plant Science Center, Yokohama 230-0045, Japan

ABSTRACT

Arabidopsis mitochondrial fatty acid synthase (mtFAS) generates precursors for the biosynthesis of lipoic acid and lipid A-like molecules. In this process, mitochondrial enoyl-ACP reductase (mtER) catalyzes the iterative reduction of enoyl-ACP to form acyl-ACP. Here we report the identification and characterization of an Arabidopsis mtER enzyme (AT3G45770). GFP transgene experiments show that the N-terminal sequence of this protein functions as a mitochondrial targeting peptide. The enoyl-ACP reductase activity was functionally identified by its ability to complement the yeast mutant strain lacking the mtER activity. *In vitro* kinetic assays on a series of enoyl-CoA substrates demonstrate that Arabidopsis mtER displays a broad substrate preference for acyl moieties of up to 16-carbon chain length. Characterization of T-DNA insertion mutant alleles establish that Arabidopsis plants lacking the mtER function exhibit no morphological phenotype and display only subtle metabolic alterations on a series of
metabolites that are altered by other mutant alleles that affect mtFAS (e.g., amino acids, fatty acids, hydroxyfatty acids, and lipids). These results indicate that the mtER gene is functionally redundant with respect to a yet unknown enzyme for the enoyl-ACP reductase activity in Arabidopsis mitochondria.

**INTRODUCTION**

In eukaryotic cells, fatty acid biosynthesis takes place in multiple subcellular compartments: in the cytosol, in mitochondria, and in plastids of plants. These events play distinct biological roles and do not reflect just metabolic redundancy (1-3). Based on the well characterized mitochondrial FAS (mtFAS) in yeast and humans, it is a Type II system that is composed of four discrete enzymes, namely mitochondrial β-ketoacyl-ACP synthase (mtKAS), β-ketoacyl-ACP reductase (mtKR), 3-hydroxyacyl-ACP dehydratase (mtHD), and enoyl-ACP reductase (mtER), which sequentially catalyze iterative reactions that produce fatty acids (2). This mitochondrial system resembles the Type II FAS that occurs in bacteria and plant plastids (3,4), but is in contrast to the Type I FAS that occurs in the cytosol of mammals and fungi, which is a multifunctional protein that contains all of the reaction centers utilized to produce fatty acids (1). In yeast, all of the mtFAS components have been characterized from *Saccharomyces cerevisiae* (5-8). *In silico* analyses and genetic screens have led to the identification of all mtFAS components in humans (9-12). In plants, two mtFAS components, mtKAS and mtHD, have been investigated from *Arabidopsis thaliana* at both the biochemical and physiological levels (13) (Chapter 2 and Chapter 3).

The mtFAS system generates octanoyl-acyl carrier protein (ACP), which is used as the substrate for the biosynthesis of lipoic acid in yeast, mammals, and plants.
(2,14,15). Lipoic acid is the cofactor that is essential for pyruvate dehydrogenase (PDH), β-ketoglutarate dehydrogenase (KGDH), and glycine decarboxylase complex (GDC) (16). However, because all of the mtFAS components that have been characterized in vitro accept substrates containing acyl moieties of more than 8 carbon atoms, the mtFAS system may also support the biosynthesis of longer chain fatty acids (9-11,17) (Chapter 2 and Chapter 3). Indeed, the Arabidopsis mtFAS system has been characterized to donate 3-hydroxymyristyl-ACP for the generation of lipid A-like molecules (Chapter 2). In addition, it has also been proposed that the long-chain acyl-ACP molecules produced by the mtFAS system are required for the remodeling of cardiolipins in fungi (18,19) and plants (20,21).

In yeast, knockout of any gene encoding the enzyme of mtFAS leads to the loss of cytochromes (5-8), the inability to grow on the non-fermentable carbon source (5-8), and a deficiency in mitochondrial RNA processing (22). The understanding of the physiological functions of the mtFAS system in mammals is still elusive. However, down-regulation of mitochondrial ACP, the substrate carrier protein involved in this pathway, compromises lipoylation and respiratory complex in human cell cultures (23). Similarly, knockdown of the expression of Mus musculus mitochondrial malonyl-CoA ACP transacylase (mtMCAT), an enzyme that generates the malonyl-ACP building block for the mtFAS system, results in the defective protein lipoylation and energy equilibrium (24). In plants, knockout of the mtKAS gene and knockdown of the mtHD gene in Arabidopsis thaliana lead to the diminishment of the lipoylation of the H-subunit of GDC, an essential enzyme in photorespiration, resulting in the deficiency in photorespiration (25). However, the lipoylation of PHD and KGDH is only slightly
affected by the mutations in the mtFAS system (25); this may indicate that there’s another lipoic acid generation system in plants, other than mitochondria (26,27). In addition to the defective lipoic acid synthesis, mutations in the Arabidopsis mtFAS system deplete lipid A-like molecules, a class of mitochondrial lipids whose physiological functions are currently unknown (Chapter 2 and Chapter 3).

Here, we demonstrate the identification and characterization of Arabidopsis mtER, the enzyme that is responsible for the enoyl-ACP reductase reaction in mitochondria. Loss-of-function mutant study shows that mutations in the \textit{mtER} gene do not lead to the typical morphological and metabolic phenotypes associated with the deficiency in mtFAS, suggesting the occurrence of the gene redundancy for this enzymatic function in Arabidopsis mtFAS.

RESULTS

1. Identification of the Arabidopsis \textit{mtER} candidate

Using \textit{S. cerevisiae} \textit{ETR1} sequence as the query, BLAST searches of protein database reveal a single \textit{mtER} candidate (AT3G45770) in Arabidopsis genome (total score 168, query cover 76%, and E value 3e-47); this candidate is the only Arabidopsis gene with a score above 40. The amino acid sequence alignment of Arabidopsis \textit{mtER} and \textit{S. cerevisiae} \textit{ETR1} exhibits 42.2% similarity and 26.6% identity. In addition, Arabidopsis \textit{mtER} also shares significant homology with its \textit{H. sapiens} ortholog (54.3% similarity and 41.2% identity) and \textit{C. tropicalis} ortholog (43.7% similarity and 30.9% identity). However, these gene products do not share sequence homology at N-termini, which are rich in basic amino acids and lack acidic residues (Fig 1). These features are characteristics of mitochondrial targeting sequence elements, and indeed, all of the
Arabidopsis orthologs in *S. cerevisiae*, *C. tropicalis*, and *H. sapiens* were characterized to be mitochondrially localized proteins with N-terminal mitochondria transit peptides (8,12,17,28). The Arabidopsis AT3G45770 gene product is predicted to be a mitochondrially localized protein by MitoProt II (27 N-terminal residues with a score of 0.9078) (29) and PSORT (with a score of 0.613) (30), however, its subcellular localization is predicted to be plastids by Target P (44 N-terminal residues with a score of 0.599) (31,32).

2. Subcellular localization of the protein encoded by AT3G45770

To determine the subcellular localization of the Arabidopsis AT3G45770 gene product, a C-terminal GFP fusion was constructed by cloning the 5’ terminus of AT3G45770 that encodes 100 amino acid residues in-frame with GFP-coding sequence, and stably expressed in Arabidopsis under the control of 35S promoter. The transgenic plants were examined by confocal microscopy. In mesophyll cells, GFP signals localize to small but distinct areas, and they do not overlap with chlorophyll auto-fluorescence (Fig 2A), revealing that the fusion proteins do not reside in plastids. Root cells of these transgenic plants were further investigated to avoid interference from chlorophyll auto-fluorescence. MitoTracker Orange was applied in parallel as the mitochondrial marker, and its signal was recorded simultaneously with the GFP signal. The wild-type control does not exhibit any GFP fluorescence (Fig 2B). The P35S::GFP control localizes to the cytosol and nucleus, which are distinct from mitochondria (Fig 2C). GFP fluorescence of the AT3G45770-GFP fusion protein overlaps with MitoTracker Orange fluorescence (Fig 2D), demonstrating that the N-terminal sequence of the protein encoded by AT3G45770
directs GFP to Arabidopsis mitochondria. We therefore conclude that AT3G45770 encodes a protein that is mitochondrially localized.

3. Biochemical identification of the protein encoded by AT3G45770

To test whether AT3G45770 encodes a protein that is a functional equivalent of the *S. cerevisiae* Ybr026p/Etr1p protein, *S. cerevisiae* ybr026c mutant strain was transformed with the YEp351 empty vector, *ETR1*-YEp351, *AT3G45770*-YEp351, and *AT3G45770*-YEp351M (Fig 3). The negative control, YEp351 empty vector, does not rescue the growth deficiency of the mutant strain, whereas the positive control, *ETR1*-YEp351, reversed the normal growth of the mutant on the glycerol containing media. In parallel, the mutant strains can also grow on the media containing glycerol as the sole carbon source when expressing *AT3G45770*-YEp351 (full length protein) or *AT3G45770*-YEp351M (N-terminally truncated protein with a yeast *COQ3* mitochondrial targeting peptide), demonstrating that AT3G45770 encodes a mitochondrial enoyl-ACP reductase (mtER) component of the mtFAS system.

4. Kinetic property of the recombinant mtER protein

*In vitro* kinetic assays were performed with the recombinant Arabidopsis mtER enzyme to evaluate its acyl chain specificity. A truncated *mtER* gene was expressed in *E. coli* without the N-terminal 32 amino acid residues that constitute the putative mitochondrial targeting peptide. The His-tagged protein was purified by nickel affinity column chromatography. *In vitro* enzymatic assays with the recombinant mtER proteins in *S. cerevisiae, C. tropicalis*, and *H. sapiens* suggested that these enzymes can utilize enoyl-CoA in stead of enoyl-ACP as substrates (8,12,17,28). Arabidopsis mtER enzyme also exhibits activities on enoyl-CoA substrates containing acyl moieties of 4, 6, 8, 10,
and 16 carbon atoms. We therefore used these enoyl-CoA molecules as substrates for the acyl chain specificity assays. The recombinant mtER enzyme reduces all of the enoyl-CoA substrates in a NADPH dependent manner (Fig 4A). In comparison, when NADH is used as the reducing reagent, reduction of the substrates is near the detection limit of the assay system used (data not shown). These data demonstrate that Arabidopsis mtER is NADPH dependent, which is in keeping with its orthologs in *S. cerevisiae*, *C. tropicalis*, and *H. sapiens* (8,12,17,28), but is distinct from *E. coli* fabI enzyme, a functional equivalent that utilizes NADH as the cofactor (2). The $K_m$ and $V_{max}$ values for the recombinant Arabidopsis mtER enzymes on the substrates tested (Fig 4B) are comparable to those from *S. cerevisiae* (8), *C. tropicalis* (8,28), and *H. sapiens* (12,17). The catalytic efficiency ($k_{cat}/K_m$) on all the enoyl-CoA substrates tested is ranked in the following orders: C10 > C8 > C16 > C4 > C6 (Fig 4B). Collectively, the Arabidopsis mtER enzyme exhibits broad acyl chain specificity. This is in agreement with the specificities of the Arabidopsis mtKAS enzyme and mtHD enzyme, which are responsible for the 1\textsuperscript{st} and 3\textsuperscript{rd} reactions of mitochondrial fatty acid synthesis, respectively (Chapter 2 and Chapter 3), suggesting that Arabidopsis mtFAS may support the biosynthesis of both medium- and long-chain fatty acids.

5. Expression profiles of the *mtER* gene

Quantitative RT-PCR was performed with RNA templates to determine the spatial and temporal expression of the *mtER* gene in aerial organs of young seedlings, roots, flowers, siliques, and rosette leaves of wild-type Arabidopsis (Fig 5). These data show that *mtER* gene is expressed in all organs tested. The highest expression was found in rosette leaves, followed by aerial organs of young seedlings, flowers, roots, and siliques.
The near ubiquitous expression pattern of the mtER gene is consistent with the public microarray data visualized by Arabidopsis eFP-Browser (33,34). The near ubiquitous expression pattern of the mtER gene is also in parallel with the expression of the mtKAS gene and mtHD gene, which are involve in the mtFAS system (Chapter 2 and Chapter 3).

