Subcellular localization and protein-protein interactions of two methyl recycling enzymes from *Arabidopsis thaliana*

by

Sanghyun Lee

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Biology

Waterloo, Ontario, Canada, 2010

AUTHOR' S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

ABSTRACT

This thesis documents the subcellular localization and protein-protein interactions of two methyl recycling enzymes. These two enzymes, adenosine kinase (ADK) and *S*-adenosyl-L-homocysteine hydrolase (SAHH), are essential to sustain the hundreds of *S*-adenosyl-L-methionine (SAM)-dependent transmethylation reactions in plants. Both ADK and SAHH are involved in the removal of a competitive inhibitor of methyltransferases (MTs), *S*-adenosyl-L-homocysteine (SAH), that is generated as a by-product of the each transfer of a methyl group from SAM to a substrate. This research focused on understanding how SAH is metabolized in distinct cellular compartments to maintain MT activities required for plant growth and development.

Localization studies using green fluorescent protein (GFP) fusions revealed that both ADK and SAHH localize to the cytoplasm and the nucleus, and possibly to the chloroplast, despite the fact that the primary amino acid sequence of neither protein contains detectable targeting signals. This suggested the possibility that these methyl-recycling enzymes may be targeted by specific protein-protein interactions. Moreover, deletion analysis of SAHH1 indicated that the insertion region (IR) of 41 amino acids (Gly^{150}-Lys^{190}), which is present only in plants and parasitic protozoan SAHHs among eukaryotes, is essential for nuclear targeting. This result suggested that the surface-exposed IR loop may serve as a binding domain for interactions with other proteins that may direct SAHH to the nucleus.

To investigate protein-protein interactions, several methods were performed including co-immunoprecipitation, bimolecular fluorescence complementation, and pull-down

assays. These results not only revealed that ADK and SAHH possibly interact through the IR loop of SAHH *in planta*, but also suggested that this interaction is either dynamic or indirect, requiring a cofactor/another protein(s) or post-translational modifications. Moreover, possible interactions of both ADK and SAHH with a putative Arabidopsis mRNA cap methyltransferase (CMT), which is localized predominantly in the nucleus, were also confirmed. These results support the hypothesis that the nuclear targeting of both SAHH and ADK can be mediated by the interaction with CMT. In addition, purification of Strep-tagged SAHH1 expressed in Arabidopsis identified a novel interaction between SAHH and aspartate-semialdehyde dehydrogenase (ASDH), an enzyme that catalyzes the second step of the aspartate-derived amino acid biosynthesis pathway. Analysis of ASDH-GFP fusions revealed that ASDH localizes to the chloroplast and the stromule-like structure that emanates from chloroplasts. Moreover the mutation in three amino acids (Pro^{164}-Asp^{165}-Pro^{166}) located within the IR loop of SAHH disrupted its binding to ASDH which affected the plastid localization of SAHH, suggesting that the interaction between SAHH and ASDH is required for plastid-targeting of SAHH.

Taken together, this thesis demonstrated that the localization of ADK and SAHH in or between compartments is possibly mediated by specific protein interactions, and that the surface-exposed IR loop of SAHH is crucial for these interactions.

ACKNOWLEDGEMENTS

There are so many people who have helped me and have made my time at the UW so enjoyable and productive. The work presented in this thesis could not have been accomplished without their help and support.

First and foremost, I would like to thank my supervisor Dr. Barb Moffatt for her guidance, support, encouragement, and infinite patience during the past five years. I would also like to express my sincere appreciation to my committee Dr. Susan Lolle and Dr. Simon Chuong for sharing their valuable expertise and advice. Also thank to my external committee members Dr. Susanne Kohalmi and Dr. Marc Aucoin for their critical reading and comments on my thesis.

I am grateful to the faculty, staff, and students of the Department of Biology for the excellent support and advice. A special thanks to Linda, Dale, Lynn, Karen, Andrew, Zhenyu, Dr. Kim, Dr. B. Duncker and his lab members, not only for all kinds of technical input and assistance, but also for their kindness and friendship.

I would also like to thank all of the members of the Moffatt and Simon's Lab: Tony, Katja, Sarah, Yong, Ishari, Makoto, Terry, Jen, Dr. Kondo, Malay, Rick and many other past and present members. I have been very lucky to meet and get along with you; working with you was a wonderful experience that I will never forget.

Lastly, I would like to thank my parents for their never-ending love and support throughout my entire life. And special thanks to my wife Yoon and my kids, Daniel and Amie, for their support, patience, and understanding.

TABLE OF CONTENTS

AUTHOR' S DECLARATION ..... **ii**

ABSTRACT ..... **iii**

ACKNOWLEDGEMENTS ..... **v**

TABLE OF CONTENTS ..... **vi**

LIST OF FIGURES ..... **x**

LIST OF TABLES ..... **xii**

LIST OF ABBREVIATIONS ..... **xiii**

INTRODUCTION **1**

1. Arabidopsis as a model system **1**

2. Activated methyl cycle **5**

2.1. Methylation 5

2.2. SAM-dependent transmethylation 6

2.3. Adenosine salvage pathways 11

3. Adenosine Kinase and *S*-adenosyl-L-Homocysteine Hydrolase **13**

3.1. ADK 13

3.2. ADK in Arabidopsis 15

3.3. SAHH 17

3.4. Dynamic protein structure of SAHH 18

3.5. Insertion region in SAHH 22

3.6. SAHH in plants 26

3.7. ADK and SAHH in cytokinin metabolism 29

3.8. Subcellular localization of ADK and SAHH 31

4. Mechanism of protein transport **35**

4.1. Nuclear targeting 35

4.2. Chloroplast targeting 39

5. Useful tools for studying plant sciences **44**

5.1. Fluorescent proteins 44

5.2. Reverse genetics 45
5.3. Protein-protein interactions 50
5.3.1. Yeast two-hybrid 52
5.3.2. Pull-down assays 53
5.3.3. Bimolecular fluorescence complementation 56

6. Objectives of this research 59

CHAPTER 1: SUBCELLULAR LOCALIZATION OF METHYL RECYCLING ENZYMES 60

1. Introduction 60
2. Materials and Methods 64
2.1. Chemicals and reagents 64
2.2. Plant material and growth conditions 64
2.3. Vector constructions for ADK1 and SAHH1 localization 65
2.4. Cloning of yeast *SAHH* and control vectors for plastid and nuclear localization 67
2.5. Deletion analysis for SAHH1 localization 69
2.6. Arabidopsis transformation 71
2.7. Transient expressions in tobacco plants 72
2.8. Protoplast isolation and transfection 73
2.9. Immunoblots 74
2.10. Inverse PCR 75
2.11. Confocal microscopy 76
3. Results 79
3.1. Subcellular localization of ADK and SAHH 79
3.2. Nuclear localization of ADK1 and SAHH1 86
3.3. Deletion analyses of ADK1 and SAHH1 89
3.4. The IR affects nuclear localization of SAHH1 95
3.5. Chloroplast localization of ADK1 and SAHH1 104
3.6. Plants arose from the overexpression of GFP-fusion constructs 113
4. Discussion 117
4.1. Determination of subcellular location using fluorescent proteins 117

4.2. Methyl recycling enzymes localize to multiple compartments 119
4.3. How are methyl-recycling enzymes localizing to multiple compartments? . 121
4.4. Insertion region of SAHH1 may control its subcellular targeting..... 126
4.5. Protein interactions via its IR may direct SAHH to the destination 128

CHAPTER 2: LOCATION OF METHYL RECYCLING ENZYMES IS MEDIATED BY SPECIFIC PROTEIN INTERACTIONS 131

1. Introduction 131
2. Materials and Methods 135
2.1. Source of chemicals 135
2.2. Plant growth 135
2.3. Stable transformation of Arabidopsis 136
2.4. Transient expressions in tobacco plants and Arabidopsis protoplasts 137
2.5. Subcellular localization of CMT and ASDH 139
2.6. Bimolecular fluorescence complementation (BiFC) 141
2.7. Strep tag purification 142
2.8. Protein identification by Mass spectrometry 144
2.9. Yeast two-hybrid assay 145
2.10. Generation of antibodies 146
2.11. Immunoprecipitation and other pull-down assays 147
2.12. SDS-PAGE and Western blotting 149
2.13. Screening T-DNA insertion lines 151
2.14. Artificial microRNA constructs for ADK and SAHH 153
3. Results 156
3.1. Interaction between ADK1 and SAHH1 156
3.2. Bimolecular fluorescence complementation for ADK1-SAHH1 interaction . 163
3.3. Involvement of cap methyltransferase in ADK1-SAHH1 interactions 168
3.4. Subcellular localization of CMT 172
3.5. Protein-protein interaction of CMT 179
3.6. Identification of other SAHH protein interactors: Isolation of aspartate semialdehyde dehydrogenase 183
3.7. Subcellular localization of ASDH 188

3.8. Protein interaction of ASDH and SAHH 192
3.9. Protein interaction of ASDH and SAHH variants 196
3.10. Protein interaction with APT 201
3.11. T-DNA insertion mutants of CMT and ASDH 205
3.12. Creation of artificial microRNA silenced lines for ADK and SAHH 209
4. Discussion and conclusions 221
4.1. ADK interacts with SAHH in planta 221
4.2. Dynamic interaction of ADK and SAHH 222
4.3. CMT may direct ADK and SAHH to the nucleus via protein-protein interactions 224
4.4. The interaction with ASDH is required for plastid localization of SAHH 226
SUMMARY 233
APPENDICES 237
Appendix I. The pSAT vector system 238
Appendix II. Docking model for ADK:SAHH interaction 239
Appendix III. Purification of ADK1-Strep 240
Appendix IV. Yeast two-hybrid assay 242
Appendix V. The aspartate pathway 243
Appendix VI. Docking model for SAHH:ASDH interaction 244
REFERENCES 245

LIST OF FIGURES

INTRODUCTION

Figure 1. *Arabidopsis thaliana* 3
Figure 2. SAM-dependent transmethylation cycle and related pathways..... 9
Figure 3. Three-dimensional structure of *Plasmodium falciparum* SAHH tetramer..... 20
Figure 4. Amino acid sequence alignment of SAHHs..... 24
Figure 5. Methylation in multiple subcellular compartments of a plant cell..... 34
Figure 6. Nuclear import machinery by active transport 37
Figure 7. Mechanism of chloroplast protein import..... 42

CHAPTER 1

Figure 1. Subcellular localization of ADK1 and SAHH1 in Arabidopsis protoplasts 81
Figure 2. Immunoblot analysis of ADK1-EGFP and SAHH1-EGFP 83
Figure 3. Various EGFP-fusion constructs of ADK and SAHH in Arabidopsis and tobacco plants 84
Figure 4. Nuclear localization of double GFP fusions of ADK1 and SAHH1 87
Figure 5. Deletion analyses of ADK1 and SAHH1..... 91
Figure 6. Nuclear localization of SAHH1 deletions..... 93
Figure 7. Comparison of the deduced amino acid sequences of SAHHs from various organisms 98
Figure 8. The affection of the IR for nuclear localization of SAHH1 100
Figure 9. Mutation within the IR of SAHH1..... 102
Figure 10. ADK1 and SAHH1 GFP fusions showing chloroplast localization..... 107
Figure 11. Subcellular localization of the small subunit of Rubisco complex, ADK1 and SAHH1..... 109

Figure 12. Chloroplast localization of SAHH1 in Arabidopsis protoplasts 111
Figure 13. Plants generated by the expression of various GFP-fusion constructs..... 115
Figure 14. Predicted gene models of *ADK1* and *SAHH1* based on EST collections 124

CHAPTER 2

Figure 1. Protein interaction between ADK1 and SAHH1 159
Figure 2. ADK1-Strep purification and immunoprecipitation to examine ADK1-SAHH1 interaction..... 161
Figure 3. BiFC assay for ADK1-SAHH1 interaction..... 166
Figure 4. Comparison of the deduced amino acid sequences of CMTs found in various organisms..... 170
Figure 5. Nuclear localization of CMT..... 175
Figure 6. Possible plastidic localization of CMT 177
Figure 7. Protein interaction of CMT with other proteins 181
Figure 8. SAHH1-Strep purification and ASDH identification 185
Figure 9. Subcellular localization of ASDH..... 190
Figure 10. Interaction between SAHH and ASDH 194
Figure 11. Interaction between ASDH and SAHH variants 199
Figure 12. Interaction of APT with ADK and SAHH..... 203
Figure 13. T-DNA insertion lines for CMT and ASDH..... 207
Figure 14. Phenotype of amiADK plants..... 215
Figure 15. Phenotype of amiSAHH plants..... 217
Figure 16. Immunoblot analyses of amiRNA lines..... 219
Figure 17. Working model for ADK and SAHH localization 231

LIST OF TABLES

CHAPTER 1

Table 1. List of primers used in this research 77

CHAPTER 2

Table 1. Primers used for cloning and T-DNA screening 152

Table 2. Primers used to construct amiRNAs..... 155

Table 3. Protein identification by Mass spectrometry 187

LIST OF ABBREVIATIONS

ADA	Adenosine deaminase
Ade	Adenine
ADK	Adenosine kinase
ADN	Adenosine nucleosidase
ADP	Adenosine diphosphate
Ado	Adenosine
amiRNA	Artificial microRNA
AMP	Adenosine monophosphate
APT	Adenine phosphoribosyltransferase
ASDH	Aspartate-semialdehyde dehydrogenase
ATP	Adenosine triphosphate
BiFC	Bi-molecular fluorescence complementation
BRET	Bioluminescent resonance energy transfer
BSA	Bovine serum albumin
CaMV 35S	Cauliflower mosaic virus 35S promoter
CAS	Cellular apoptosis susceptibility protein
Chr	Chromosome
CMT	Cap Methyltransferase
DDT	Dithiothreitol
DMSO	Dimethylsulfoxide
DsRed2	*Discosoma sp.* red fluorescent protein
EDTA	Ethylene diamine tetraacetic acid

EGFP	Enhanced green fluorescent protein
EMS	Ethyl methane sulphonate
EST	Expressed sequence tag
FRET	Fluorescence resonance energy transfer
GFP	Green fluorescent protein
GST	Glutathione S-transferase
GTP	Guanosine triphosphate
GUS	β-glucuronidase
HA	Influenza Hemagglutinin
Hcy	Homocysteine
His-tag	Histidine-tag
Hog	Homology-dependent gene silencing
HPLC	High performance liquid chromatography
Hsp	Heat shock protein
IgG	Immunoglobulin G
Ino	Inosine
iP	Isopenetyl-adenine
iPA	Isopentenyl-adenosine
iPRMP	Isopentenyladenine riboside monophosphate
IPTG	Isopropyl-D-thiogalactopyranoside
IR	Insertion region
kDa	Kilodaltons
LB	Luria-Bertani
MES	4-morpholineethanesulfonic acid
Met	Methionine
mRNA	Messenger RNA

miRNA	MicroRNA
MS	Murashige-Skoog media
MT	Methyltransferase
MW	Molecular weight
NAD	Nicotinamide adenine dinucleotide
NES	Nuclear export signal
NPC	Nuclear pore complex
NLS	Nuclear localization signal
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PDB	Protein data bank
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PMSF	Phenylmethyl sulfonyl fluoride
RbcS	Small subunit of rubisco complex
RE	Restriction enzyme
RFP	Red fluorescent protein
RNAi	RNA interference
SAH	*S*-adenosylhomocysteine
SAHH	*S*-adenosylhomocysteine hydrolase
SAM	*S*-adenosylmethionine
siRNA	Small interferring RNA
SMM	*S*-methylmethione
SDS	Sodium dodecyl sulfate
TAIR	The Arabidopsis information resource

TAP	Tandem affinity purification
T-DNA	Transfer DNA
TEV	Tobacco etch virus
Tic	Translocon at the inner envelope membrane of chloroplasts
TILLING	Targeting Induced Local Lesions IN Genomes
Toc	Translocon at the outer envelope membrane of chloroplasts
tZ	Trans-zeatin
URH	Uridine ribohydrolase
UTR	Untranslated region
YC	C-terminal fragments of YFP
YFP	Yellow fluorescent protein
YN	N-terminal fragments of YFP

INTRODUCTION

1. Arabidopsis as a model system

Arabidopsis thaliana, a member of the mustard family (Brassicaceae), is a small flowering plant (Figure 1). It was discovered by Johannes Thal in the Harz Mountains, in Germany in the 1500's, and first named as *Pilosella siliquosa*.

Over 750 accessions of Arabidopsis have been collected from different environments around the world and among these, Columbia and Landsberg erecta are the two ecotypes which are most commonly used. Although Arabidopsis lacks agricultural significance, in the past 40 years it has risen to become one the most studied and characterized organisms used by thousands of plant science researchers around the world (The Arabidopsis Genome Initiative, 2000; The Arabidopsis Information Resource (TAIR), 2010).

Arabidopsis has been used extensively for studies of plant molecular genetics, development, metabolism, physiology and biotechnology. The main reasons for this are as follows.

First, Arabidopsis has a short life cycle, about 6 weeks, resulting in the production of 5,000-10,000 seeds.

Second, its genome structure makes it well suited for functional genomic studies including small genome size (about 125 Mb) relative to other plants, relatively low amount of interspersed repeat sequence DNA, small introns and few gene families. These traits facilitate the use of molecular techniques for gene mapping, isolation and characterization (The_Arabidopsis_Genome_Initiative, 2000; Wortman et al., 2003).

As well, the complete genome has been sequenced and multiple databases of DNA sequences have been compiled and are available online. Up until now, 31,407 genes have been reported, and among these numbers, 26,751 are annotated as protein-coding genes (TAIR, 2010). In addition to the useful online databases, the stock centers providing thousands of cloned genes and a large number of well-characterized Arabidopsis mutants including point mutants, insertion mutants and activation-tagged lines are available, and facilitate a quick analysis of genes of interest (TAIR, 2010; The_Arabidopsis_Genome_Initiative, 2000).

Lastly, production of transgenic plants is relatively easy by using the bacterium *Agrobacterium tumefaciens* and various methods for transformation of Arabidopsis have been developed (Bent, 2000; Zhang et al., 2006). Among these, the floral dip method is most widely used for producing transgenic Arabidopsis plants due to its simplicity and efficiency (Clough and Bent, 1998). For example, transformation can be accomplished by dipping Arabidopsis inflorescences for a few seconds, and a transformation frequency of at least 1% can be obtained within about 2 months (Zhang et al., 2006). These numerous advantages led to the use of Arabidopsis in the research outlined in this thesis which examines the subcellular localization and protein-protein interaction of methyl recycling enzymes.

Figure 1. ***Arabidopsis thaliana*****.**

A five-week-old wild-type Arabidopsis plant (the Columbia accession) grown under standard conditions (16-hour day length at 22℃) is shown (A). The mature Arabidopsis plant consists of five main parts: roots, shoots, leaves, flowers and siliques. The rosette leaves (E) form at the base of the plant, and the axillary buds of these leaves develop into shoots from which cauline leaves (D) arise. The shoot produces flowers (B; young floral buds and flowers) after about 3 weeks, and the flowers become siliques (C) upon self-pollination.


A
B
C
D
E
B
C
D
E

2. Activated methyl cycle

2.1. Methylation

Methylation including the addition or substitution of a methyl group on various substrates is a process necessary for many biological functions. It is known to be part of various biological processes such as: the regulation of gene expression and gene silencing (Finnegan et al., 2000; Turker, 2002); the regulation of protein functions involved in signal transduction, stress responses, protein trafficking, and protein-protein interactions (Dolzhanskaya et al., 2006; Kim et al., 1998; Mowen et al., 2004); and the modification of mRNAs by 5' cap methylation, which affects transcript stability, subsequent mRNA processing, and initiation of translation (Tourriere et al., 2002).

In plants, methylation is a very pervasive reaction since a number of compounds acting in response to various environmental stimuli, throughout plant development require the addition of a methyl group for their functionality (Weretilnyk et al., 2001). For example, secondary metabolites such as flavonoids and anthocyanins which contain methyl groups are synthesized in response to UV light, drought, salt stress, or pathogen (Dixon and Paiva, 1995; Ibrahim et al., 1998). The biosynthesis and function of pectin and lignin, which are the major components of cell wall structure, are dependent on methylation/demethylation reactions (Boudet, 1998; Bourlard et al., 1997). Many other compounds including plant hormones are also methylated for their synthesis or function: glycine betaine (Summers and Weretilnyk, 1993); chlorophyll (Von Wettstein et al., 1995); phospholipids (Datko and Mudd, 1988); nicotine, caffeine, and other alkaloids (Facchini, 2001; Koshiishi et al., 2001); and plant hormones such as cytokinins,

gibberellic acid and brassinosteroids (Fujita et al., 1977; Levi et al., 1978). All these methylation reactions depend on the availability of methyl donors.

2.2. SAM-dependent transmethylation

In cells, there are several common methyl donors such as *S*-adenosyl-L-methionine, 5-methyltetrahydrofolate, trimethylglycine, methylsulfonylmethane and methylcobalamin. Among these, the major methyl donor for all biological methylation reactions is *S*-adenosyl-L-methionine (SAM), which is a high energy intermediate in the cellular metabolism of both prokaryotes and eukaryotes (Edwards, 1996; Lu, 2000). SAM is a small molecule composed of methionine and an adenosyl group which is synthesized by SAM synthetase (SAMS; EC 2.5.1.6) from L-methionine and adenosine triphosphate (ATP) (Chou and Talalay, 1972; Schröder et al., 1997). SAM functions not only as an active methyl donor but it is also used as a key molecule for many biological processes in plants: regulation of the biosynthesis of methionine and amino acids of the aspartate family, inhibition of an isozyme of aspartate kinase (Azevedo et al., 1997), activation of threonine synthase (Curien et al., 1998), and regulation of the activities of spermidine and spermine synthases as a substrate of decarboxylases (Pegg, 1986). In addition, it is used as the source of 5′-deoxyadenosyl radicals which leads to the production of methionine and 5′-deoxyadenosyl radical intermediate (Roje, 2006). It also acts as precursor for the biosynthesis of some compounds including polyamines, nicotianamine and ethylene (Roje, 2006).

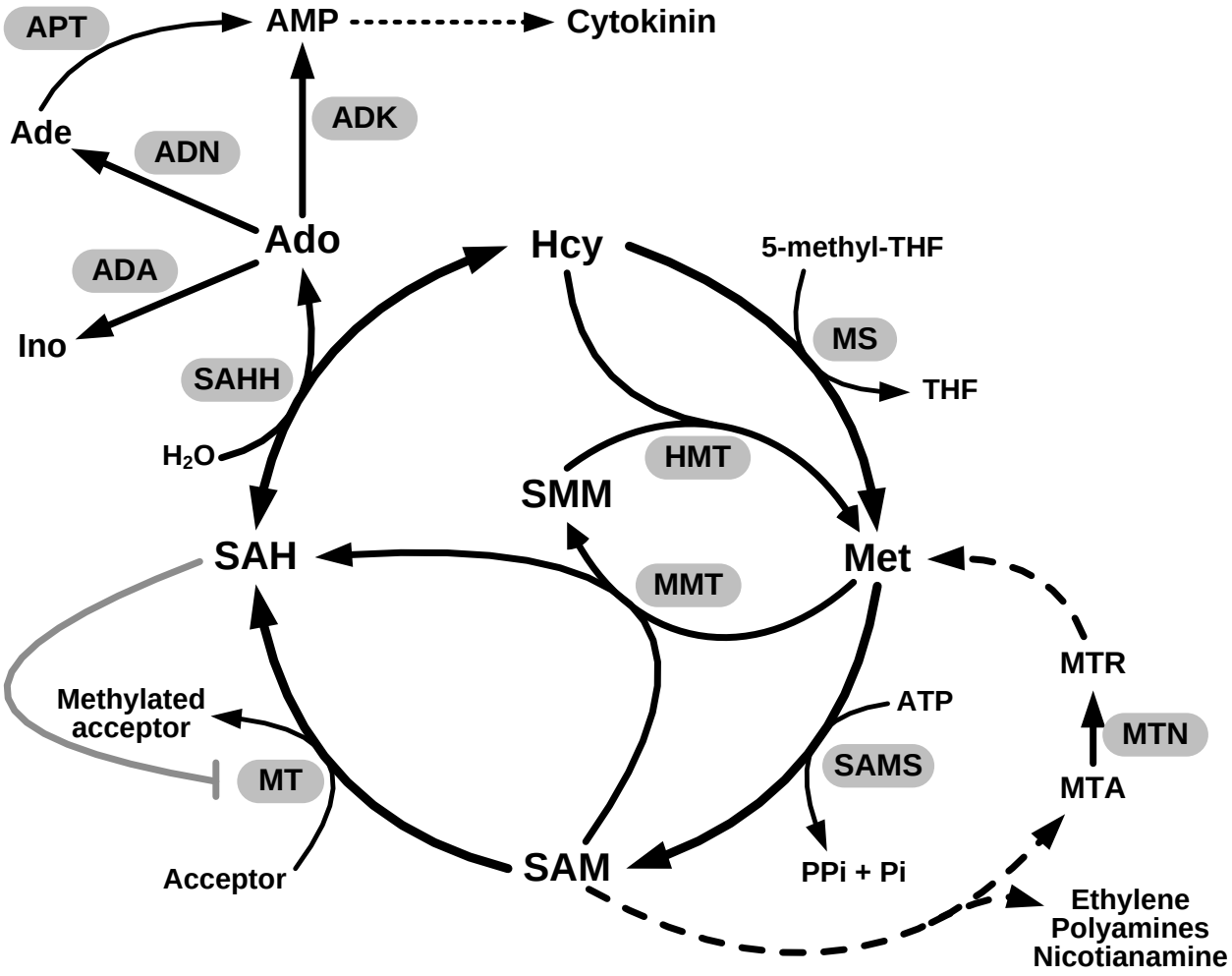
Since over 90% of SAM is used by many different methyltransferases (MT; EC 2.1.1) for methylation reactions (Clarke et al., 2003), a steady supply of SAM for the activated methyl cycle is essential for the majority of MT activities which are involved in a variety of biological functions. These SAM-dependent MT reactions are catalyzed by substrate-specific MTs that transfer a methyl group from the methyl donor SAM to various methyl-accepting compounds. It is thought that a different MT acts on each acceptor, which leads to the conclusion that hundreds of different MTs exist in eukaryotic cells. As a by-product, of these reactions a molecule of *S*-adenosyl-L-homocysteine (SAH) is produced for each methyl group transferred from SAM to a methyl acceptor (Poulton, 1981). Since MT activities are essential for modification or biosynthesis of numerous compounds such as DNA, RNA, protein, pectin, lignin, phosphatidylcholine and plant growth regulators, SAH is continuously produced as a result of SAM-dependent transmethylation (Moffatt and Weretilnyk, 2001; Poulton, 1981). The activated methyl cycle is shown in Figure 2.

Since SAH is similar in structure to SAM it acts as a competitive inhibitor of SAM-dependent MT activities, which could lead to blocking essential transmethylation reactions (Poulton and Butt, 1975). SAH must therefore be metabolized to maintain SAM-dependent transmethylation. Despite the diversity of MTs that exist in eukaryotic cells, only one enzyme, *S*-adenosyl-L-homocysteine hydrolase (SAHH; EC 3.3.1.1), is capable of SAH removal (De La Haba and Cantoni, 1959). SAHH breaks down the SAH into homocysteine (Hcy) and adenosine (Ado), allowing the MT activities to continue (De La Haba and Cantoni, 1959). This reaction is critical to determining the balance between SAM and SAH (known as methylation ratio, methylation potential or methyl index) in cells (Loehrer et al., 1998).

The SAHH-catalyzed reaction is reversible and the direction of SAH synthesis is favored thermodynamically, thus the two end-products of this metabolism, Hcy and Ado, must also be removed (De La Haba and Cantoni, 1959; Poulton, 1981). Hcy is converted to methionine (Met) by addition of a methyl group either from 5-methyl-tetrahydrofolate which is catalyzed by Met synthase (EC 2.1.1.13) or from *S*-methylmethionine (SMM) by Hcy *S*-methyltransferase (EC 2.1.1.10). Met is subsequently recycled back into SAM or SMM by SAM synthetase (EC 2.5.1.6) or Met *S*-methyltransferase (EC 2.1.1.12), respectively (Hanson et al., 1994). SMM is known to have a role in the regulation of Met and SAM levels and in the transport of sulphur from leaves to seeds via the phloem, and as a Met storage form or as a plant-specific methyl group donor (Bourgis et al., 1999; Kocsis et al., 2003; Mudd and Datko, 1990). On the other hand, Ado is metabolized by three different routes (Figure 2): hydrolysis of Ado to adenine (Ade) by Ado nucleosidase (ADN; EC 3.2.2.9), deamination of Ado to inosine (Ino) by Ado deaminase (ADA; EC 3.5.4.4), and the conversion of Ado into Ado monophosphate (AMP) via phosphorylation by Ado kinase (ADK; EC 2.7.1.20).

Figure 2. SAM-dependent transmethylation cycle and related pathways.

SAH, which acts as an inhibitor of MTs, is produced as a by-product of the methyl-transfer from SAM to an acceptor by a MT and subsequently cleaved into Hcy and Ado by SAHH. While Hcy is recycled to Met either by MS and 5-methyl-THF as a methyl donor or by HMT via SMM cycle, Ado is metabolized by three different routes as indicated. SAH inhibits MT activities. SAMS catalyzes the conversion of Met to SAM that serves not only as methyl donor but also as a precursor in the biosynthesis of ethylene, polyamines and nicotianamine. MTA, a key by-product of polyamine biosynthesis, can be recycled to Met. The dotted arrow indicates the involvement of AMP in cytokinin biosynthesis and interconversion. ADA, adenosine deaminase; Ade, adenine; Ado, adenosine; ADK, adenosine kinase; ADN, adenosine nucleosidase; AMP, adenosine monophosphate; APT, adenine phosphoribosyltransferase; Hcy, homocysteine; HMT, homocysteine *S*-methyltransferase; Ino, inosine; Met, methionine; MMT, methionine *S*-methyltransferase; MS, methionine synthase; MT, methyltransferase; MTA, methylthioadenosine; MTN, methylthioadenosine nucleosidase; MTR, methylthioribose; SAH, *S*-adenosylhomocysteine; SAHH, *S*-adenosylhomocysteine hydrolase; SAM, *S*-adenosylmethionine; SAMS, *S*-adenosylhomocysteine synthase; SMM, *S*-methylmethionine; THF, tetrahydrofolate


APT
AMP
Cytokinin
Ade
ADN
ADK
Ado
ADA
Ino
SAHH
H2O
Hcy
5-methyl-THF
MS
THF
HMT
SMM
SAH
Met
MMT
MTR
ATP
MTN
Methylated acceptor
MT
SAMS
MTA
Acceptor
SAM
PPi + Pi
Ethylene
Polyamines
Nicotianamine

2.3. Adenosine salvage pathways

Since the synthesis of Ado is energetically expensive requiring the hydrolysis of 7 ATP or GTP molecules, Ado is salvaged by three different reactions (Figure 2). In mammalian cells, ADN has a critical role in Ado metabolism but is thought to be less predominant in plants. ADN catalyzes the irreversible hydrolysis of Ado to Ade, with the Ade being subsequently phosphoribosylated to AMP by Ade phosphoribosyltransferase (APT; EC 2.4.2.7). Although ADN activity has been found in several plant crude extracts such as barley (Guranowski and Schneider, 1977), wheat (Chen and Kristopeit, 1981), and Arabidopsis (Auer, 1999), in only a few cases is ADN thought to be a major contributor to Ado metabolism. For example, caffeine biosynthesis in tea leaves (*Camellia sinensis* L.) has been shown to involve ADN-produced AMP (Koshiishi et al., 2001) and leaves of *Avicennia marina* are reported to have 25-fold higher ADN activity than ADK activity under non-stressed conditions (Suzuki et al., 2003). Riewe et al. (2008) reported that the potato (*Solanum tuberosum*) ADN is involved in salvage of extracellular ATP and is highly specific for Ado. Gas chromatography-mass spectrometry (GC-MS) analysis using the extracts prepared from potato tubers showed the production of ribose by ADN in the presence of Ado. Moreover, this enzyme was characterized as a cell wall-bound ADN, since it is not released during extraction procedures including a non-ionic detergent (Riewe et al., 2008). Recently, a uridine ribohydrolase (URH) that is highly homologous to nucleoside hydrolases from protists and yeast was identified from Arabidopsis (Jung et al., 2009). It was shown that URH has a key role in pyrimidine degradation and utilizes uridine as a preferred substrate, with Ino, Ado and cytokinin ribosides being less well accepted substrates (Jung et al., 2009). The two isoforms of

Arabidopsis ADN, ADN1 and ADN2 were further studied recently using a spectrophotometric assay (Engel, 2009). Again no substrate was found for ADN1 while ADN2 preferred uridine. Taken together, it is clear that while ADN activity is certainly present in plants, a gene encoding an enzyme specific for Ado is not yet evident.