6. Mutations in the mtER gene do not lead to significant phenotypic and metabolic alterations

To understand the physiological significance of the mtER gene, we identified and characterized eight T-DNA-tagged mutant alleles by searching the SALK Insertion Sequence Database (http://signal.salk.edu) (35). Plant genomic DNA sequences flanking the T-DNA border were amplified and sequenced, and the molecular structure of the mutant allele was deduced by aligning these flanking DNA sequences with the Arabidopsis genome sequence. Of the eight mter alleles that were analyzed, only 4 (mter-1 (SALK_056770), mter-2 (SALK_085297), mter-3 (SALK_033308), and mter-4 (SALK_130583)) were confirmed to have T-DNA insertions residing in the coding area of the mtER gene (Fig 6A). The insertion in mter-1 carries a chimeric T-DNA that deletes the mtER gene sequence from the 6th intron to the 8th exon. In the mter-2 allele, the T-DNA is inserted in the mtER coding area with the deletion of about 3,500bp DNA sequence, spanning from the upstream gene to the 7th exon of the mtER gene. The mter-3 and mter-4 mutant alleles designate the same mutant with the insertion located in the last intron. RT-PCR assays cannot detect the expression of mtER in mter-1 and mter-2 mutant alleles (Fig 6B), revealing the loss of mtER functionalities in these two mutant strains. Details of the other mutant alleles that were investigated but proved not to be useful in this experiment are not shown.
Visually no detectable morphological difference could be ascribed to the *mter-1* and *mter-2* mutants relative to their wild-type siblings in the standard laboratory growth conditions used in our studies (Fig 7). This is distinct from the morphological phenotype associated with the deficiency in the mtFAS system, such as the *mtKAS* mutants and *mtHD* mutants (25) (Chapter 2 and Chapter 3).

Metabolic profiling was performed on the aerial organs of *mter-1* and *mter-2*, however, alterations are barely detected in these *mtER* mutants. Although Gly levels are slightly elevated in the mutants, to about 2-fold as compared to the wild-type plants, alterations in the levels of other amino acids are barely detected (Fig 8A). Fatty acid compositions of the mutants resemble that of the wild-type control (Fig 8B). In addition, contents of lipids (e.g., MGDG, DGDG, SQDG, DAG, PC, PE, PG, PI, and TAG) in the two mutants are similar to those of the wild-type plants (Fig 8C). Chlorophyll levels are not altered in the *mter* mutants (Fig 8C). Finally, the levels of 3-hydroxymyristic acid, the fatty acid marker for lipid A-like molecules, are only slightly reduced in the mutants, by about 25% of the wild-type level (Fig 8D). These slight changes in metabolic homeostasis in the *mter* mutant alleles are distinct from the marked alterations associated with the deficiency in mtFAS, such as those characterized in plants carrying mutations in the *mtkas* or *mthd* alleles.

Taken together, the normal morphological appearance and the subtle metabolic alteration of the *mter* mutant alleles suggest that mtFAS only exhibits minor changes as a result of the mutations in the *mtER* gene. This is likely due to a functional equivalent of an enoyl-ACP reductase in Arabidopsis mitochondria that can complement the loss of mtER functionalities.
DISCUSSION

Arabidopsis mtFAS system

Plant cells appear to have at least two de novo fatty acid forming systems, which occur in distinct subcellular compartments: plastids and mitochondria. Plastidial FAS (ptFAS) is a Type II system that utilizes dissociated components to produce fatty acids. The existence of a mitochondrial FAS (mtFAS) system was first suggested by the discovery of a mitochondrially located ACP in Neurospora (36). All of the mtFAS components have been biochemically identified in *S. cerevisiae* and *H. Sapiens* (2). By contrast, in Arabidopsis only three mtFAS components have been biochemically characterized, including mtACP (37), mtKAS (13), and mtHD (Chapter 3). Although many of the Arabidopsis components have yet to be identified, the role of mtFAS in providing octanoyl-ACP, the precursor for lipoic acid production has been established (25). Accordingly, mutations in the Arabidopsis mtFAS system display a typical photorespiration-deficient appearance and glycine accumulation, due to the loss of the lipoylation of the H-subunit of glycine decarboxylase complex, an essential enzyme complex in photorespiration (25). Acyl-lipids metabolism is also altered as a secondary effect of the deficiency in photorespiration (Chapter 2). In addition, a novel metabolic destination of the mtFAS system in supplying 3-hydroxymyristyl-ACP for the synthesis of lipid A-like molecules has been discovered, and the deficiency in mtFAS leads to the depletion of 3-hydroxymyristic acid, a hallmark of lipid A-like molecules (Chapter 2).

Arabidopsis mtER enzyme

The last step of the mtFAS acyl chain elongation cycle is the iterative reduction reactions that drive the formation of saturated acyl-ACP molecules from enoyl-ACP
precursors harboring a double bond at the trans-2 position. These reactions are catalyzed by enoyl-ACP reductase.

Plant mitochondria are known to have enoyl-ACP reductase activity (38), however, the enzyme involves in this process has yet to be identified. This work provides the evidence that Arabidopsis mtER, an enzyme encoded by AT3G45770, is responsible for the mitochondrial enoyl-ACP reductase reaction. Specifically, amino acid sequence alignment shows that the Arabidopsis protein is homologous to those biochemically identified mtER enzymes, namely Etr1p in \textit{S. cerevisiae}, Etr1p and Etr2p in \textit{C. tropicalis}, and MECR in \textit{H. sapiens}. The N-terminal extension sequence of Arabidopsis mtER is characterized to be a mitochondrial transit peptide that directs the mtER-GFP fusion protein to Arabidopsis mitochondria. In addition, heterologous expression of the Arabidopsis mtER enzyme in mitochondria of the \textit{S. cerevisiae} ybr026c mutant strain restores its deficiency in Etr1p functions, leading to the complement of the growth on media containing glycerol as sole carbon source. Finally, the recombinant Arabidopsis mtER enzyme is enzymatically active that it catalyzes the NADPH-dependent reduction of trans-2-enoyl-CoA to acyl-CoA, a reaction that is also catalyzed by its orthologs in \textit{S. cerevisiae}, \textit{C. tropicalis}, \textit{H. sapiens} (8,12,17,28).

From the evolutionary point of view, Arabidopsis mtER enzyme and its orthologs in other eukaryotes are distinct from the rest of known enoyl-ACP reductases in Type II systems, since the mitochondrial enzymes are members of the medium chain alcohol dehydrogenase / reductase (MDR) super family (39), whereas the other known enoyl-ACP reductase belong to either short chain alcohol dehydrogenase / reductase (SDR) super family or triose phosphate isomerase (TIM) super family (39,40). For example, at
least four different types of enoyl-ACP reductase exist in bacteria, namely fabI, fabL, fabV, and fabK. FabI in *E. coli* and *Bacillus subtilis* (41,42), fabL in *Bacillus subtilis* (42), and fabV in *Vibrio cholera* (43), are members of the SDR superfamily, while fabK in *Streptococcus pneumoniae* (44) belongs to the TIM superfamily. In eukaryotes, the only known enoyl-ACP reductase in the Type II FAS system, with the exception of Arabidopsis mtER and its orthologs, is the plant plastidial enoyl-ACP reductase (ptER) (45,46), which is the plant ortholog of the *E. coli* fabI enzyme. Arabidopsis mtER does not share significant amino acid sequence similarity with fabI, fabL, fabV, fabK, and ptER, and it is an example of functional convergence during evolution. In addition, Arabidopsis mtER displays distinct cofactor preferences as compared to the other enoyl-ACP reductases. Specifically, Arabidopsis mtER reduces the double bond only in the presence of NADPH, which is consistent with its orthologs in *S. cerevisiae*, *C. tropicalis*, *H. sapiens* (8,12,17,28), as well as fabL in *Bacillus subtilis* (42). In comparison, fabI enzymes in *E. coli* and *Bacillus subtilis* are NADH-dependent (42); fabV in *Vibrio cholera* recognizes both NADH and NADPH, but prefers NADH over NADPH (43); fabK in *Streptococcus pneumoniae* requires FMN as a second co-factor in addition to NADH (44).

*In vitro* kinetic assays with the recombinant Arabidopsis mtER enzyme demonstrate that it can catalyze the reductions of a broad range of trans-2-enoyl substrates to the corresponding acyl thioesters, which is in keeping with the broad substrate preferences of MECR, its ortholog in *H. sapiens* (17). The crystal structure of the *H. sapiens* MECR enzyme suggests a deep curved pocket that extents from the catalytic site to the interior of the catalytic domain, and this pocket is long enough to
accommodate acyl moieties of up to 16 carbon atoms (17). The broad acyl chain
specificity of Arabidopsis mtER is also in agreement with that of Arabidopsis mtKAS
and mtHD, two enzymes of the same acyl chain elongation cycle (Chapter 2 and Chapter
3). Collectively, the acyl chain specificity data of the recombinant Arabidopsis mtER
enzyme support the multiple metabolic destinations of the mtFAS intermediates, namely
octanoyl-ACP for lipoic acid synthesis and 3-hydroxymyristyl-ACP for the synthesis of
lipid A-like molecules.

**Functional redundancy of Arabidopsis mtER**

Arabidopsis *mter* alleles do not exhibit the marked morphological phenotypes as
the *mtkas* alleles and *mthd* alleles (Chapter 2 and Chapter 3). In addition, metabolic
alterations are barely detected in the *mter* alleles in the levels of amino acid, fatty acid,
lipid, and 3-hydroxymyristic acid; these alterations are hallmarks of the deficiency in
mtFAS (Chapter 2 and Chapter 3). The subtle changes in morphological and metabolic
phenotypes in the *mter* alleles show that functionalities of the mtFAS system is not
greatly affected by mutations in the *mtER* gene, suggesting a functionally equivalent
enzyme that can complement the loss of mtER functions in the mutants. The functional
redundancy of Arabidopsis mtER is distinct from its *S. cerevisiae* ortholog, Etr1p, which
is the only enzyme responsible for the mitochondrial enoyl-ACP reductase activity, and
knockout of the *ETR1* gene leads to a deficiency in mtFAS (8).

The phenomenon of functional redundancy is widespread. In the case of enoyl-
ACP reductase of fatty acid biosynthesis, the redundancy could be due to the presence of
protein isoforms, for instance the two mitochondrial enoyl-ACP reductase isoforms in *C.
tropicalis*, Etr1p and Etr2p, which are encoded by different genes and differ only three
amino acid residues (12). The redundancy could also be a result of an equivalent enzyme without significant sequence homology, for example the two enoyl-ACP reductases in *Bacillus subtilis*, fabI and fabL, which share low similarity and display different cofactor preferences, catalyze the same reaction (42). After BLAST analysis, we exclude the possibility of a second mtER isoform in Arabidopsis genome. As an alternative hypothesis, the enoyl-ACP reductase reaction in mitochondria of the mtER mutant alleles could be catalyzed by a structurally unrelated functional equivalent of Arabidopsis mtER. This hypothesis receives supports from the fact that *S. cerevisiae* Etr1p enzyme can be functionally replaced by the *E. coli* fabI enzyme (8), suggesting that the enzymatic activity of enoyl-ACP reductase is more important than the structural properties of the mtFAS complex in the physiological environment. In this respect, many candidate genes are selected for the mtER functions based on the sequence homology to the known enoyl-ACP reductases in SDR and TIM superfamilies. One candidate is the ptER enzyme encoded by AT2G05990, an ortholog of the *E. coli* fabI enzyme (34.9% similarity and 21.9% identity); this enzyme belongs to the SDR superfamily and is proposed to be transported to Arabidopsis mitochondria other than plastids due to the possible alternative splicing. When *Bacillus subtilis* fabL is used as the query to search Arabidopsis genome, many Arabidopsis homologous genes are found; they are members of the SDR superfamily. In addition, Arabidopsis contains a single ortholog of *Streptococcus pneumoniae* fabK enzyme (44.0% similarity and 28.7% identity), which is encoded by AT5G64250; this Arabidopsis enzyme belongs to the TIM superfamily. To discover the functional equivalent of Arabidopsis mtER, we have initiated experiments to test the biochemical functions and subcellular localizations on a few selected candidate genes.
METHODS

Plant materials and growth conditions

Arabidopsis thaliana ecotype Columbia (Col-0) and mutant Arabidopsis genetic stocks were obtained from the Arabidopsis Biological Resource Center (Columbus, OH). All of the mutant alleles are in Col-0 background. Seeds were sterilized and sown on Murashige and Skoog agar medium as describe in Chapter 2. After seeds were sown, plates were placed at 4°C for 4 days to break seed dormancy, and plants were grown in ambient air, where the temperature was maintained at 23 ± 2°C with continuous illumination (100 µmol m⁻² s⁻¹).

RNA extraction and quantitative RT-PCR

RNA was extracted from 50 mg fresh tissues using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA). A 2µg aliquot of RNA was used for reverse transcription using the SuperScript III First-Strand Synthesis Super Mix (Life Technologies, Carlsbad, CA) with oligo (dT)₂₀ primer. Quantitative RT-PCR was performed on StepOnePlus Real-Time PCR System using SYBR Select Master Mix (Life Technologies, Carlsbad, CA). Actin-2 (AT3G18780) was used as the reference RNA for relative quantification. The gene specific primers are E1 (TCCGTCACCAGATTGGTGAATC) - E2 (CGGATACACACCTTCAATTCGA) for mtER and E3 (TCGTACAACCGGTATTGTGCTG) - E4 (AGGTCACGTCCAGCAAGGTCAA) for Actin-2.

Recombinant DNA constructn

Arabidopsis cDNA encoding the mtER gene was cloned from reverse transcribed RNA preparations from 16-day-old aerial organs of the wild-type Arabidopsis plants. The
5’-terminal coding sequence, which encodes 100 amino acid residues was amplified and subcloned into pEarleyGate103 (47) for subcellular localization study.

For yeast molecular genetic assays, the yeast *ETRI* gene was amplified from yeast genomic DNA extracts, while the full-length Arabidopsis *mtER* was amplified from cDNA preparation. These two genes were cloned into the yeast expression vector YEp351 (48) under the control of PGK promoter; the two constructs were named *ETRI*-YEp351 and *mtER*-YEp351, respectively. A truncated mtER-coding sequence was amplified, removing 5’-terminal sequence that encodes the putative 32 amino acid transit peptide. This PCR product was cloned into the yeast mitochondrial expression vector, YEp351M (Chapter 3), resulting in a construct named *mtER*-YEp351M.

For protein expression, the *mtER* ORF was amplified, removing the transit peptide sequence. The PCR product was cloned into pET30b vector (EMD Millipore, Billerica, MA), and the resulting construct was named *mtER*-pET. The *mtER*-pET construct expresses a recombinant mtER enzyme containing a His-tag located at C-terminus.

**Plant transformation and selection**

Arabidopsis transformation was performed using the *Agrobacterium tumefaciens*-mediated infiltration method (49) to generate transgenic plants that express the P35S::*mtER-GFP* fusion transgene. Selection of the transformants was performed as described in Chapter 3.

**Confocal microscopy**

T2 generation of transgenic plants, which express the P35S::*mtER-GFP* were used for the subcellular localization study by confocal microscopy at Confocal
Microscope Facility, Iowa State University. Seven-day-old seedling grown on MS plates were harvested, stained using MitoTracker Orange CMTMRos (Life Technologies, Carlsbad, CA), and analyzed using a Leica TCS NT confocal microscope system as described in Chapter 3.

**Yeast molecular genetic assays**

Yeast *ETR1* knockout strain (YBR026C, BY4741 background, mating-type A) was obtained from Thermo Scientific (Waltham, MA). The deletion of the *ETR1* gene in the mutant was confirmed by PCR and sequencing. The yeast mutant strain was transformed with yeast expression constructs (i.e., YEp351, *ETR1*-YEp351, *mtER*-YEp351, and *mtER*-YEp351M) using MicroPulser Electroporator (BioRAD) following manufacturer’s instruction. Transgenic yeast selection and genetic complementation were performed on solid SC medium supplemented with 2% glucose and 3% glycerol as sole carbon source, respectively, as described in Chapter 3.

**In vitro kinetic assays**

The *mtER*-pET construct was transformed into *E. coli* BL21* strain (Life Technologies, Carlsbad, CA) for recombinant protein production. His-tagged mtHD was purified using Probond Nickel-Chelating Resin (Life Technologies, Carlsbad, CA) following the manufacturer’s instruction. Enoyl-ACP reductase activity of the recombinant mtER was measured using NADPH and trans-2-enoyl-CoA containing acyl-moieties of 4, 6, 8, 10, and 16 carbon atoms. These enoyl-CoA molecules were chemically synthesized from free fatty acids and CoA, and were purified using HPLC as previously described (50) (Chapter 3). The reaction was monitored using a spectrometric method, in which the reduction in NADPH was assayed at 340nm as previously described.
Kinetic parameters were calculated using Prism Version 5.0 (GraphPad Software, La Jolla CA).

**Metabolomics analysis**

Metabolites (i.e., amino acids, fatty acids, hydroxyfatty acids, lipids) were extracted from 16-day-old Arabidopsis aerial organs and analyzed using HPLC, GC-MS, or LC-MS as described in Chapter 2 and Chapter 3.

**REFERENCES**


thioester reductases essential for mitochondrial respiratory competence.  
*Molecular and cellular biology* 21, 6243-6253


insights into its substrate recognition properties. *Journal of molecular biology* 379, 830-844


express alpha-amylase and glucoamylase separately or as bifunctional fusion proteins. *Applied microbiology and biotechnology* **43**, 1067-1076


FIGURES

Figure 1. Sequence alignment of Arabidopsis mtER and its counterparts.
Figure 2. Subcellular localization of Arabidopsis mtER.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mesophyll cells</strong></td>
<td>WT</td>
<td>P35S::GFP</td>
<td>P35S::mtER-GFP</td>
</tr>
<tr>
<td><strong>P35S::mtER-GFP</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GFP</td>
<td>GFP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chlorophyll auto-fluorescence</td>
<td>MitoTracker Orange</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merged</td>
<td>Merged</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 3. Genetic complementation in yeast.

\[ \text{ybr026c + YE}\text{p-351} \]

\[ \text{ybr026c + } E\text{TRI-YE}\text{p-351} \]

\[ \text{ybr026c + AT3G45770-YE}\text{p-351} \]

\[ \text{Ybr026c}\]

\[ + \text{AT3G45770-YE}\text{p-351M} \]

Serial dilution
Figure 4. Acyl chain specificity of the recombinant Arabidopsis mtER enzyme. (A) Graph of the substrate concentration dependence. (B) Kinetic parameters.

<table>
<thead>
<tr>
<th>Enoyl-CoA</th>
<th>Km (μM)</th>
<th>Vmax (μmol/min/mg)</th>
<th>$k_{cat}/K_m$ (μM$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C4</td>
<td>35.80 ± 7.26</td>
<td>4.20 ± 0.29</td>
<td>0.18</td>
</tr>
<tr>
<td>C6</td>
<td>24.57 ± 5.60</td>
<td>2.79 ± 0.20</td>
<td>0.17</td>
</tr>
<tr>
<td>C8</td>
<td>11.52 ± 1.92</td>
<td>4.60 ± 0.20</td>
<td>0.60</td>
</tr>
<tr>
<td>C10</td>
<td>8.80 ± 0.91</td>
<td>4.79 ± 0.12</td>
<td>0.82</td>
</tr>
<tr>
<td>C16</td>
<td>12.10 ± 2.37</td>
<td>4.40 ± 0.23</td>
<td>0.55</td>
</tr>
</tbody>
</table>
Figure 5. Expression profiles of Arabidopsis mtER in various organs.
Figure 6. Molecular characterization of *mter* mutant alleles. (A) Genome structure. (B) Confirmation of the knockout of gene expression.
Figure 7. Morphological appearances of the *mter* mutant alleles.
Figure 8. Metabolic profiles of the _mter_ mutant alleles. (A) Profiles of amino acids. (B) Profiles of fatty acids. (C) Profiles of galactolipids, sulfolipids, phospholipids, DAG, TAG, and chlorophyll-like molecules. (D) Profiles of 3-hydroxymyristic acid.
Figure 8 continued.

C

- MGDG
- DGDG
- SQDG
- DAG
- PC
- PE
- PG
- PI
- TAG

Chlorophylls and their degradation intermediates
Figure 8 continued.
CHAPTER V: REVERSE GENETIC CHARACTERIZATION OF A PHOSPHOPANTETHEINYL TRANSFERASE FOR ACTIVATING ACYL CARRIER PROTEINS IN ARABIDOPSIS MITOCHONDRIA

Xin Guan and Basil J. Nikolau

Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, Iowa 50011

ABSTRACT

Mitochondrial acyl carrier protein (mtACP) is the carrier of acyl intermediates in the mitochondrial fatty acid synthase (mtFAS) system. MtACP is generated as an apoprotein, and is activated by the transfer of a phosphopantetheine cofactor from CoA to a conserved serine residue to form holo-mtACP. In this study, we report the characterization of an Arabidopsis mitochondrial phosphopantetheinyl transferase (mtPPT), which is responsible for the activation of apo-mtACP. The activity of this enzyme was identified by its ability to activate the Arabidopsis apo-mtACP1 protein in an in vitro assay. GFP subcellular localization experiments show that the N-terminal sequence of the mtPPT protein functions as a mitochondrial targeting peptide, demonstrating that the mtACP activation process occurs in Arabidopsis mitochondria. Using an RNAi transgene, we generated a series of Arabidopsis lines with a range of down-regulated mtPPT expression levels. Plants with the most decreased levels of the mtPPT expression display a typical deficiency in the mtFAS system, namely miniatuized organs, slow growth, hyperaccumulation of glycine, and reduced 3-hydroxymyristic acid levels. These morphological and metabolic alterations, with the exception of 3-hydroxymyristic acid
reductions, are a consequence of the deficiency in photorespiration, because they are reversed when the plants are grown in a non-photorespiratory condition (in an atmosphere of elevated CO₂ levels). The reduced 3-hydroxymyristic acid levels are a direct result of the deficiency in the mtFAS system. These results demonstrate that mtPPT is responsible for the mtACP activation in Arabidopsis mitochondria, and is important for the mtFAS-related physiological functions.

INTRODUCTION

*De novo* fatty acid biosynthesis in plant cells occurs in plastids and mitochondria. Both plastidial fatty acid synthase (ptFAS) and mitochondrial fatty acid synthase (mtFAS) are Type II systems, which are comprised of discrete, monofunctional enzymes that coordinately act to produce fatty acids using malonyl-ACP as the source of 2-carbon units to elongate a primer substrate (1,2). These Type II systems are in contrast to the Type I fatty acid synthase system that occurs in the cytosol of mammals and yeasts; these Type I enzymes are multifunctional and contain all of the reaction centers to produce a fatty acid (3).

The ptFAS system generates the majority of fatty acids in a plant cell, which are the precursors for the biosynthesis of a plethora of lipid molecules, including phospholipids, galactolipids, sulfolipids, diacylglycerol, and triacylglycerol (4). In addition, fatty acids produced by the ptFAS system serve as the substrates for the fatty acid elongation process in the ER membrane, leading to the production of very-long-chain fatty acids, which further support the biosynthesis of surface lipids, distinct membrane lipid pools, and the ceramide moiety of spingolipids (5,6). In comparison, the mtFAS system contributes the biosynthesis of lipoic acid, the cofactor that is essential for
glycine decarboxylase of the photorespiration process, as well as pyruvate dehydrogenase and β-ketoglutarate dehydrogenase of the TCA cycle (7). Furthermore, the mtFAS system supports the biosynthesis of lipid A-like molecules (Chapter 2).

Acyl carrier protein (ACP) is recruited as the carrier of acyl intermediates for the acyl chain elongation process in the ptFAS and mtFAS systems. It shuttles the acyl intermediates between the active sites of different enzymes (1,2). Arabidopsis genome appears to have at least 8 ACP-coding genes, 5 of which encode plastidial ACP (ptACP) isoforms, whereas 3 genes encode mitochondrial ACP (mtACP) isoforms (8-11). ACP is synthesized as an inactive apo-protein, and is subsequently activated to the holo-form by the addition of a phosphopantetheine cofactor from CoA to a conserved serine residue of apo-ACP, a reaction catalyzed by phosphopantetheinyl transferase (PPTase). Holo-ACP is used as a carrier of the intermediates of the fatty acid biosynthesis processes, where the growing acyl chain is covalently attached to the thiol group of the phosphopantetheine cofactor during the iterative reaction cycles.

There are two groups of PPTase that are responsible for ACP activation. Group I PPTase is typified by E. coli ACP synthase, and they exhibit narrow substrate specificity and mainly recognize ACP of fatty acid biosynthesis (12). Group II PPTase is typified by the Bacillus subtilis Sfp protein, which generally accepts a broad set of protein substrates, including ACP, as well as other carrier proteins, such as nonribosome polypeptide synthesis and polyketide synthesis (13,14). In eukaryotes, Saccharomyces cerevisiae appears to have a single Group I PPTase; it is mitochondrially localized and specifically active on the mtACP component of the mtFAS system (15). By contrast, the human AASDHPPPT gene encodes a Group II PPTase, which is a cytosolic enzyme that can
activate both the mtACP component of mtFAS system and the ACP domain of cytosolic Type I fatty acid synthase (16). In plants however, the PPTase responsible for the phosphopantetheinylation of ptACP and mtACP isoforms, as well as the subcellular localization of these activation processes, have yet to be characterized.