ADA catalyzes the deamination of Ado to Ino and of deoxyadenosine to deoxyinosine (Fox and Kelley, 1978). Although ADA has been found in most organisms including mammals, invertebrates, bacteria, and fungi where it is thought to play a key role in Ado salvage exert pathway (Cristalli et al., 2001), it has been generally considered that ADA activity is not present in plants (Brady and Hegarty, 1966; Dancer et al., 1997). Recently, the functionality of the only putative ADA coding sequence in Arabidopsis (At4g04880) was investigated (Engel, 2009). Based on the recombinant protein assays and functional complementation, this gene product does not possess ADA activity and is not essential for Arabidopsis plant growth (Engel, 2009). Thus, if ADA-mediated Ado salvage exists, it must be encoded by a different gene.

In contrast to ADN and ADA, recent evidence indicates that ADK activity is the predominant route for the removal of Ado in plants, based on its high transcript abundance and kinetic activity in plants (Moffatt et al., 2002). Moreover, the abnormal phenotype of ADK-deficient plants, which correlates with the accumulation of Ado, also supports the importance of ADK activity for plant growth (Moffatt et al., 2002).

3. Adenosine Kinase and *S*-adenosyl-L-Homocysteine Hydrolase

3.1. ADK

ADK is a typical housekeeping enzyme, being constitutively expressed in and essential to all cells based on its contribution to various cellular processes (Moffatt et al., 2000). It is considered to be the most abundant nucleoside kinase in mammalian cells (Mathews et al., 1998). The lethality of ADK knock-out mice indicates its critical role in eukaryotic organisms (Boison et al., 2002). In the livers of these mice AMP and ADP were decreased by 60% and ATP was reduced by 65% compared to wild type, while SAH was increased by 2.3-fold. The ADK-deficient mice exhibited microvesicular hepatic steatosis within 4 days after birth and they ultimately died within 14 days with fatty livers. The Ade nucleotide deficiency, including ATP, in the liver is thought to cause mitochondrial dysfunction as well as deficiencies in fatty acid transport and oxidation. Although the accumulation of SAH in ADK-deficient mice is consistent with a shift in the equilibrium of the SAHH reaction toward accumulation of SAH, the SAH/SAM ratio is not significantly altered compared to that of wild type so transmethylation reactions are likely not affected (Boison et al., 2002).

In wild-type mammalian cells, ADA mediates the major route for Ado recycling during embryogenesis with prominent expression levels in the placenta (Blackburn et al., 1995). ADK is present in most tissues and highly expressed in the postnatal liver where ADA is expressed at very low levels (Andres and Fox, 1979; Guranowski et al., 1981). Thus, Ado is mainly metabolized to Ino by ADA, during embryonic development, and

converted to AMP by ADK, after birth. At this point, ADA-expression shifts to cells lining the alimentary canal and the regulation of the Ado level in most tissues becomes dependent on ADK-expression (Chinsky et al., 1990). ADA knock-out mice are embryo lethal due to the formation of cytotoxic derivatives of 2'-deoxyadenosine, which are potent inhibitors of SAHH. In these mutants, the level of SAH and SAM are increased 5-6-fold and 2-fold, respectively, compared to wild-type controls (Migchielsen et al., 1995; Wakamiya et al., 1995).

While only three prokaryotic ADKs have been reported from *Mycobacterium tuberculosis* (Long et al., 2003), *Streptomyces lividans* (Rajkarnikar et al., 2007), and *Xanthomonas campestris* (Lu et al., 2009), ADKs have been isolated and characterized from many eukaryotes, including human (Spychala et al., 1996) and other mammalian organisms (Fisher and Newsholme, 1984; McNally et al., 1997), yeast (Caputto, 1951; Kornberg and Pricer, 1951), parasitic protozoa (Darling et al., 1999; Datta et al., 1987), moss (von Schwartzenberg et al., 1998) and several plant species (Chen and Eckert, 1977; Faye and LeFloch, 1997; Guranowski, 1979; Moffatt et al., 2000).

In mammalian cells, two isoforms of ADK exist; one of the two contains additional 20–21 amino acids at the N-terminus. Both isoforms are believed to have the same functions based on their similar biochemical properties (Sakowicz et al., 2001; Spychala et al., 1996). ADKs have an optimum pH between 4.8 and 8.0 and a temperature range between 30 and 41°C depending on the organism from which they have been isolated (Schomburg et al., 2004). *M. tuberculosis* ADK exists as a dimer (Long et al., 2003), but all other characterized ADKs act as monomers, with an average molecular weight of 34 to 56 kDa. The crystal structures of ADK for the enzyme isolated from *Homo sapiens*,

Toxoplasma gondii, *M. tuberculosis* have been determined at high resolution (Cook et al., 2000; Mathews et al., 1998; Reddy et al., 2007).

ADKs have been characterized from various plants including wheat (Chen and Eckert, 1977), peach (Faye and LeFloch, 1997), yellow lupin (Guranowski, 1979), Arabidopsis (Moffatt et al., 2000), and tobacco (Laukens et al., 2003). A strong structural similarity between human and plant ADKs has been proposed based on the similarity of their deduced amino acid sequences and three-dimensional structure modeling (Moffatt et al., 2000; von Schwartzenberg et al., 1998).

3.2. ADK in Arabidopsis

In Arabidopsis, there are two genes encoding ADK, designated *ADK1* (At3g09820) and *ADK2* (At5g03300) (Moffatt et al., 2000). These enzymes have a molecular mass of about 38 kDa, and share 56% amino acid sequence identity with human ADK (Moffatt et al., 2000; Spychala et al., 1996). The two isoforms share significant sequence similarity at both the amino acid (92%) and nucleotide (89%) levels. Moreover, both *ADK* genes are constitutively expressed throughout plant development, although *ADK1* transcripts are generally more abundant than those of *ADK2* (Moffatt et al., 2000). Current transcript profiles documented in the online database Genevestigator (https://www.genevestigator.com/) indicate the *ADK1* mRNA levels are 1.5-fold higher than those of *ADK2* in all tissues throughout development, with the exception of the germinated seed stage in which *ADK2* transcript levels are slightly higher (Hruz et al., 2008).

In an effort to determine the importance of ADK in plants, Arabidopsis transgenic plants with reduced ADK activity were generated (Moffatt et al., 2002). These plants, designated as sADK for sense silencing of ADK, were made by siRNA transgene silencing. The individual lines vary in their residual ADK activity, from 7 % to 70 % of that in the wild type (Moffatt et al., 2002). The sADK plants display developmentally abnormal phenotypes including clustered inflorescences, wrinkled leaves, and delayed senescence (Moffatt et al., 2002). Compared to wild type, sADK leaves with less than 20% residual ADK activity have a 40-fold increase in SAH, almost 2-fold decrease in SAM (Moffatt et al., 2002) and 3.5-fold increase in Ado (N. Emery, S. Schoor and B Moffatt, manuscript in preparation). The increased SAH and Ado levels are consistent with the scenario that accumulated Ado leads to SAH synthesis instead of SAH hydrolysis resulting in an accumulation of SAH that inhibits MT activities. Both seed coat pectin methylation and DNA methylation are reduced in these lines. It is possible that the wavy leaf morphology of the sADK lines may in part be due to altered pectin methyl-esterification, which changes how tightly adjacent cell walls are associated with each other (Pereira et al., 2006).

Taken together, these studies provide evidence that ADK is essential for maintaining MT activities in Arabidopsis and likely other plants (Moffatt et al., 2002; Pereira et al., 2006). This conclusion is further supported by the increase of ADK activity levels in response to methylated osmolyte glycine betaine (*N*,*N*,*N*-trimethylglycine) accumulation under salt stress conditions (Suzuki et al., 2003; Weretilnyk et al., 2001). In addition, ADK activity and transcript abundance are coordinately regulated with those of SAHH

during plant development, suggesting a possible relationship between these two enzymes to sustain MT activities (Pereira et al., 2007).

3.3. SAHH

Since SAH acts as an inhibitor of MTs (Gibson et al., 1961), SAH metabolism is essential to maintain the intracellular balance of SAM:SAH ratio in a cell (Kramer et al., 1990; Loehrer et al., 1998). In most bacterial species, SAH is metabolized by 5'-methylthioadenosine (MTA) nucleosidase (EC 3.2.2.16) or SAH nucleosidase (EC 3.2.2.9) to Ade and *S*-ribosylhomocysteine but in eukaryotes only SAHH removes SAH (De La Haba and Cantoni, 1959). Whereas the SAHH reaction is reversible and its equilibrium lies in the biosynthetic direction to form SAH, the rapid removal of Ado and Hcy under normal conditions favors the catabolic direction of the reaction (Palmer and Abeles, 1979).

The importance of SAH metabolism is evidenced by a number of mutant studies. A reduction in human SAHH activity to 3-10% of the control results in the accumulation of SAM and SAH by 30-fold and 150-fold, respectively, accompanied by severe growth and developmental defects including myopathy and mental retardation (Barić et al., 2004; Buist et al., 2006); the complete loss of activity results in embryonic lethality in mice (Miller et al., 1994).

SAHH activity is a target for antiviral and antiparastic drugs. Since transmethylation is essential for the replication of these pathogens and they rely on the host SAHH for SAH hydrolysis, these drugs are particularly effective (Robins et al., 1998; Yang and Borchardt,

2000). They act by the following mechanism: inhibition of viral or parasitic SAHH → elevation of SAH levels → inhibition of MT activity → reduction in mRNA caps on viral transcripts → suppression of viral or parasitic growth. For example, anti-SAHH inhibitors are the only known effective inhibitors of the Ebola virus (Yuan et al., 1999). Since SAHH is a highly conserved protein (Parker et al., 2003), most parasitic enzymes are nearly identical to human SAHH in terms of their amino acid sequences and protein structures, including the cofactor binding site (Hu et al., 1999; Tanaka et al., 2004; Turner et al., 1998).

3.4. Dynamic protein structure of SAHH

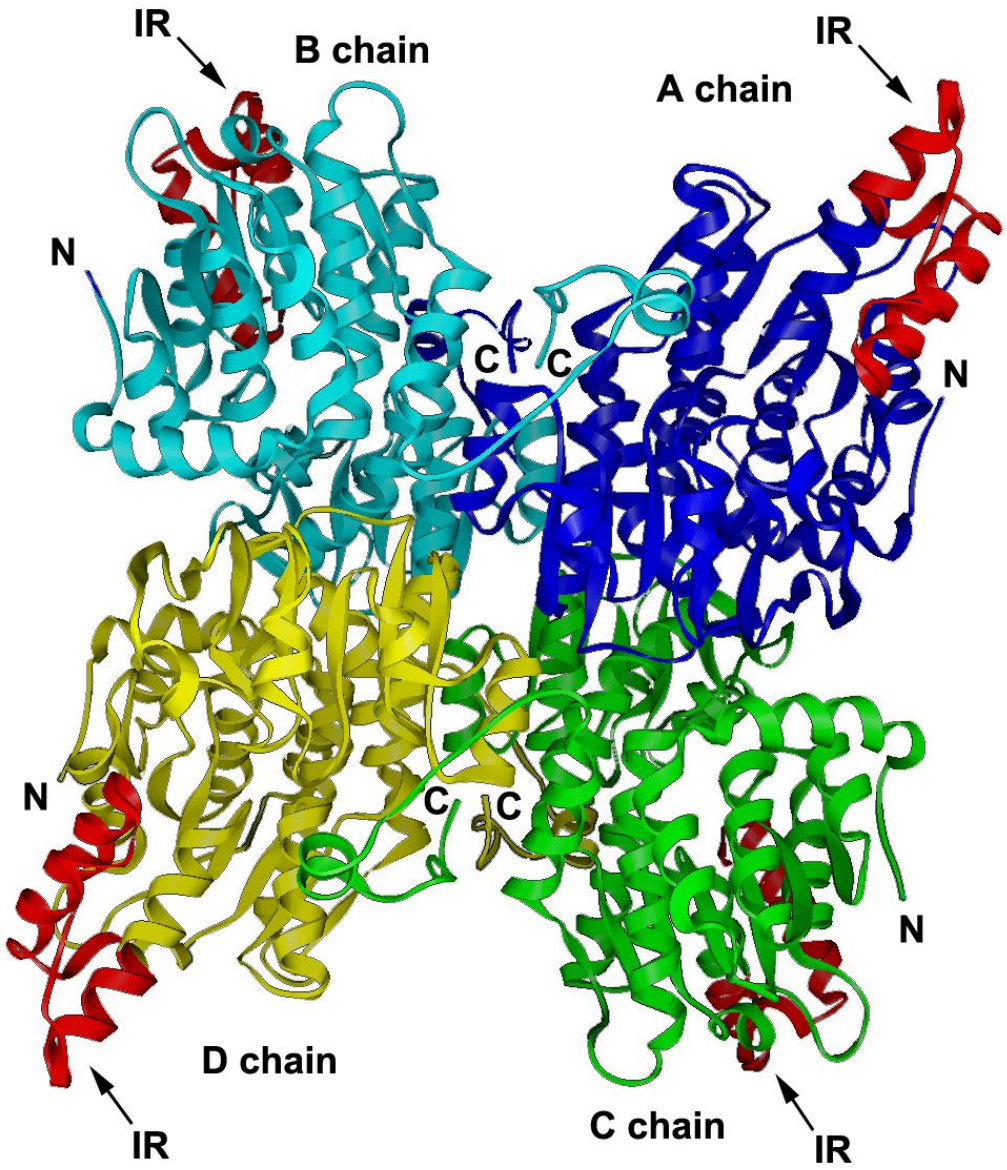
The crystal structures of SAHH purified from several organisms including human (Turner et al., 1998), rat (Hu et al., 1999), parasitic protozoa (*Plasmodium falciparum*) (Tanaka et al., 2004), and pathogenic bacteria (*M. tuberculosis*) (Reddy et al., 2008) have been determined. Except for one case, that of yellow lupin (*Lupinus luteus*) seeds, all other SAHHs exist as highly conserved homotetramic enzymes designated A, B, C and D (Guranowski and Pawelkiewicz, 1977; Turner et al., 1998). Each subunit contains one tightly bound nicotinamide Ade dinucleotide (NAD) molecule in its active site (Turner et al., 2000). Each monomer consists of three domains: an N-terminal substrate-binding domain, a cofactor-binding domain, and a small C-terminal tail that is important for tetramer stability (Ault-Riché et al., 1994; Turner et al., 1998). The two adjacent subunits A and B form a dimer in a reciprocal manner by integrating their small C-terminal tails, thereby covering the cofactor-binding site of the other partner. This dimerization also

occurs between subunits C and D. The A-B and C-D dimers bind tightly to each other to form the tetramer, which results in the formation of central core surrounded by the four cofactor-binding domains (Turner et al., 1998)(Figure 3).

SAHH undergoes a conformational change upon substrate binding, changing from an "open form" to a "closed form" (Wang et al., 2005; Yang et al., 2003). This transition is important for the enzymatic catalysis as well as substrate recognition and binding. The open form allows the substrate easy access to the active site; and the transition from open to closed form occurs once the substrate binds. The closed form provides a closer contact between substrate and cofactor (NAD^+/NADH) within the enzyme, which is supportive for catalysis (Wang et al., 2005; Yang et al., 2003; Yin et al., 2000).

Figure 3. Three-dimensional structure of *Plasmodium falciparum* SAHH tetramer.

The 3-D structure of the human malaria parasite *Plasmodium* SAHH tetramer (1V8B) obtained from the protein data bank (PDB; http://www.pdb.org/pdb/home/home.do) and refined using 3D Molecule viewer of VectorNTI (Invitrogen). The background color was changed from black to white using Adobe Photoshop. The four subunits (A-D chains) are shown in different colors with a solid ribbon. The insertion region (IR) of each subunit is shown in red. C, C-terminus; IR, Insertion region; N, N-terminus.


IR
B chain
IR
A chain
N
C
C
N
N
C
C
N
D chain
C chain
IR
IR

3.5. Insertion region in SAHH

SAHH is one of the most conserved enzymes in terms of the sequences and protein structural similarities reported to date (Sganga et al., 1992; Turner et al., 2000). As mentioned earlier, SAHH forms a homotetramer and catalyzes the conversion of SAH to Ado and Hcy (De La Haba and Cantoni, 1959; Poulton, 1981; Turner et al., 1998). The estimated molecular weight of each subunit ranges between 53-57 kDa for plant SAHH (e.g. 53 kDa with 485 residues in Arabidopsis)(Guranowski and Pawelkiewicz, 1977; Mitsui et al., 1993), while it is ~48 kDa in mammalian organisms (47.5 kDa with 432 residues in humans)(Bethin et al., 1995; Turner et al., 2000). The difference in the molecular weights between plant and mammalian SAHHs is mainly due to the presence of 40-45 amino acids in the middle of the polypeptide (positions 150-190 in Arabidopsis SAHH)(Figure 3). This insertion region (IR) is not found in SAHHs from other eukaryotic organisms (Figure 4) with the exception of several parasitic protozoa SAHHs: *Entamoeba histolytica*, *Trichomonas vaginalis*, *Cryptosporidium parvum*, *T. gondii*, *Theileria annulata*, and *P. falciparum* (Stepkowski et al., 2005). It has been proposed that the loss of the IR in some eukaryotes perhaps preceded the divergence of their lineages in a limited number of species (Stepkowski et al., 2005).

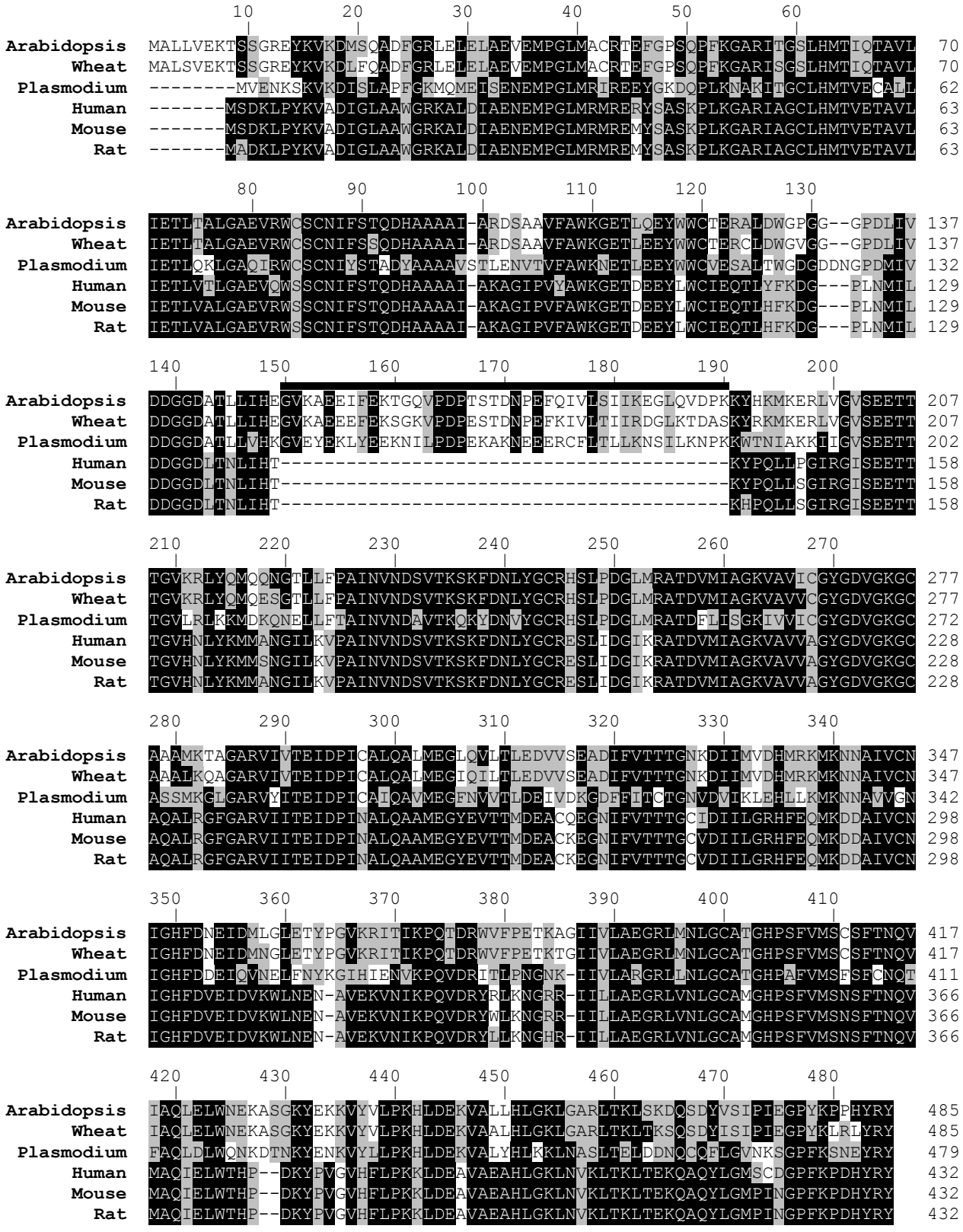
Recently, two crystal structure studies resolved the IR structure of the human malaria parasite *P. falciparum* that contains a 40-residue insertion at position 146-186 (Tanaka et al., 2004) and the pathogenic bacteria *M. tuberculosis* that contains a 37-residue insertion at 167-203 (Reddy et al., 2008). The two IRs appear similar to each other, forming strand-turn-helix (*Mycobacterium*) or helix-loop-helix (*Plasmodium*) structures on the surface of the substrate-binding domain of each subunit. Since the IR is separated from

the main body of the tetramer and does not contact other subunits, it does not affect the formation of SAHH tetramer (Reddy et al., 2008; Tanaka et al., 2004).

Reddy and colleagues (Reddy et al., 2008) also determined the crystal structure of the SAHH:SAH complex to better understand the binding mode of Hcy and the involvement of the IR. Their results led to the proposal that the presence of the IR changes the entrance of the access channel to the Hcy-binding region from a shallow depression, to a deep pocket with a narrow opening. This deep pocket-shaped entrance may allow His363 to act as a switch that opens up the solvent access channel to accept the substrate, and closes down (restricts solvent access channel) after release of the free Hcy (Reddy et al., 2008). Since all plant SAHHs contain the unique segment IR that is not present in most of other eukaryotes, its biological and physiological significance in plant cells remains to be elucidated.

Figure 4. Amino acid sequence alignment of SAHHs.

The alignment was performed using AlignX of VectorNTI with default setting and edited in Microsoft Word. Conserved residues are black-shaded, and similar residues are in gray. Numbers on right represent amino acid sequence of each SAHH indicated; numbers above the alignment represent amino acid sequence of Arabidopsis SAHH. The black line above the sequence alignment represents the insertion region (IR; Gly^{150}-Lys^{190}). The amino acid sequences used in the alignment are (Swiss-Prot accession numbers): Arabidopsis (*Arabidopsis thaliana*; O23255.1), wheat (*Triticum aestivum*; P32112.1), Plasmodium (*Plasmodium falciparum*, P50250.2), human (*Homo sapiens*; P23526.4), mouse (*Mus musculus*; P50247.3), rat (*Rattus norvegicus*; P10760.3).


Arabidopsis
Wheat
Plasmodium
Human
Mouse
Rat

3.6. SAHH in plants

For decades, the majority of the research on SAHH has been focused on mammalian or parasitic enzymes due to their importance as potential drug or therapeutic targets (Parker et al., 2003; Yang and Borchardt, 2000). However since SAHH plays a key role in metabolic processes essential for plant development, it has also been studied in various plant species using different approaches.

First, the enzyme activity of SAHH, with its fluctuating pattern in relation to metabolic and developmental processes that requires methylation, has been characterized in several plants including tobacco (Sebestova et al., 1984), spinach (Poulton and Butt, 1976), parsley (Kawalleck et al., 1992), and alfalfa (Edwards and Dixon, 1991). In addition, SAHH enzyme activity was assessed using crude extracts prepared from different organs of five-week-old wild-type Arabidopsis plants (Todorova, 2002). The highest specific activity was measured in extracts of roots, while lower levels of activity were detected in buds, flowers, siliques, and upper stem (Todorova, 2002).

Second, *SAHH* gene expression has been documented in several development instances. *SAHH* transcripts are predominantly located in stems of lucerne (*Medicago sativa*), with a significant decrease occurring in response to wounding and various hormone treatments, due to a decrease in transcription (Abrahams et al., 1995). In parsley leaves, highly abundant *SAHH* transcripts are detected in floral buds and stems, whereas lower levels are found in roots and leaves (Kawalleck et al., 1992). Moreover, the markedly increased expression was observed in response to pathogen-derived elicitor as a defense mechanism (Kawalleck et al., 1992). In addition, an increase in *SAHH* expression

is observed in parsley exposed to UV light (Logemann et al., 2000). In Arabidopsis, SAHH is constitutively expressed throughout the plant including roots, stems and siliques with the highest transcript abundance, and buds and flowers with relatively lower mRNA abundance (Todorova, 2002).

Lastly, the characterization of plants deficient in SAHH activity supports the essential role of this enzyme in plant growth and development. For example, petunia plants, in which *SAHH* gene expression is suppressed 2-10 fold by antisense gene expression, have profuse branching, delayed flowering, increased leaf size and higher seed yield, while Arabidopsis plants overexpressing *SAHH* (*HOMOLOGY-DEPENDENT GENE SILENCING1; HOG1*) exhibit early flowering with a significantly reduced plant biomass and number of leaves (Godge et al., 2008). Moreover, a 2-20 fold reduction of *SAHH* transcript levels in tobacco plants expressing an SAHH antisense transgene, lead to a 10-fold higher level of SAH relative to the wild type (Masuta et al., 1995), and exhibit abnormal morphological changes in floral organs with stunting of growth, loss of apical dominance, delayed senescence, and hypomethylation of repetitive DNA (Tanaka et al., 1997). These traits are similar to those observed in Arabidopsis sADK plants as mentioned earlier in Section 3.2 (Moffatt et al., 2002), suggesting that ADK and SAHH deficiencies affect similar physiological and developmental processes (Moffatt et al., 2002; Pereira et al., 2006; Tanaka et al., 1997).

The Arabidopsis genome carries two highly similar genes encoding SAHH activity. The two isoforms, designated *SAHH1* (At4g13940) and *SAHH2* (At3g23810), have a 92% nucleic acid and 95% amino acid sequence identity. Both are constitutively expressed throughout plant development, while *SAHH1* transcript levels are generally higher than

those of *SAHH2* (Li et al., 2008; Pereira et al., 2007). Microarray data compiled in Genevestigator (Hruz et al., 2008) reveals *SAHH1* mRNA levels are 1.6 to 3.5-fold higher than those of *SAHH2* in all tissues throughout development, with highest levels of expression in the germinated seed stage for *SAHH1* and the mature flowering stage for *SAHH2*, respectively (https://www.genevestigator.com/; as of Oct. 2010). Recently, several SAHH mutants have been isolated and characterized from Arabidopsis by different research groups. EMS-induced point mutations in Arabidopsis *SAHH1 (HOG1)* gene result in developmental abnormalities including growth retardation, low fertility and reduced seed germination, with a decrease in DNA (total cytosine is reduced to 10%) methylation (Rocha et al., 2005). However, a complete knockout created by a transposon (*hog1-4*) and a T-DNA (*hog1-5*) insertion in the *HOG1* gene results in zygotic lethality (Rocha et al., 2005). Arabidopsis plants, with a decrease in *SAHH1* gene expression, created by a T-DNA insertion in the *SAHH1* gene have defects in root-hair development along with the other phenotypes observed in *hog1* plants (Wu et al., 2009).

Taken together with the difference in the abundance of *SAHH1* and *SAHH2* transcripts, the SAHH1 mutant studies indicate that SAHH1 is predominant over SAHH2 and these two isoforms may have distinct rather than redundant roles in Arabidopsis. This is also supported by the observation that a T-DNA insertion in the *SAHH2* gene displays indistinguishable phenotypes from its wild-type parent with normal growth and fertility, while an *SAHH1* gene knockout causes zygotic lethality as mentioned above (Rocha et al., 2005). In addition, the key phenotypes observed in the antisense *SAHH1* gene suppression lines suggest the involvement of its gene product in cytokinin signaling (Godge et al., 2008).

3.7. ADK and SAHH in cytokinin metabolism

Cytokinin, a class of plant hormones, has several important roles in aspects of plant growth and development including cell division, cell and organ enlargement, senescence, sink strength, and bud dormancy (Hoth et al., 2003; Werner et al., 2003). The predominant cytokinins in higher plants are isopentenyladenine (iP) and *trans*-zeatin (tZ) (Werner et al., 2003). Both iP and tZ are derivatives of N^6-prenylated Ade and are thought to be the biologically active compounds in Arabidopsis (Inoue et al., 2001; Yamada et al., 2001). In the first step of cytokinin biosynthesis, the isoprenoid group is transferred to AMP by adenylate isopentenyltransferase (adenylate-IPT; EC 2.5.1.27), resulting in the formation of isopentenyladenine riboside monophosphate (iPRMP). This reaction happens not only on AMP but also on ADP and ATP, which allow the formation of isopentenyladenine riboside (di- or tri-) phosphate (iPRDP or iPRTP, respectively). Subsequently, the conversion of iPRMP to iP is catalyzed by 5′-nucleotidase and Ado nucleosidase. The conversion of iP into tZ is catalyzed by P450 monooxygenases (Takei et al., 2004). In the iPRMP-independent pathway, as an alternative pathway for tZ biosynthesis, the side chain precursor dimethylallyl diphosphate (DMAPP) is directly transferred to AMP to produce zeatin monophospate (ZMP) (Kudo et al., 2010).

Several lines of evidence lead to the conclusion that ADK is involved in cytokinin metabolism. *In vitro* assays of recombinant Arabidopsis ADKs and wheat homogenates containing ADK activity showed that these enzymes are able to use both Ado and isopentenyladenosine (iPA) as substrates (Chen and Eckert, 1977; Moffatt et al., 2000). In the case of Arabidopsis ADKs, their affinity for Ado is 10 times higher than for iPA (Moffatt et al., 2000). *In vivo* feeding of chloronemal tissues of the moss *Physcomitrella*

patens with cytokinins also demonstrated that the ADK activity is important for the conversion of cytokinin ribosides to the corresponding nucleotides based on *in vivo* feeding of iPA (von Schwartzenberg et al., 1998). More recently, kinetic analysis of the four purified tobacco ADK isoforms of BY-2 cells showed they have strong affinity for the three predominant cytokinin ribosides, iPA, zeatin riboside, and dihydrozeatin riboside *in vitro* (Kwade et al., 2005). Consistent with this, the apoptosis induced by iPA in tobacco BY-2 cells depends on ADK activity that catalyzes the conversion of iPA to iPA monophosphate, since the caspase activities and the number of dead cells are markedly decreased when ADK activity is inhibited (Mlejnek and Prochazka, 2002).

There is also evidence that supports SAHH's role as a cytokinin-related protein (Mitsui et al., 1993). It was identified as a component of cytokinin-binding protein complex in tobacco leaves, following purification of the complex by a cytokinin-affinity chromatography (Mitsui et al., 1993). Moreover, SAHH-deficient tobacco lines with decreased SAHH activity have about three times higher levels of cytokinin than those of wild-type plants (Masuta et al., 1995) and the abnormal phenotype of these plants correlates with their increased cytokinin levels (Masuta et al., 1995; Tanaka et al., 1997). More recently, Godge and colleagues (2008) purified an epitope-tagged SAHH from transgenic Arabidopsis plants and found that it had a strong affinity for cytokinins based on thermal calorimetry measurements (Godge et al., 2008). Finally, the cytokinin profile of Arabidopsis *sahh1* mutant roots was determined by high performance liquid chromatography (HPLC) (Li et al., 2008). The T-DNA insertion mutant analyzed has markedly reduced *SAHH1* gene expression and exhibits several phenotypic changes, including delayed flowering and senescence, fewer and shorter roots, and increased

chlorophyll levels compared to wild type, with normal fertility. The mutant roots have a three-fold higher level of zeatin than those in the wild-type Arabidopsis roots. Gene expression levels of both *SAHH1* and *ADK1* were increased upon an addition of iP, however none of the cytokinins tested affected SAHH enzyme activity *in vitro* (Li et al., 2008).

Taken together, it is clear that both ADK and SAHH are not only important for maintaining in SAM-dependent transmethylation, but also likely to be involved in cytokinin metabolism. This conclusion is supported by the observation that cytokinin distribution patterns in Arabidopsis floral meristems (Corbesier et al., 2003) are consistent with the transcript abundance of both *ADK* and *SAHH* in different meristematic regions (Pereira et al., 2007).

3.8. Subcellular localization of ADK and SAHH

As mentioned earlier, transmethylation is essential for biosynthesis or modification of numerous methyl acceptor compounds in plant cells including DNA and mRNAs, pectin, lignin, phosphatidylcholine, and others (Moffatt and Weretilnyk, 2001) (Figure. 2). To fulfill all these processes, SAM-dependent transmethylation must occur throughout the cell in specific subcellular locations, depending upon the substrate, as shown in Figure 5. Clearly, the enzymes involved in these reactions, including SAHH and ADK need to be located in the same cellular compartments to maintain MT activities. However, both ADK and SAHH are regarded as cytosolic enzymes since no specific localization or conventional N-terminal signal sequences have been found by various prediction

programs designed to determine the subcellular localization of proteins, including PA-SUB (Lu et al., 2004), MultiLoc2 (Blum et al., 2009), WoLF PSORT (Horton et al., 2007), NOCtree (Nair and Rost, 2005), ChloroP (Emanuelsson et al., 1999), and TargetP (Emanuelsson et al., 2000). Moreover, the cytoplasm is proposed to be the primary site where the methylation reactions occur (Hanson et al., 2000; Hanson and Roje, 2001). Consistent with these predictions, both SAHH and ADK are positioned in the cytosol by immunoblot analysis using different fractions prepared from spinach leaves (Weretilnyk et al., 2001).

However, ADK and SAHH may also reside in the nucleus. For example, research using *Xenopus leavis* oocytes showed the accumulation of SAH in the nucleus and reduced mRNA methylation due to the inhibition of SAHH activity in cell cultures (Radomski et al., 1999). In a follow-up study, these authors reported evidence of SAHH translocation from the cytoplasm to the nucleus, mediated via a specific protein interaction with the MT involved in mRNA capping (Radomski et al., 2002). Recently, the localization of the two isoforms of ADK in various mammalian cells including human HT-1080, human HeLa, mouse LMTK-, and Chinese hamster CHO cells was reported (Cui et al., 2009). In this study, the immunofluorescence labeling showed that the long form of ADK that contains an additional 20 amino acids at its N-terminus is localized in the nucleus, whereas the short form resides in the cytoplasm. Moreover, the additional N-terminal sequence of long form was identified as a nuclear localization signal (NLS) based on its ability to target a fused green fluorescent protein (GFP) to this compartment (Cui et al., 2009).

Earlier work from the Moffatt lab using immunogold labeling and β-glucuronidase (GUS) fusions suggested that Arabidopsis ADK and SAHH are present in the nucleus and the cytoplasm, and possibly in the chloroplast as well (Pereira, 2004; Schoor, 2007). Thus, part of the research described in this thesis was designed to test these observations. Moreover, it examines how both enzymes migrate between the cytoplasm and other organelles. In order to understand this, background information on protein transport systems in the nucleus and the chloroplast is provided in the next section.

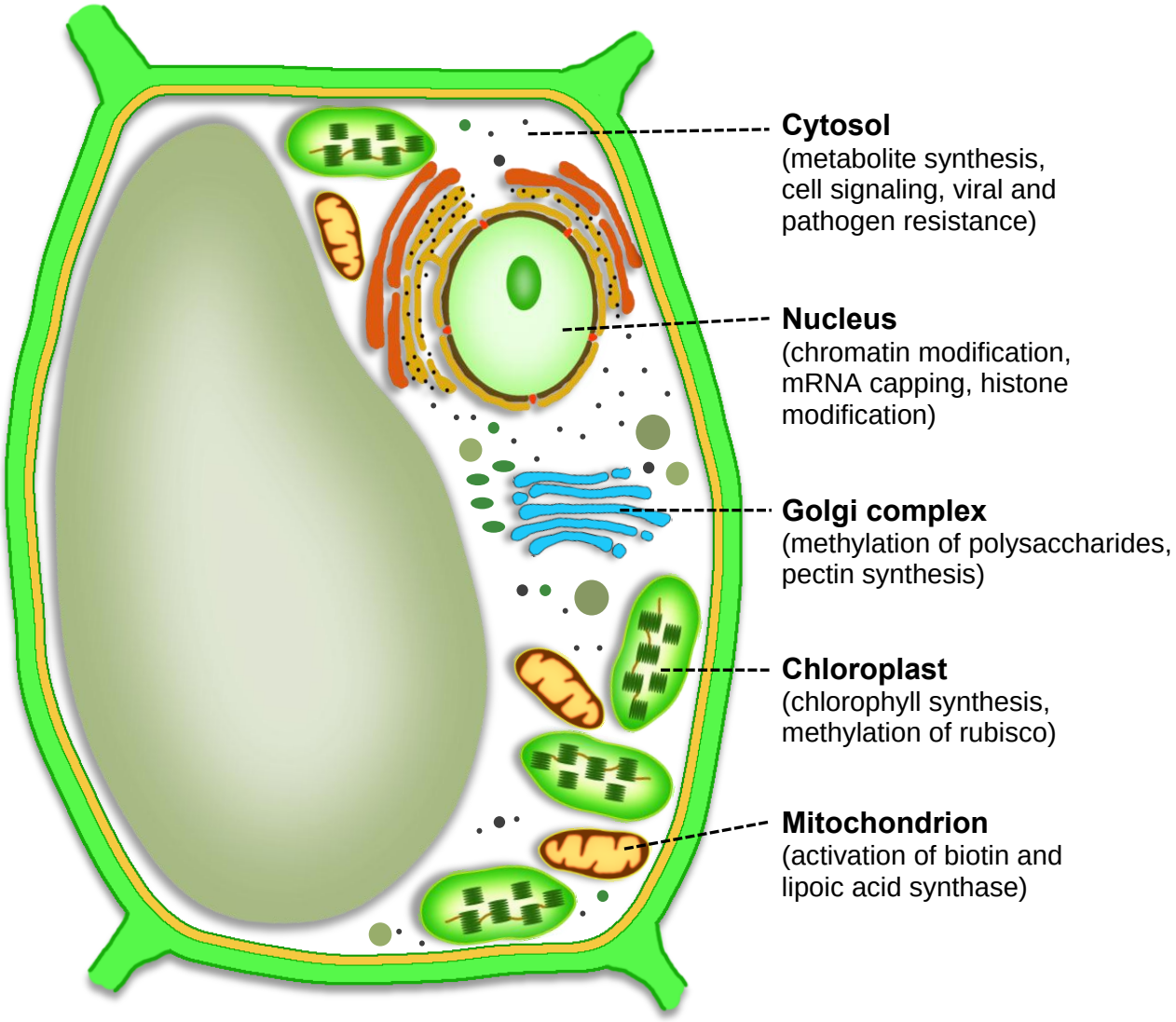


Figure 5. Methylation in multiple subcellular compartments of a plant cell.

Methylation reactions take place every second and everywhere in plant cells for many important metabolic functions. The plant cell image was originally adapted from (Silva-Filho, 2003) and redrawn using Adobe Photoshop.

4. Mechanism of protein transport

4.1. Nuclear targeting

The nucleus is a dynamic organelle that contains the majority of the genetic information of an organism. The nucleus is also the site for various processes, such as DNA replication and repair (Lisby et al., 2003; Nakayasu and Berezney, 1989), RNA transcription and splicing (Fu and Maniatis, 1990; Wansink et al., 1993), and other modifications including mRNA capping (Shafer et al., 2005) and chromatin modification (Richards and Elgin, 2002). To fulfill all these essential processes, a myriad of factors and proteins (which are synthesized in the cytosol), must enter and exit the nucleus throughout the cell cycle (Weis, 2003).

The nucleus is surrounded by a nuclear envelope consisting of two lipid bilayers containing numerous nuclear pore complexes (NPC) (Tran and Wente, 2006; Weis, 2003). The nuclear envelope not only provides a physical barrier that separates the nuclear genome from the intermediary metabolism and signaling systems of the cytoplasm and other organelles, but also allows macromolecular trafficking in and out of the nucleus through the NPC (Merkle, 2003; Tran and Wente, 2006; Weis, 2003).

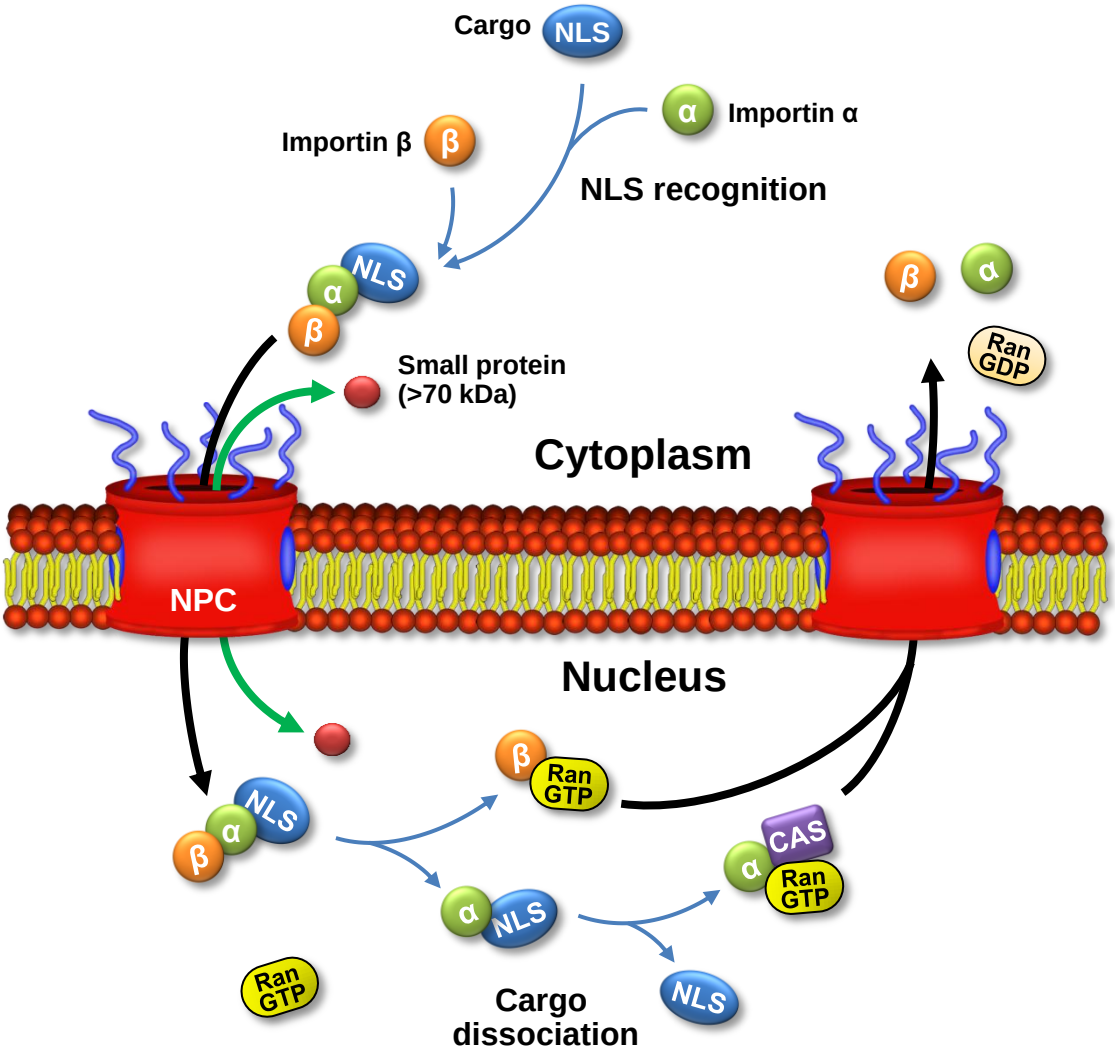
Molecules smaller than 20-40 kDa such as water, ions, and small proteins can passively diffuse through the NPC (Bohnsack et al., 2002; Fried and Kutay, 2003), whereas macromolecules >~70 kDa require an energy-dependent transport machinery (active transport) (D'Angelo et al., 2009). Active transport is mediated by transport receptors (e.g. importin α/β) and signal peptides containing nuclear localization signals (NLSs) or nuclear export signals (NESs) (Harel and Forbes, 2004). These are present in

most, but not all nuclear targeting proteins, as nuclear import can also be mediated by interaction with another protein that does have an NLS (Nigg, 1997).

Import of proteins into the nucleus by active transport involves several steps (Figure 6). Importin β binds to the cargo via an adaptor protein, importin α, that recognizes the NLS on the cargo (NLS-cargo) in the cytoplasm. The import complex (importin-α:β heterodimer-cargo or cargo:carrier complex) is then translocated to the cytoplasmic side of the NPC and subsequently to the nucleoplasmic side of the NPC. In the nucleus, RanGTP (the GTP-bound form of the G protein Ran) binds to importin β, resulting in the dissociation of the complex into the NLS-cargo and an importin β-RanGTP complex. Then, both importins separately recycle back to the cytoplasm complexed with either importin β-RanGTP or importin α-RanGTP-CAS (CAS, Cellular Apoptosis Susceptibility protein; the export cofactor for importin α), leaving the NLS-cargo in the nucleus. In the cytoplasm, RanGAP catalyzes the conversion of RanGTP to RanGDP, resulting in the release of importins in the cytoplasm for another round of nuclear transport. The exportin-mediated transport of NES-cargo to the cytoplasm occurs in similar, but reverse manner. The directionality of nucleocytoplasmic transport is effectively controlled by the nucleotide state of Ran (between GTP- and GDP-bound states) that regulates cargo binding and release of import and export receptors. The Ran GTPase provides the energy for this nuclear transport (Merkle, 2003; Stewart, 2007; Tran and Wente, 2006).

Figure 6. Nuclear import machinery by active transport.

Nuclear import process includes five distinct steps: (1) NLS recognition in the cytoplasm, the cargo binds to the receptor (importins); (2) Translocation of importins-cargo complex into the nucleus, the complex enters through the nuclear pore complex (NPC); (3) Cargo dissociation in the nucleus, RanGTP dissociates importin β from cargo and importin α binds to CAS (Cellular Apoptosis Susceptibility protein); (4) Recycling of importins to the cytoplasm, an exportin CAS forms a complex consisting of CAS, RanGTP and importin α; (5) Importin dissociation in the cytoplasm, the conversion of RanGTP to RanGDP releases importins. Small molecules (<70 kDa) can passively diffuse through the NPC (green arrow). The image was originally adapted from (Xylourgidis and Fornerod, 2009) and redrawn using Adobe Photoshop.


Cargo
NLS
Importin α
α
Importin β
β
NLS recognition
NLS
α
β
Small protein
(>70 kDa)
β
α
Ran
GDP
Cytoplasm
NPC
Nucleus
β
Ran
GTP
NLS
α
β
CAS
α
Ran
GTP
α
NLS
Ran
GTP
Cargo
dissociation
NLS

4.2. Chloroplast targeting

Chloroplasts, the photosynthetic organelles of green plants and algae, are the most prominent members of the plastid family that includes leucoplasts, amyloplasts, and chromoplasts. Plastids, along with mitochondria, are of prokaryotic origin (cyanobacteria and proteobacteria, respectively) and entered the eukaryotic lineage through endosymbiosis (Gould et al., 2008; Reyes-Prieto et al., 2007). Chloroplasts contain the green pigment chlorophyll and are hence responsible for photosynthetic activities including light-harvesting and carbon-fixation reactions (Gould et al., 2008). It is estimated that about 18% of Arabidopsis nuclear genes (about 4,500 of ~25,000 genes) were acquired from a cyanobacterial ancestor of the plastid by endosymbiotic gene transfers (Martin et al., 2002); among these only a small fraction of genes (15-209) are known to be encoded in the chloroplast genome (Keeling and Palmer, 2008). As a result, chloroplasts greatly rely on sophisticated transport systems to ensure that essential proteins encoded by the nuclear DNA are imported from the cytoplasm where the proteins are synthesized. In Arabidopsis, approximately 3,000 nuclear encoded proteins are estimated to be targeted to chloroplasts (Li and Chiu, 2010). The majority of these proteins are synthesized in precursor form with a cleavable N-terminal transit peptide that is required for targeting to the organelle by an active, post-translational targeting process (Li and Chiu, 2010).

Transport of nucleus-encoded proteins from the cytoplasm to the chloroplast is mediated by two membrane protein complexes called translocon components that are located either in the outer membrane (Toc; translocon at the outer envelope membrane of chloroplasts) or in the inner membrane (Tic; translocon at the inner envelope membrane

of chloroplasts) of chloroplasts, respectively (Schnell et al., 1997)(Figure 7). The Toc complex (Toc159, Toc34 and Toc75) which is exposed at the chloroplast surface, mediates the initial recognition and translocation of preproteins across the outer membrane requiring ATP and GTP (Li and Chiu, 2010). Although it is not clearly known yet how preproteins are diverted from the Toc complex into the inter-membrane space, a possible role of a cytosolic guidance complex, composed of an Hsp70 (70 kDa heat shock protein) chaperone and a 14-3-3 protein, is proposed for directing preproteins to the Toc or Tic complex (May and Soll, 2000). At this point, the Tic complex (Tic20, Tic21 and Tic110), which forms the inner membrane translocation channel, physically associates with the Toc complex, and the translocation proceeds simultaneously across both membranes. In addition, the interaction of Toc/Tic complexes with a set of chaperones (Hsp93, 93 kDa heat shock protein; Cpn60, 60 kDa chaperonin protein) appears to drive translocation of preproteins into the stroma (Nielsen et al., 1997) and facilitate the folding and assembly of newly imported proteins (Kessler and Blobel, 1996). Upon translocation across the Toc/Tic complexes, a soluble stromal processing peptidase (SPP) cleaves the transit peptide from the preprotein (Richter and Lamppa, 1998).

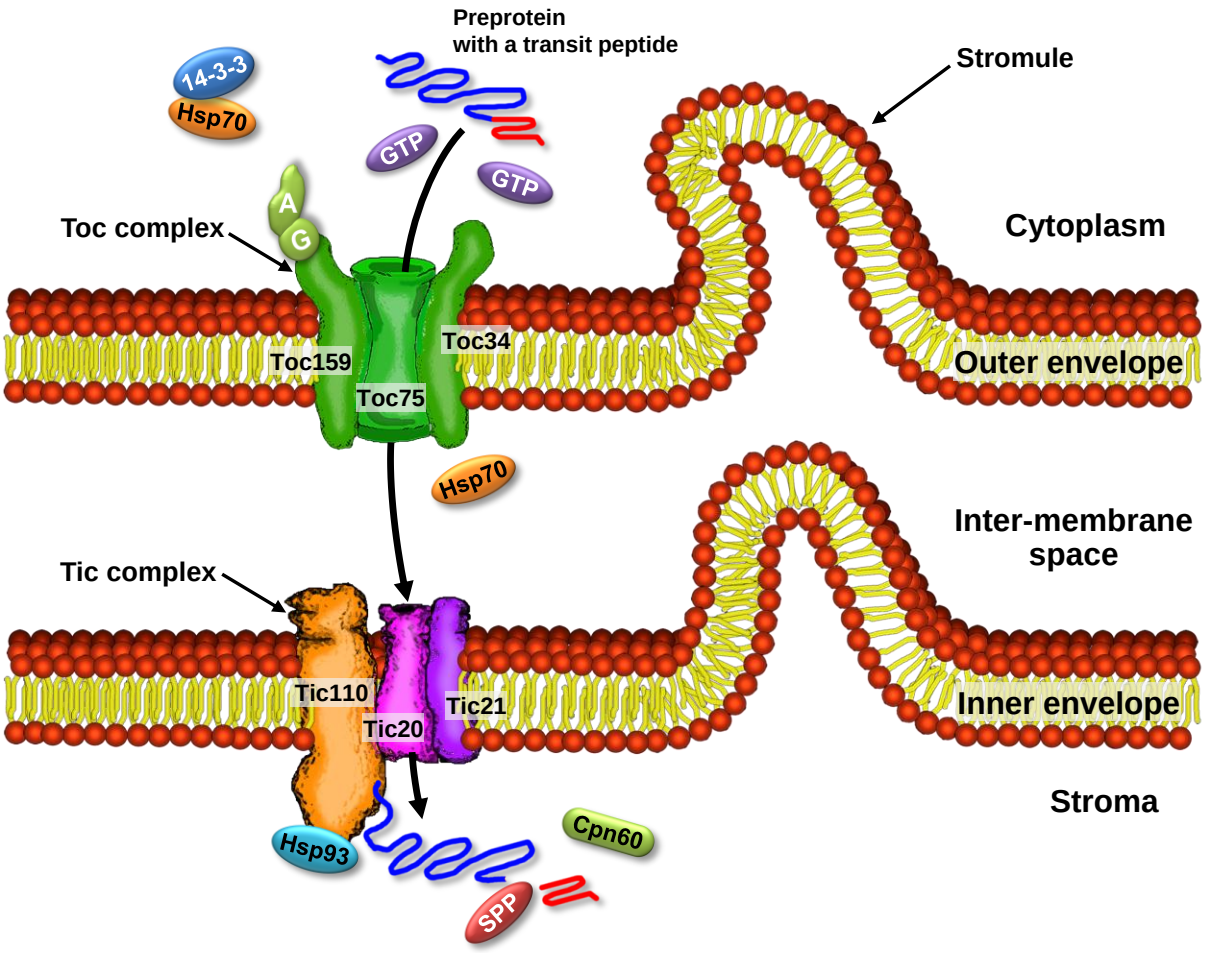
In addition, while there is no direct evidence, it is proposed that an alternative pathway for chloroplast import may exist: protein trafficking through the endomembrane system or unusual translocation of chloroplast proteins lacking cleavable transit peptides (Jarvis and Robinson, 2004). Recent microscopy studies using GFP show that the plastid envelope has profound tubular structures, termed stromules (for stroma-filled tubules) (Kwok and Hanson, 2004). Although their exact mechanism of action including their formation and function is not clearly established yet, it is suggested that stromules play a

role in exchanging materials between plastids and other organelles (Kwok and Hanson, 2004). Moreover, from the observation of an increase of the surface area of the plastid envelope membrane by extrusion of stromules, it is further suggested that stromules may also facilitate protein import from the cytoplasm into plastids (Hanson and Sattarzadeh, 2008).

To characterize the mechanisms of organelle protein transport, much work described above has been done using various experimental tools, including GFP, mutant plants and other approaches for protein interaction detection. These observations revealed the location and association between Toc/Tic complexes and/or other factors involved in the machinery. My research also applies these approaches to understanding how MT activities are maintained throughout the cell, including the nucleus and the chloroplast. Thus, the next section discusses several useful tools for studying the aspects of plant science.

Figure 7. Mechanism of chloroplast protein import.

A guidance complex (14-3-3/Hsp70) directs preproteins that contain a transit peptide to the chloroplast surface. GTP supports the binding of the transit peptide of preproteins to the Toc complex (most likely at the G domain of Toc159). Preprotein translocation occurs through the outer and inner membranes simultaneously. Upon the change of ATP concentration (a low concentration (less than 100 μM) requires at this point) and proper temperature, the translocation of preprotein proceeds across the outer envelope membrane through the Toc75 channel, and at the same time, the preprotein is completely translocated into the stroma through the Tic complex of the inner envelope membrane. Upon arrival in the stroma, the transit peptide is cleaved by SPP, and the processed mature protein is then completely released from the Tic complex by Hsp93. The chaperonin Cpn60 likely assists the folding of mature proteins. Stromules may contribute to share molecules such as metabolic intermediates and proteins among plastids or between plastids and other compartments. However, the exact mechanism of their action or structure is unknown. The image was originally adapted from (Li and Chiu, 2010) and (Jarvis and Robinson, 2004), and redrawn using Adobe Photoshop.


Preprotein
with a transit peptide
Stromule
14-3-3
Hsp70
GTP
GTP
A
G
Toc complex
Cytoplasm
Toc159
Toc34
Toc75
Outer envelope
Hsp70
Inter-membrane
space
Tic complex
Tic110
Tic21
Tic20
Inner envelope
Stroma
Hsp93
Cpn60
SPP

5. Useful tools for studying plant sciences

5.1. Fluorescent proteins

GFP from the jellyfish *Aequorea victoria* was first discovered in the middle of the 1900's (Shimomura et al., 1962) and three decades later the gene encoding this protein was cloned (Prasher et al., 1992). GFP is 238 amino acids long and has a molecular weight of 27 kDa. The cloning of the GFP sequence trigged a revolution in cell biology, particularly since GFP can fluoresce in living cells without any additional substrate or cofactors, even when fused to other proteins of interest (Chalfie et al., 1994).

Since its discovery, GFP has been used for hundreds of different applications (Wang et al., 2008) including intracellular protein trafficking and localization studies, as well as for studying cell apoptosis (Li and Horwitz, 1997), protein-protein interactions (Hu and Kerppola, 2003), and as a reporter for protein folding (Waldo et al., 1999). Moreover, a variety of GFP-homologous fluorescent proteins that produce different colors have been isolated and characterized (Chudakov et al., 2005), and further variants created by mutation studies resulting in improved brightness, folding efficiency, and expression properties (Pedelacq et al., 2006). For example, there is a variety of different fluorescent proteins including BFP (blue), CFP (cyan), YFP (yellow), RFP (red) and other enhanced forms of fluorescent protein (e.g. EGFP), which allow its widespread use as a fluorescent protein tag and even its multicolor imaging in a single cell (Giepmans et al., 2006). In addition, photoactivatable fluorescent proteins (PA-FPs), whose spectral properties can be changed in response to irradiation with specific light, enable specific marking of

proteins of interest within a living cell in a spatial and temporal manner (Lukyanov et al., 2005).

In the research described in this thesis, fusion proteins of GFP and RFP linked to ADK, SAHH, and other proteins are used to investigate their subcellular localization and protein trafficking. The fusion proteins were expressed in several different hosts including transient expression in Arabidopsis protoplasts (PEG methods), transient expression in tobacco (*Nicotiana benthamiana*; leaf infiltration method) plants, and stable transformation of Arabidopsis plants. Likewise, combining with the transgenic approach a fluorescent protein not only can enable scientists to tag proteins of interest and to visualize their cellular activity, but also provides a more powerful tool when applied with other techniques such as reverse genetics and protein interaction assays. For example, when a fluorescent protein is used with other detection systems it allows *in vivo* studies of protein interactions in living cells.

5.2. Reverse genetics

Part of my research also uses genetic-based tools to create various T-DNA insertion lines and to study mutant plants. Genetics is the study of how phenotype and genotype are passed down from one generation to the next. Molecular genetics has been rapidly developed through two different approaches: forward and reverse genetics. Forward genetics attempts to determine the genetic basis for recognizable phenotypic variations induced by environmental stimuli or chemical mutagens (e.g., ethyl methanesulfonate; EMS) (Maple and Moller, 2007; Snow et al., 1984). Reverse genetics on the other hand,

seeks to discover the function of a gene, by creating phenotypes using the gene sequence. Since numerous genome sequences (5483 as of Sep. 2010; NCBI) are now available, reverse genetics has emerged as a powerful analysis tool to understand gene function and regulation. A DNA sequence of interest is engineered to disrupt the function of a target gene by a modification or a mutation, and this genetic material is introduced into the organism; the resulting phenotypes caused by the gene modification are then assessed (Bouche and Bouchez, 2001). Several genetic tools have been developed to generate loss- and/or gain-of-function mutations in plants.

First, chemical mutagenesis is an effective way to create random genetic variation. The most commonly used chemical mutagen for Arabidopsis research is EMS which induces transition point mutations (G/C to A/T) by methylation of guanine bases (Maple and Moller, 2007; Snow et al., 1984). EMS mutagenesis can be used for both forward and reverse genetic studies. It is a particularly powerful tool when it is combined with TILLING (Targeting Induced Local Lesions IN Genomes). This is a relatively recent reverse genetics tool based on the use of a mismatch-specific cleavage enzyme to identify target gene mutations (Henikoff et al., 2004). TILLING was originally developed for Arabidopsis (McCallum et al., 2000) and has been successfully applied to the analysis of other plant as well as animal species (Slade et al., 2005; Winkler et al., 2005). This allows the rapid and inexpensive high-throughput screening of a large mutant population to identify those with changes in a target gene (Henikoff et al., 2004). For Arabidopsis TILLING, DNA samples extracted from 8 M2 plants derived from the seeds treated with EMS are pooled and PCR-amplified using gene-specific primers for a targeted region of the genome. To detect mutations, PCR reaction pools are heated and cooled to allow

heteroduplexes formation (mismatched base pairs between mutated and wild-type DNA), which are then screened using denaturing HPLC that allows the detection of the mismatches in the heteroduplexes as an extra peak in the chromatogram. When an alteration is detected, DNAs from individual mutant plants are PCR amplified and the mutant plants are ultimately identified by sequencing the PCR product carrying the alteration (McCallum et al., 2000).

Homologous recombination is a method to create site-directed mutations or null mutations by substituting a wild-type gene sequence with a modified gene sequence (Capecchi, 1989). It has been extensively used to produce transgenic animals including knockout mice using embryonic stem cells (Mansour et al., 1988; Smithies et al., 1985), but is not widely used in plants. Recently, zinc-finger nucleases, which are modified to create DNA double-strand breaks, have been used to cleave and stimulate mutations at specific genomic sites in Arabidopsis (Lloyd et al., 2005). Moreover, the successful use of zinc-finger nucleases for high-frequency specific-gene modification in tobacco suggests the possibility of its use for plant genome engineering through homologous recombination in the future (Townsend et al., 2009; Wright et al., 2005).

My research has relied, in part, on a third type of mutagenesis, insertional mutagenesis which is a powerful tool for identifying new genes and their functions (Azpiroz-Leehan and Feldmann, 1997; Martienssen, 1998). The most commonly used insertional mutagens in plants are transposons and T-DNA vectors, which can disrupt the expression of the gene into which it is inserted and also acts as a marker to identify the mutation (Krysan et al., 1999). The T-DNA of *Agrobacterium tumefaciens* creates random and stable insertions in the plant genome that are maintained in subsequent generations (Azpiroz-

Leehan and Feldmann, 1997). In the case of Arabidopsis, this tool is even more powerful as large collections of T-DNA insertion mutants have been made publically available, along with sequence information for the location of the T-DNA insert in each mutant (TAIR, 2010). These collections allow for the easy identification of mutants affected in genes of interest by screening databases of the sequences flanking the insertion sites (Azpiroz-Leehan and Feldmann, 1997).

I have also utilized transgene-mediated gene silencing by RNA interference (RNAi) and artificial microRNAs (amiRNA) in my research. These types of gene regulation were first described relatively recently (Fire et al., 1998) but are surprisingly very active in eukaryotic cells. The expression of gene targets may be silenced at either the transcriptional or post-transcriptional level; the factors determining the level at which silencing occurs are not completely understood (Fahlgren et al., 2007). In the case of RNAi, short 21-26 nucleotide RNAs (small interferring RNAs or siRNAs) are produced by a RNase III-like enzyme (Dicer) cleavage of double-stranded RNAs (Elbashir et al., 2001). These 21-to 26-nucleotide single-stranded RNAs, which are the primary silencing RNAs, target long RNA molecules after binding to a RNase H-like protein of the Argonaute class. In plants, the targeted RNA is then converted to long double-stranded RNA by the RNA-directed RNA polymerases (RdRPs) and secondary siRNAs are generated by Dicer cleavage (Baulcombe, 2007). The siRNAs then assemble into endoribonuclease-containing complexes (an RNA-induced silencing complex; RISC), unwinding in the process. The single-stranded siRNA subsequently guide RISC to complementary RNA molecules and ultimately cleave and destroy specific target mRNAs

resulting in suppression of gene expression or inhibition of translation (Carthew and Sontheimer, 2009; Ghildiyal and Zamore, 2009).