Here, we demonstrate the identification and characterization of Arabidopsis mitochondrial phosphopantetheinyl transferase (mtPPT). This enzyme is mitochondrially localized and activates the mtACP components. Knockdown in the expression of mtPPT gene leads to the typical morphological and metabolic phenotypes associated with the deficiency in mtFAS, confirming the role of mtPPT enzyme in the mtFAS system, rather than the ptFAS system.

**RESULTS**

1. **Identification and cloning of the Arabidopsis PPT candidate**

   The *E. coli ACPS* gene (12), which encodes a typical member of group I PPT, was used as the query to test if a group I PPT candidate can be identified in Arabidopsis. BLAST searches of the protein database, however, indicate that no significant Arabidopsis ortholog of *E. coli ACPS* exists. The *Bacillus subtilis Sfp* gene (13), which encodes a typical member of group II PPT, was used as the query to discover Arabidopsis candidates for the group II PPT. BLAST searches of the protein database reveal two Arabidopsis candidates for the group II PPT, AT3G11470 (total score 82.0, query cover 67%, and E value 4e-18) and AT2G02770 (total score 68.2, query cover 49%, and E value 1e-12).

   AT3G11470 is predicted to encode a protein that belongs to 4’-phosphopantetheinyl transferase superfamily; shares 30.6% similarity and 19.9% identity
at the amino acid sequence level relative to *Bacillus subtilis* *Sfp*. The ORF for AT3G11470 was PCR amplified from RNA extracted from 16-day-old aerial organs of the wild-type Arabidopsis. The protein encoded by AT3G11470 is 300 amino acids long and contains a N-terminal extension of 30 residues as compared to the *Bacillus subtilis* Sfp protein (Fig. 1A). In addition, AT3G11470 exhibits 29.9% similarity and 16.8% identity with the *human AASDHPPT* gene (16), which encodes a cytosolic phosphopantetheinytransferase, and the Arabidopsis protein contains a N-terminal extension of 14 amino acid residues as compared to its human homolog (Fig. 1A). The N-terminal extending sequence of the protein encoded by AT3G11470 is rich in basic amino acids and lack acidic residues; these features are characteristics of mitochondrial targeting sequence elements. Indeed, this Arabidopsis protein is predicted to be mitochondrially localized by PSORT (with a score of 0.716) (17) and TargetP (12 N-terminal residues with a score of 0.652) (18,19). However, prediction by MitoProt II (30 N-terminal residues with a score of 0.6289) (20) suggests a relatively low possibility for the AT3G11470 gene product to be targeted to mitochondria.

AT2G02770 is predicted to encode a PPT-like protein of 661 amino acid residues, which is composed of the N-terminal COP1 interacting protein domain and C-terminal PPT domain. Its PPT domain shares 35.4% similarity and 23.0% identity with the *Bacillus subtilis* Sfp protein, and 65.7% similarity and 60.6% identity with AT3G11470 (Fig. 1B). However, transcription of the full length PPT-like gene or the sequence encoding the PPT domain could not be detected by RT-PCR using several AT2G02770-specific primer sets and the mRNA templates extracted from many different organs, including aerial organs (3, 5, 8, 16-day-old), roots, flowers, siliques, and rosette leaves.
The failure to detect the AT2G02770 mRNA demonstrates that this gene may be a pseudogene or expressed only under the induction of certain discrete environmental conditions. This hypothesis is supported by the fact that a \textit{PPT-like} homolog, which is composed of a fused COP1 interacting protein domain sequence and a PPT domain sequence, is not found in the genomes of other plants, including \textit{Zea mays}, \textit{Oryza sativa}, \textit{Glycine max}, \textit{Vitis vinifera}, and \textit{Populus trichocarpa} (data not shown).

\textbf{2. Subcellular localization of the protein encoded by AT3G11470}

The subcellular localization of the protein encoded by the Arabidopsis AT3G11470 gene was determined with a C-terminal GFP fusion transgene constructed by cloning the 5' terminus of the AT3G11470 gene encoding 80 amino acid residues, in-frame with the GFP-coding sequence. The transgenic Arabidopsis plants stably expressing the GFP transgene under the control of 35S promoter were examined by confocal microscopy. In mesophyll cells, GFP signals localize to the areas that are distinct from chloroplasts that are indicated by chlorophyll auto-fluorescence (Fig 2A), showing that the fusion protein does not localize in chloroplasts. Root cells of these transgenic plants were further investigated to avoid interference from chlorophyll auto-fluorescence. MitoTracker Orange was applied in parallel as the mitochondrial marker, and its fluorescence was recorded simultaneously with the GFP fluorescence. The wild-type control does not exhibit any GFP fluorescence (Fig 2B). The P35S::GFP control localizes to cytosol and nucleus, which are distinct from mitochondria (Fig 2C). GFP signal of the AT3G11470-GFP fusion protein overlaps with the MitoTracker Orange signal (Fig 2D), demonstrating that the N-terminal sequence of the protein encoded by
AT3G11470 directs GFP to Arabidopsis mitochondria. We therefore conclude that the Arabidopsis AT3G11470 gene encodes a mitochondrially localized protein.

3. Biochemical identification of the protein encoded by AT3G11470

To test whether Arabidopsis AT3G11470 encodes a protein that is capable of catalyzing the phosphopantetheinyl transferase reactions, an *in vitro* assay was performed with the recombinant mature protein encoded by AT3G11470. Based on the transit peptide cleavage site predicted by TargetP (18,19), a truncated AT3G11470 protein was expressed in *E. coli* without the N-terminal 12 amino acid residues that constitute the putative mitochondrial transit peptide. The other substrates used in this assay are the recombinant Arabidopsis mtACP1 mature protein and CoA. Mitochondrial ACP 1 (mtACP1) protein is encoded by AT2G44620 (10,11). Sequence alignment of AT2G44620 and the *E. coli* ACPP gene (i.e., a gene encodes ACP) (21) shows an extension on the 5’ terminus of the AT2G44620 encoding sequence, which probably represents the putative mitochondrial transit peptide (Fig. 3A). Therefore, a truncated AT2G44620 was expressed in *E. coli* without these N-terminal 40 amino acid residues. The His-tagged proteins were purified by nickel affinity column chromatography. As indicated by MALDI-QTOF analysis, this *E. coli* expressed protein preparation is only composed of apo-mtACP1, other than holo-mtACP1 (Fig 3B). By contrast, in the presence of the protein encoded by AT3G11470, all of the apo-mtACP1 is converted to its holo-form (Fig 3C). These data demonstrate that the protein encoded by AT3G11470 is a functional phosphopantetheinyl transferase. Considering its mitochondrial localization, this enzyme is designated as mitochondrial phosphopantetheinyl transferase (mtPPT).
4. Expression of the mtPPT gene

Quantitative RT-PCR was conducted with RNA templates to determine the spatial and temporal expression of the mtPPT gene in aerial organs of young seedlings, roots, flowers, siliques, and rosette leaves of wild-type Arabidopsis (Fig 4). These data reveal that the mtPPT gene is expressed in all organs tested. The highest expression was detected in 16-day-old aerial organs, flowers, siliques, and rosette leaves, followed by roots and 3, 5, 8-day-old aerial organs. The near ubiquitous expression of the mtPPT gene is in agreement with the public microarray data visualized by the Arabidopsis eFP-Browser (22,23). In addition, the expression pattern of the mtPPT gene is in parallel with the expression of the mtKAS, mtHD, and mtER genes, which are involved in the acyl chain elongation cycle of the Arabidopsis mtFAS system (Chapter 2, Chapter 3, and Chapter 4).

5. Compromised growth and development of the mtppt-rnai mutant alleles

To investigate the physiological significance of the mtPPT gene, we identified six T-DNA-tagged mutant alleles by searching the SALK Insertion Sequence Database (http://signal.salk.edu) (24). Plant genomic DNA sequences flanking the T-DNA border were amplified and sequenced, and the molecular structures of these mutant alleles were deduced by aligning these flanking DNA sequences with the Arabidopsis genome sequence. Of the six mtppt alleles that were analyzed, only 4 (mtppt-1 (SALK_O74379C), mtppt-2 (SALK_O89901), mtppt-3 (SAIL_240_H05), and mtppt-4 SALK_126601) were confirmed to contain a T-DNA insertion, in which mtppt-1 and mtppt-4 have T-DNA resides in the same location. In all four of these mutant alleles, the T-DNA insertions reside in the 5’ UTR area of the mtPPT gene (Fig. 5A). Quantitative RT-PCR analyses
show that the levels of the \textit{mtPPT} transcript are not affected in these alleles (data not shown), revealing that the T-DNA insertions in the 5' UTR areas of the \textit{mtPPT} gene do not impair the \textit{mtPPT} gene expression. Indeed, in laboratory conditions, no visually detectable morphological difference could be detected between the four \textit{mtppt} alleles and wild-type siblings (data not shown).

As an alternative strategy, RNAi transgenic Arabidopsis plants expressing the double-strand \textit{mtPPT} RNA fragment under the control of 35S promoter were generated and characterized (Fig. 5B). As indicated by PCR amplification of the P35S:\textit{mtPPT} RNAi transgene, we recovered 32 independent lines that carries the transgene (data not shown). About one fourth of these transgenic lines exhibit significantly reduced size of aerial organs in 16-day-old plants relative to the wild-type controls. Two transgenic lines with the most severe phenotypic differences are designated as \textit{mtppt-rnai-1} and \textit{mtppt-rnai-2} (Fig. 5C). These two alleles express 20±7\% and 23±9\% of the \textit{mtPPT} transcript levels, respectively, as compared to the wild-type plants. The morphological differences associated with the expression of the RNAi transgene could be reversed when the mutants are grown in an atmosphere enriched in CO$_2$ (Fig. 5D), suggesting that the compromised growth is a result of the deficiency in photorespiration.

6. Metabolic alterations associated with the \textit{mtppt-rnai} mutant alleles

Metabolic profiling was performed on the 16-day-old aerial organs of \textit{mtppt-rnai-1} and \textit{mtppt-rnai-2}. The most dramatic alteration is the 30-fold increase in Gly levels in the mutants (Fig 6A); this is in agreement with the previous finding that mutations in the \textit{mtFAS} system diminish the lipoylation of the H-subunit of glycine decarboxylase complex, leading to the inability to metabolize photorespiratory Gly (7). Several other
amino acids exhibit slight changes (Fig 6A), and they are likely to be a result of the
defective nitrogen metabolism. Glycolate is the metabolite upstream of Gly in the
photorespiration process. Its level also increases in the \textit{mtppt} mutant alleles, to
approximately 1.5-fold relative to the wild-type plants (Fig 6C). Sucrose is the transport
form of sugar in Arabidopsis, and it is considered to be insufficient in the mutants with a
deficiency in photorespiration due to the inefficient carbon fixation and recovery (25). In
the \textit{mtppt-rnai} mutant alleles, the sucrose levels drop to 35% of the normal level (Fig.
6D), indicating the defective carbon fixation due to the deficiency in the mtPPT
functions; this is consistent with the sucrose depletion in other mutants deficient in the
mtFAS system, including the \textit{mtkas} alleles (7) (Chapter 2) and \textit{mthd} alleles (Chapter 3).
The metabolic alterations in Gly, glycolate, and sucrose are a consequence of the
deficiency in photorespiration, because they are reversed when the plants are grown in an
elevated CO$_2$ atmosphere (Fig. 6B-6D). In contrast to the slightly altered fatty acid
composition in the \textit{mtkas} alleles and \textit{mthd} alleles (Chapter 2 and Chapter 3), fatty acid
composition is not affected in the \textit{mtppt} alleles (Fig. 6E); it is likely because the
metabolic alterations in the two \textit{mtppt} mutant alleles are not as severe as the \textit{mtkas} alleles
and \textit{mthd} alleles. Finally, the levels of 3-hydroxymyristic acid, the fatty acid hallmark for
lipid A-like molecules in Arabidopsis, are reduced to about 60% of the wild-type level
(Fig. 6F), suggesting the defective biosynthesis of lipid A-like molecule in the mutants.
Collectively, these metabolic differences detected in the \textit{mtppt-rnai} alleles resemble those
in the \textit{mtkas} alleles and \textit{mthd} alleles, suggesting that mtPPT, mtKAS, and mtHD are all
involved in the same metabolic process, mitochondrial fatty acid synthesis.
7. A mutant allele for the PPT-like gene (AT2G02770)

A T-DNA tagged mutant (SALK_152625C) was identified for the PPT-like gene AT2G02770 by searching the SALK Insertion Sequence Database (http://signal.salk.edu) (24). This allele contains a T-DNA insertion in the putative coding area of the N-terminal COP1 interacting protein domain. Visually no detectable morphological difference could be ascribed to the T-DNA insertion in this allele. The lack of morphological phenotype in this mutant supports our hypothesis that AT2G02770 may not normally express or is only expressed under certain unknown inductive conditions.