One of the differences between miRNAs and most siRNAs is the accuracy of their ends. The miRNAs have highly exact ends, whereas siRNAs tend to be more heterogeneous in end composition. This feature of miRNAs probably allows them to interact with greater specificity on substrate mRNAs without a large overlap (Carthew and Sontheimer, 2009). In addition, precursors of siRNAs form fully complementary long double-stranded RNAs, while miRNAs are derived from single-stranded transcripts that include imperfect foldbacks (Ossowski et al., 2008).

Since previous studies have shown that several nucleotide substitutions within the stem-loop structure of the miRNA have no effect on its functionality, amiRNAs can be designed by modifying miRNA sequences to target endogenous genes resulting in post-transcriptional silencing of these target genes (Schwab et al., 2006; Vaucheret et al., 2004; Warthmann et al., 2008). Because plant amiRNAs have high specificity, similar to that of endogenous miRNAs, sequences can be simply designed for a highly specific silencing of target transcripts without affecting the expression of other transcripts (Ossowski et al., 2008). A web-based bioinformatic tool for amiRNA design has been developed for over 250 plant species (as of Sep. 2010) and is now available at http://wmd3.weigelworld.org/ (Schwab et al., 2006). However, although this online tool provides a list of possible amiRNA sequences, it does not design candidates with 100% efficiency. Thus, selecting two or more amiRNAs in a different region of a target gene from the list are recommended to effectively reduce the time required to obtain the plants with successful suppression of gene expression (Schwab et al., 2006).

5.3. Protein-protein interactions

Proteins are key components essential for all biological processes in a cell. The traditional approach to understand protein function is to define the activities of purified proteins. We now know that the cell is a compilation of networks and many proteins function only when associated with partners or as components of large multiprotein complexes (Gingras et al., 2007). The complete sequence of a number of eukaryotic genomes, including humans (International_Human_Genome_Sequencing_Consortium, 2001), Arabidopsis (The_Arabidopsis_Genome_Initiative, 2000), rice (International_Rice_Genome_Sequencing_Project, 2005), and many others, provides an estimation of the primary sequences and the number of proteins encoded by each genome. However, these sequences do not reveal the functional linkages and the complexity of the proteome. In order to explain all cellular functions, it is necessary to understand how these proteins assemble into complexes and to form the molecular machines that perform the various functions essential for cell viability (Eisenberg et al., 2000; Stelzl et al., 2005). Thus, many proteins properly function only when bound to other cellular components such as nucleic acids, phospholipids, and other proteins. All these protein interactions are connected into an extensive network to cover the diversity of biological functions, from a limited number of original components (Gavin et al., 2002; Ideker et al., 2001; Ito et al., 2001). For instance, plant proteomes that contain approximately 40,000 proteins (Sterck et al., 2007) are estimated to have an 75,000-150,000 interaction pairs (Morsy et al., 2008). Based on an understanding of this dynamic and complex interacting network, the functions of proteins in a cell can ultimately be defined by how their biochemical activities are regulated and utilized in related biological processes (von Mering et al.,

2002). Moreover, many cell functions are carried out by multi-protein complexes that consist of more than one protein assembly, which can be referred to as a set of large protein machines resulting from the interactions between protein assemblies (Eisenberg et al., 2000; Ideker et al., 2001). Since protein-protein interactions are so critical, it is essential to appreciate their diversity and functionality in order to have a reasonable understanding of cellular metabolism.

Although protein-protein interactions can be classified in several different ways based on their distinctive interaction patterns, they can be generally categorized into two major groups: stable and transient interactions. Transiently interacting proteins can be easily removed from a protein complex during the purification process, whereas stably interacting protein complexes are relatively straightforward to recover and analyze (Bauch and Superti-Furga, 2006; Kaake et al., 2010). While they are more difficult to detect, transient interactions are important contributors to cellular function since many of cellular processes are controlled by these interactions (Nooren and Thornton, 2003b). Transient interactions can be strong or weak, exist in a temporally or spatially specific manner, and can be affected by conditions that promote or modify the interaction. These transient protein complexes are thought to be involved in various protein processes including protein folding, transport, and signaling (Nooren and Thornton, 2003a; Nooren and Thornton, 2003b).

Protein-protein interactions can be analyzed by a number of experimental methods including: the yeast two-hybrid system (Fields and Song, 1989), β-galactosidase complementation (Rossi et al., 1997), split ubiquitin system (Dunnwald et al., 1999), G-protein fusion system (Ehrhard et al., 2000), optical spectroscopy (Lakey and Raggett,

1998), tandem affinity purification (TAP) (Gavin et al., 2002), immunoprecipitation (Williams, 2000), glutathione S-transferase (GST)-pull down (Smith and Johnson, 1988), fluorescence resonance energy transfer (FRET) (Ozawa et al., 2001), bioluminescent resonance energy transfer (BRET) (Xu et al., 1999), and bi-molecular fluorescence complementation (BiFC) (Hu et al., 2002).

Among these, this research relied on the yeast two-hybrid system, fluorescence based methods (BiFC), and epitope tag-based affinity purifications (TAP, immunoprecipitation, GST-pull down, etc) for protein-protein interaction studies. Each method will be further described below.

5.3.1. Yeast two-hybrid

The yeast two-hybrid system was introduced in 1989, and it has been widely adopted and adapted for large scale analyses in many organisms such as viruses, bacteria, and eukaryotes (Fields and Song, 1989). It exploits the two distinct domains present in the transcription factors: a DNA-binding domain and a transcription activation domain. The DNA-binding domain targets the activator to a specific *cis* element, and the activation domain contacts other proteins of the transcriptional machinery to enable transcription to occur. In the two-hybrid system, one open reading frame (ORF) is fused to a DNA-binding domain (bait) and the other ORF is fused to an activation domain (prey) and expressed in a yeast strain containing a reporter gene downstream of the target *cis* element. If the bait and prey fusion proteins are expressed in this yeast host strain and interact, functional transcription factor is produced, which triggers transcription of the

reporter gene. The expression of the reporter gene can be detected histochemically and used to score for the interaction of the bait and prey proteins (Fields and Song, 1989). The most commonly used reporter gene is the *Escherichia coli* ß-galactosidase (*lacZ*) gene (Fields and Song, 1989), and other reporters include the yeast histidinol-phosphate aminotransferase (*HIS3*) (Durfee et al., 1993) or the yeast leucine biosynthesis gene (*LEU2*) (Zervos et al., 1993). The yeast two-hybrid system has also been successfully adapted for other organisms including *E. coli* (Bartel et al., 1996), *C. elegans* (Walhout et al., 2000), *Helicobacter pylori* (Rain et al., 2001), mammalian cells (Dang et al., 1991), and hepatitis C virus (HCV) (Flajolet et al., 2000).

There are two key limitations to the two-hybrid assay for protein-protein interactions. The detection of the interactions is restricted in the nucleus and thus preventing the detection of interactions with proteins in other subcellular compartments. Therefore, this system is not suitable to detect interactions of membrane proteins (Ito et al., 2002). In addition, this system is known to miss a large portion of true interactions and produces a considerable number of false positives (Uetz, 2002; von Mering et al., 2002). This is particularly likely in the case of the detection of plant protein interactions that require post-translational modifications or other cofactors. In this case, the two-hybrid system can easily miss true interactions, since yeast may not contain the specific factor(s) that are present in plants.

5.3.2. Pull-down assays

There are several types of pull-down assays available which allow the recovery of protein complexes using an antibody specific for one component of a complex. If a

polyclonal or monoclonal antibody is available for a target protein, one of the conventional methods, immunoprecipitation (IP), can be the simplest method to test the interaction between two proteins (Williams, 2000). To perform this technique, an antibody specific for a component of the complex and a protein A-conjugated resin that has high affinity for immunoglobulin G (IgG) are added to cell lysates expressing a target protein. Binding of the IgG-conjugated protein A resin to the target protein allows its recovery from the extract, possibly in association with its protein interactors. Then, the components of the complex can be analyzed by various methods depending on the type of information desired. Native gel electrophoresis and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) can be used to detect the molecular weight (MW) of individual components. Immunoblotting can be used to distinguish individual components using an antibody specific for another possible complex component; mass spectrometry of the precipitate may identify other protein components.

If no antibody is available for one of the components of a putative complex then it may be possible to add an epitope tag to a target protein of interest. The epitope-tagged recombinant protein is prepared in a bacterial host (most often *E .coli*) by overexpression and purification. Commonly used protein tags are glutathione S-transferase (GST) (Smith and Johnson, 1988), 6x histidine-tag (His-tag) (Hochuli et al., 1987), the Myc epitope or the Flag-tag (Hopp et al., 1988), as well as the maltose-binding protein (Bedouelle and Duplay, 1988). Among these, GST and His-tags have been most widely used for pull-down assays because of the availability of commercial vectors and antibodies that facilitate the easy and rapid cloning of ORFs and the detection of the fusion protein. With these tags, it is possible to test any target protein of interest by using proper resins that

allow interacting partners to be efficiently recovered and eluted. However, adding these tags to proteins can affect their activities or folding that may cause interruption of certain protein interactions (Arnau et al., 2006).

Tandem affinity purification (TAP) can be used on either a small or large scale to detect components of multi-protein complexes. For example, 232 distinct protein complexes were examined in yeast (Gavin et al., 2002) and 40 novel 14-3-3-binding proteins were identified in mice (Angrand et al., 2006).

This technique is based on two sequential steps of affinity purification using two specific fusion tags: protein A and calmodulin binding protein (Gavin et al., 2002). The first step involves a gentle purification of protein complexes linked to the TAP-tagged fusion protein (bait protein) by IP with protein A beads; the complex is subsequently released from the protein A beads by cleavage with the tobacco etch virus (TEV) protease at a recognition site located between the two tags. A second purification step uses calmodulin binding beads to reduce non-specific protein recovery. The final protein complex is ultimately released by ethylene glycol tetraacetic acid (EGTA) treatment (Gavin et al., 2002; Rigaut et al., 1999). The components of the recovered proteins can be analyzed using the same methods as outlined for immunoprecipitation.

Although TAP tagging is a relatively rapid and simple method for detecting protein-protein interactions, many previous studies have shown that this method is not sufficient to examine weak protein interactions or low-abundant proteins in complex eukaryotic cells. Its two-step purification procedure with a relatively large size of TAP tag may cause incomplete release of protein from the column; furthermore protein degradation

may result during TEV protease cleavage (Drakas et al., 2005; Witte et al., 2004). Nevertheless, with a number of modifications or improvements, this TAP method has been successfully used in various organisms such as *Drosophila* (Veraksa et al., 2005), *E. coli* (Gully and Bouveret, 2006), mammalian cells (Drakas et al., 2005), and plants (Rohila et al., 2004; Rohila et al., 2006; Rubio et al., 2005).

Recently, an epitope-tag based purification technique for plants was introduced. This involves the addition of an eight-amino acid (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) StrepII epitope tag to the protein of interest, followed by one-step purification (Witte et al., 2004). The StrepII epitope tag binds to the biotin binding site of streptavidin (StrepTactin™, Novagen) and is efficiently eluted from the resin by washing with biotin or desthiobiotin (Korndorfer and Skerra, 2002). The small size of the StrepII tag provides a quick and effective protein purification but suffers from the possibility of co-purifying transient protein interactors (Witte et al., 2004).

5.3.3. Bimolecular fluorescence complementation

Bimolecular fluorescence complementation (BiFC) is a recently developed technique for monitoring protein-protein interactions in living cells using fluorescent proteins. The BiFC assay is based on the reconstitution of two non-fluorescent fragments of a yellow fluorescent protein (YFP) (Hu et al., 2002). In this assay, two split fragments of the YFP, N-terminal and C-terminal of YFP, are each fused to potential interacting partners. If the fusion proteins come within approximately 100 Å̊ngstroms of each other, the YFP is reconstituted sufficiently to produce fluorescence which is commonly detected by

fluorescent or confocal microscopy (Hu et al., 2002; Kerppola, 2006). Thus, the reconstitution of a YFP fluorescent signal provides direct evidence not only for interactions between the partners but often the subcellular location where the interaction takes place.

BiFC was first introduced to examine interactions between the basic-region leucine zipper (bZIP)-domain and Rel-family proteins in mammalian cells (Hu et al., 2002). Since then, this technique has been extensively used for other protein-protein interaction studies such as the dimerization of Myc-, Max-, and Mad-family transcription factors (Grinberg et al., 2004), and the complex formation of SMAD-family transcription factors for transforming growth factor (TGF) β signaling (Remy et al., 2004).

The animal cell based BiFC vectors had to be modified for use in plants. In order to allow expression of two split fragments of YFP in plants, the corresponding sequences are cloned between the cauliflower mosaic virus promoter (CaMV 35S) and the nopaline synthase terminator (NOS). The entire expression cassettes are subsequently subcloned into T-DNA base plant binary vectors (Bracha-Drori et al., 2004). These have since been used for many studies. For example, this method was used to analyze the dimer formation of Arabidopsis transcription factors and a tobacco 14-3-3 protein (Walter et al., 2004), and to identify the interaction between a host protein and a viral protein in tobacco (Caplan et al., 2008).

Several recent improvements have been made to these vectors. New vectors have been developed to aid the cloning and transformation steps (Citovsky et al., 2006; Gehl et al., 2009); other vectors facilitate the use of BiFC for protoplast transient expression (Chen et

al., 2006). Finally a multicolor BiFC system that allows detection of multiple protein interactions in a single cell is now available (Waadt et al., 2008).

Not surprisingly, BiFC also suffers from possible artifacts. The association kinetics of the YFP fragments and the low reversibility of YFP fluorophore formation may result in the failure of capturing the changes in the protein association state (Hu et al., 2002). Moreover, depending on the kinetics of transient protein-protein interactions, the reconstituted YFP signal can be artificially stabilized, resulting in the production of false-positive signals (Fricker et al., 2006; Walter et al., 2004; Zamyatnin et al., 2006; Zhong et al., 2008). Previous studies have shown that BiFC fluorescence can be produced by the self-association between the two YFP fragments without specific interacting partners in plant cells (Zamyatnin et al., 2006; Zhong et al., 2008). Addition of tag (N- or C- terminal fragment of YFP) or even its orientation when fused to the proteins of interest may interfere with their potential interactions, resulting in the weak reconstituted fluorescent signal. In addition, special care must be taken when BiFC is used in plant cells because of the chlorophyll and cell wall autofluorescence that may cause difficulty to differentiate from the true reconstituted fluorescence (Bracha-Drori et al., 2004).

Nevertheless, BiFC is powerful method allowing quick and easy detections of fluorescent complex in the living cells without specialist equipment or software for complex ratio-metric quantification of signal intensities required for FRET (Ozawa et al., 2001) and BRET (Xu et al., 1999). BiFC can be performed using a conventional epifluorescence microscopy that is a relatively inexpensive compared to confocal microscopy (Citovsky et al., 2008; Fricker et al., 2006).

6. Objectives of this research

The research outlined in this thesis examines the subcellular localization and the protein interactions of the key enzymes mediating methyl recycling in plants. Several research objectives were addressed:

- Determine the subcellular localization of ADK and SAHH in Arabidopsis plants.
- Define the role of the insertion region (IR) of SAHH in subcellular targeting.
- Determine if the dynamic movement and trafficking of these methyl recycling enzymes between multiple subcellular compartments occurs via protein-protein interactions.
- Discover any other proteins or complexes involved in these protein-protein interactions to address other functions of methyl recycling enzymes.
- Create target-specific gene silencing lines against both *ADK* and *SAHH* genes using an artificial microRNA to study their functional significance in plant development.

CHAPTER 1: SUBCELLULAR LOCALIZATION OF METHYL RECYCLING ENZYMES

1. Introduction

S-adenosyl-L-methionine (SAM), a small molecule synthesized from L-methionine and ATP, is a major methyl donor for all biological methylation reactions (Lu, 2000). In the activated methyl cycle, a steady supply of SAM is essential to maintain the activity of methyltransferases (MTs; EC 2.1.1) that transfer a methyl group from SAM to methyl-accepting compounds (Poulton, 1981). These SAM-dependent MT activities are crucial for modification or biosynthesis of various compounds including DNA, RNA, protein, pectin, lignin, choline, as well as plant growth regulators (Moffatt and Weretilnyk, 2001). Upon each transmethylation reaction, one molecule of *S*-adenosyl-L-homocysteine (SAH), which is a potent inhibitor of MTs, is produced from SAM as a by-product (Poulton, 1981). The two key enzymes involved in SAH metabolism in plants are *S*-adenosyl-L-homocysteine hydrolase (SAHH; EC 3.3.1.1) and adenosine kinase (ADK, adenosine 5'-phosphotransferase; EC 2.7.1.20) (Moffatt et al., 2000). SAHH breaks down SAH into L-homocysteine (Hcy) and adenosine (Ado) (De La Haba and Cantoni, 1959), and ADK catalyzes the conversion of the by-product Ado into adenosine monophosphate (AMP) by phosphorylation using one molecule of ATP (Moffatt et al., 2000).

Since SAM-dependent transmethylation reactions occur throughout the cell, SAH is also generated in these same locations (Poulton, 1981). Thus, the SAH produced in various subcellular compartments must be metabolized to maintain these methylation

reactions. However, both ADK and SAHH have been widely thought to be as cytosolic enzymes, since the primary location of methyl recycling reactions is the cytoplasm, where the methyl donor SAM is synthesized (Hanson et al., 2000; Hanson and Roje, 2001; Ravanel et al., 1998).

To sustain transmethylation activities occurring in different subcellular compartments locations, two possible solutions can be considered. The first possibility is that SAH metabolism occurs in the cytoplasm and SAH produced in different locations is transported into the cytoplasm where ADK and SAHH reside. The other possibility is that both ADK and SAHH reside in multiple compartments.

The transport of SAH into the cytoplasm is evidenced by the presence of a SAM transport system (or a SAM/SAH exchanger) in mitochondria and plastids. The SAM transport system was first found in mitochondria of rat liver, in which cytosolic SAM is taken up into mitochondria via a carrier-mediated system (Horne et al., 1997). Later, other mitochondrial transporters were identified in yeast (Marobbio et al., 2003) and human cells (Agrimi et al., 2004). In both of these cases, the transporters exchange cytosolic SAM for mitochondrial SAH thereby maintaining the SAM/SAH ratio within the organelle. In plants, the study of the uptake of cytosolic SAM by isolated spinach (*Spinacia oleracea*) chloroplasts first demonstrated the presence of a specific transport system, which exchanges SAH produced in plastids for SAM synthesized in the cytoplasm (Ravanel et al., 2004). More recently, similar SAM transporters were identified and characterized in both Arabidopsis plastids (Bouvier et al., 2006) and mitochondria (Palmieri et al., 2006). In both cases, the transporters are capable of exchanging SAM/SAH between the cytoplasm and the organelle. In particular, gene

silencing of the plastid transporter gene *SAMT1* in Arabidopsis and tobacco (*Nicotiana benthamiana*) plants affect the chlorophyll pathway with decreased prenyllipids and plastid pigments (Bouvier et al., 2006). It is proposed that this may due to decreased SAM flux into plastids via the limited plastid transporter, which disrupts a key methylation reaction in the chlorophyll pathway (Bouvier et al., 2006).

On the other hand, SAH metabolism in other organelles is supported by a few reports on nucleus or chloroplast localization studies of ADK and SAHH. Since various essential processes such as chromatin modification (Baubec et al., 2010) and mRNA capping (Shafer et al., 2005) take place in the nucleus and these processes rely on transmethylation, both ADK and SAHH need to be capable of entering the nucleus to maintain these reactions. A recent study examining the Arabidopsis SAHH mutant *hog1-7* as well as the effects of an SAHH inhibitor, dihydroxypropyladenine (DHPA), clearly demonstrates that SAHH plays a crucial role in the maintenance of chromatin structure by DNA/histone methylation in the nucleus (Baubec et al., 2010). Additional evidence for the nuclear localization of SAHH is demonstrated by its role in mRNA cap methylation (Radomski et al., 1999). This research shows that *Xenopus* SAHH (xSAHH) localizes to the nucleus of *Xenopus leavis* oocytes during times of increased transcription (Radomski et al., 1999), and the movement of xSAHH from the cytoplasm to the nucleus occurs during the maturation period of the oocytes (Radomski et al., 2002). Mutation analysis demonstrates that both N- and C-termini of xSAHH are required for its efficient nuclear translocation (Radomski et al., 1999). This xSAHH translocation is mediated by a specific protein interaction with mRNA (guanine-7-) methyltransferase (Radomski et al., 2002).

A recent study of the presence of ADK in the nucleus was verified by the immunofluorescence labeling study using various mammalian cell lines including HT-1080, HeLa, LMTK, and CHO cells (Cui et al., 2009). Mammalian cells express two different versions of ADK, a "long form" ADK that contains an additional 20 amino acids at its N-terminus and a "short form" ADK, lacking this extension. The additional N-terminal sequence of human ADK functions as a nuclear localization signal sequence (Cui et al., 2009).

The evidence for chloroplast localization of both ADK and SAHH is also suggested by previous studies. For example, the enzymatic activity assay using fractionated leaf samples showed that the SAHH activity is predominantly detected in the chloroplast fraction of tea leaves (Koshiishi et al., 2001). The presence of ADK in the chloroplast was also reported by two proteomics studies under different growth conditions: in the chloroplast of Arabidopsis grown under normal growth conditions (Kleffmann et al., 2004) and in the chloroplast of the green alga *Dunaliella salina* grown under high salinity conditions (Liska et al., 2004). In addition, earlier work performed in the Moffatt's lab showed that the presence of both ADK and SAHH in the chloroplasts by immunogold labeling, β-glucuronidase (GUS) fusions, and immunoblots of purified chloroplast extracts, although these results are not consistent, nor conclusive. However, results obtained from the earlier work provides more convincing evidence for their localization in the nucleus and the cytoplasm (Pereira, 2004; Schoor, 2007).

Taken together, all of this evidence suggests that both enzymes ADK and SAHH may exist in multiple compartments. This research aims to clarify this for Arabidopsis and examine the possible mechanisms involved.

2. Materials and Methods

2.1. Chemicals and reagents

All chemicals used in the study outlined in this thesis were purchased from Sigma-Aldrich (Oakville, Canada), Bioshop (Burlington, Canada), or BioBasic (Markham, Canada), unless otherwise indicated. Restriction enzymes (REs) and T4 DNA ligase were purchased from New England Biolabs (Pickering, Canada) or Fermentas (Burlington, Canada). Taq polymerase used for colony PCR was generated from Moffatt lab and the one used for cloning was purchased from New England Biolabs (Finnzyme's Phusion DNA Polymerase) or Invitrogen (High Fidelity DNA polymerase; Burlington, Canada). All primers used were obtained from Sigma-Aldrich. All constructs obtained were confirmed by sequencing (The Centre for Applied Genomics (TCAG); http://www.tcag.ca/; Toronto, Canada).

2.2. Plant material and growth conditions

Arabidopsis thaliana (accession Columbia) seeds were grown on soil or agar media. For growth on soil, seeds were sown on a 50:50 soil mixture of Sunshine LC1 Mix and Sunshine LG3 Germination Mix (SunGro, Canada) in individual pots (25 cm^2) or large pots (530 cm^2). These pots placed in flats (1600 cm^2) were covered with plastic domes and incubated in the dark at 4°C for 3 days to synchronize seed germination, then transferred to a growth chamber under long day conditions with 16 hours light (150 ± 20 $\mu mol\ m^{-2}\ s^{-1}$ photo-synthetically active radiation [PAR]) at 22°C. The plants were watered

every two days and fertilized with 20:20:20 fertilizer (Plant-Prod, Plant Products Co Ltd, Canada) once a week until they reached maturity (about 6 weeks). For growth on media, about 0.1 mL of seeds in an open microfuge tube was sterilized using chlorine gas produced by adding 3 mL of concentrated HCL to a beaker containing 100 mL of 3-6% (v/v) sodium hypochlorite (Bleach) in a desiccator jar in a fume hood with a vacuum condition for 1 hour. Sterile seeds were then sown on 1/2 Murashige-Skoog (MS) salt media (Murashige and Skoog, 1962) supplemented with B5 vitamins (2.5 mM 4-morpholineethanesulfonic acid (MES), 30 g/L sucrose, 8 g/L agar, pH 5.7-5.8). The plates (140 mm diameter x 20 mm deep) were sealed and incubated for 24-72 hours in the dark at 4°C, then transferred to a tissue culture chamber (TC7, Conviron) with 24 hour light (150 ± 20 µmol m^{-2} s^{-1} PAR) at 20°C.

2.3. Vector constructions for ADK1 and SAHH1 localization

Standard molecular biology techniques were used to create fusion constructs. To examine subcellular localizations of ADK and SAHH, pSAT6-EGFP-N1 or pSAT6-EGFP-C1 and pSAT6-DsRed2 vectors were used (Tzfira et al., 2005) (Appendix I). These vectors contain the constitutive cauliflower mosaic virus (CaMV) 35S promoter, followed by the tobacco etch virus (TEV) leader to enhance translation efficiency and the CaMV 35S terminator. For green and red fluorescent tags, the enhanced GFP (EGFP) and the DsRed2 (modified red fluorescent protein from *Discosoma sp.*) were used, respectively. For plant transformation, the entire expression cassette of pSAT6 vectors were subcloned into PI-*Psp*I rare cutting intron-encoded endonuclease recognition site of

the T-DNA region of pPZP-RCS2 binary vector (Tzfira et al., 2005), which carries plant selectable marker coding for resistance to the herbicide bialaphos (*bar* gene).

For construct pSAT6-ADK1-EGFP, the coding sequence for ADK1 was amplified using a primer set (AK1cDNcF and ADK1cR-Sma; Table 1). The PCR product was digested with *Nco*I and *Xma*I and cloned into the pSAT6-EGFP-N1 vector. For pSAT6-EGFP-ADK1, the *ADK1* gene was amplified using a primer set (ADK1cF-Kpn and ADK1cR-Sma; Table 1), digested with *Kpn*I and *Xma*I, and cloned into the pSAT6-EGFP-C1 vector. For pSAT6-ADK1p:ADK1-EGFP, the 1-kb upstream promoter region of the *ADK1* gene was amplified using a primer set (SAHH2pF-Age and ADK1 lux lower; Table 1). The PCR product was digested with *Age*I and *Nco*I and cloned into the pSAT6-ADK1-EGFP vector to substitute the 35S promoter with the *ADK1* native promoter. For pSAT6-ADK1-(no linker)-EGFP, the coding sequence for ADK1 was amplified using a primer set (AK1cDNcF and AK1cDNcR; Table1) and cloned into the *Nco*I site of the pSAT6-EGFP-C1 vector. For pSAT6-ADK1-(24aa linker)-EGFP, the full-length *ADK1* coding sequence was obtained from pSAT6-ADK1-(no linker)-EGFP by digestion with *Nco*I and subsequently inserted into the pSAT6-EGFP-N1 vector.

For construct pSAT6-SAHH1-EGFP, the coding sequence for SAHH1 was amplified using a primer set (SAHH1cDNcF and SAHH1cR-Sma; Table 1). The PCR product was digested with *Nco*I and *Xma*I and cloned into the pSAT6-EGFP-N1 vector. For pSAT6-EGFP-SAHH1, the *SAHH1* gene was amplified using a primer set (SAHH1cF-Kpn and SAHH1cR-Sma; Table 1), digested with *Kpn*I and *Xma*I, and cloned into the pSAT6-EGFP-C1 vector. For pSAT6-SAHH1p:SAHH1-EGFP, the 1-kb upstream promoter region of the *SAHH1* gene was amplified using a primer set (SAHH1pF-Age and

SAHH1Nc-R; Table 1). The PCR product was digested with *Age*I and *Nco*I and cloned into the pSAT6-SAHH1-EGFP vector to substitute the 35S promoter with the *SAHH1* native promoter. For pSAT6-SAHH2-EGFP, the coding sequence for SAHH2 was amplified using a primer set (SAHH2for and SAHH2cR-Sma; Table 1). The PCR product was digested with *Nco*I and *Xma*I, and then cloned into a pSAT6-EGFP-N1 vector. For pSAT6-SAHH2p:SAHH2-EGFP, the 1.2-kb upstream promoter region of the *SAHH2* gene was amplified using a primer set (SAHH2pF-Age and SAHH2 Nc-R; Table 1). The PCR product was digested with *Age*I and *Nco*I and cloned into a pSAT6-SAHH2-EGFP vector to substitute the 35S promoter with the *SAHH2* native promoter.

To verify the nuclear localization of ADK1 and SAHH1, double GFP constructs were generated in the basic pSAT6 vector by replacing the EGFP with *Nco*I and *EcoR*I fragments of 2xGFPs from the 3x sGFP vector (Kim et al., 2005). This ligation resulted in the pSAT6-2xGFP vector. For pSAT6-ADK1-2xGFP and pSAT6-SAHH1-2xGFP, the coding sequence for ADK1/SAHH1 was cloned into *Nco*I site of the pSAT6-2xGFP vector. The expression cassette of all pSAT6 vectors was inserted into the PI-*Psp*I site of the pPZP-RCS2 binary vector for plant transformation.

2.4. Cloning of yeast *SAHH* and control vectors for plastid and nuclear localization

To clone yeast *SAHH* gene, the 1,359-bp DNA fragment containing the coding region of yeast SAHH (P39954) was isolated from yeast genomic DNA by PCR reaction using a primer set (ySAHH-F1 and ySAHH-R1; Table 1). The PCR product was subcloned into

pJET1.2 (Fermentas Canada Inc.) and used for further cloning after confirming sequence integrity by sequencing. For pSAT6-ySAHH-EGFP, the full-length yeast *SAHH* gene was PCR amplified from a pJET-yeastSAHH vector using a primer set (ySAHH-F Xho and ySAHH-R Xma; Table 1) and inserted into the *Xho*I and *Xma*I site of the pSAT6-EGFP-N1 vector. For pSAT6 -EGFP-ySAHH, the coding sequence for yeast SAHH was amplified using a primer set (ySAHH-F Xho2 and ySAHH-R Xho; Table 1). The PCR product was digested with *Xma*I and cloned into the pSAT6-EGFP-C1 vector.

For a nuclear targeting control construct pSAT6-SV40NLS-DsRed2, the NLS sequence of Simian virus 40 (SV40) large T antigen (PKKKRKVEDP) (Goldfarb et al., 1986) was PCR amplified using a primer set (NLS-DsRed2F and DsRed2R-Nco; Table 1). The amplified fragments were cloned into the *Nco*I site of the pSAT6-DsRed2-N1 vector, which resulted in the pSAT6-SV40NLS-DsRed2. To allow co-expression of pSAT6-NLS-DsRed2 with either SAHH1-2xGFP or ADK1-2xGFP fusion in a single vector, the *Age*I and *Not*I fragment of the pSAT6-SV40NLS-DsRed2 was substituted with that of pSAT4 basic vector to create pSAT4-SV40NLS-DsRed2. The expression cassette of this pSAT4 vector was then transferred into I-*Sce*I unique recognition site of the pPZP-RCS2-ADK1-2xGFP or pPZP-RCS2-SAHH1-2xGFP for plant transformation.

For a control construct for plastid localization, the small subunit of rubisco complex (*RbcS*) gene (At1g67090) was isolated from a cDNA prepared from wild-type Arabidopsis plants. For this, the full-length coding region of *RbcS* (543-bp) was PCR amplified using a primer set (SSU-F-Bam and SSU-R-Bam; Table 1) and inserted into the *Bam*HI site of pSAT6-EGFP-N1 or pSAT6-DsRed2-N1 vector. For co-expression of RbcS-DsRed2 with other EGFP fusion vectors in a single vector, the 1,041-bp DNA

fragment containing the sequence of *RbcS* coding region along with a part of DsRed2 sequence was inserted into the *Nco*I site of pSAT4-SV40NLS-DsRed2 to replace SV40NLS with RbcS. The expression cassette of the pSAT4-RbcS-DsRed2 vector was then inserted into the I-*Sce*I site of pPZP-RCS2-ADK1-EGFP or pPZP-RCS2-SAHH1-EGFP for plant transformation.

2.5. Deletion analysis for SAHH1 localization

EGFP fusion constructs for SAHH1 deletion analysis were prepared by multiple PCR reactions. For pSAT6-SAHH1(Δ1-28)-EGFP, the 1377-bp of *SAHH1* gene fragment was amplified using a primer set (SAHH1-D29F and SAHH1cR-Sma; Table 1) and inserted into the *Nco*I and *Xma*I sites of the pSAT6 vector. For pSAT6-SAHH1(Δ1-62)-EGFP, the 1275-bp *SAHH1* gene fragment was amplified using a primer set (SAHH1-D63F and SAHH1cR-Sma; Table 1) and inserted into the *Nco*I and *Xma*I sites of the pSAT6 vector. For pSAT6-SAHH1(Δ411-485)-EGFP, the 1230-bp of *SAHH1* gene fragment was amplified using a primer set (SAHH1cDNcF and SAHH1-D410R; Table 1) and inserted into the *Nco*I and *Xma*I sites of the pSAT6 vector. For pSAT6-SAHH1(Δ150-190)-EGFP, the 447-bp of 5' coding sequence for SAHH1 was amplified using a primer set (SAHH1cDNcF and SAHH1-D149R; Table 1) and cloned into the *Nco*I and *Pst*I sites of the pSAT6 vector to generate a N-terminal part of SAHH1 fusion (pSAT6-SAHH(1-149)-EGFP). This vector was then ligated with the 885-bp of *Pst*I and *Nco*I fragment (C-terminal part of SAHH1; residues 191-485) that was amplified using a primer set (SAHH1-D191F and SAHH1cR-Sma; Table 1) from a *SAHH1* gene sequence. For

pSAT6-SAHH1(Δ150-190)-2xGFP, the SAHH1(Δ150-190) fragment obtained from the pSAT6-SAHH1(Δ150-190)-EGFP vector was inserted into the *Nco*I/*Xma*I sites of the pSAT6-2xGFP vector. For pSAT6-SAHH1(Δ1-28, Δ411-485)-EGFP, the 1152-bp *SAHH1* gene fragment was amplified using a primer set (SAHH1-D29F and SAHH1-D410R; Table 1) and cloned into *Nco*I/*Xma*I sites of the pSAT6 vector. For pSAT6-SAHH1(Δ1-62, Δ411-485)-EGFP, the 1050-bp of *SAHH1* gene fragment was amplified using a primer set (SAHH1-D63F and SAHH1-D410R; Table 1) and cloned into *Nco*I/*Xma*I sites of the pSAT6 vector. For pSAT6-SAHH1(Δ1-62, Δ411-485)-2xGFP, the 1044-bp of *SAHH1* gene fragment was amplified using a primer set (SAHH1-D63F and SAHH1-410RN; Table 1) and inserted into the *Nco*I site of the pSAT6-2xGFP vector. For pSAT6-SAHH1(150-190)-2xGFP, the 126-bp of *SAHH1* gene fragment was amplified using a primer set (SAHH1-149F and SAHH1-190R; Table 1) and inserted into the *Nco*I site of the pSAT6-2xGFP vector.

In order to create mutations within the insertion region (IR) of SAHH, several PCR reactions were performed using specific primers that allow amplification of the mutated *SAHH1* gene sequence. For pSAT6-SAHH1(Δ150-177)-EGFP, the 924-bp of 3'-coding sequence for SAHH1 (residues 178-485) was amplified using a primer set (SAHH D178-F and SAHH1cR-Sma; Table 1) and cloned into the *Pst*I and *Xma*I sites of the pSAT6-SAHH(1-149)-EGFP vector. For pSAT6-SAHH1(Δ178-190)-EGFP, the 531-bp of 5'-coding sequence for SAHH1 (residues 1-177) was amplified using a primer set (SAHH1cDNcF and SAHH D177-R; Table 1) and cloned into the *Nco*I and *Pst*I sites of the pSAT6-SAHH1(Δ150-190)-EGFP vector.

To substitute three amino acids, Pro^{163}-Asp^{164}-Pro^{165} with three alanine residues, the 521-bp partial fragment of *SAHH1* gene containing a PDP163AAA mutation (SAHH1ΔPDP) was amplified using a primer set (SH1-PDPto3A and SAHH1-inR; Table 1). The PCR product was then replaced the wild-type *SAHH1* gene sequence with SAHH1ΔPDP by inserting the PCR fragment into the *Bgl*II and *Age*I sites of the pXCS-SAHH1-Strep vector (see Chapter 2). To mutate Thr^{166}-Ser^{167}-Thr^{168}, the 521-bp partial fragment of *SAHH1* gene containing a TST166AAA mutation (SAHH1ΔTST) was amplified using a primer set (SH1-TSTto3A and SAHH1-inR; Table 1). The PCR product was inserted into the *Bgl*II and *Age*I sites of the pXCS-SAHH1-Strep vector to replace the wild-type *SAHH1* gene sequence with SAHH1ΔTST. To create GFP fusion constructs with both SAHH1 mutations, the 1176-bp partial DNA fragment cut with *Xho*I from pXCS-SAHH1ΔPDP-Strep or pXCS-SAHH1ΔTST-Strep vector was replaced with the 1176-bp *Xho*I fragment of pSAT6-SAHH1-EGFP vector to generate pSAT6-SAHH1ΔPDP-EGFP and pSAT6-SAHH1ΔTST-EGFP, respectively.

The expression cassette of all SAHH-deletion or substitution fusion vectors was then transferred into PI-*Psp*I site of pPZP-RCS2 binary vector for plant transformation.

2.6. Arabidopsis transformation

Transformation of Arabidopsis plants was carried out by the floral dip method (Clough and Bent, 1998). To do this, *Agrobacterium tumefaciens* GV3101 (Koncz and Schell, 1986) strain carrying the transgene was grown overnight in 3 mL of LB media containing appropriate antibiotics at 28℃. This culture was diluted 100-fold in 200 mL fresh LB

medium and incubated for further 16 hours. Bacterial cells were then collected by centrifugation, and the pellet was resuspended in 5% sucrose and 0.05% Silwet L-77 (Lehle Seeds, Cat. No. VIS-01) at an OD_{600} of 0.8. The prepared solution was used to dip the floral part of 4-5 week-old Arabidopsis plants. After floral dipping, the plants were covered with plastic dome to maintain high humidity and kept at room temperature for 24 hours in the dark condition. The plants were then transferred back to the growth chambers until seed set. The seeds collected from dipped-plants were sown and selected on soil by spraying 0.067% (v/v) (40 mg/mL) of the herbicide glufosinate ammonium (Basta; "Wipeout" from Wilson Laboratories Inc.; Dundas, Canada), when the plants have fully expanded first leaves (~2 weeks). The "Wipeout" was applied once a day until non-transformed seedlings turned yellow (1-2 weeks). The Basta-resistant T1 plants with healthy true leaves were transplanted into individual pots.

2.7. Transient expressions in tobacco plants

Transient expression assays of fluorescent fusion proteins were performed using tobacco (*Nicotiana benthamiana*) plants as described previously (Voinnet et al., 2003). For this, *A. tumefaciens* strains carrying the transgene were grown overnight in 3 mL of LB media containing appropriate antibiotics at 28℃. This culture was diluted 100-fold in 6 mL fresh LB media supplemented with 10 mM MES (pH 5.6), 20 μM acetosyringone, and appropriate antibiotics, then incubated for further 16 hours. Bacterial cells were then collected by centrifugation and resuspended in 10 mM $MgCl_2$, 10 mM MES, 150 μM acetosyringone at an OD_{600} of 1.0. The resuspended cells were mixed with equal volume

of a culture of P19 strain at an OD_{600} of 1.0. The P19 strain expresses the P19 protein of tomato bushy stunt virus, which acts as a suppressor of post-transcriptional gene silencing in plants (Voinnet et al., 2003). The mixtures were incubated at room temperature for at 2-3 hours and pressure infiltrated onto the underside of 4-6 weeks old tobacco leaves. Fluorescent signals were monitored 3-4 days after infiltration using confocal laser scanning microscope.

2.8. Protoplast isolation and transfection

Transient expression in Arabidopsis protoplasts was carried out using the PEG-mediated transfection method (Sheen, 1996). Protoplasts were isolated from 3-4 weeks old Arabidopsis leaves grown on ½ MS media. Leaf samples were incubated for 3 hours at 22℃ in darkness in a 50 ml tube containing 1% (w/v) cellulase R-10 (Yakult Honsha, Japan), 0.25% (w/v) macerozyme R-10 (Calbiochem, Germany), 20 mM MES, 0.1% (w/v) bovine serum albumin (BSA), 400 mM mannitol, 10 mM $CaCl_2$, 20 mM KCl, and KOH (pH 5.6). The digested sample mixture was filtered through a 100 μm nylon mesh and transferred to a tube containing 21% (w/v) sucrose. To remove the dead protoplasts and cellular debris, the tube was centrifuged at 98 x g for 10 minutes and floating intact protoplasts at the top part of the solution were transferred to a tube containing W5 solution (154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, 5 mM glucose, 1.5 mM MES, KOH pH5.6). Protoplasts were carefully washed twice with W5 solution and resuspended in 400 mM mannitol, 15 mM $MgCl_2$, 5 mM MES, KOH (pH 5.6) at a final concentration of 4×10^5 protoplasts/mL.

For the transfection, 100 µL of protoplasts (2 x 10^4 protoplasts) were mixed with 10 µL of plasmid DNA (20 µg) and incubated with 110 µL of polyethylene glycol (PEG) solution (40% (w/v) PEG-MW8000, 400 mM mannitol, 100 mM $CaCl_2$) at room temperature for 30 minutes. Finally, the protoplast mixture was washed twice with W5 solution to remove PEG and resuspended in 1 mL of fresh W5 solution. After 24 hours incubation in the dark or light condition, protoplasts were examined by confocal laser scanning microscopy.

2.9. Immunoblots

Protein extracts were prepared by homogenizing 50-100 mg of leave tissues in three volumes of extraction buffer (50 mM Tris pH 7.5, 1 mM ethylenediaminetetraacetic acid [EDTA] pH 8.0, 8 mM $MgCl_2$, 1% (v/v) β-mercaptoethanol, 1mM phenylmethylsulfonyl fluoride [PMSF]). The extracts were centrifuged at 10,000 x *g* for 10 minutes at 4℃ to remove insoluble plant material and save the supernatant containing total protein extracts. Protein quantification was determined by the Bradford assay (Bio-Rad). Protein samples were then boiled in SDS loading buffer for 5 minutes and subjected to SDS-PAGE. Protein samples were applied to the SDS-PAGE gel (12.5% separating and 5% stacking gels) along with a prestained protein marker (Fermentas, PageRuler SM0671). The samples were electrophoresed at 80–100 V in a running buffer (0.1% (w/v) SDS, 192 mM glycine, and 25 mM Tris base) until the 18 kDa prestained marker ran off the gel.

After electrophoresis, the gel was washed twice for 15 minutes in a transfer buffer (48 mM Tris, 190 mM glycine, 0.0375% (w/v) SDS, and 20% (v/v) methanol). The separated

proteins were transferred onto nitrocellulose membranes using a semi-dry electroblotting device (Bio-Rad) at 20 V for 45 minutes followed by staining with Ponceau S (3 % (w/v) trichloroacetic acid [TCA], 0.2 % (w/v) Ponceau S [sodium salt; Sigma-Aldrich]). After washing with PBS-T (1% (w/v) nonfat dry milk powder, 0.3% (v/v) Tween 20 in PBS), the membrane was then blocked with 0.1% (w/v) polyvinyl alcohol (PVA; Sigma-Aldrich) in PBS for 30 seconds and incubated with the primary antibody diluted in PBS-T (anti-ADK, 1:3,500; anti-SAHH, 1:5,000; both were previously generated from Moffatt lab) overnight at 4℃. After washing three times with PBS-T, the membrane was incubated with an alkaline phosphatase-conjugated secondary antibody (anti-rabbit IgG, 1:3,000; Sigma-Aldrich) for 1 hour at room temperature. After washing three times with PBS-T, chemifluorescent detection was performed using enhanced-chemifluorescence (ECF) substrate and the resulting image was acquired using the Typhoon 9400 laser scanning system (Amersham Bioscience; Pittsburgh, PA).

2.10. Inverse PCR

To isolate sequences flanking T-DNA, inverse PCR was carried out using genomic DNA extracted from young leaves of an Arabidopsis plant expressing 35S::SAHH1-EGFP (Line No. 839-2-23). One µg of genomic DNA was digested with *Pst*I restriction endonuclease for 12 hours. The reaction was inactivated by incubating at 65 ℃ for 15 minutes, and then self-ligated by T4 DNA ligase at 4 ℃ for 16 hours. The ligation products were cleaned by ethanol precipitation, and the DNA was resuspended in 10 µL of water. The PCR was performed with an initial 5 minute denaturation at 95 ℃, followed by 35 cycles (each cycle: 95 ℃, 30 seconds; 56 ℃, 30 seconds; and 72 ℃, 2

minutes), then a final 10 minutes at 72 ℃. The primers 839-2 Pst (forward; Table 1) and 839-2 LB (reverse; Table 1) were used for the inverse PCR. The 2.3-kb DNA band obtained from the PCR reaction was cloned into pJET1.2 vector and sequenced to determine T-DNA insert location in the genome.

2.11. Confocal microscopy

Laser scanning confocal microscopy (Carl Zeiss LSM 510; Jena, Germany) was used to examine various fusion constructs expressed in different plant samples including transiently transformed Arabidopsis protoplasts and *Nicotiana benthamiana*, and stably transformed Arabidopsis plants. For imaging the expression of GFP constructs, excitation lines of an argon ion laser of 488 nm were used with a 505/530 nm bandpass (BP) filter in the single-track facility. Red fluorescence of DsRed2 was excited at 543 nm (BP543) and detected from 560 to 615 nm of emission wavelength. Chlorophyll autofluorescence was visualized using an optical filter LP650 (excitation 488 nm, emission 650 nm).

Table 1. List of primers used in this research

Name	Sequence (5' to 3')	RE	Tm
AK1cDNcF	CATccatggCTTCCTCTGATTTC	*Nco*I	56
ADK1cR-Sma	ATcccgggGTTGAAGTCTGGTTTCTC	*Xma*I	64
ADK1cF-Kpn	GTTggtaccATGGCTTCCTCTGATTTCG	*Kpn*I	64
SAHH2pF-Age	CAaccggtCCTTAAATTGCAGGAAC	*Age*I	60
ADK1 lux lower	GTACGccatggATGATGATGTTAGAAGG	*Nco*I	61
AK1cDNcR	GAccatggTGAAGTCTGGTTTCTC	*Nco*I	55
SAHH1cDNcF	TCAAccatggCGTTGCTCGTC	*Nco*I	57
SAHH1cR-Sma	TTTcccgggGTACCTGTAGTGAGGAGGCTT	*Xma*I	56
SAHH1cF-Kpn	TAAggtaccATGGCGTTGCTCGTCGAGAAG	*Kpn*I	68
SAHH1pF-Age	CATaccggtGATCCACCGACGTTCCTAC	*Age*I	66
SAHH1 Nc-R	GCAACGccatggTTGAGCTAGATC	*Nco*I	60
SAHH2 for	CAAccatggCTTTGCTTGTAGAGAAAACC	*Nco*I	63
SAHH2cR-Sma	GTTcccgggGTACCTGTAGTGAACAGGCTTG	*Xma*I	69
SAHH2pF-Age	CAaccggtCCTTAAATTGCAGGAAC	*Age*I	60
SAHH2 Nc-R	GCAAAGccatggTTGAATAGTTTC	*Nco*I	55
ySAHH-F Xho	AActcgagAATGTCTGCTCCAGCTCAAAAC	*Xho*I	64
ySAHH-R Xma	TTCCCGGGATATCTGTAGTGGTCGGCCTTG	*Xma*I	70
ySAHH-F Xho2	AActcgagACATGTCTGCTCCAGCTC	*Xho*I	61
ySAHH-R Xho	AActcgagTCAATATCTGTAGTGGTCGG	*Xho*I	59
SSU-F-Bam	CAAggatccTCATGGCTTCCTCTATGCTCTC	*Bam*HI	66
SSU-R-Bam	GTTggatccAACCGGTGAAGCTTGGTGGCTTG	*Bam*HI	74
NLS-DsRED2F	ATccatggTGCCAAAAAAGAAGAGAAAGGTAG AAGACCCCGTG**ATG**GCCTCCTCCGAGAACG*	*Nco*I	86
DsRED2R-Nco	AGCccatggTCTTCTTCTGCATC	*Nco*I	58
SAHH1-D29F	GATccatggATCTCGAGCTCGCCGAAGTTG	*Nco*I	70
SAHH1-D63F	CATccatggATATGACCATCCAAACCGCCGTAC	*Nco*I	71

Table 1. Continued

Name	Sequence (5' to 3')	RE	Tm
SAHH1-D410R	GTTcccgggAGACATCACGAAACTTGGGTG	*Xma*I	71
SAHH1-D149R	CAActgcagCTCATGAATCAAAAGAGTAG	*Pst*I	59
SAHH1-D191F	GAGctgcagAAGTACCACAAGATGAAGGAG	*Pst*I	63
SAHH1-410RN	ATccatggACATCACGAAACTTGGGTG	*Nco*I	64
SAHH1-149F	AAccatggAGGGTGTTAAAGCTGAG	*Nco*I	58
SAHH1-190R	ATccatggTCTTAGGATCAACTTGAAG	*Nco*I	57
SAHH D178-F	GAGctgcagTTGTCTATTATCAAGGAAG	*Pst*I	57
SAHH D177-R	GAActgcagCACGATCTGAAACTCAG	*Pst*I	60
SH1-PDPto3A	AGGagatctTTGAGAAGACTGGTCAAGTTGCTG CTGCTACTTC**	*Bgl*II	73
SAHH1-inR	GTTaccggtGGTGGTGAC	*Age*I	47
SH1-TSTto3A	AGGagatctTTGAGAAGACTGGTCAAGTTCCTG ATCCTGCTGCTGCTGAT**	*Bgl*II	79
839-2 Pst	TCTTTATGCGGACACTGACG	N/A	51
839-2 LB	TGCAGGTCAAACCTTGACAG	N/A	51

The lowercased sequences represent RE sites indicated. *NLS sequence is underlined; start codon of DsRed2 is bolded. **Underlined nucleotide sequences indicate the nucleotide substitutions to create three alanine residues. RE, restriction enzyme; Tm, melting temperature.

3. Results

3.1. Subcellular localization of ADK and SAHH

The subcellular localization of ADK and SAHH were examined first by transient expression of Enhanced GFP (EGFP) fusion proteins in protoplasts. This is a red-shifted variant of wild-type GFP which is particularly suited to plant cell research because it provides increased brightness of fluorescence in plant cells through its more efficient translation or folding (Cormack et al., 1996). The EGFP was fused to the C-terminus of ADK1 and SAHH1 and expressed under the control of the constitutive 35S promoter. These constructs were introduced into Arabidopsis protoplasts by PEG-mediated transformation and the localization of EGFP was examined by confocal microscopy. The fluorescence of both EGFP fusion proteins was observed in both the nucleus and cytoplasm (Figure 1A and B).

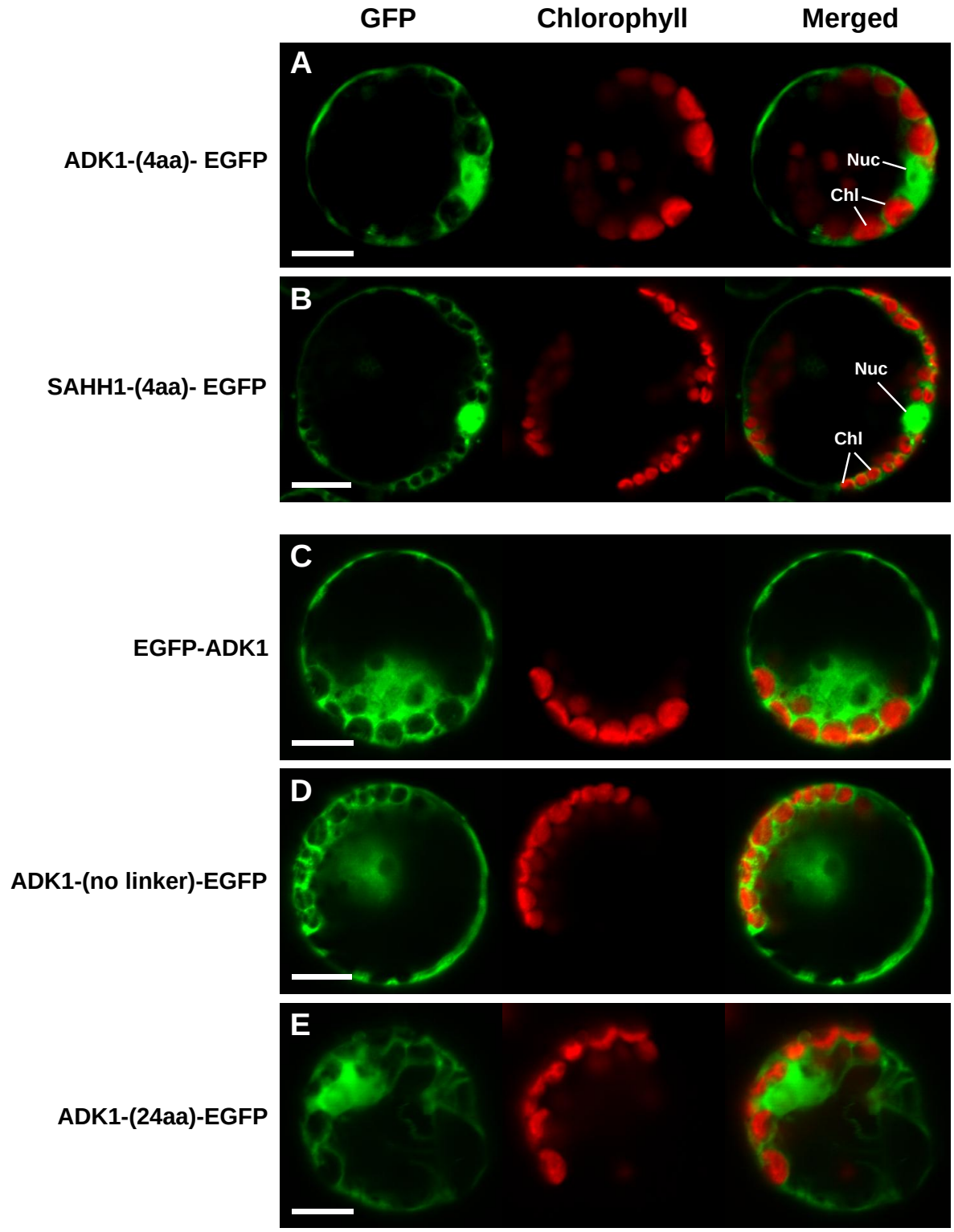
Additional constructs were generated to test whether the expression/localization of the ADK-EGFP fusion was affected by changing the number of amino acids between the two proteins (linker) or by the different orientations of the target protein-EGFP fusions (Figure 1C-E). All the tested EGFP fusions including those with no linker (Figure 1D) or linkers of 4 amino acids (Figure 1A), or 24 amino acids (Figure 1E) between ADK and EGFP, as well as the EGFP fusion to N-terminus of ADK (Figure 1D) showed similar fluorescent signals. In all cases ADK localized to both the nucleus and cytoplasm. These results indicate that the localization of the ADK-EGFP fusion proteins is not influenced by the flexibility of the linkers between the two proteins or whether the EGFP is fused to the N- or C- terminus.

Seven different EGFP fusion constructs for ADK and SAHH were examined in transgenic Arabidopsis plants to verify these transient expression results. At least twenty independent T1 generation transgenic lines generated by *Agrobacterium*-mediated transformation were obtained for each construct. The expression of the ADK1-EGFP and SAHH1-EGFP fusion proteins (each with 4 amino acids linkers) in the transgenic plants was examined by immunoblot analyses using either anti-ADK or anti-SAHH specific antibodies, respectively (Figure 2). ADK1-EGFP protein (Figure 2A) and SAHH1-EGFP protein (Figure 2B) were successfully detected at 65 kDa and 80 kDa in all cases; the corresponding native proteins were observed at 38 kDa and 53 kDa. Thus EGFP fusions of ADK1 (Figure 3A and B) and SAHH1 (Figure 3C and D) expressed from the 35S promoter in transgenic plants showed similar localization patterns in the cytoplasm and nucleus, as had been observed by transient expression.

Transgenic lines expressing ADK1 (Figure 3H) and SAHH1 (Figure 3F) under the control of their endogenous promoters displayed the similar localization patterns as those with the constitutive 35S promoter. These results indicate that the use of the 35S promoter did not affect the localization of either EGFP fusion protein. Overall, the SAHH1 fusions were more strongly targeted to the nucleus than the cytoplasm (Figure 3C-F) as compared to the ADK1 fusions (Figure 3A, B, and H). The other isoform of SAHH, SAHH2, when expressed from its own promoter had a localization pattern indistinguishable from that of SAHH1 (Figure 3G). In addition, tobacco leaves transiently expressing EGFP fusions of ADK1, SAHH1, and SAHH2 also showed fluorescence in the cytoplasm and nucleus of epidermal cells (Figure 3I-L).

Figure 1. Subcellular localization of ADK1 and SAHH1 in Arabidopsis protoplasts.

Protoplasts transformed with ADK1-(4aa)-EGFP (A), SAHH1-(4aa)-EGFP (B), EGFP-ADK1 (C), ADK1-(no linker)-EGFP (D), and ADK1-(24aa)-EGFP (E) were imaged using confocal microscopy. In all cases the fusion proteins were expressed from the CaMV 35S promoter. Green and red colours represent EGFP and chlorophyll autofluorescent signals, respectively. Chl, chloroplast; Nuc, nucleus. Scale bars represent 10 μm.


GFP
Chlorophyll
Merged
A
ADK1-(4aa)- EGFP
Nuc
Chl
B
SAHH1-(4aa)- EGFP
Nuc
Chl
C
EGFP-ADK1
D
ADK1-(no linker)-EGFP
E
ADK1-(24aa)-EGFP

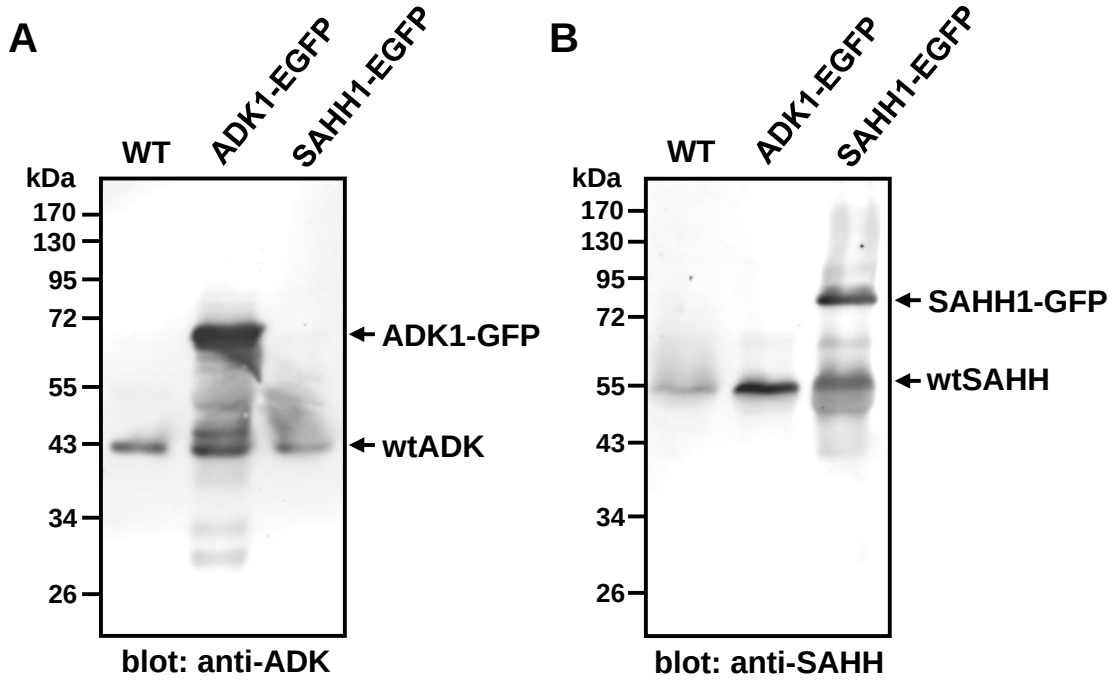
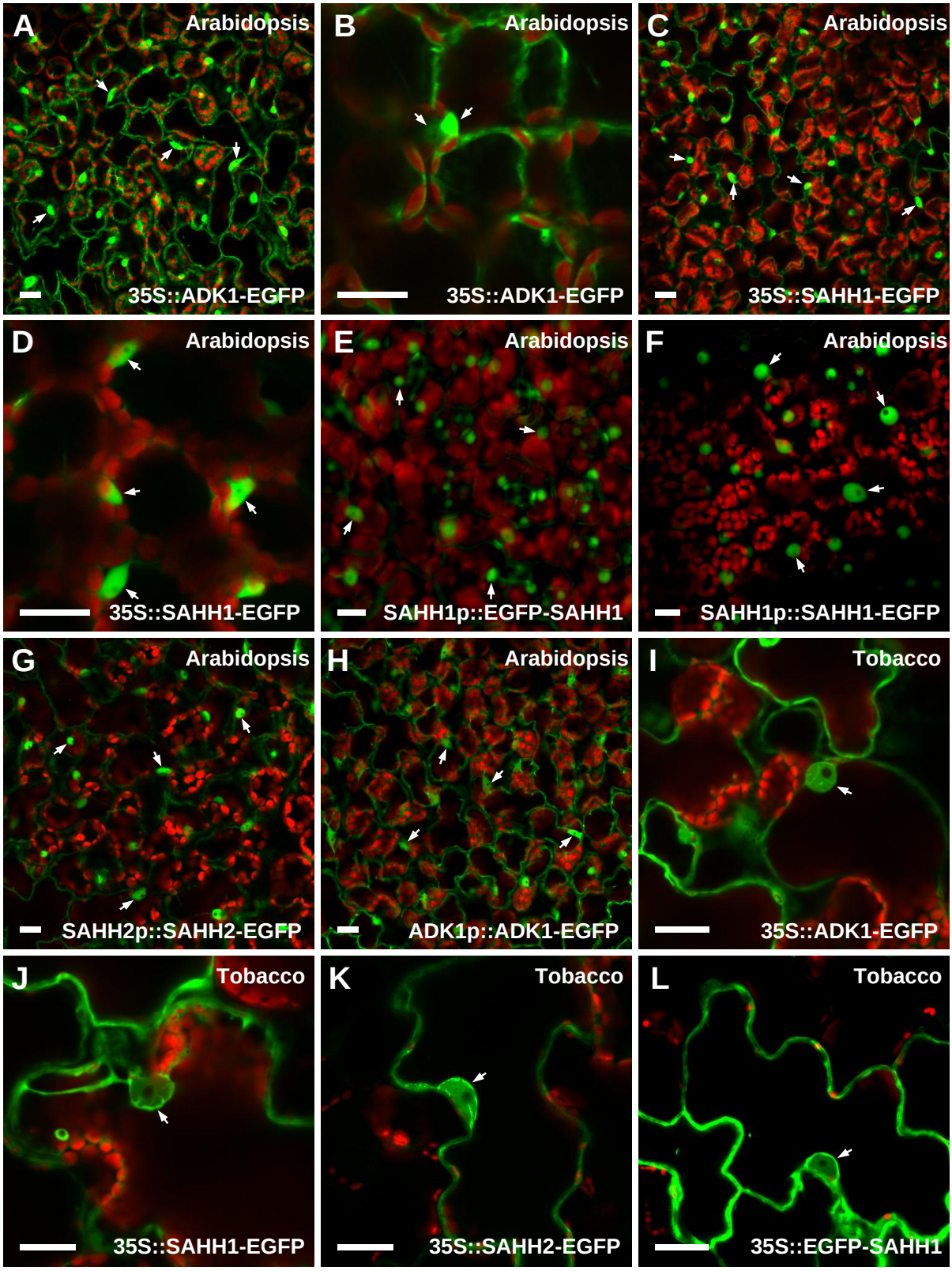


Figure 2. Immunoblot analysis of ADK1-EGFP and SAHH1-EGFP.