8. Mutant alleles for the mitochondrial ACP-coding genes

To understand the physiological significance of the mtACP components, which are activated by the mtPPT enzyme, we identified and characterized several T-DNA-tagged mutant alleles for the genes that encode the mtACP isoforms. Arabidopsis contains at least three mitochondrial ACP components, encoded by *mtACP1* (AT2G44620), *mtACP2* (AT1G65290), and *mtACP3* (AT5G47630) (11). Of the two *mtacp1* alleles that were analyzed, none of them have T-DNA inserted in the coding area. Of the two *mtacp2* alleles, the *mtacp2-1* allele (SALK_073185) carries a T-DNA in the intron area. In addition, of the two *mtacp3* alleles characterized in this study, the *mtacp3-1* allele (SALK_127678C) contains the T-DNA insertion in the first exon. No alteration in the morphological appearance could be detected in *mtacp2-1* and *mtacp3-1* alleles, demonstrating that *mtACP2* and *mtACP3* are redundant for the Arabidopsis mtFAS system.
DISCUSSION

Acyl carrier protein (ACP) of Arabidopsis fatty acid synthesis

Plant cells harbor two de novo fatty acid forming systems, which occur in distinct subcellular compartments: plastids and mitochondria. Both plastidial FAS (ptFAS) and mitochondrial FAS (mtFAS) are Type II systems, which are catalyzed by discrete, monofunctional enzymes to produce fatty acids. In both plastids and mitochondria, acyl carrier protein (ACP) is utilized as the carrier of the intermediates for the acyl chain elongation process; it shuttles the acyl intermediates between the active sites of different FAS enzymes. The ptFAS system recruits homologous ptACP isoforms encoded by five genes (26). The ptACP4 is the most abundant ptACP isoform (8); ablation of the ptACP4 expression leads to a chlorotic appearance and altered fatty acid composition (9,27).

The Arabidopsis mtFAS system utilizes three mtACP isoforms (11). In this study, T-DNA-tagged mutant alleles for mtACP2 and mtACP3 were segregated, but no visually detectable morphological alteration could be ascribed to the mtacp2 and mtacp3 mutants. It has yet to be elucidated whether mutations in the mtACP1 gene leads to any growth phenotype. The redundancy for the mtACP components in Arabidopsis is distinct from yeast and humans, in which a single mtACP is responsible for supporting the mtFAS process (28,29). Even though the physiological role of the Arabidopsis mtACP components cannot be explored through a reverse genetic approach, the roles of the mtACP components in supporting the mtFAS system to contribute to the biosynthesis of lipoic acid and lipid A-like molecules has been demonstrated (7) (Chapter 2 and Chapter 3). Accordingly, mutations in the Arabidopsis mtFAS system display a typical photorespiration-deficient appearance and glycine accumulation, due to the loss of the
lipoylation of the H-subunit of glycine decarboxylase complex, an essential enzyme in
the photorespiration process (7). In addition, the levels of 3-hydroxymyristic acid, a
hallmark of lipid A-like molecules, are diminished as a consequence of the deficiency in
mtFAS (Chapter 2 and Chapter 3).

**Biochemical and physiological functions of the Arabidopsis mtPPT enzyme**

ACP is synthesized as an inactive apo-protein, and is subsequently activated to
the holo-form by transferring the phosphopantetheine cofactor from CoA to a conserved
serine residue by PPTase. The holo-ACP is used in the FAS process, in which the
growing acyl chain is attached to the thiol group of phosphopantetheine cofactor during
the iterative reaction cycles. In Arabidopsis, the PPTase responsible for the ACP
activation have yet to be discovered. This study provides the evidence that Arabidopsis
mtPPT enzyme is encoded by AT3G11470, which is responsible for the
phosphopantetheinyl transferase reaction in mitochondria. Specifically, amino acid
sequence alignment reveals that the Arabidopsis protein is a Group II PPTase, which is
homologous to the biochemically identified PPTases in *Bacillus subtilis* and humans. The
N-terminal extension sequence of the Arabidopsis mtPPT is characterized to be a
mitochondrial transit peptide that directs the mtPPT-GFP fusion protein to mitochondria.
In addition, the recombinant Arabidopsis mtPPT enzyme is enzymatically active, and can
catalyze the activation of the recombinant Arabidopsis mtACP1 protein, but its activities
on the other mtACP isoforms (e.g., mtACP2 and mtACP3) have yet to be evaluated.

In addition to mtPPT, Arabidopsis genome is predicted to encode a PPT-like
protein, which is encoded by AT2G02770. The protein encoded by this gene appears to
be bipartite and contains an N-terminal domain homologous to COP1 (i.e., a E3 ubiquitin
ligase) and a C-terminal domain that is homologous to PPTase. However, considering that the expression of this gene can not be detected by RT-PCR, it may be a pseudogene or a gene expressed under certain discrete inductive conditions. In contrast to the two Group II PPTase candidates in Arabidopsis, BLAST searches do not reveal any Group I PPTase candidates. The *E. coli* ACPS and yeast Ppt2p are homologs of the Type I PPTase. Therefore, Arabidopsis appears to contain a single functional mtPPTase homolog, a Type II enzyme.

Down-regulation of the Arabidopsis mtPPT leads to a defective mtFAS system. Specifically, the *mtppt-rnai* mutant alleles display typical phenotypes that are associated with the deficiency in photorespiration, namely growth retardation, glycine over-accumulation, and sucrose depletion. These morphological and metabolic alterations are reversed when the plants are grown in an atmosphere enriched in CO₂. Furthermore, knockdown of the *mtPPT* expression results in the decreased levels of 3-hydroxymyristic acid, which are a consequence of the reduced accumulation of lipid A-like molecules that are synthesized from the mtFAS intermediate 3-hydroxymyristyl-ACP. These physiological consequences further confirm that mtPPT contributes to the posttranslational modification of mtACP in mitochondria.

It is noticeable that the morphological and metabolic phenotypes associated with down-regulation of the *mtPPT* gene are not as strong as those exhibited by the *mtkas* alleles and *mthd* alleles (Chapter 2 and Chapter 3). That’s probably due to the fact that the RNAi-induced down-regulation does not represent a null allele, and the remaining mtPPT activity can still partially phosphopantetheinylate the mtACP components for mtFAS process.
It is noteworthy that no morphological or metabolic alteration can be detected when the \textit{mtppt-rnai} mutant alleles are grown in an elevated CO$_2$ atmosphere. This observation suggests that mtPPT is unlikely to be responsible for the modification of the ptACP components, because if mtPPT contributes to the phosphopantetheinylation of the ptACP components, the phenotypic alterations associated with the inactivation of ptACP4 should be detected, irrespective of whether the mutants are grown in ambient air or the elevated CO$_2$ atmosphere. It is likely that an as yet undiscovered, novel type of PPTase is responsible for the phosphopantetheinylation of the Arabidopsis ptACP components.

The mitochondrial localization of mtPPT reveals that the mtACP components in Arabidopsis appear to be activated after import into mitochondria. This result is consistent with the mitochondrially localized yeast Ppt2p enzyme, which is responsible for the mtACP phosphopantetheinylation, but it is non-homologous to Arabidopsis mtPPT (15). Indeed, Arabidopsis mitochondria contain the CoA substrate, which is produced in the cytosol and transported into mitochondria by CoA transporters (30). By contrast, human AASDHPPT, an Arabidopsis mtPPT ortholog and the only known PPTase in humans, lacks the mitochondrial transit peptide and localizes to the cytosol, implying that mtACP in humans is posttranslationally modified prior to its import into mitochondria (16). The discrepancy has yet to be explained between the mitochondrial localization of Arabidopsis mtPPT and the cytosolic localization of human AASDHPPT. Alternatively, a yet undetected transcript variant that encodes a mitochondrially localized PPTase could exist in humans. Conversely, a yet undetected transcript splice variant, which encodes a cytosolically located PPTase, may be expressed in Arabidopsis. To test
the hypotheses on the alternative localizations of Arabidopsis mtPPT and *Homo sapiens* AASDHPPT, further investigations are required to determine the 5’ UTR and 5’ coding sequences of these two genes.

**Implication of Arabidopsis mtPPT for polyketide and nonribosomal peptide syntheses**

From the evolutionary point of view, the Arabidopsis mtPPT enzyme belongs to the Group II PPTase family, which is made up of enzymes that are commonly found to be associated with nonribosomal peptide synthesis. Enzymes in this group generally accept a wide range of protein substrates for phosphopantetheine transferase reactions (31). For instance, *Bacillus subtilis* Sfp, which is required to produce the peptide antibiotic surfactin (13), also catalyzes the phosphopantetheinylation reactions for several other carrier proteins *in vitro* (14). Another example is human AASDHPPT, which is active on the modification of a set of protein substrates, including ACP domain of human cytosolic FAS, human mtACP, carrier protein domain of yeast α-aminoadipate semialdehyde dehydrogenase, *Bacillus brevis* peptidyl carrier protein, and *Bacillus subtilis* ACP (16,32). The biochemical property of the Group II PPTase is distinct from that of Group I PPTase, which is typified by *E. coli* ACPS and yeast mtPPT, since the Group I PPTases generally have narrow substrate specificity (31). In nature, several examples are identified in both prokaryotic and eukaryotic organisms that Group I PPTase is missing, and a single Group II PPTase is used for activating carrier proteins of both primary and specialized metabolic pathways (33,34). In this respect, the Arabidopsis mtPPT enzyme has the potential to modify both the mtACP components of mtFAS and the carrier proteins of specialized metabolic pathways of polyketide and nonribosomal peptide
syntheses. However, because of the poor understanding on the synthetic pathways of polyketide and nonribosomal peptide in Arabidopsis and other plant species, we are not able to construct hypothesis concerning the potential Arabidopsis protein substrates for mtPPT. As an alternative approach, we have initiated experiments to test the \textit{in vitro} activities of mtPPT on several protein substrates of bacterial polyketide and nonribosomal peptide syntheses.

**METHODS**

**Plant materials and growth conditions**

\textit{Arabidopsis thaliana} ecotype Columbia (Col-0) and mutant Arabidopsis genetic stocks were obtained from the Arabidopsis Biological Resource Center (Columbus, OH). All of the mutant alleles are in Col-0 background. Seeds were sterilized and sown on Murashige and Skoog agar medium as describe in Chapter 2. After seeds were sown, plates were placed at 4°C for 4 days to break seed dormancy, and plants were grown in ambient air or the non-photorespiratory condition (1\% CO$_2$ in a growth chamber), where the temperature was maintained at 23 ± 2°C with continuous illumination (100 µmol m$^{-2}$ s$^{-1}$).

**RNA extraction and quantitative RT-PCR**

RNA was extracted from 50 mg fresh tissues using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA). A 2µg aliquot of RNA was used for reverse transcription using the SuperScript III First-Strand Synthesis Super Mix (Life Technologies, Carlsbad, CA) with oligo (dT)$_{20}$ primer. Quantitative RT-PCR was performed on StepOnePlus Real-Time PCR System using SYBR Select Master Mix (Life Technologies, Carlsbad, CA). \textit{Actin-2} (AT3G18780) was used as the reference RNA for relative quantification. The
gene specific primers are T1 (ATGGGAAGCCTGAGGTAGATTG) - T2
(TCGATACCAACCGGAACGTGTA) for mtPPT and T3
(TCGTACAACCGGTATTGTGCTG) - T4 (AGGTCACGTCCAGCAAGGTCAA) for Actin-2.

**Recombinant DNA construction**

Arabidopsis cDNA encoding the mtPPT gene was cloned from reverse transcribed RNA preparations from 16-day-old aerial organs of the wild-type Arabidopsis plants. The 5’-terminal coding sequence, which encodes 80 amino acid residues, was amplified and subcloned into pEarleyGate103 (35) for subcellular localization study. A mtPPT-specific sequence was amplified using primers T5 (CACCCGAAGCTTATCGATCGAATTGCC) - T6 (GCAAGCAAAGCATTCTTCTTAAGCTC) and subcloned into pB7GWIWG2(II) (36) for the RNAi experiment.