Total protein extracts were prepared from wild-type and transgenic Arabidopsis plants expressing either 35S::ADK1-EGFP or 35S::SAHH1-EGFP. Ten µg of each protein extract were electrophoresed through 10% SDS-PAGE gels and transferred onto nitrocellulose membranes. The two identical membranes were probed with either anti-ADK (A) or anti-SAHH (B) antibody. Molecular masses are indicated on the left in kilodaltons (kDa). WT, protein extracts obtained from wild-type Arabidopsis plants; wtADK, endogenous ADK; wtSAHH, endogenous SAHH.

Figure 3. Various EGFP-fusion constructs of ADK and SAHH in Arabidopsis and tobacco plants.

(A-H) Leaves of transgenic Arabidopsis plants expressing 35S::ADK1-EGFP (A and B), 35S::SAHH1-EGFP (C and D), SAHH1p::EGFP-SAHH1 (E), SAHH1p::SAHH1-EGFP (F), SAHH2p::SAHH2-EGFP (G), and ADK1p::ADK1-EGFP (H) were imaged using confocal microscopy. (I-L) Tobacco leaves transiently transformed with 35S::ADK1-EGFP (I), 35S::SAHH1-EGFP (J), 35S::SAHH2-EGFP (K), and 35S::EGFP-SAHH1 (L) were imaged. The green and red colours represent GFP and chlorophyll autofluorescent signals, respectively. Arrows indicate nuclei. Scale bars represent 10 μm.


A
Arabidopsis
35S::ADK1-EGFP
B
Arabidopsis
35S::ADK1-EGFP
C
Arabidopsis
35S::SAHH1-EGFP
D
Arabidopsis
35S::SAHH1-EGFP
E
Arabidopsis
SAHH1p::EGFP-SAHH1
F
Arabidopsis
SAHH1p::SAHH1-EGFP
G
Arabidopsis
SAHH2p::SAHH2-EGFP
H
Arabidopsis
ADK1p::ADK1-EGFP
I
Tobacco
35S::ADK1-EGFP
J
Tobacco
35S::SAHH1-EGFP
K
Tobacco
35S::SAHH2-EGFP
L
Tobacco
35S::EGFP-SAHH1

3.2. Nuclear localization of ADK1 and SAHH1

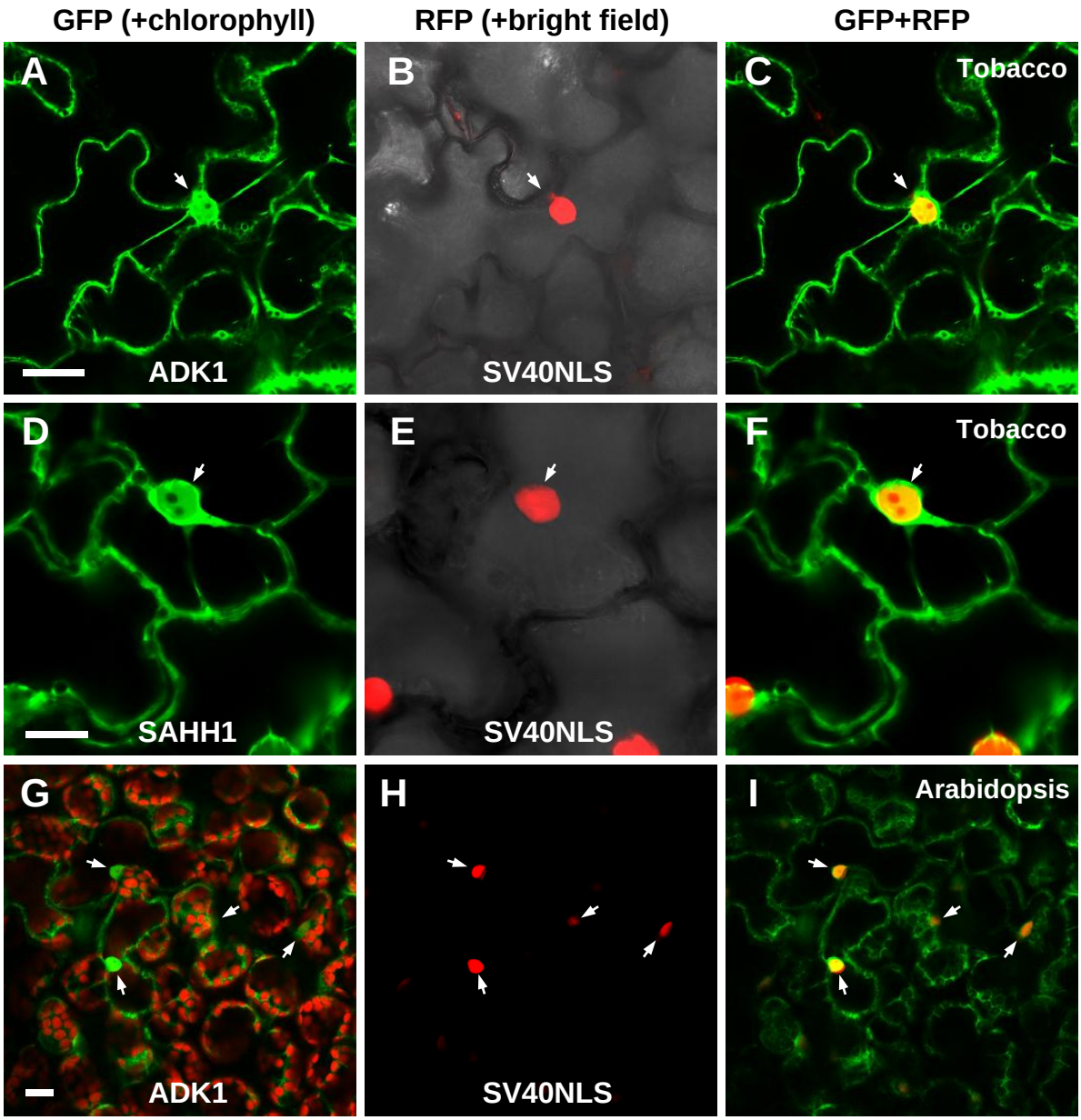
Given that both ADK1 and SAHH1 were strongly localized in the nucleus but no obvious nuclear localization signal is detected in either protein by various bioinformatic tools including ScanProsite (Gasteiger et al., 2003) and WoLF PSORT (Horton et al., 2007), the question arose as how they are targeted to the nucleus. The most obvious solution to this is that both ADK1-EGFP (65 kDa) and SAHH1-EGFP (80 kDa) enter the nucleus by simple diffusion, since both fusion proteins are below or close to the exclusion limit of the nuclear pore (70 kDa) (D'Angelo et al., 2009).

To test this, both ADK1 and SAHH1 were fused to two concatenated GFPs to increase the size of the fusion proteins to 92 kDa (ADK1-2xGFP) and 107 kDa (SAHH1-2xGFP), respectively. These proteins should not be able to simply diffuse into the nucleus based on current estimates of the nuclear pore exclusion limit. As a nuclear localization control, the NLS sequence of simian virus 40 (SV40) large T antigen (PKKKRKVEDP) (Goldfarb et al., 1986) fused to *Discosoma sp* red fluorescent protein (DsRed2) was introduced into plants (Figure 4J). Both ADK1-2xGFP and SAHH1-2xGFP were co-infiltrated into tobacco leaves along with the SV40NLS-DsRed2 control; the ADK1 and SV40NLS constructs were co-transformed into wild-type Arabidopsis plants.

The GFP signals of the ADK2xGFP and SAHH1-2xGFP fusion proteins clearly overlapped the nuclear DsRed2 signals of SV40NLS (Figure 4C, F, and I), indicating that both fusion proteins are also targeted to the nucleus. These results suggested that the two enzymes are imported into the nucleus by active transport machinery, rather than by simple diffusion.

Figure 4. Nuclear localization of double GFP fusions of ADK1 and SAHH1.

(A-F) Transient expression of ADK1-2xGFP (A) and SAHH1-2xGFP (D) were examined in tobacco plants by co-infiltration with SV40NLS-DsRed2 (B and E). (G-I) Transgenic Arabidopsis leaves expressing both ADK1-2xGFP (G) and SV40NLS-DsRed2 (H) were imaged using confocal microscopy. (J) Nucleotide and amino acid sequences of SV40NLS used for the nuclear localization control. Note the CaMV 35S promoter directs expression of all these constructs. The green and red colours represent GFP and RFP signals, respectively. The overlap of the GFP and RFP signals is shown in yellow (C, F, and I), and the autofluorescent signal of chlorophyll is depicted in red (G). Arrows indicate nuclei. Scale bars represent 10 μm.



J

atggtgccaaaaaagaagagaaaggtagaagaccccgtgatggcctcctccgagaacgtc

M V P K K K R K V E D P V M A S S E N V

SV40NLS

DsRed2

3.3. Deletion analyses of ADK1 and SAHH1

In order to search for subcellular localization signals or important regions for nuclear targeting, a series of deletions was constructed for both ADK1 and SAHH1 (Figure 5). First, six distinct ADK1 deletions were created by PCR-based mutagenesis as illustrated in Figure 5A. All these deletions were fused to EGFP and transiently expressed in tobacco leaves using the infiltration method (Voinnet et al., 2003). All deletion fusions with EGFP showed similar fluorescent signals in the nucleus and cytoplasm (Shah, 2006; 499 project, data not shown). Since these fusion proteins were below the diffusion limit of the nuclear pore, they may have simply diffused into the organelle. Thus, a second set of fusions to GUS (70 kDa) were generated. Surprisingly all the GUS fusions were excluded from the nucleus. Moreover the nuclear envelopes of plants expressing either ADK1Δ49-344-GUS or ADK1Δ49-344-GUS fusion protein were stained blue by the histochemical GUS assay (Jefferson et al., 1987)(Shah, 2006; 499 project, data not shown).

Seven different deletion constructs for SAHH1 were created to evaluate the importance of the N-terminus, C-terminus, and internal sequences in nuclear targeting (Figure 5B). Surprisingly, deletion of an internal region of SAHH (IR; Δ150-190) lacked any nuclear localization signal (Figure 6E), while all other deletions fused to EGFP showed similar patterns of fluorescent signals in the nucleus and cytoplasm of transiently expressed tobacco leaves (Figure 6A-L).

The localization of the IR deletion fusion protein was further examined in transgenic Arabidopsis plants. Consistent with the result of transient expression in tobacco, the

transgenic Arabidopsis leaves expressing SAHH1ΔIR-EGFP from the 35S promoter showed no fluorescent signals in the nucleus (Figure 6N). However, strong nuclear signals were detected in plants expressing a larger deletion: SAHH1Δ1-62 ΔIR Δ411-485-EGFP (61.2 kDa; Figure 6P). This may have been due to the diffusion of the fusion protein into the nucleus. Plants expressing SAHH1Δ1-62 ΔIR Δ411-485-2xEGFP (88 kDa) also displayed nuclear signals, but the level of fluorescence in the nucleus was typically weaker than that of full-length SAHH1-EGFP fusion in terms of the comparison of nuclear versus cytoplasmic fluorescent signals (Figure 6R). Taken together, these results suggested that the IR of SAHH1 is important for its nuclear localization.

Figure 5. Deletion analyses of ADK1 and SAHH1.

(A) A schematic diagram showing deletion constructs of ADK1. All ADK1 deletions were fused to EGFP and GUS and transiently expressed in tobacco plants for examination. The wtADK1 indicates full-length ADK1 protein (344 amino acids). The coloured boxes represent the region that is present in the ADK1 fusion proteins, while the horizontal lines represent deleted regions. The numbers correspond to amino acid residues of ADK1. These constructs were created by Malay Shah as part of his 499 project (Shah, 2006). (B) A schematic diagram showing deletion constructs of SAHH1. The wtSAHH1 in the left of amino acid sequences indicates the wild-type SAHH1 sequence, while periods below the original SAHH1 sequence represent deleted sequences. The extra amino acid sequences generated by PCR primers are underlined. A part of the EGFP sequence linked to one of the SAHH1 deletions is shown in italics. The wtSAHH1 on the left of the boxes indicates full-length SAHH1 protein (485 amino acids). The coloured boxes represent the region that is present in the SAHH1 fusion proteins, while the horizontal lines represent deleted region. The numbers represent amino acid residues of SAHH1.

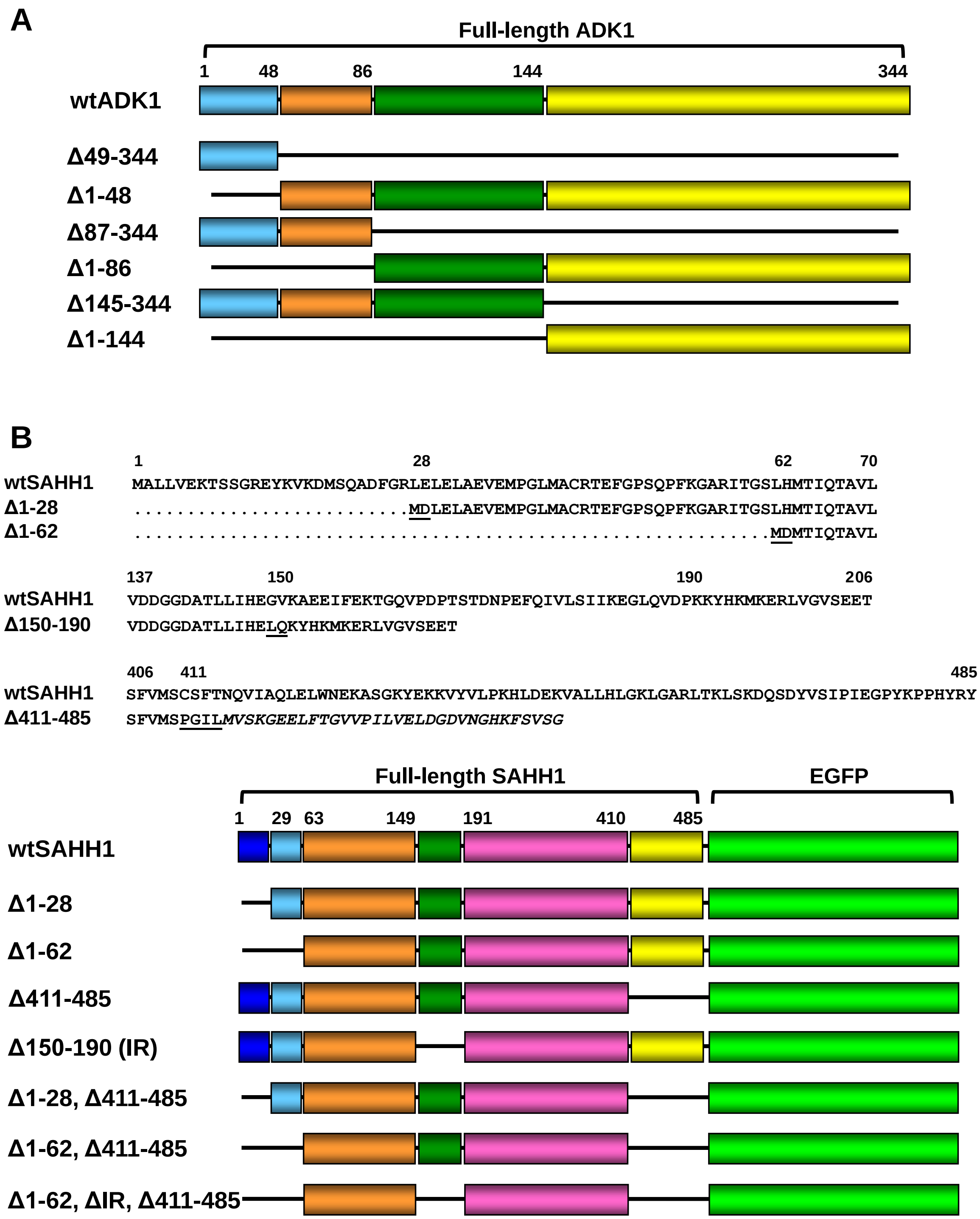
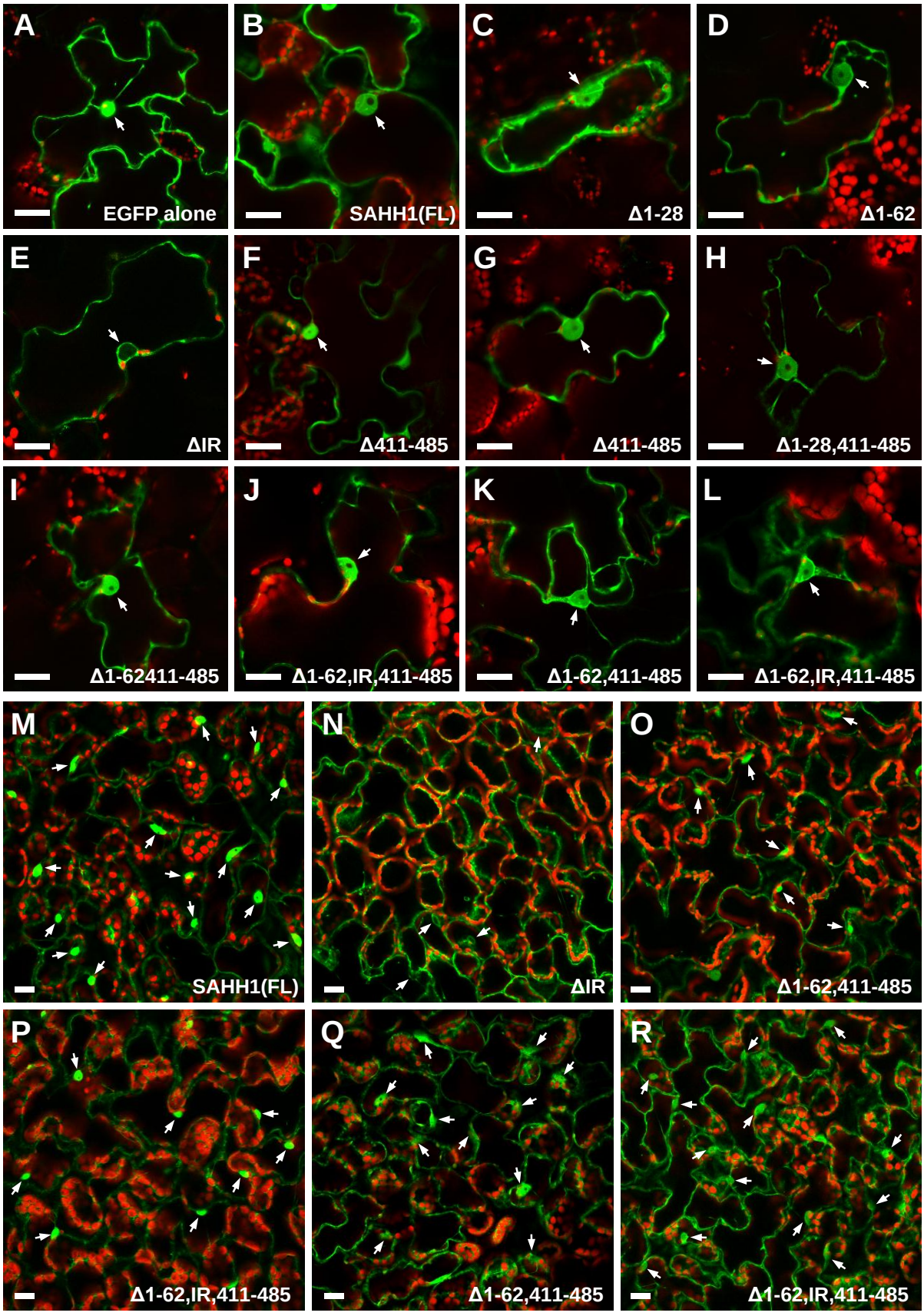

A
Full-length ADK1
1
48
86
144
344
wtADK1
Δ49-344
Δ1-48
Δ87-344
Δ1-86
Δ145-344
Δ1-144
B
1
28
62
70
wtSAHH1 MALLVEKTSSGREYKVKDMSQADFGRLELELAEVEMPGLMACRTEFGPSQPFKGARITGSLHMTIQTAVL
Δ1-28 ...........................MDLELAEVEMPGLMACRTEFGPSQPFKGARITGSLHMTIQTAVL
Δ1-62 ....MDMTIQTAVL
137
150
190
206
wtSAHH1 VDDGGDATLLIHEGVKAEEIFEKTGQVPDPTSTDNPEFQIVLSIIKEGLQVDPKKYHKMKERLVGVSEET
Δ150-190 VDDGGDATLLIHELQKYHKMKERLVGVSEET
406
411
485
wtSAHH1 SFVMSCSFTNQVIAQLELWNEKASGKYEKKVYVLPKHLDEKVALLHLGKLGARLTKLSKDQSDYVSIPIEGPYKPPHYRY
Δ411-485 SFVMSPGIL*MVSKGEELFTGVVPILVELDGDVNGHKFSVSG*
Full-length SAHH1
EGFP
1
29
63
149
191
410
485
wtSAHH1
Δ1-28
Δ1-62
Δ411-485
Δ150-190 (IR)
Δ1-28, Δ411-485
Δ1-62, Δ411-485
Δ1-62, ΔIR, Δ411-485

Figure 6. Nuclear localization of SAHH1 deletions.

All deletion constructs were examined by transient expressions in tobacco leaves (A-L) and some of them were further examined in transgenic Arabidopsis plants (M-R). (A) EGFP control, (B) Full-length SAHH1-EGFP, (C) SAHH1Δ1-28-EGFP, (D) SAHH1Δ1-62-EGFP, (E) SAHH1ΔIR-EGFP, (F) SAHH1Δ411-485-EGFP, (G) EGFP-SAHH1Δ411-485, (H) SAHH1Δ1-28Δ411-485-EGFP, (I) SAHH1Δ1-62Δ411-485-EGFP, (J) SAHH1Δ1-62ΔIRΔ411-485-EGFP, (K) SAHH1Δ1-62Δ411-485-2xGFP, (L) SAHH1Δ1-62ΔIRΔ411-485-2xGFP, (M) Full-length SAHH1-EGFP, (N) SAHH1ΔIR-EGFP, (O) SAHH1Δ1-62Δ411-485-EGFP, (P) SAHH1Δ1-62ΔIRΔ411-485-EGFP, (Q) SAHH1Δ1-62Δ411-485-2xGFP, (R) SAHH1Δ1-62ΔIRΔ411-485-2xGFP. Note the 35S promoter directs expression of all these constructs. The green and red colours represent GFP and autofluorescent signal of chlorophyll, respectively. Arrows indicate nuclei. Scale bars represent 10 μm.


A
EGFP alone
B
SAHH1(FL)
C
Δ1-28
D
Δ1-62
E
ΔIR
F
Δ411-485
G
Δ411-485
H
Δ1-28,411-485
I
Δ1-62411-485
J
Δ1-62,IR,411-485
K
Δ1-62,411-485
L
Δ1-62,IR,411-485
M
SAHH1(FL)
N
ΔIR
O
Δ1-62,411-485
P
Δ1-62,IR,411-485
Q
Δ1-62,411-485
R
Δ1-62,IR,411-485

3.4. The IR affects nuclear localization of SAHH1

Based on an analysis of multiple alignments of SAHH sequences, the 41 amino acids of the IR (Gly^{150}-Lys^{190}) are present in many SAHHs including those from photosynthetic bacteria in prokaryotes, but it is present only in a few eukaryotic organisms, including plants and parasitic protozoans (Figure 7). Although all SAHHs found in plants contain this specific amino acid IR segment, its biological and physiological significance is unknown.

Thus, to verify whether the IR segment is essential for the nuclear localization of SAHH, further experiments were performed using IR-related constructs (Figure 8). First, fusion of GFPs with an IR-deleted SAHH1 failed to localize the protein to the nucleus as its GFP signals did not overlap with RFP signals of the NLS control (Figure 8A-C). These results suggest that the IR may be essential for the nuclear localization of SAHH1 in plants. Next, the 41 amino acid IR segment was fused to double GFP (MW: 58.9 kDa) and examined in both tobacco and Arabidopsis plants to test whether it could act as a NLS. The GFP signals clearly overlapped with the NLS control signals in the nucleus, and also strongly localized in the cytoplasm of both tobacco and Arabidopsis plants (Figure 8D-F), suggesting that IR does not act independently as a nuclear localization signal. Next, the localization of the yeast SAHH (P39954) protein that shares 63% amino acid sequence identity with Arabidopsis SAHH1 but does not have IR sequence was investigated. The yeast SAHH was fused to either N- or C-terminus of EGFP and expressed transiently in tobacco and constitutively in Arabidopsis plants to examine whether another eukaryote SAHH lacking the IR could localize to the nucleus in plants (Figure 8J-M). Although both tobacco and Arabidopsis plants expressing the yeast

SAHH-EGFP fusion protein (C-terminal fusion) showed strong fluorescent signals in the nucleus and cytoplasm (Figure 8J and L), only weak or even no fluorescent signals of the EGFP-yeast SAHH fusion protein (N-terminal fusion) were detected in the nucleus (Figure 8K and M). These results suggested that the lack of the IR affected nuclear targeting of yeast SAHH in plant cells.

Further deletion analyses were performed to search for residues within the IR that are important for SAHH nuclear targeting. Three parts of the IR, both ends and the middle, were examined. First, the ΔIR150-177-EGFP fusion which lacks the first 28 amino acids of the IR and the ΔIR178-190-EGFP fusion which lacks the 13 amino acids terminal residues of the IR (Figure 9A) were examined by transient expression in tobacco plants; their localization in stable transformants of Arabidopsis was also evaluated. Both fusions expressed in tobacco plants showed GFP fluorescent signals in both the nucleus and cytoplasm (Figure 9B and C), while the Arabidopsis transgenic plants expressing either fusion protein displayed limited fluorescent signals in the nucleus (Figure 9D and E). When comparing the two, however, the fluorescent signals of ΔIR178-190-EGFP were weakly detected in the nucleus, whereas those of ΔIR150-177-EGFP were almost undetectable. Although further study is needed to assess their expression levels in the nucleus, these results suggested that both parts of the IR may contribute to the nuclear localization of SAHH but the residues 178-190 are less important than the residues 150-177 within the IR of SAHH.

Next, two small regions, $PDP^{164-166}$ and $TST^{167-169}$, from this essential region were substituted with three alanine residues (Figure 9A). These regions were chosen based on sequence analysis using the bioinformatic tools provided by the ELM server

(http://elm.eu.org/) and the NetPhos server (http://www.cbs.dtu.dk/services/NetPhos/). The TST residues are predicted to be phosphorylation sites and PDP is an important motif for mediating protein-protein interactions in other proteins. Both tobacco and Arabidopsis plants expressing either GFP-fusion protein showed cytoplasmic and nuclear localization signals (Figure 9F-I). However, an unusual punctate fluorescent pattern was detected from the tobacco plants expressing the $PDP^{164-166}$mutated EGFP fusion protein (PDP164AAA-EGFP; Figure 9F; arrowheads). The punctate signals of PDP164AAA-EGFP were also detected in transiently transformed protoplasts (Figure 9J and K; arrowheads), and similar punctate signals were observed in the protoplasts expressing the ΔIR-EGFP fusion protein (Figure 9M). Taken together, these results suggest that the 41 amino acids of full-length IR may be required for nuclear localization of SAHH, and the $PDP^{164-166}$ residues in the IR may be particularly important for this targeting.

Figure 7. Comparison of the deduced amino acid sequences of SAHHs from various organisms.

The alignment was performed using AlignX (Vector NTI, Invitrogen) and edited in Microsoft Word. Conserved residues are black-shaded, and similar residues are in gray. Numbers represent amino acid sequence of Arabidopsis SAHH1 (ARATH1). The amino acid sequences used in the alignment are as following: ARATH1 (O23255.1), ARATH2 (Q9LK36.1), LUPLU (Q9SP37.1), CATRO (P35007.1), PETCR (Q01781.2), TOBAC (P68173.1), SOLLC (Q9SWF5.1), MEDSA (P50246.1), PHASS (P50249.1), MESCR (P93253.1), WHEAT (P32112.1), PLAF7 (P50250.2), PLACH (Q4XZZ5.1), PLAYO (Q7RKK8.1), TRIVA (P51540.1), DICDI (P10819.2), DROME (P50245.2), ANOGA (O76757.2), SCHPO (O13639.1), YEAST (P39954.1), CANAL (P83783.2), PNECA (Q12663.1), XENLA1 (P51893.1), XENLA2 (O93477.1), RAT(P10760.3), MOUSE1(P50247.3), MOUSE2 (Q68FL4.1), HUMAN1 (P23526.4), HUMAN2 (Q96HN2.1), MACFA (Q4R596.3), PONAB (Q5R889.1), BOVIN (Q3MHL4.3), PIG (Q710C4.3), HALHL (A1WXM7.1), GRABC (Q0BW40.1), COXBU (Q83A77.2), PSEAB (Q02TY0.1), ACIAD (Q6FA43.1), RHOBA (Q7TTZ5.1), LEPBJ (Q04NN6.1), BACTN (Q8A407.1), SOLUE (Q01VU1.1), SYNSC (Q3ANF4.1), PROMA (Q7V9P3.1), PROVI (A4SF77.1), PELPB (B4SD43.1), CHLCH (Q3AQC2.1), METCA (Q60CG8.1), CHRVO (Q7NZF7.1), POLNS (B1XSH8.1), AZOSE (Q5P6B7.1), CUPTR (B2AGG2.1), RALEJ (Q476T8.1), BORPA (Q7W1Z7.1), GEOSL (P61617.1), CAUCR (Q9ABH0.1), SINMW (A6UEJ1.1), AGRRK (B9JG53.1), PELCD (Q3A392.1), DESDG (Q30WL8.1), HYPNA (Q0C427.1), RHOPA (Q6N2N5.1), BRAJA (Q89HP6.1), STRAA (Q936D6.1), CORDI (P61456.1), COREF (Q8FRJ4.1), MYCSS (Q1BCD6.1), NOCFA (Q5YQS7.1), MYCBO (Q7TWW7.3).

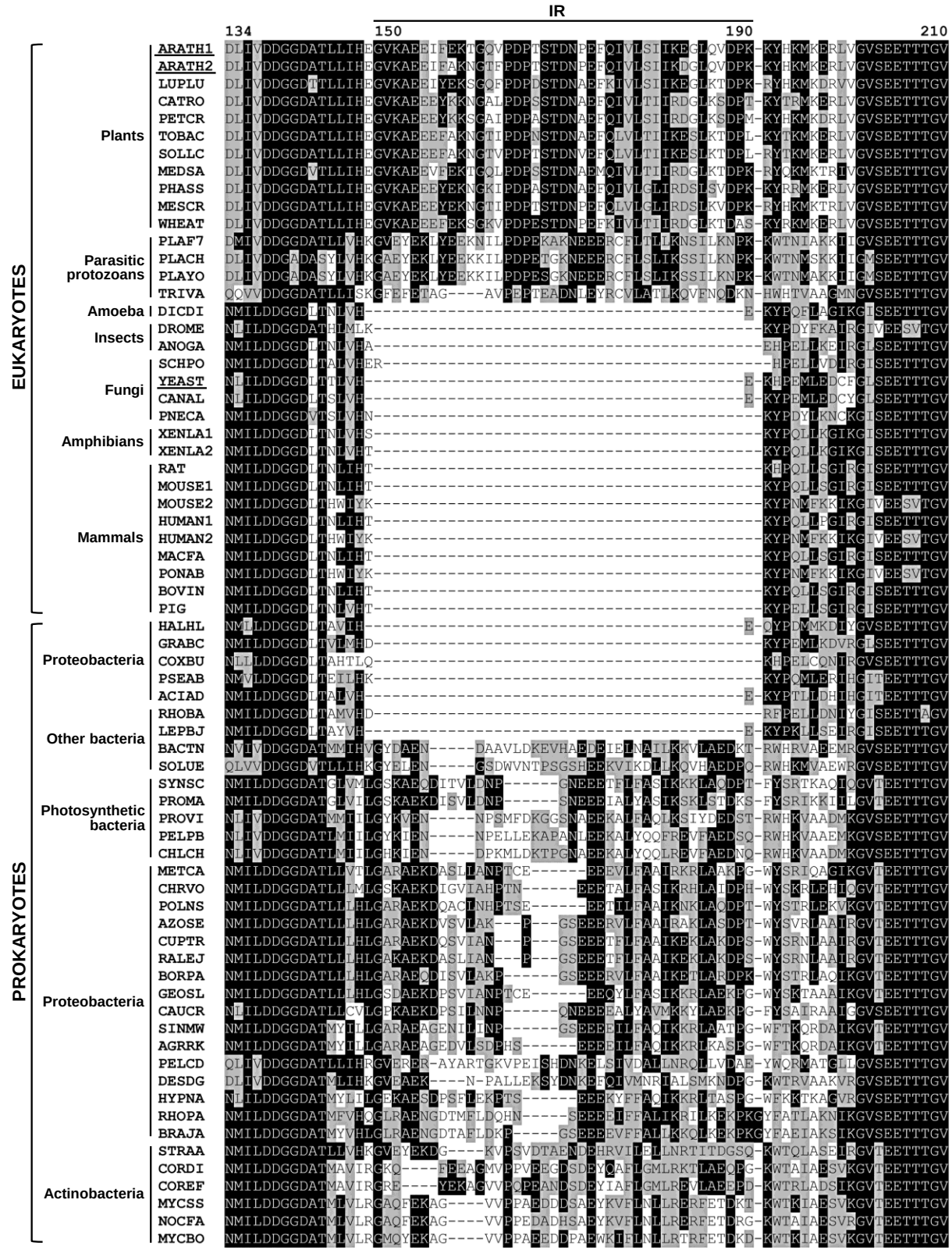

IR
134
150
190
210
EUKARYOTES
PROKARYOTES
Plants
Parasitic protozoans
Amoeba
Insects
Fungi
Amphibians
Mammals
Proteobacteria
Other bacteria
Photosynthetic bacteria
Proteobacteria
Actinobacteria
ARATH1
ARATH2
LUPLU
CATRO
PETCR
TOBAC
SOLLC
MEDSA
PHASS
MESCR
WHEAT
PLAF7
PLACH
PLAYO
TRIVA
DICDI
DROME
ANOGA
SCHPO
YEAST
CANAL
PNECA
XENLA1
XENLA2
RAT
MOUSE1
MOUSE2
HUMAN1
HUMAN2
MACFA
PONAB
BOVIN
PIG
HALHL
GRABC
COXBU
PSEAB
ACIAD
RHOBA
LEPBJ
BACTN
SOLUE
SYNSC
PROMA
PROVI
PELPB
CHLCH
METCA
CHRVO
POLNS
AZOSE
CUPTR
RALEJ
BORPA
GEOSL
CAUCR
SINMW
AGRRK
PELCD
DESDG
HYPNA
RHOPA
BRAJA
STRAA
CORDI
COREF
MYCSS
NOCFA
MYCBO

Figure 8. The affection of the IR for nuclear localization of SAHH1.

(A-C) Arabidopsis leaves expressing both IR-deleted SAHH1 fused with double GFP (A) and NLS-RFP control (B) fusion proteins were examined using confocal microscopy, and the GFP and RFP images were merged (C). (D-I) The 40 amino acids of IR fragment fused with double GFP (D and G) was co-expressed with the NLS-RFP control (E and H) in both tobacco (D-F) and Arabidopsis plants (G-I), and the GFP and RFP images were merged (F and I). The green and red colours represent GFP and RFP signals, respectively. The overlap of the GFP and RFP signals is shown in yellow in the right panel. (J-M) Yeast SAHH (P39954) gene was cloned from yeast (*Saccharomyces cerevisiae*) genomic DNA and used to create GFP constructs: (J and L) yeastSAHH-EGFP, (K and M) EGFP-yeastSAHH. Both constructs were examined in tobacco (J and K) and Arabidopsis plants (L and M). Note the CaMV 35S promoter directs expression of all these constructs. The GFP signal is shown in green and the autofluorescent signal of chlorophyll is red (J-M). Arrows indicate nuclei. Scale bars represent 10 μm.

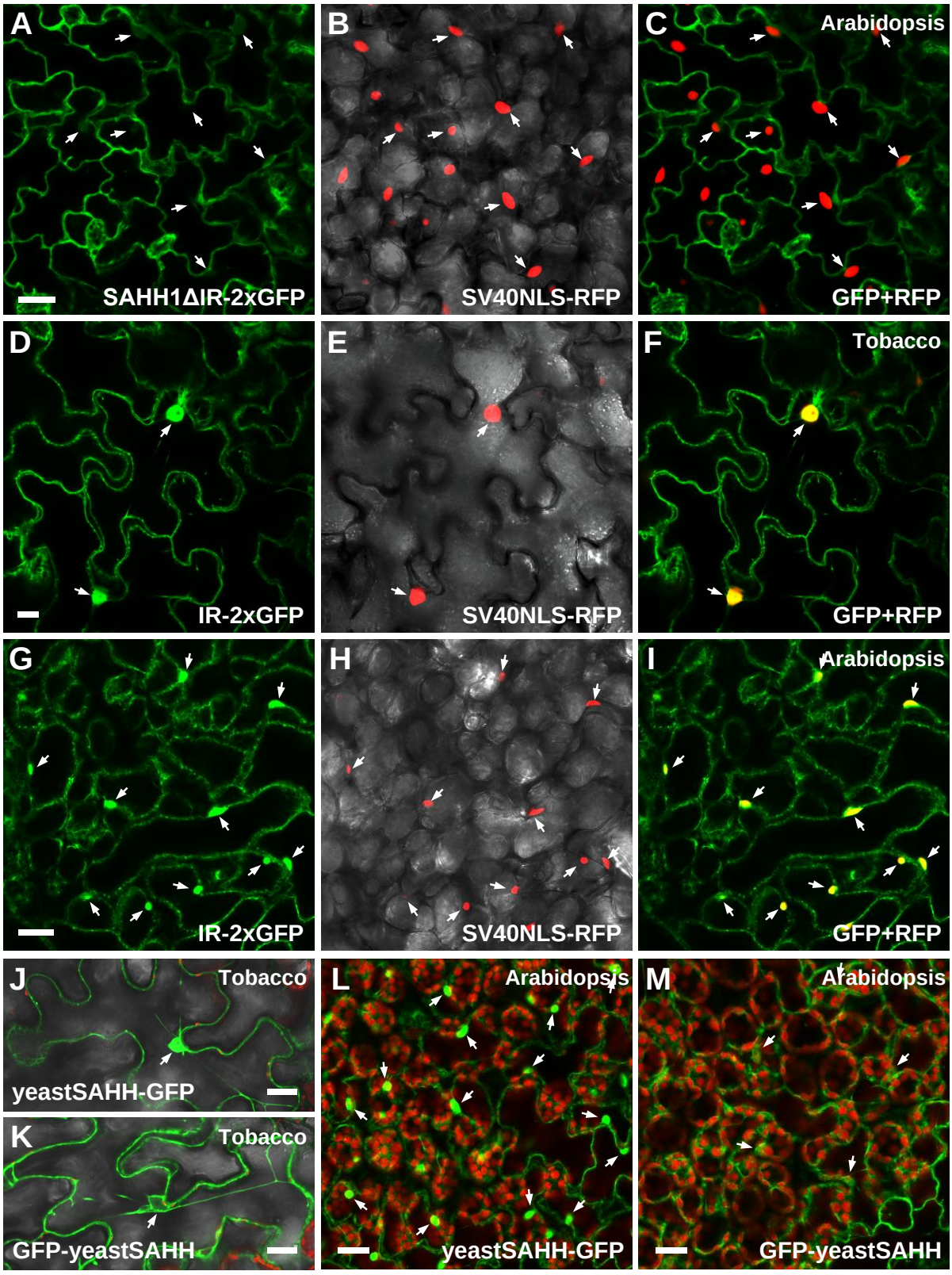
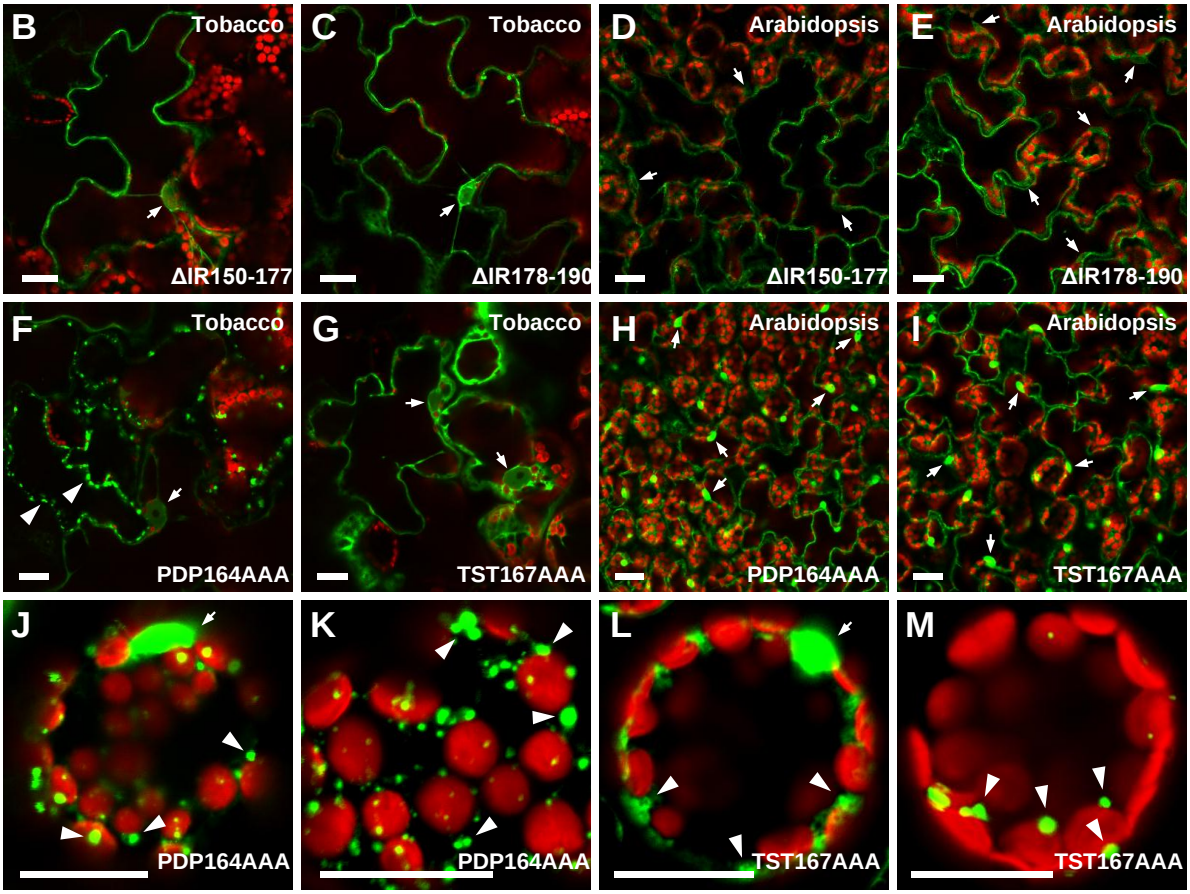

A
SAHH1ΔIR-2xGFP
B
SV40NLS-RFP
C
Arabidopsis
GFP+RFP
D
IR-2xGFP
E
SV40NLS-RFP
F
Tobacco
GFP+RFP
G
IR-2xGFP
H
SV40NLS-RFP
I
Arabidopsis
GFP+RFP
J
Tobacco
yeastSAHH-GFP
K
Tobacco
GFP-yeastSAHH
L
Arabidopsis
yeastSAHH-GFP
M
Arabidopsis
GFP-yeastSAHH

Figure 9. Mutation within the IR of SAHH1.

(A) A schematic diagram showing the sequence of deleted or substituted regions prepared for this study of SAHH1. The wtSAHH1 represents native Arabidopsis SAHH1 sequence. The periods indicate the deleted amino acid residues. The substituted regions (PDP to AAA or TST to AAA) are underlined. The numbers correspond to the amino acid residues of wtSAHH1. (B and D) The ΔIR150-177-EGFP fusion in tobacco and Arabidopsis plants, respectively. (C and E) The ΔIR178-190-EGFP fusion in tobacco and Arabidopsis plants, respectively. (F and H) The PDP164AAA-EGFP fusion in tobacco and Arabidopsis plants, respectively. (G and I) The TST167AAA-EGFP fusion in tobacco and Arabidopsis plants, respectively. (J-K) PDP164AAA-EGFP fusion in protoplasts. (L) TST167AAA-EGFP fusion in protoplasts. (M) ΔIR-EGFP fusion in protoplasts. The 35S promoter was used to direct the expression of all constructs. The green and red colours represent GFP and autofluorescent signal of chlorophyll, respectively. Arrows and arrowheads indicate nuclei and punctate signals, respectively. Arrows indicate nuclei. Scale bars represent 20 μm.

A

```
                 140        150                                                 190        200
wtSAHH1          GGDATLLIHEGVKAEEIFEKTGQVPDPTSTDNPEFQIVLSIIKEGLQVDPKKYHKMKERLV
ΔIR150-177       GGDATLLIHE.........................LSIIKEGLQVDPKKYHKMKERLV
ΔIR178-190       GGDATLLIHEGVKAEEIFEKTGQVPDPTSTDNPEFQIV.............KYHKMKERLV
PDP164AAA        GGDATLLIHEGVKAEEIFEKTGQVAAATSTDNPEFQIVLSIIKEGLQVDPKKYHKMKERLV
TST167AAA        GGDATLLIHEGVKAEEIFEKTGQVPDPAAADNPEFQIVLSIIKEGLQVDPKKYHKMKERLV
```



3.5. Chloroplast localization of ADK1 and SAHH1

Although all previous results using EGFP fusions with full-length SAHH or ADK showed the consistent localization of both proteins in the nucleus and cytoplasm, their fluorescent signals were often detected in or on the surface the chloroplast, as well. The green fluorescence in transiently expressed protoplasts with ADK1- and SAHH1-EGFP fusion proteins often closely overlapped with the red chlorophyll autofluorescence of a few chloroplasts (Figure 10A and B; arrowheads). Occasionally GFP fluorescence was observed in the nuclear-like structure (nucleoid) of chloroplasts (Figure 10A and B; arrows). In addition, Arabidopsis transgenic plants expressing either ADK1 or SAHH1 fusion proteins also displayed green fluorescent signals within a few chloroplasts (Figure 10E and F), whereas this was never observed in wild-type and transgenic plants expressing an EGFP control protein (Figure 10C and D).

To investigate the possible chloroplast localization of ADK and SAHH, Arabidopsis transgenic plants co-expressing a C-terminal DsRed2 fusion of small subunit of rubisco complex (RbcS) along with ADK- or SAHH-EGFP fusion were created. For this, full-length *RbcS* (At1g67090) was cloned from a cDNA prepared from wild-type Arabidopsis plants and examined in tobacco and Arabidopsis plants after fusion with EGFP and DsRed2 (Figure 11A-F). Both fluorescent signals of RbcS-EGFP and RbcS-DsRed2 fusion proteins were successfully detected in tobacco (Figure 11A and B) and Arabidopsis transgenic plants (Figure 11C). In addition, irregularly-spaced punctate-like signals of RbcS fusion protein were observed in tobacco epidermal cells. These bodies appear to be completely overlapping with the autofluorescence of chlorophyll (Figure 11D-F; arrows).

Arabidopsis plants co-expressing RbcS-DsRed2 and either ADK1-2xGFP or SAHH1-2xGFP displayed some green fluorescent signals in chloroplasts, although these signals did not completely overlap with RFP signals of RbcS fusion protein (Figure 11G-L; arrowheads). These results suggested that both ADK and SAHH may also be transiently associated with chloroplasts, and localization is limited to the surface of the organelle in most cases.

The $PDP^{164-166}$ and $TST^{167-169}$ mutated SAHH-GFP fusions were further examined by transient expression in protoplasts to examine whether these IR mutations affect chloroplast localization of SAHH as they did for the nuclear localization of SAHH. The protoplasts expressing the $TST^{167-169}$ mutated SAHH-GFP fusion had fluorescent signal patterns indistinguishable from those of wild-type SAHH-GFP fusion (i.e. in the cytoplasm and surrounding the nuclei of chloroplasts) (Figure 12A and C). However, the $PDP^{164-166}$ mutated SAHH-GFP expressing protoplasts displayed only a punctate-like GFP pattern as shown in Figure 9, with almost no signal detected in the cytoplasm and chloroplasts (Figure 12B). Interestingly, these punctate signals often appeared to be detected near the nucleus of chloroplast in most of the protoplasts viewed, but they did not completely overlap. Perhaps these punctate GFP signals indicate the presence of cellular protein aggregates generated by the mislocalization of SAHH that is supposed to localize in the chloroplasts, or they may represent localization to other organelles such as mitochondria or peroxisomes based on their size and patterns.

Together these results show that both ADK and SAHH localize not only to the nucleus and cytoplasm, but may also be targeted to the chloroplast, although their localization

signals in the chloroplast are not as consistent. Moreover, the three amino acids PDP[164-166] in the IR may contribute to chloroplast localization of SAHH.

Figure 10. ADK1 and SAHH1 GFP fusions showing chloroplast localization.

(A and B) Expression of ADK1-EGFP and SAHH1-EGFP fusion proteins in protoplasts. (C-F) EGFP expression in wild-type (C) and transgenic Arabidopsis plants harboring EGFP alone (D), ADK1-EGFP (E), and SAHH1-EGFP (F) fusions. The CaMV 35S promoter directed the expression of each fusion protein. Merged images of the GFP and chlorophyll autofluorescence are shown on the right. GFP and chlorophyll autofluorescence are represented by green and red, respectively. Arrowheads and arrows indicate chloroplast localization and the nuclei of chloroplasts, respectively. Scale bars represent 10 μm.

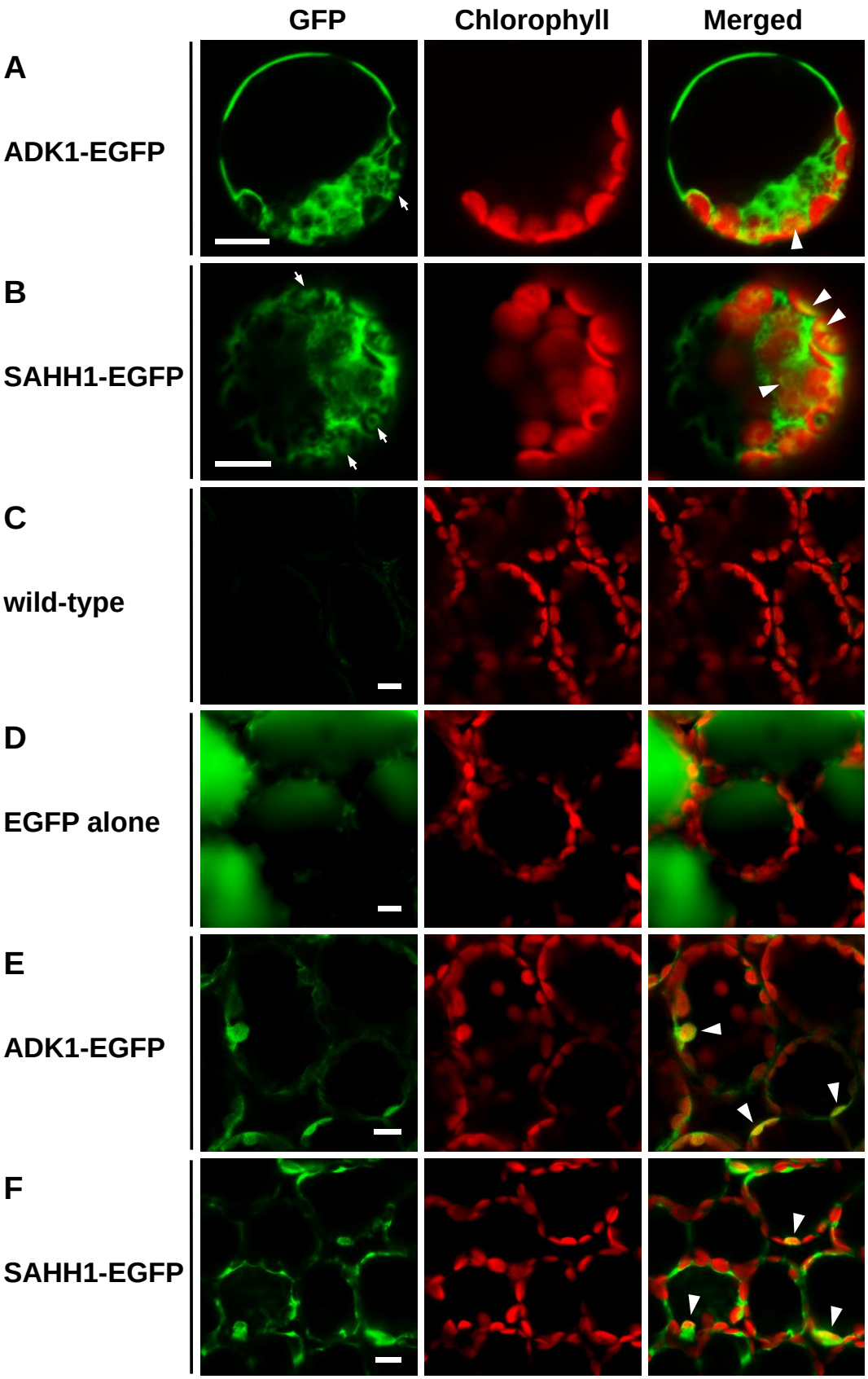

GFP
Chlorophyll
Merged
A
ADK1-EGFP
B
SAHH1-EGFP
C
wild-type
D
EGFP alone
E
ADK1-EGFP
F
SAHH1-EGFP

Figure 11. Subcellular localization of the small subunit of Rubisco complex, ADK1 and SAHH1.

(A-C) Chloroplast localization of RbcS fused to EGFP (A and B; in tobacco plants) and DsRed2 (C; in Arabidopsis). (D-F) The plastidic localization of RbcS-EGFP in tobacco epidermal cells. (G-L) Transient expression of ADK1-2xGFP and SAHH1-2xGFP along with RbcS-DsRed2 in tobacco plants. Note the CaMV 35S promoter directs expression of all these constructs. GFP and chlorophyll autofluorescent signals (A, B, E, and F) are represented by green and red, respectively; DsRed2 is shown in red (C, G, and J); Merged images of the GFP and chlorophyll autofluorescence (A, B, and F) and of the GFP and DsRed2 (I and L) are shown yellow. Arrows indicate plastid localization of GFP and chlorophyll in the tobacco epidermal cells; arrowheads indicate colocalization of EGFP fusions with RbcS. Scale bars represent 10 μm.

.

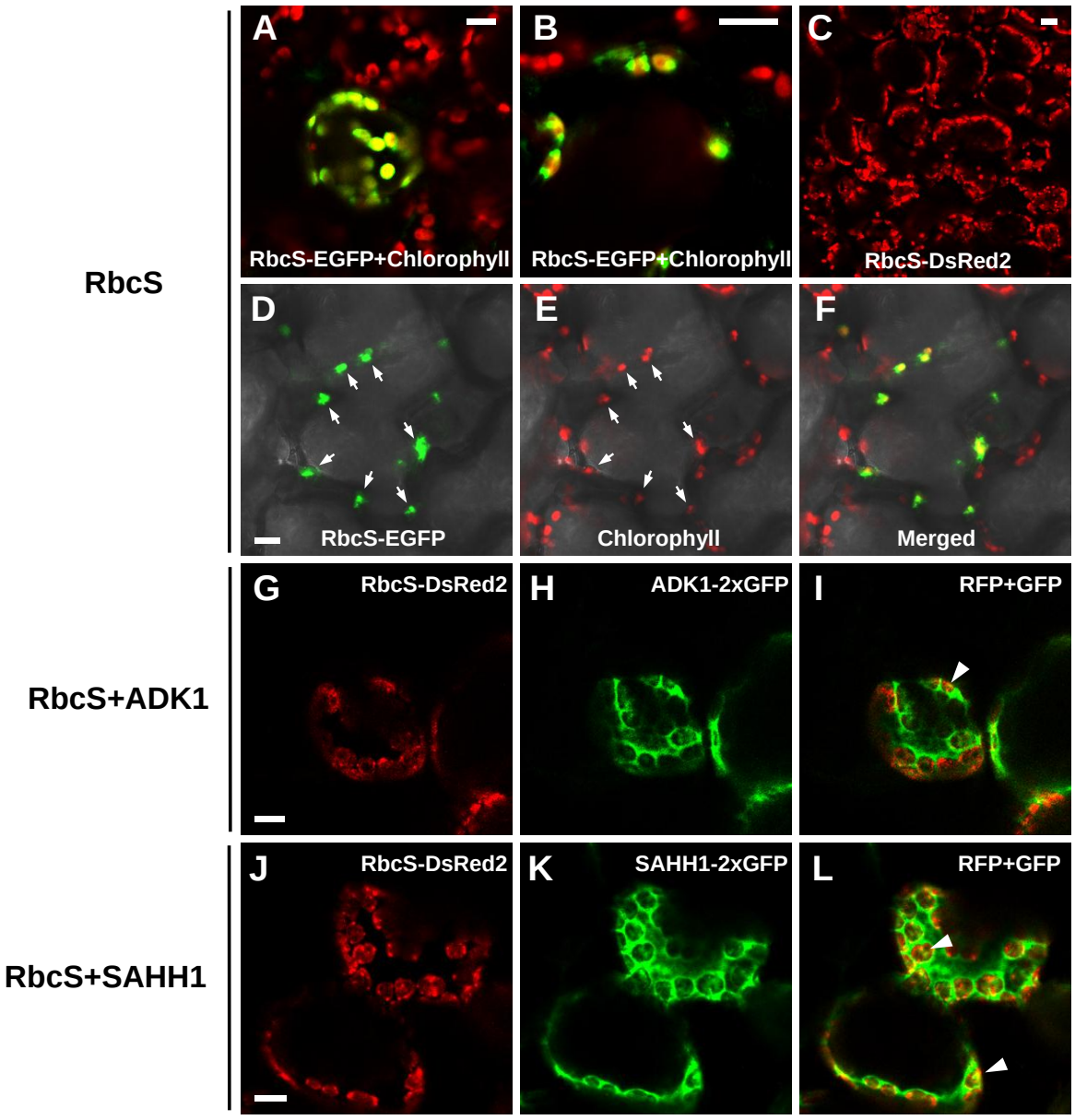
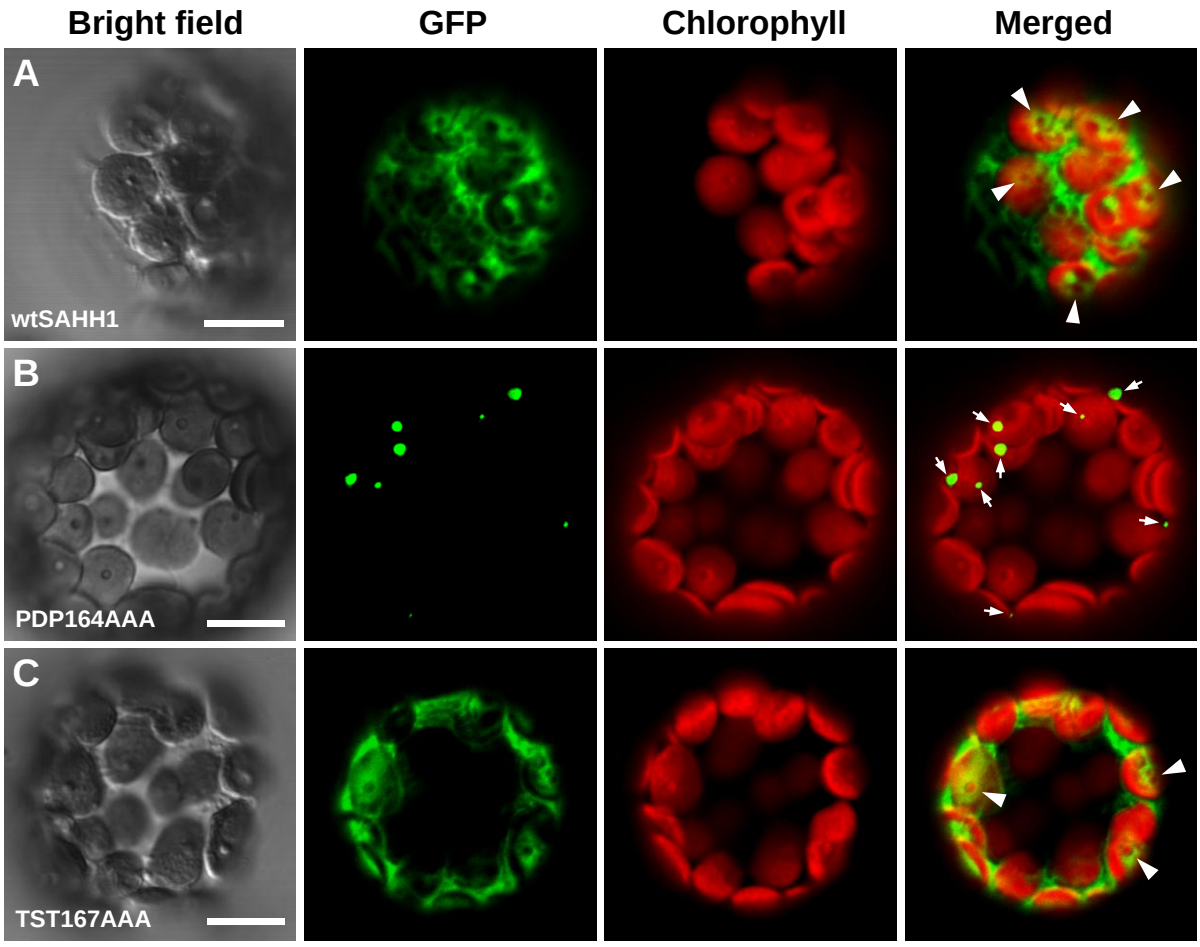

RbcS
RbcS+ADK1
RbcS+SAHH1
A
RbcS-EGFP+Chlorophyll
B
RbcS-EGFP+Chlorophyll
C
RbcS-DsRed2
D
RbcS-EGFP
E
Chlorophyll
F
Merged
G
RbcS-DsRed2
H
ADK1-2xGFP
I
RFP+GFP
J
RbcS-DsRed2
K
SAHH1-2xGFP
L
RFP+GFP

Figure 12. Chloroplast localization of SAHH1 in Arabidopsis protoplasts.

(A) Transient expression of the full-length of SAHH1-EGFP fusion protein in protoplasts. (B) Expression of the $PDP^{164-166}$ mutated SAHH1-EGFP fusion protein showing the punctate pattern of GFP signals. (C) Expression of the $TST^{167-169}$ mutated SAHH1-EGFP fusion protein. Note the CaMV 35S promoter directs expression of all these constructs. GFP and chlorophyll autofluorescent signals are represented by green and red, respectively; Bright field images of protoplasts are shown on the left; Merged images of the GFP and chlorophyll autofluorescence are shown on the right. Arrows indicate punctate GFP signals; arrowheads indicate nuclei of chloroplasts. Scale bars represent 10 μm.