For protein expression, the ORFs, which encode Arabidopsis mtPPT and mtACP1, were chemically synthesized by GenScript (Piscataway, NJ) after codon optimization for expression in E. coli. The mtPPT ORF and mtACP1 ORF were amplified, respectively, removing the organelle targeting peptide sequence. The PCR products were cloned into pET30b vector (EMD Millipore, Billerica, MA), and the resulting constructs were named mtPPT-pET and mtACP1-pET, respectively. The mtHD-pET and mtACP1-pET constructs express recombinant enzymes containing a His-tag located at C-terminus.

**Plant transformation and selection**

Arabidopsis transformation was performed using the *Agrobacterium tumefaciens*-mediated infiltration method (37) to generate transgenic plants that express the GFP
fusion transgene and the \textit{mtPPT} RNAi transgene. Selection of the transformants was performed as described in Chapter 3.

\textbf{Confocal microscopy}

T2 generation of transgenic plants, which express the P\textit{35S::mtPPT-GFP} were used for the subcellular localization study by confocal microscopy at Confocal Microscope Facility, Iowa State University. Seven-day-old seedling grown on MS plates were harvested, stained using MitoTracker Orange CMTMRos (Life Technologies, Carlsbad, CA), and analyzed using a Leica TCS NT confocal microscope system as described in Chapter 3.

\textbf{The phosphopantetheinyl transferase assay}

The \textit{mtPPT-pET} construct and \textit{mtACP1-pET} construct were transformed into \textit{E. coli} BL21* strain (Life Technologies, Carlsbad, CA) for recombinant protein production. His-tagged \textit{mtPPT} and \textit{mtACP1} were purified using Probond Nickel-Chelating Resin (Life Technologies, Carlsbad, CA) following the manufacturer’s instruction. Molecular weight of the purified \textit{mtACP1} was measured by MALDI-QTOP analysis at Proteomics Facility, Iowa State University.

Phosphopantetheinyl transferase activity of the recombinant \textit{mtPPT} enzyme was measured by mixing 20\textmu{M} \textit{mtACP1}, 0.2\textmu{M} \textit{mtPPT}, 100\textmu{M} CoA, 10mM MgCl\textsubscript{2}, and BisTris-HCl (pH=6.5), and incubating at 37°C for 1h. The reaction mixture was then analyzed by MALDI-Q-TOF analysis.
Metabolomics analysis

Metabolites (i.e., amino acids, fatty acids, hydroxyfatty acids, sucrose, and glycolate) were extracted from 16-day-old Arabidopsis aerial organs and analyzed using HPLC or GC-MS as described in Chapter 2.

REFERENCES


FIGURES

Figure 1. Sequence alignment. (A) The protein sequence encoded by AT3G11470 and its orthologs. (B) The PPT domain of the protein sequence encoded by AT2G02770 and its orthologs.
Figure 2. Subcellular localization of the protein encoded by AT3G11470.
Figure 3. *In vitro* biochemical identification of the recombinant enzyme encoded by AT3G11470. (A) Sequence of mtACP1 (AT2G44620) and *E. coli* ACPP. (B) The negative control of the *In vitro* biochemical assay. (C) The *In vitro* biochemical assay.
Figure 4. Expression profiles of the Arabidopsis *mtPPT* gene in various organs.
Figure 5. Morphological appearances of the mtppt-rnai mutant alleles.
Figure 6. Metabolic profiles of the *mtppt-rnai* mutant alleles. (A) Alterations in amino acids in ambient air. (B) Alterations in amino acids in the elevated CO$_2$ atmosphere. (C) Alterations in glycolate. (D) Alterations in sucrose. (E) Profiles of fatty acids in ambient air. (F) Alterations in 3-hydroxymyristic acid in the elevated CO$_2$ atmosphere.
Figure 6 continued.
Figure 6 continued.
CHAPTER VI: MITOCHONDRIAL MALONYL-COA GENERATION BY MALONYL-COA SYNTHETASE IN ARABIDOPSIS

Manuscript in preparation and to be submitted to *Plant Physiology*

Xin Guan and Basil J. Nikolau

Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, Iowa 50011

**ABSTRACT**

Malonyl-CoA is the precursor for all of the known fatty acid forming systems. In plant plastids and cytosol, malonyl-CoA is generated via the carboxylation of acetyl-CoA, however, in mitochondria its origin is still elusive. In this study, we report the characterization of a mitochondrial malonyl-CoA synthetase (mtMCS), encoded by AT3G16170. This enzyme was previously characterized to be a cytosolic MCS. We demonstrate a novel MCS transcript containing a 5’-terminal extension, which encodes a putative mitochondrial transit peptide. A T-DNA-tagged mutant allele of AT3G16170 displays the severe morphological and metabolic phenotypes associated with other mutations in mitochondrial fatty acid synthase (mtFAS), namely arrested growth, hyperaccumulation of glycine, and disturbed fatty acid composition. As with other mutations in mtFAS, these alterations are confirmed to be a consequence of the deficiency in photorespiration, because they are reversed when the plants are grown in a non-photorespiratory condition (an atmosphere of elevated CO₂ levels). Our results demonstrate that the mtMCS enzyme is responsible for the formation of mitochondrial malonyl-CoA, the precursor of the mtFAS system, which is essential for plant growth and development.
INTRODUCTION

One of the metabolic fates of malonyl-CoA is as the extension substrate common to all fatty acid forming systems. Plants contain three distinct fatty acid forming systems, which occur in distinct subcellular compartments: plastids, the ER membrane, and mitochondria. In plastids, malonyl-CoA is generated via the ATP-dependent carboxylation of acetyl-CoA catalyzed by the heteromeric acetyl-CoA carboxylase (htACCase) (1). The plastidial htACCase is an enzyme complex that is comprised of four protein subunits, including biotin-carboxyl carrier protein, biotin carboxylase, and alpha and beta subunits of carboxyl transferase (1). The malonyl group is then transferred from CoA to ACP, leading to the formation of malonyl-ACP, the donor of 2-carbon elongation unit for plastidial fatty acid biosynthesis. In the cytosol, malonyl-CoA is produced in a similar carboxylation reaction of acetyl-CoA, but catalyzed by the homomeric acetyl-CoA carboxylase (hmAACase), which is a multifunctional enzyme that possesses all of the reaction centers to catalyze the carboxylation reaction (1). Cytosolic malonyl-CoA serves directly as the extender in the fatty acid elongation process (1).

Discrepancy has been found in eukaryotic cells, regarding the biochemical biogenesis of the mitochondrial malonyl-CoA. In yeast, a mitochondrial hmACCCase, Hfa1p, has been characterized to generate malonyl-CoA in a carboxylation reaction, which mimics the plant cytosolic hmACCase enzyme (2). By contrast, human ACSF3, a mitochondrial malonyl-CoA synthetase (mtMCS), catalyzes the formation of malonyl-CoA from free malonate and CoA (3). In plants, even though hmACCases have been identified in mitochondria of grasses, it appears to be absent from Arabidopsis mitochondria (4).
Despite the fact that the source of mitochondrial malonyl-CoA is unclear, it is proposed to support the generation of mitochondrial malonyl-ACP, which serves as the initiation building block and 2-carbon donor for the iterative acyl chain elongation process of fatty acid biosynthesis. Specifically, it is proposed that the initiation reaction of mtFAS is the condensation reaction between two molecules of malonyl-ACP catalyzed by mitochondrial β-ketoacyl-ACP synthase (mtKAS), driving the formation of acetoacetyl-ACP. The subsequent extension reactions of mtFAS are a condensation of malonyl-ACP with an existing acyl-ACP precursor, again catalyzed by mtKAS, resulting in the production of β-ketoacyl-ACP (5). The acetoacetyl-ACP and β-ketoacyl-ACP are then sequentially acted upon by mitochondrial β-ketoacyl-ACP reductase, mitochondrial 3-hydroxyacyl-ACP dehydratase (mtHD), and mitochondrial enoyl-ACP reductase (mtER), leading to the formation of an acyl-ACP containing an elongated acyl chain. All of the mtFAS components have been biochemically identified from yeast and human (6).

In plants, three mtFAS components of the acyl chain elongation cycle (mtKAS, mtHD, and mtER), and mitochondrial ACP (mtACP) components and mitochondrial phosphopantetheinyl transferase (mtPPT) that activates mtACP, have been characterized from Arabidopsis at both the biochemical and physiological levels (5,7) (Chapter 2, 3, 4, and 5).

The primary role of the mtFAS system appears to be the generation of octanoyl-ACP, which is used as the substrate for the biosynthesis of lipoic acid (8). Lipoic acid is the cofactor that is essential for pyruvate dehydrogenase (PDH), β-ketoglutarate dehydrogenase (KGDH), and glycine decarboxylase (GDC) (9). In Arabidopsis, knockout of the *mtKAS* gene and knockdown of the *mtHD* and *mtPPT* genes result in the reduction
in the lipoylation of the H-subunit of GDC, an essential enzyme in the photorespiration process, leading to the deficiency in photorespiration (7) (Chapter 2, 3, and 5). However, the lipoylation of PHD and KGDH is only slightly affected by the mutations in the mtFAS system (7), which may be indicative of another lipoic acid generating system in plants, other than mitochondria (10,11). In addition to the generation of medium-chain fatty acid, plant mtFAS system may also support the biosynthesis of longer chain fatty acids, because all of the Arabidopsis mtFAS components that have been characterized in vitro accept substrates containing acyl moieties of up to 16 carbon atoms (Chapter 2, 3, and 4). Indeed, the Arabidopsis mtFAS system has been characterized to support the biosynthesis of fatty acids that are used in the assembly of lipid A-like molecules, by providing the 3-hydroxymyristyl-ACP substrate (Chapter 2). In addition, it has been proposed that the long-chain acyl-ACP molecules produced by the mtFAS system are required for the remodeling of cardiolipins in plants (12,13).

Here, we demonstrate the identification and characterization of the Arabidopsis mtMCS, the enzyme that generates mitochondrial malonyl-CoA. We show that a novel 5’ extension coding sequence for this gene, as compared to the previously cloned protein coding sequence (14), has been cloned, and that this sequence encodes a putative mitochondrial transit peptide. In addition, loss-of-function mutant studies show that mutations in the mtMCS gene lead to the typical morphological and metabolic phenotypes associated with the deficiency in the mtFAS system. Our findings demonstrate the biochemical biogenesis of mitochondrial malonyl-CoA, the precursor molecule of the mtFAS system.
RESULTS

1. Cloning of an alternative MCS transcript

The protein encoded by AT3G16170 has been characterized as a cytosolic malonyl-CoA synthetase (MCS) (14). However, the human homolog of MCS is encoded by the ACSF3 gene that codes for a mitochondrial MCS (3); BLAST analysis indicates that the two proteins are highly homologous (total score 377, query cover 91%, and E value 7e-122). Upon examining the genome sequence of the previously identified MCS start codon (14), we detected a putative exon sequence upstream of the currently annotated MCS gene that would add an N-terminal extension to the MCS protein. Using a primer flanking the putative upstream start codon and a primer flanking the stop codon of the MCS ORF, RT-PCR was conducted on the RNA template isolated from 16-day-old aerial organs. Cloning and sequencing of this product lead to the identification of an alternative MCS transcript with a 5’ extension relative to the previously annotated MCS gene. The sequence of this newly identified transcript (MCS-2) (Fig 1A) shows that an intron is spliced out to form the extra 5’ coding sequence (Fig 1B). The protein encoded by the MCS-2 transcript contains 64 amino acid residues extended at the N-terminus relative to the MCS protein (Fig 1C).

This protein sequence extension is rich in basic amino acids and lack acidic residues, features that are characteristics of mitochondrial targeting sequence elements. Indeed, the protein encoded by the MCS-2 transcript is predicted to be a mitochondrially localized protein by MitoProt II (36 N-terminal residues with a score of 0.9841) (15) and PSORT (with a score of 0.581) (16), however, Target P predicts to be localized in the plastids (85 N-terminal residues with a score of 0.732) (17,18). To evaluate the organelle
targeting capability of the newly identified N-terminal extension of the MCS-2 protein, we have initiated an Arabidopsis transgene experiment, in which the 5’ coding sequence of the *MCS*-2 transcript is cloned in frame with the GFP-coding sequence and stably transformed into Arabidopsis under the control of 35S promoter. These transgenic plants will be investigated by confocal microscopy.

2. **Complementation of the morphological alterations associated with the mcs-1 allele**

Arabidopsis *mcs-1* mutant allele, which has T-DNA resided in the 13th intron of the *MCS* gene (Fig 2A), leads to a compromised growth, however, the biochemical reason has yet to be discovered (14). If the MCS-2 protein is mitochondrially localized, it would generate malonyl-CoA as the substrate for the mtFAS system. Therefore, mutations in the *MCS* gene would lead to the deficiency in the mtFAS system, which would further compromise the lipoylation of the H-subunit of glycine decarboxylase complex, leading to a deficiency in photorespiration.

To test if the morphological phenotype is a consequence of the deficiency in photorespiration, *mcs-1* mutant allele was grown in both ambient air and an atmosphere enriched in CO2. In the ambient atmosphere under constant illumination, the growth of *mcs-1* mutant is arrested at the cotyledon stage (Fig 2B). In addition, the cotyledons exhibit a yellowish leaf color. By contrast, when the plants are grown in an elevated CO2 atmosphere, visually no morphological difference could be detected in the mutant (Fig 2B). The compromised growth in ambient air and normal growth in the elevated CO2 atmosphere are a typical consequence of the deficiency in photorespiration (19); these morphological phenotypes of *mcs-1* also resemble those of the mutants with deficiency in
the mtFAS system, such as mtkas alleles (7) (Chapter 2), mthd-rnai alleles (Chapter 3), and mtppt-rnai alleles (Chapter 5).

To further confirm the phenotypic observations made with the mcs-1 mutant allele, we identified a second T-DNA-tagged mutant allele from the SALK Insertion Sequence Database (http://signal.salk.edu) (20) (SALK_138017), designated as mcs-2 (Fig 2A). Plant genomic DNA sequences flanking the T-DNA borders were amplified and sequenced, and the molecular structure of the mutant allele was deduced by aligning these flanking DNA sequences with the Arabidopsis genome sequence. The insertion in the mcs-2 allele carries a T-DNA insert in the 4th exon of the MCS gene (Fig 2A). We have been segregating the homozygous mcs-2 seeds for the further study.

3. Alterations in amino acid levels in the mcs-1 allele

Amino acid profiling was performed on the aerial organs of mcs-1 mutant and compared to the wild-type plants. The most dramatic alteration is a 258-fold increase in Gly levels in the mutant (Fig 3A). This is in agreement with the previous finding that mutations in the mtFAS system diminish the lipoylation of the H-protein subunit of glycine decarboxylase complex, leading to an inability to metabolize photorespiratory Gly (7). Several other amino acids exhibit significant changes (Fig 3A), which are likely to be a consequence of the defective nitrogen metabolism in the mutants.

The alterations in all of the amino acids, with the exception of Gly, Ala, and Glu, return to the normal level when the plants are grown in an elevated CO₂ atmosphere (Fig 3B), revealing that their alterations are a result of the deficiency in photorespiration. Even though the hyperaccumulation of Gly levels are greatly reduced when photorespiration is inhibited in the elevated CO₂ atmosphere, Gly levels are still about 29-fold higher in the
mutant, indicating that photorespiration is still occurring, but at a lower rate in plants grown in the non-photorespiratory conditions. For this reason, significantly altered levels of Ala and Glu seem to be a result of the altered nitrogen metabolism in the mutants.

4. Alterations in fatty acid compositions in the mcs-1 allele

Fatty acid composition was analyzed on the 16-day-old aerial organs of mcs-1 mutant and the wild-type control. The mcs-1 mutant displays markedly altered fatty acid composition, namely elevated levels of 16:0, 18:1, and 18:2, and reduced levels of 16:3 and 18:3 (Fig 4A). These alterations in the fatty acid composition are reversed when the plants are grown in an atmosphere enriched in CO₂ (Fig 4B), demonstrating that the fatty acid alterations are a result of the deficiency in photorespiration. The decreased level of 16:3 is in agreement with the reduction in the level of this fatty acid in other mutants with the deficiency in the mtFAS system, such as mtkas alleles (Chapter 2) and mthd alleles (Chapter 3). The reduced level of 16:3 indicates that the mcs-1 mutant allele may exhibit a defective thylakoid membrane system, since 16:3 is unique to this membrane system of chloroplasts (21).

DISCUSSION

Arabidopsis mtFAS system

Plant cells have at least three fatty acid forming systems, which occur in distinct subcellular compartments: plastids, cytosol, and mitochondria. All of them are organized as Type II fatty acid synthase (FAS) or fatty acid elongase, in which four individual enzymes are utilized to catalyze the iterative reactions of fatty acid biosynthesis.

All of the mtFAS components have been biochemically identified from yeast and humans (6). In plants, such as Arabidopsis, many mtFAS components have been
biochemically characterized, including the ACP components (22,23), the enzymes (e.g., mtKAS (5) (Chapter 2), mtHD (Chapter 3), and mtER (Chapter 4)) that catalyze 3 of the 4 reactions of the acyl chain elongation cycle, and mtPPT that activates the ACP components (Chapter 5). Based on the loss-of-function genetic investigations on those known mtFAS components, the role of the mtFAS system in providing octanoyl-ACP, the precursor of lipoic acid production, has been established (7). Accordingly, mutations in the Arabidopsis mtFAS system display a typical photorespiration-deficiency and glycine over-accumulation, due to the loss of lipoylation of the H-subunit of glycine decarboxylase complex, an essential enzyme in the photorespiration process (7) (Chapter 2, 3, and 5). In addition, a novel metabolic destination of the mtFAS products has been discovered that it supplies 3-hydroxymyristyl-ACP for the synthesis of lipid A-like molecules (Chapter 2).

**Biosynthesis of mitochondrial malonyl-CoA by MCS**

Malonyl-CoA is the precursor to produce malonyl-ACP, which serves as the donor of 2-carbon unit for the acyl chain elongation cycle of the mtFAS system. The canonical mode of malonyl-CoA generation in plastids and the cytosol is via the ATP-dependent carboxylation of acetyl-CoA catalyzed by acetyl-CoA carboxylase (ACC). The plastidial ACC is an enzyme complex that is comprised of biotin-carboxyl carrier protein (BCCP), biotin carboxylase (BC), and carboxyl transferase (CT), whereas the cytosolic ACC is a multifunctional enzyme that possesses all of the reaction centers to catalyze the ACC reactions (1). In yeast mitochondria, malonyl-CoA is also generated by an ACC enzyme, which is encoded by *HFA1* and is homologous to plant cytosolic AAC (2). By contrast, human mitochondria appear to produce malonyl-CoA via an alternative route, in
which a mitochondrially localized malonyl-CoA synthetase (MCS) encoded by the 
ACSF3 gene catalyzes the formation of malonyl-CoA from free malonate and CoA (3). In 
plant, a mitochondrial ACC protein has been identified in mitochondria of grasses, 
however, it appears to be absent from the mitochondria of Arabidopsis (4). Considering 
that plant mitochondria are known to have malonyl-CoA synthetase activity (24), 
Arabidopsis is likely to generate mitochondrial malonyl-CoA from malonate by MCS 
enzyme.

This study provides the evidence that Arabidopsis MCS, which is previously 
characterized to be cytosolically located, is responsible for the mitochondrial MCS 
reaction. Specifically, the severely compromised growth of mcs-1 mutant allele in the 
ambient air (14) is reversed when the plants are grown in an elevated CO₂ atmosphere, 
where the photorespiration process is inhibited. In addition, dramatic accumulation in 
glycine, a hallmark of the deficiency in the mtFAS system, is detected when the mutant is 
grown in the ambient air, and the glycine elevation is greatly reduced when the mutant is 
treated with enriched CO₂. Furthermore, mcs-1 allele affects fatty acid composition in 
ambient air, specifically, the marked reduction in 16:3 resembles those observed in other 
mtFAS mutants (Chapter 2 and Chapter 3). Finally, a 5’ extension of the MCS coding 
sequence is cloned from Arabidopsis mRNA; an Arabidopsis transgene experiment has 
been initiated to directly test if the extension sequence encodes a mitochondrial transit 
peptide.

It is noteworthy that mcs-1 mutant is almost lethal under constant illumination and 
that its growth is arrested after the expansion of cotyledons. In addition, the glycine 
accumulation is stronger in mcs-1 allele than that in the mtkas alleles under the same
growth conditions (Chapter 2). The stronger morphological and metabolic alterations associated with the mcs-1 allele are in agreement with the observations of the mtkas-2/mthd-rnai double mutants (Chapter 3). It suggests that the lipoylation of the H-subunit of glycine decarboxylase is likely to be exclusively dependent on the mtFAS system and, thus, the remaining lipoylation of the H-subunit of glycine decarboxylase in the mtkas alleles (7) is probably due to the residual mitochondrial β-ketoacyl-ACP synthase activity.

The origin of malonate in Arabidopsis

Malonate is found in plants, especially in roots and nodules of legumes (25), however, its origin is still unclear. Several biochemical processes could contribute the malonate generation; these processes include the oxidative decarboxylation of oxaloacetate, the degradation of pyrimidines, and dehydrogenation of malondialdehyde (14). In this respect, mitochondrial malonate could be recruited from endogenous sources or taken up by mitochondria via the yet unknown transporters.

METHODS

Plant materials and growth conditions

Arabidopsis thaliana ecotype Columbia (Col-0) and two T-DNA insertion mutants (mcs-1 (SALK_083785) and mcs-2 (SALK_138017)) were obtained from the Arabidopsis Biological Resource Center (Columbus, OH). Seeds were sterilized and sown on Murashige and Skoog agar medium as describe in Chapter 2. After seeds were sown, plates were placed at 4°C for 4 days to break seed dormancy, and plants were grown for 16 days in ambient air or the non-photorespiratory condition (1% CO2 in a
growth chamber), where the temperature was maintained at 23 ± 2°C with continuous illumination (100 μmol m^{-2} s\(^{-1}\)).

**Cloning of an alternative MCS transcript**

RNA was extracted from 50 mg fresh leaf tissues using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA). A 2μg aliquot of RNA was used for reverse transcription using the SuperScript III First-Strand Synthesis Super Mix (Life Technologies, Carlsbad, CA) with oligo (dT)\(_{20}\) primer. Primers MS1 (CACCATGACCGCTACGACAACATTAAAG) - MS2 (TTATTCTTGATTTTCCAGAGATTTCTTTAGC) were used to clone a ORF encoding Arabidopsis MCS with a 5’ extension coding sequence from the cDNA mix.

**Metabolomics analysis**

Amino acids and fatty acids were extracted from 16-day-old Arabidopsis aerial organs and analyzed using HPLC and GC-MS, respectively, as described in Chapter 2.

**REFERENCES**


FIGURES

Figure 1. An alternative MCS coding sequence. (A) Sequence of the MCS-2 transcript. (B) Genome structure of the MCS transcript and MCS-2 transcript. (C) Sequence of the protein encoded by the MCS-2 transcript.
Figure 1 continued.

C

<table>
<thead>
<tr>
<th>1</th>
<th>MTATTLKSF NYLSLHRL NNHSYAILSS PLLPR3HPS A3SFNSGFRF</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td>FQSNHLFSQ SGSLMEVFKA AFSEASNSCD RIAIKADGKS YSYGQLTSSA</td>
</tr>
<tr>
<td>101</td>
<td>LRISKLFLKD DIINGQEQIK KYEGFGSLKG ARIGIVKPS AEFVAGVLGT</td>
</tr>
<tr>
<td>151</td>
<td>WFSGGVAVEL ALSYPEAELL HMVNDSDL SLSTEDHSET MKTIAAKSGA</td>
</tr>
<tr>
<td>201</td>
<td>RFHLIPFVVN STSETVACQ FQD8SFEAEG KFLDDPALIV YTSGTTGKFP</td>
</tr>
<tr>
<td>251</td>
<td>GVVHTHNSIN SQVRLMLTEAW EYTSAHDHFLH CLPLHVHGL FNALFAPLYA</td>
</tr>
<tr>
<td>301</td>
<td>RSLVEFPLPKF SVSGLRWR EYSPVNEKTI NDSIIVFGV PMYTRLIQG</td>
</tr>
<tr>
<td>351</td>
<td>YEAMDKREMQT SSASARLKR LMSGSSALP RPVMRQGESI TGHRLLERYG</td>
</tr>
<tr>
<td>401</td>
<td>MTFVMAMSIN PLRGNAGT VGKLPGLVEA KIKEDENDAN GVGEICVKS</td>
</tr>
<tr>
<td>451</td>
<td>SLFKEYWNLP ETVKESFTED GYKGDAGR VDEDGYYVIL GRNASDMMK</td>
</tr>
<tr>
<td>501</td>
<td>GGYKLGSAME ETELHFTPV AECCVSLTID NDYGEAATIA AIASEAAKRR</td>
</tr>
<tr>
<td>551</td>
<td>RDEESNPVIL LEELCGWAKD KLAPYKLPIR LLIWESLPRN AMGKVNNKEL</td>
</tr>
<tr>
<td>601</td>
<td>KKSLENQ*</td>
</tr>
</tbody>
</table>
Figure 2. Morphological phenotypes. (A) Genome structures of two mcs mutant alleles. (B) Morphological alterations of the mcs-1 mutant allele.
Figure 3. Amino acid analysis. (A) Alterations in amino acids in ambient air. (B) Alterations in amino acids in the elevated CO$_2$ atmosphere.
Figure 4. Fatty acid analysis. (A) Alterations in fatty acids in ambient air. (B) Alterations in fatty acids in the elevated CO\textsubscript{2} atmosphere.
CHAPTER VII: GENERAL CONCLUSIONS

In this dissertation, five components of Arabidopsis mitochondrial fatty acid synthase (mtFAS) have been characterized. These enzymes include mitochondrial β-ketoacyl-ACP synthase (mtKAS), mitochondrial 3-hydroxyacyl-ACP dehydratase (mtHD), mitochondrial enoyl-ACP reductase (mtER), mitochondrial phosphopantetheinyl transferase (mtPPT), and mitochondrial malonyl-CoA synthetase (mtMCS). MtKAS, mtHD, and mtER catalyze 3 of the 4 reactions of the fatty acyl chain elongation cycle; mtPPT supports the activation of apo-mtACP isoforms; and mtMCS generates malonyl-CoA, the precursor for the acyl chain elongation process. These dissociated, monofunctional enzymes support fatty acid biosynthesis, which mimics the canonical Type II FAS in bacteria and plant plastids (1-3). The biochemical and physiological characterization of these enzymes demonstrate the occurrence of the mtFAS system in Arabidopsis and its functional significance to plant growth and development.

Evolutionary insights into the Arabidopsis mtFAS system

From the evolutionary point of view, mtFAS may have been introduced into plant mitochondria along with endosymbiotic bacteria during evolution. For this reason, sequence homology still exists between some mtFAS components (mtKAS and mtPPT) and their bacterial counterparts. Specifically, mtKAS is highly homologous to bacterial KAS I and KAS II, and is moderately homologous to KAS III (4) (Chapter 2), while mtPPT displays sequence similarity with the bacterial Group II phosphopantetheinyl transferase (PPTase) (Chapter 5). However, sequence homology is absent between the Arabidopsis enzymes (e.g., mtHD and mtER) and their bacterial counterparts (Chapter 3 and 4), suggesting the divergent evolution of the fatty acid biosynthesis in plant
mitochondria and bacteria. In addition, Arabidopsis mtMCS catalyzes a distinct reaction for the biogenesis of malonyl-CoA (Chapter 6), a reaction that is in contrast to the classical carboxylation of acetyl-CoA, revealing an evolutionary process that is unique to plant mitochondria.

Within the eukaryotic mtFAS systems, Arabidopsis mtFAS is likely to be distinct from that in yeast (e.g., *Saccharomyces cerevisiae*). This is due to the observation that sequence homology is not found between the Arabidopsis mtFAS components (e.g., mtHD and mtPPT) and their yeast counterparts (Chapter 3 and 5). Furthermore, the biochemical mechanism that generates mitochondrial malonyl-CoA in Arabidopsis mitochondria differs from that occurs in yeast mitochondria, which appears to resemble the bacterial mechanism (Chapter 6). In comparison, Arabidopsis mtFAS system is organized in the same biochemical mode as that in mammals (e.g., *Homo sapiens*) (Chapter 2-6). Specifically, all of the 5 Arabidopsis mtFAS components share homology to their counterparts in humans; in addition, the Arabidopsis enzymes and their human orthologs exhibit the similar in vitro enzymatic properties (Chapter 2-6). The consistency and discrepancy among mtFAS systems in Arabidopsis, mammals, and yeast demonstrate that the genetic incorporation of mtFAS in eukaryotic mitochondria may occur at an early stage of evolution, and after that, mtFAS endure divergent evolution. For this reason, the sequence properties of Arabidopsis mtFAS components are more closely related to those of the higher eukaryotes (e.g., mammals), which diverged from plants at a relatively late evolutionary stage, as compared to lower eukaryotes (e.g., yeast).
Two missing mtFAS components in Arabidopsis

All of the Arabidopsis mtFAS enzymes have been identified, with the exception of mitochondrial β-ketoacyl-ACP reductase (mtKR) and mitochondrial malonyl-CoA ACP transacylase (mtMCAT).

MtKR is responsible for the 2nd step of the acyl chain elongation process, the reduction of β-ketoacyl-ACP to 3-hydroxyacyl-ACP. In bacteria, fabG catalyzes this reaction and utilizes NADPH or NADH as the cofactors (5). Oar1 is the yeast mitochondrial ortholog of bacterial fabG, and it is NADPH-dependent (6). Humans appears to have two mtKR homologs (i.e., 17β-HSD8 and CBR4), which shares sequence similarity with bacterial fabG and yeast Oar1; the two proteins are organized in the form of heterotetramer, which supports the β-ketoacyl-ACP reductase reaction in a NADH-dependent manner (7). Crystal structures of *E. coli* fabG reveal a classical Rossmann dinucleotide binding fold (8,9), which is also likely to be possessed by yeast Oar1 and human mtKR. When *E. coli* fabG, yeast Oar1, and human mtKR are used as queries, BLAST searches of Arabidopsis protein database suggest AT1G24360 to be the most significant homolog. However, based on the gene expression analysis using GeneCAT (10), the expression of AT1G24360 is highly correlated with that of AT2G05990, the gene encodes a plastidial form of enoyl-ACP reductase (11), revealing that AT1G24360 may encode a plastidial β-ketoacyl-ACP reductase (data not shwon). But we can not exclude the possibility for the existence of a yet undetected transcript variant of AT1G24360, which encodes a mitochondrially localized enzyme. In addition, BLAST searches suggest many other mtKR candidates with lower scores; these candidates are annotated as “NAD(P)-binding Rossmann-fold superfamily protein”, which contain a
structural moiety that is common to that in *E. coli* fabG (data not shown). These hypotheses need to be experimentally tested in the future work.

MtMCAT transfers a malonyl moiety from CoA to the thiol group of the phosphopantetheinyl cofactor of holo-ACP, resulting in the formation of malonyl-ACP, the donor of 2-carbon units for the acyl chain elongation cycle in the mtFAS system. This enzyme was first characterized from *E. coli* as fabD (12,13). Based on the sequence similarity to *E. coli* fabD, the mtMCAT components were identified in yeast and humans, respectively (6,14). Using the known sequences of fabD and mtMCAT as queries, BLAST searches of Arabidopsis protein database reveal a single candidate (AT2G30200). However, in silico expression analysis using GeneCAT shows that the expression of AT2G30200 is highly correlated with AT2G05990, the gene encodes a plastidial enoyl-ACP reductase of fatty acid biosynthesis (11), illustrating that AT2G05990 may encode a plastidial form of malonyl-CoA ACP transacylase (data not shown). However, AT2G30200 could yield a yet undetected transcript splice variant, which encodes a mitochondrially located mtMCAT in Arabidopsis.

**Novel physiological functions of the Arabidopsis mtFAS system**

In yeast and humans, the only known metabolic destination of mtFAS is that it provides octanoyl-ACP for the synthesis of lipoic acid. The physiological functions of mtFAS is best described in yeast, where the deletion of any member of this pathway leads to devoid lipoic acid, deficiency in respiration, depletion in cytochrome, and deficiency in mitochondrial RNA processing, demonstrating that the mtFAS system is essential for yeast mitochondrial functions (3). In humans, even though the physiological significance
of the mtFAS system is poorly understood, its function in contributing lipoic acid biosynthesis has been established (15,16).

In Arabidopsis, the mtFAS system also provides octanoyl-ACP for the generation of lipoic acid (17-19). Unlike yeast and humans, where lipoic acid biosynthesis is exclusively dependent on mtFAS, Arabidopsis lipoic acid from the mtFAS system is likely to be partially redundant for the lipoylation of E2 subunit of pyruvate dehydrogenase and E2 subunit of β-ketoglutarate dehydrogenase (20). However, lipoylation of H-subunit of glycine decarboxylase complex, an essential enzyme in the photorespiration process, is likely to be exclusively dependent on the mtFAS system (Chapter 2, 3, and 6). Therefore, mutations in mtFAS lead to a typical deficiency in photorespiration (Chapter 2, 3, and 6). In addition, defective metabolism of acyl-lipids (e.g., surface lipids, galactolipids, sulfolipids, and triacylglycerol) is characterized to be associated with mutations in mtFAS, as a secondary effect of the deficiency in photorespiration (Chapter 2 and 3). Moreover, the Arabidopsis mtFAS system supports the biosynthesis of lipid A-like molecules by providing 3-hydroxymyristyl-ACP as the substrate; this is a second metabolic destination that has been discovered for the mitochondria-derived fatty acids in eukaryotes (Chapter 2).

It is noteworthy that all of these impaired biological processes are unique to plant processes. Specifically, photorespiration is initiated by the oxygenase activity of RuBisCO in the photosynthesis process, and is responsible for the fixation and recovery of the photorespiratory carbon (21). Surface lipids form the barrier between plants and abiotic environments (22). Galactolipids and sulfolipids make up the monolayer membrane system of the thylakoid photosynthetic center (23). Plastidial triacylglycerol is
associated with plant senescence processes (24). Finally, lipid A-like molecules are proposed to be involved in chloroplast functions, signal transduction, and plant defense responses (25). Alterations in these biological processes demonstrate the plant-specific physiological functionalities of the Arabidopsis mtFAS system.

**Implications in the production of fatty acid-derived fuels and chemicals**

Fatty acid biosynthesis is a pathway that has been being engineered for the production of fatty acid-based fuels and chemicals (26). Based on the unique biochemical properties and compartmentalization, mtFAS in eukaryotes provides potentials for the metabolic engineering of novel metabolic processes.

For instance, acyl carrier protein (ACP) is the carrier of acyl intermediates in fatty acid biosynthesis. Current efforts in overexpressing ACP in *E. coli* do not lead to a significant improvement to the yield of fatty acid-derived products, because of the inhibited *E. coli* lipid metabolism and subsequently compromised cell growth introduced by apo-ACP (27). However, cell growth appears to be normal when Arabidopsis mitochondrial ACP (mtACP) isoforms are overexpressed in *E. coli* (Chapter 5), indicating the compatibility between *E. coli* cell growth and the overexpression of the Arabidopsis mtACP components. In addition, Arabidopsis mtACP components display the preference for the biosynthesis of medium-chain fatty acids, rather than long-chain fatty acids (28). This acyl-chain preference may help alter the chain-length distribution of fatty acids and their derivatives that an ACP-dependent fatty acid biosynthetic engineering system can produce.

Another example is the human mtKR isoforms, which catalyzed a β-ketoacyl-ACP reductase reaction, which utilizes NADH as the cofactor, other than NADPH (7).
Considering that NADH is much more abundant than NADPH in \textit{E. coli}, these novel mtKR isoforms display the potential to improve the efficiency of fatty acid biosynthesis. Indeed, the application of a bacterial $\beta$-ketoacyl-ACP reductase, which prefers NADH over NADPH, has shown an increased yield in the production of fatty acids and their derivatives in \textit{E. coli} (5).

Finally, the compartmentalization of the mtFAS system sheds lights on a potential to recruit yeast mitochondria as bioreactors that produce fatty acid-derived chemicals. Compartmentalization in mitochondria has many benefits, namely greater local enzyme concentrations and increased availability of fundamental building blocks (e.g., NADH and acetyl-CoA). Indeed, recent efforts of the metabolic engineering of yeast mitochondria have been proven to be successful in the production of plant terpenoids and branched chain alcohols (29,30). Regarding the fact that several mtFAS components (e.g., mtKR, mtHD, and mtER) recognize both acyl-ACP and acyl-CoA as substrates for fatty acid biosynthesis (3) (Chapter 3 and 4), novel synthetic pathways could be reconstituted in yeast mitochondria employing those mtFAS components and a variety of condensation enzymes and thioesterases in CoA- or ACP-dependent manners (26,31).

\textbf{REFERENCES}