Bright field
GFP
Chlorophyll
Merged
A
wtSAHH1
B
PDP164AAA
C
TST167AAA

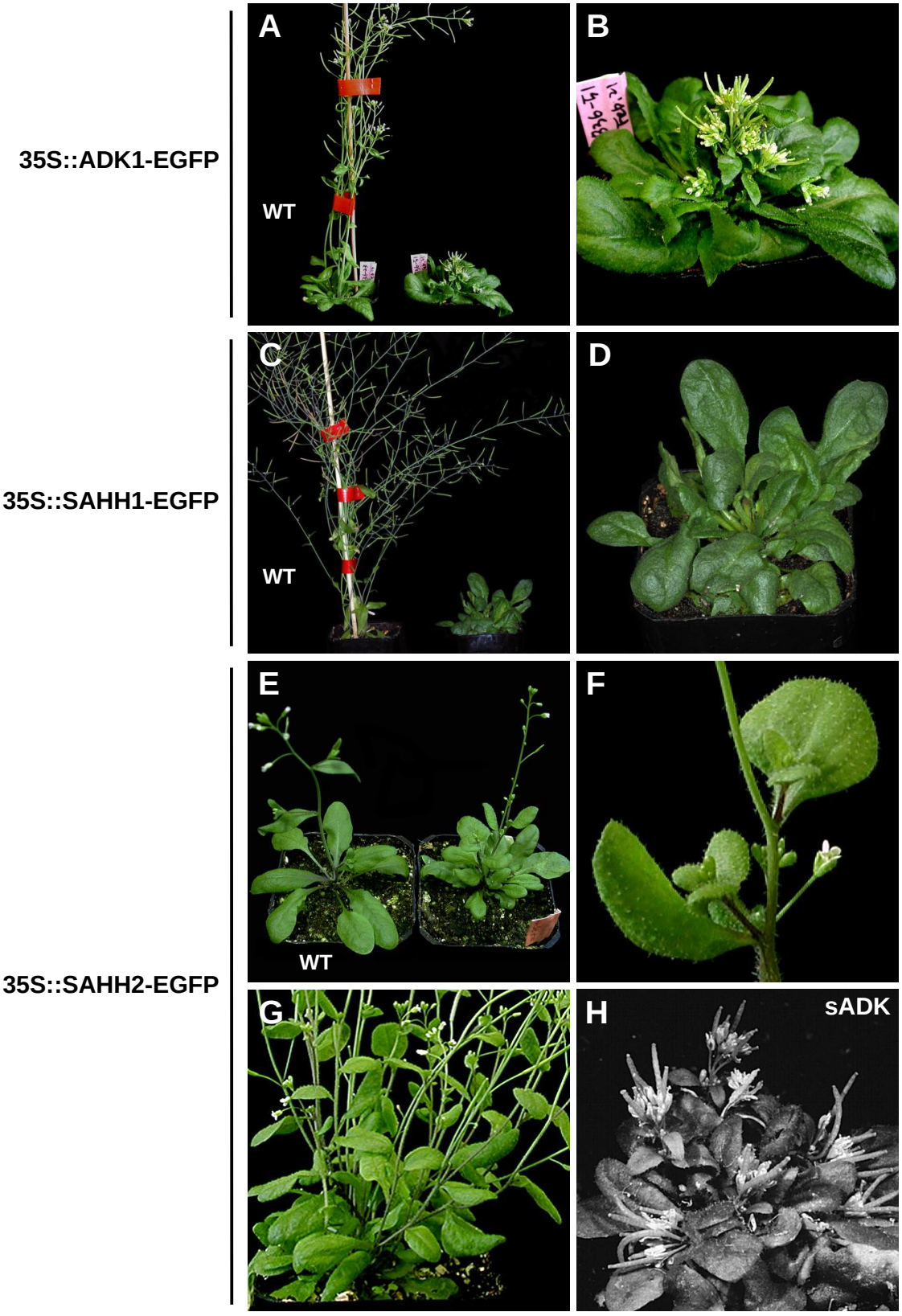
3.6. Plants arose from the overexpression of GFP-fusion constructs

Three different lines overexpressing ADK1-EGFP, SAHH1-EGFP, or SAHH2-EGFP from the constitutive 35S promoter were created to study the subcellular localization of ADK and SAHH. In each case about 5-10% of the lines displayed abnormal phenotypes as shown in Figure 13. ADK1-EGFP plants displayed developmentally abnormal phenotypes including short internode length, clustered inflorescences, wrinkled leaves, and delayed senescence (Figure 13A and B). These phenotypes are identical to those previously observed in sADK plants that overexpress the ADK1 cDNA (Figure 13H)(Moffatt et al., 2002). Developmentally abnormal phenotypes were also observed in SAHH1-EGFP plants (Figure 13C and D). These plants showed delayed senescence and abnormal leaf development with a large number of rosette leaves. They never induced shoots and inflorescences so that the plants failed to set seed. Lastly, one of SAHH2-EGFP T1-plant lines exhibited abnormal leaf and shoot development (Figure 13E-G). This plant had no cauline leaves in the primary shoot until they were 7-9 weeks old, whereas a number of rosette leaves were formed (Figure 13E). The cauline leaves formed in the lateral shoots were rounded unlike those of wild-type plants (Figure 13F). In the late plant growth stages, this plant displayed a bushy phenotype with an increased number of secondary branches (Figure 13G). In addition, no GFP fluorescence was detected from more than 60 basta resistant T1-transgenic plants expressing 35S::SAHH2-GFP examined (data not shown). This result indicates that the abnormal phenotype might be caused by silencing of SAHH2, but this line needs further exploration.

Inverse PCR was performed to verify if the abnormal phenotype of SAHH1-EGFP plants arose from a T-DNA insertional mutation (data not shown). The inverse PCR result revealed that the T-DNA containing SAHH1-EGFP had been inserted into the first exon of the *ATE1* gene (At5g05700), which encodes arginine-tRNA-protein transferase (EC. 2.3.2.8) that catalyzes the post-translational conjugation of arginine to the N-terminus of a protein in yeast and mammals (Kwon et al., 1999). *ate1* mutant plants exhibit the delayed senescence phenotype in leaves as observed in SAHH1-EGFP plants, but otherwise had a normal morphology and development (Yoshida et al., 2002). These results indicate that the abnormal phenotypes of SAHH1-EGFP plants could in part be due to the T-DNA insertion. However, the overexpression of SAHH1-EGFP might also affect the plant phenotypes since several other T1 plants also exhibited similar phenotypes including abnormal leaf and floral development as well as the delayed senescence.

Figure 13. Plants generated by the expression of various GFP-fusion constructs.

(A) Seven-week-old wild-type and T1-transgenic Arabidopsis plants containing 35S::ADK1-EGFP. (B) Enlarged image of ADK1-silenced plant shown in A. (C) 7.5-week-old wild-type and T1-transgenic Arabidopsis plant carrying 35S::SAHH1-EGFP. (D) Enlarged image of putative SAHH1-silenced plant shown in C. (E) 5-week-old wild-type and T1-transgenic Arabidopsis plant carrying 35S::SAHH2-EGFP. (F) 7-week-old SAHH2-EGFP plant showing abnormal shape of cauline leaves. (G) 12-week-old SAHH2-EGFP plant with a bushy phenotype. (H) sADK plant created by sense and antisense expression of the *ADK1* cDNA (Moffatt et al., 2002).


35S::ADK1-EGFP
35S::SAHH1-EGFP
35S::SAHH2-EGFP
A
B
WT
C
D
WT
E
F
WT
G
H
sADK

4. Discussion

4.1. Determination of subcellular location using fluorescent proteins

In this study, the subcellular localization of ADK and SAHH were examined using fluorescent fusion proteins such as EGFP and DsRed2. Fluorescent proteins have been used for localization studies in a wide variety of systems since GFP was first cloned in 1992 (Prasher et al., 1992). Fluorescent proteins are useful in labeling cells/proteins and monitoring the localization or dynamic movement of proteins without addition of exogenous substrates (Chalfie et al., 1994). However, it also has to be considered that incorrect or false localization results could be caused by many different factors, which could lead to misinterpretations of the data. For instance, the addition of the 27 kDa GFP to a protein of interest may affect protein stability or protein folding and thereby affect their functionality and localization (Marcus et al., 2001; Sedbrook, 2004).

Taken together, the determination of the correct localization of a target fusion protein with low false-positive rates, requires the confirmation of its transient overexpression by stable expression methods in multiple independent plant lines, with the GFP placed at the N- or C-terminus of its fusion partner. However, the limitation of this fluorescent tagging approach can never be completely overcome, since the use of the additional protein tag may alter the nature of the target protein in any case, and even the folding of fluorescent protein may be also disrupted by its fusion partner resulting in the weak fluorescent signal intensity (Marcus et al., 2001; Sedbrook, 2004).

To minimize these artifacts, various different constructs were created for this research including fusion proteins with different linker sizes in between GFP and a protein of interest, and those fusing GFP to either N- or C- terminus of the target sequence. Furthermore, the GFP fusion proteins were expressed from each protein's native promoter or from the constitutive 35S promoter, to test whether the degree of expression altered the subcellular localization of the GFP fusions. Finally, these different GFP fusion vectors were examined in three different plant systems including tobacco plants and Arabidopsis protoplasts using transient expression methods, and stable transgenic Arabidopsis transformants.

The majority of results from these different experiments consistently indicated that ADK and SAHH are targeted to the nucleus, cytoplasm, and possibly to the chloroplast as well. However, different patterns of GFP signals were observed based on the difference in plant systems or whether the GFP was placed either at the N- or C-terminus of the protein.

For instance, transiently expressed IR-deleted SAHH1-EGFP displayed nuclear localization in tobacco plants (data not shown), whereas no plant showing fluorescent signals in the nucleus was found from more than 20-stably transformed T1 Arabidopsis plants carrying the same construct (Figure 6N). This result indicates that the localization patterns of overexpressed GFP fusions with Arabidopsis proteins could be altered when transiently expressed in tobacco plants. Moreover, fusion of GFP to the N-terminus of yeast SAHH showed weak or no green fluorescence in the nucleus, while fusion to its C-terminus showed strong fluorescent signals in the nucleus (Figure 8J-M). This discrepancy may result from protein-folding issues with the C-terminal fusion. The

additional GFP tag on the C-terminal of yeast SAHH may have affected the target protein properties by interfering with its protein folding. On the other hand, both fusions may enter the nucleus by diffusion as the MW of N-terminal (76.8 kDa) and C-terminal (77.4 kDa) fusions is close to the diffusion limit (70 kDa) for the nuclear pore (D'Angelo et al., 2009). However, this case can be eliminated since SAHHs are active only when they are in the tetrameric mature forms (~212 kDa without additional tags), which would eliminate diffusion. Thus, the nuclear localization of C-terminal fusion of yeast SAHH could possibly have occurred by diffusion if it does not properly form tetrameric structure. However, this needs further testing with double GFP fusions to confirm the diffusion issue.

4.2. Methyl recycling enzymes localize to multiple compartments

The results using various GFP fusions showed strong fluorescent signals of both ADK and SAHH in the nucleus and cytoplasm in Arabidopsis (Figure 1 and 3). These results are consistent with those obtained using both immunogold labeling and GUS fusions for ADK and SAHH (Pereira, 2004; Schoor, 2007). In addition, the nuclear localization of ADK and SAHH are consistent with those of previous studies for human ADK (Cui et al., 2009) and *Xenopus* SAHH (Radomski et al., 1999).

Nuclear localization of ADK and SAHH was examined using fusions to two GFPs to verify if they entered the nucleus by simple diffusion that occurs when a protein is below the 70 kDa exclusion limit of the nuclear pore (D'Angelo et al., 2009). The green fluorescent signals of both ADK1-2xGFP (92 kDa) and the SAHH1-2xGFP (107 kDa for

monomer) were clearly co-localized with RFP signals of NLS control, indicating that both SAHH and ADK are targeted to the nucleus by an active localization process (Figure 4).

The ADK and SAHH double GFP fusion proteins also appeared to be occasionally associated with the chloroplast, although the fluorescent signals in chloroplasts were not as strong as those observed in the nucleus and cytoplasm (Figure 10 and 11). The signals were often detected near/on the surface of chloroplasts or near the nuclear-like structure of chloroplasts (nucleoid) (Figure 10-12). These chloroplast localization results are consistent with those obtained using immunogold labeling and GUS fusions as well as *in vitro* import assays (Schoor, 2007).

In addition, the earlier work from the Moffatt laboratory using immunogold labeling demonstrated that SAH hydrolase may also be capable of localizing to the mitochondrion, although the gold particles indicating the localization were not very abundant in the organelle (Schoor, 2007). However, no evidence for mitochondrial localization was found in the current study using fluorescently tagged SAHH.

Taken together with the results of this study, I conclude that both methyl recycling enzymes reside in the cytoplasm and nucleus, and have a transient association with plastids.

4.3. How are methyl-recycling enzymes localizing to multiple compartments?

The next question addressed was how both enzymes are localizing to multiple compartments. While most plant proteins reside in a single subcellular compartment, a number of proteins are located in more than one location, which may expand the protein's function (Mackenzie, 2005). For example, while 2,063 of the 2,446 Arabidopsis proteins are found in a single location based on mass spectrometry-derived location datasets, many reside in multiple compartments: more than 300 in two locations, more than 50 in three locations, 17 in four locations, and 1 in five different locations are claimed (Heazlewood et al., 2005).

The most common scenario for the multiple localization of proteins is dual targeting to either the nucleus and chloroplast (Schwacke et al., 2007), the peroxisome and chloroplast (Reumann et al., 2007), the endoplasmic reticulum and chloroplast (Levitan et al., 2005), or the mitochondrion and chloroplast (Carrie et al., 2009). However, it is unclear whether these dual targeted proteins have the same function in both organelles or a distinct function in each location (Mackenzie, 2005). In addition, the dual localization of isoproteins (or isoenzymes; for proteins with the same activity but different amino acid sequences) derived from two translation products is often missed by the protein localization analyses due to a highly uneven distribution of isoproteins between subcellular compartments (Regev-Rudzki and Pines, 2007). For example, one of the isoproteins that exists relatively abundantly in one location may conceal the discovery of the other isoprotein existing in the other location with its relatively low amount. In this manner the subcellular localization and functions of one isoprotein may be eclipsed by

those of the other. This phenomenon is known as an ‘eclipsed distribution’ (Regev-Rudzki and Pines, 2007).

Almost 50 proteins from seven different plant species have been experimentally demonstrated to be dual-targeted to both plastids and mitochondria (Carrie et al., 2009). Although the mechanism of dual targeting to the mitochondrion and chloroplast of plant proteins is not clearly understood yet, it is thought to occur by two different routes referred to as dual targeting signals (ambiguous presequences) and alternative transcription initiation or splicing (twin presequences) (Yogev and Pines, 2010). Based on the similarity in the mitochondrion and chloroplast, the dual targeting signals are recognized by both organelles through the specific interaction with the different receptors, Tim/Tom complexes of mitochondria and the Tic/Toc complexes of chloroplasts, which are located at the inner or outer membrane of each organelle (Carrie et al., 2009). On the other hand, distinct protein products may arise from one gene, initiating from different transcription or translation initiation sites, or produced by alternative exon splicing (Millar et al., 2006). For instance, organellar γ-type Arabidopsis DNA polymerase γ2 shows two distinct subcellular localization patterns upon the two translation start sites (Christensen et al., 2005). One, which uses the standard ATG for its translation, is targeted to chloroplasts, whereas a second form that initiates 21 nucleotides up-stream is dual-targeted to mitochondria and chloroplasts. In the latter case, the translation initiates at a CTG codon (coding for leucine) instead of at traditional ATG start codon (coding for methionine) (Christensen et al., 2005).

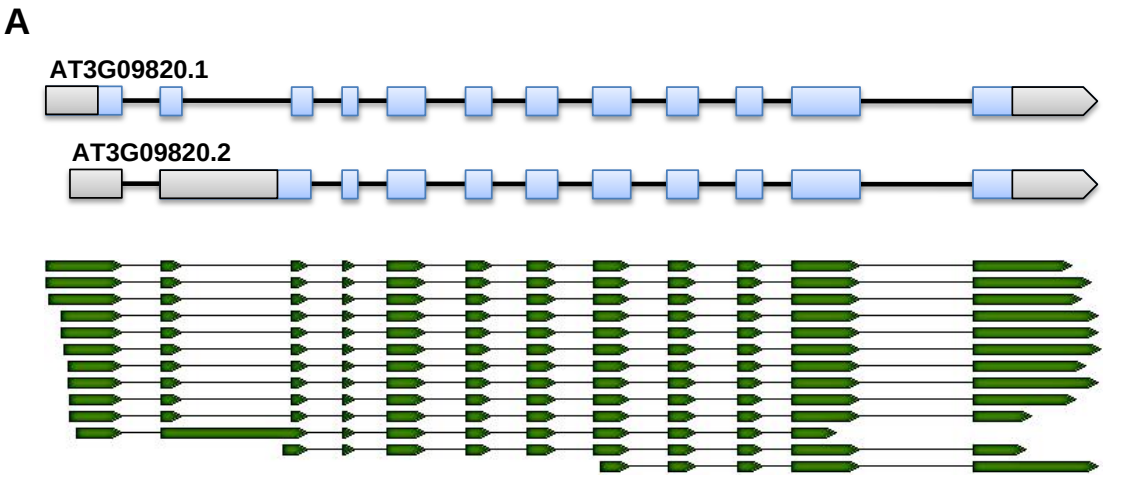
Protein-L-isoaspartate methyltransferase (PIMT) is an example of such a protein. This enzyme, which catalyzes the conversion of uncoded L-isoaspartyl residues to L-aspartate,

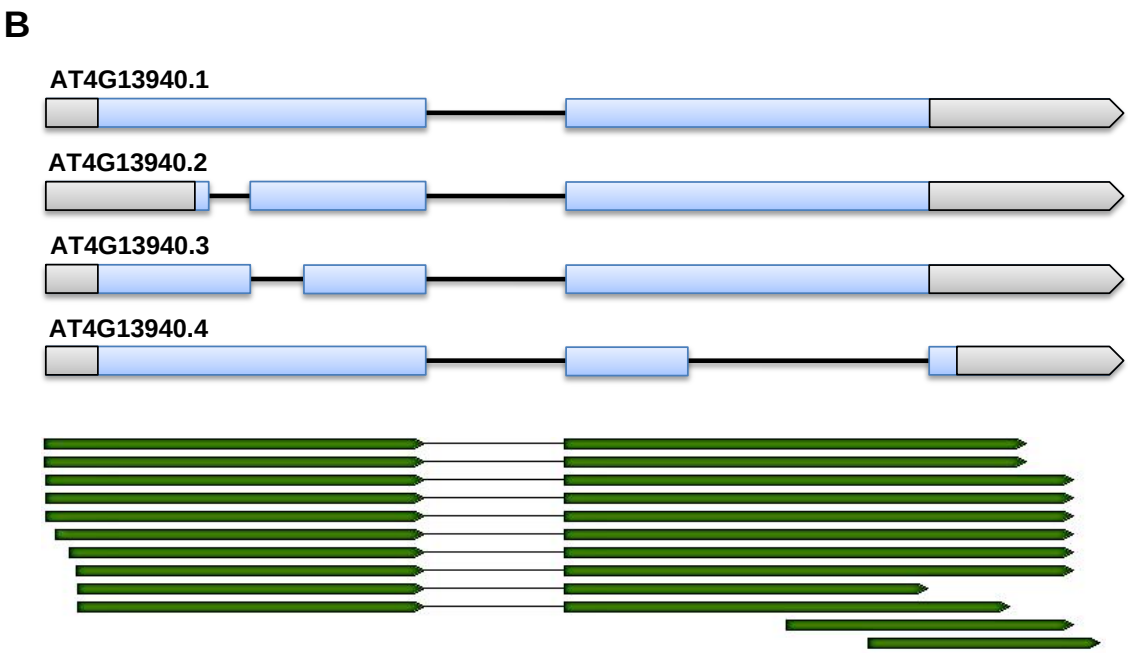
is the product of both alternative transcriptional initiation and alternative splicing (Dinkins et al., 2008). The different PIMTs are detected in distinct subcellular compartments including the cytoplasm, endo-membrane system, chloroplast, and mitochondrion (Dinkins et al., 2008).

Neither of these mechanisms fully explains multiple localizations of both ADK and SAHH, since neither enzyme contains targeting signal peptides, and alternative transcriptional/translational initiation or alternative splicing has not been found in the corresponding gene models. Both genes have multiple gene models: two for ADK1 and four for SAHH1, based on EST collections (TAIR, 2010)(Figure 14). However, the effect of these different gene models cannot be the explanation for the differential targeting because each enzyme is dual targeted when expressed as a cDNA under the control of a heterologous promoter. Instead, it is likely that the localization patterns of ADK and SAHH rely on alternate pathways such as protein-protein interactions or trafficking mediated by other proteins. These processes may act in response to the varying methyl demands of plant cells.

Figure 14. Predicted gene models of *ADK1* and *SAHH1* based on EST collections.

EST clusters for each gene are shown in green. The predicted gene models derived from each EST cluster is shown at the top with the boxes representing the exons. UTR regions are shown in gray. *ADK1* gene models have eleven or twelve exons; *SAHH1* gene models have two or three exons. AT3G09820.1 and AT4G13940.1 sequences are used in this study for ADK1 and SAHH1, respectively. Image is modified from TAIR (http://www.arabidopsis.org).


A
AT3G09820.1
AT3G09820.2


B
AT4G13940.1
AT4G13940.2
AT4G13940.3
AT4G13940.4

4.4. Insertion region of SAHH1 may control its subcellular targeting

The lack of known targeting signals in SAHH led to a search for residues important for its subcellular localization. The results of the deletion analysis demonstrated that 41 amino acids (Gly^{150}-Lys^{190}) are crucial for SAHH nuclear targeting, since no green fluorescent signal was detected in the nucleus of both tobacco and Arabidopsis plants expressing SAHH1ΔIR-EGFP fusion (Figure 6 and 8). When the IR was fused to a heterologous protein (2xGFP), the green fluorescence was seen almost everywhere in the cell including the nucleus and the cytoplasm (Figure 6E). This result indicates that the IR is not an autonomous nuclear localization signal or that the strong 35S promoter overwhelmed the system.

The alignment of known SAHH sequences revealed that the IR is only present in plant and parasitic protozoan SAHHs of eukaryotes, as well as those of most prokaryotes including cyanobacteria and proteobacteria (Figure 7). This suggests that the IR may contain information to link SAHH localization/activity with the photosynthetic/energetic capacity of plastids and mitochondria, since both prokaryotic origin organelles entered the eukaryotic lineage through endosymbiosis (Gould et al., 2008; Reyes-Prieto et al., 2007). A SAHH from *S. cerevisiae* which lacks the IR was introduced into Arabidopsis and tobacco plants with a N- or C-terminal EGFP tag to test this hypothesis. The N-terminal fusions consistently resided only in the cytosol whereas the C-terminal fusion localized to both the nucleus and cytoplasm (Figure 8J-M). The nuclear localization of C-terminal fusion of yeast SAHH might be generated by the mislocalization due to its protein folding or diffusion issue as discussed earlier. Nevertheless, assuming the N-

terminally tagged yeast SAHH is normally folded, its cytosolic localization is consistent with a requirement for an IR sequence for nuclear targeting.

Mutational analysis of residues within the IR also supported the hypothesis that this small region has a key role for the nuclear localization of SAHH (Figure 9 and 12). For this, three parts (both ends and the middle) of IR were deleted or substituted. Deletion of the first 28 residues of the IR (ΔIR150-177-EGFP) or the terminal 13 residues (ΔIR178-190-EGFP) eliminated or greatly reduced the nuclear targeting of SAHH, respectively. Substitution of three residues from the middle of IR ($PDP^{164-166}$ and $TST^{167-169}$ to three alanines) did not affect targeting to nuclei in Arabidopsis plants (Figure 9). Taken together these results indicate that both ends of the IR contribute to nuclear targeting of SAHH, with the N-terminal part being the more essential of the two; neither set of point mutations make a detectable difference in the dual targeting of SAHH.

Interestingly, both tobacco plants and Arabidopsis protoplasts transiently expressing the PDP164AAA-EGFP displayed punctate patterns of fluorescence. These signals were regularly arranged surrounding epidermal cell layer of tobacco leaves (Figure 9F) and randomly arranged around protoplasts (Figure 9K). Furthermore, GFP signals in the cytoplasm and chloroplast were barely detectable in protoplasts (Figure 12B). The punctate signals were also (but not always) detected in several protoplasts expressing SAHH1ΔIR-EGFP (Figure 9M) and TST167AAA-EGFP (data not shown).

While understanding the basis of these bodies will require more study, the punctate signals may represent different localizations of SAHH in other organelles such as mitochondria, Golgi or peroxisomes. This suggestion is based on the similarity of the

patterns to those observed from previous studies for a SAM transporter (SAMC1-GFP) in mitochondria (Palmieri et al., 2006), a dynamin-like protein (ADL6-GFP) and a manganese (AtMTP11–EYFP) in Golgi (Jin et al., 2001; Peiter et al., 2007), and a isopentenyl diphosphate isomerase (IPI1:short-GFP) in peroxisomes (Sapir-Mir et al., 2008).

Alternatively, the punctate pattern of these GFP fusion proteins may arise from aggregates of PDP164AAA-EGFP that fail to localize normally. A similar pattern was observed in a study of the ankyrin repeat protein (AKR2A)-mediated import of chloroplast outer envelope membrane (OEM) proteins in Arabidopsis (Bae et al., 2008). The authors demonstrated that AKR2A, which localizes to chloroplasts, interacts with the chloroplast OEM protein OEP7, and this interaction facilitates the targeting of OEP7 to chloroplasts. However, without this interaction OEP7 was not targeted to chloroplasts but aggregated around chloroplasts, resulting in a punctate staining pattern (Bae et al., 2008; Lee et al., 2001).

4.5. Protein interactions via its IR may direct SAHH to the destination

All the results described in this thesis demonstrate that the IR loop of SAHH has an essential role in SAHH's subcellular localization in Arabidopsis. Moreover, SAHH structural studies show that the IR appears to be located on the surface of the substrate-binding domain of each monomer, with no contact with other subunits (Reddy et al., 2008; Tanaka et al., 2004). So, the IR is separated from the main body of the tetramer and could

be accessible as a binding epitope. Thus, although the IR does not act as a nuclear localization signal (NLS) itself, it may serve as an interaction domain for proteins that have a NLS. The most likely candidate for this would be a methyltransferase specifically targeted to the nucleus, such as the mRNA cap methyltransferase (CMT), which mediated the translocation of *Xenopus* SAHH from the cytoplasm to the nucleus by a protein-protein interaction between the two (Radomski et al., 2002). Furthermore, the specific residues in the middle of IR ($PDP^{164-166}$ and $TST^{167-169}$) may also contribute to interactions directing SAHH subcelluar targeting, particularly to chloroplasts.

Small residues within the IR loop are predicted to be functionally significant by online bioinformatic tools such as the ELM server (http://elm.eu.org/) and the DILIMOT server (http://dilimot.embl.de/). In particularly, several regions are predicted as linear motifs based on the sequence similarities and patterns. Linear motifs are the short sequence patterns that play an important role in controlling highly dynamic protein-protein interactions or multiple protein complex formations (Diella et al., 2008). The short linear motifs (3-8 residues) that are usually exposed to potential binding partners allow proteins to interact with a similarly short region on the other proteins (Diella et al., 2008; Fuxreiter et al., 2007). They include various functional domains including post-translational modifications, targeting signals, cleavage sites and protein interaction modules (Diella et al., 2008; Neduva et al., 2005). Based on the results for IR sequence, the segment $VKAE^{151-154}$ is predicted as a potential sumoylation site that is recognized by small ubiquitin-like modifier 1 (SUMO-1) for post-translational modification. The sumoylation, which is similar to ubiquitination, plays a crucial role in several cellular processes such as transcriptional regulation, intracellular trafficking, and nuclear organization. All these

processes depend upon the action of sumoylation: altering protein activities, changing subcellular localizations, or providing a platform for protein-protein interactions (Geiss-Friedlander and Melchior, 2007). In addition, the segment EKTGQVPDPTSTDN$^{158\text{-}171}$ is predicted to be phosphopeptide binding domains. Within this segment, both Ser^{168} and Thr^{169} residues are potential phosphorylation sites (http://www.cbs.dtu.dk/services/NetPhos/), and the three residues, Pro^{164}-Asp^{165}-Pro^{166} are predicted as PxP motif that is important for protein channel gating and protein interaction (Labro et al., 2003). This double proline residue is recognized by proteins that contain Src homology 3 (SH3) domain, WW domain, or 14-3-3 domain for protein-protein interactions in a phosphorylation-dependent manner (Diella et al., 2008; Neduva et al., 2005; Verdecia et al., 2000). These results lead to the hypothesis that SAHH localizes to the nucleus and possibly plastids by a specific protein interaction via its IR loop region. Moreover, this interaction may be affected by the serine/threonine phosphorylation along with PxP motif and as well as by the sumoylation.

CHAPTER 2: LOCATION OF METHYL RECYCLING ENZYMES IS MEDIATED BY SPECIFIC PROTEIN INTERACTIONS

1. Introduction

Methylation is a key process in plant development and metabolism. A number of compounds, including secondary metabolites, cell wall components, and plant hormones, function upon the addition of a methyl group in response to various environmental stimuli (Weretilnyk et al., 2001). The major methyl donor for these reactions is *S*-adenosyl-L-methionine (SAM) and these SAM-dependent transmethylation reactions occur in all cellular compartments of plant cells to maintain plant growth and development (Poulton, 1981). As a by-product of these methyltransferase (MT; EC 2.1.1) activities, *S*-adenosyl-L-homocysteine (SAH) is generated after each transfer of the methyl group from SAM to an acceptor molecule (Poulton, 1981). However, SAH acts as a competitive inhibitor of MTs due to its similar structure to SAM and its higher binding affinity for MTs than for SAM. Thus, SAH must be actively removed to prevent its continuous accumulation in the cell (Poulton and Butt, 1975).

In eukaryotes, only one enzyme, *S*-adenosyl-L-homocysteine hydrolase (SAHH, EC 3.3.1.1), is capable of removing the by-product SAH by hydrolyzing it to homocysteine (Hcy) and adenosine (Ado) (De La Haba and Cantoni, 1959). However, since the hydrolysis reaction is reversible and the equilibrium lies far in the direction of SAH synthesis, both Hcy and Ado must also be metabolized to direct the reaction toward SAH degradation (De La Haba and Cantoni, 1959; Poulton, 1981). Ado can be metabolized by

three different enzymes: adenosine nucleosidase (EC 3.2.2.9; into adenine), adenosine deaminase (EC 3.5.4.4; into inosine) and adenosine kinase (ADK; EC 2.7.1.20; into AMP). Among these, ADK activity has been characterized as the major route for Ado salvage in plants, while the other two enzymes are minor or not found in plants, based on their transcript abundance and kinetic activities (Engel, 2009; Moffatt et al., 2002; Moffatt and Weretilnyk, 2001). Thus, the two enzymes SAHH and ADK are both essential to sustain SAM-dependent methylation activities in plants.

Both enzymes have been widely studied in eukaryotic organisms such as mammalian and plant cells. Their critical roles in mammal development are evidenced by the embryonic lethality in mice either by lacking SAHH activity (Miller et al., 1994) or by ADK deficiency (Boison et al., 2002). In plants, several mutants created by the deficiency of either SAHH or ADK have been characterized from various species. Tobacco plants expressing an *SAHH* antisense transgene exhibit abnormal development of floral organs with stunting of growth and delayed senescence (Tanaka et al., 1997). Point mutations in the Arabidopsis *SAHH1* (*HOG1*) gene have developmental abnormalities including slow growth, low fertility and poor germination, while a complete knockout of *HOG1* results in zygotic lethality (Rocha et al., 2005). Antisense *SAHH* gene expression in petunia causes delayed flowering, increased leaf size and higher seed yield (Godge et al., 2008). Arabidopsis SAHH mutant plants created by a T-DNA insertion in *SAHH1* gene have defects in root-hair development along with the other phenotypes observed in *hog1* plants (Wu et al., 2009). The only reported ADK-deficient plants (sADK) were created by transgene silencing and display developmentally abnormal phenotypes in leaf, shoot, and floral organs of Arabidopsis (Moffatt et al., 2002).

Protein-protein interactions are essential in regulating most cellular processes. A large number of proteins are known to interact with other proteins to accommodate and regulate the diversity of biological functions (Morsy et al., 2008). For example, the yeast protein interactome is estimated to comprise 16,000-26,000 different interaction pairs from 6,300 different yeast proteins (Grigoriev, 2003). Plant proteomes are estimated to have 75,000-150,000 interaction pairs from approximately 30,000-40,000 plant proteins (Morsy et al., 2008). In Arabidopsis, a total of 19,979 predicted interactions among 3,617 proteins were recently documented (Geisler-Lee et al., 2007). Subsequently, two other studies predicted 28,062 (Cui et al., 2008) and 224,206 (Lin et al., 2009) protein interaction pairs among Arabidopsis proteins. More recently, three protein-protein interaction databases, IntAct (http://www.ebi.ac.uk/intact/), TAIR (http:// www.Arabidopsis.org/), and BioGRID (http://www.thebiogrid.org/), compile all the experimentally validated Arabidopsis protein interactions reported in the literature. This data set reveals that a total of 4,660 interactions among 2,303 proteins have been experimentally verified in Arabidopsis to date (He et al., 2010). Studying these interactions can provide valuable insights into various mechanisms underlying protein functions and useful information of functional linkage between interacting proteins.

The subcellular localization data described in Chapter 1 of this thesis demonstrates that both ADK and SAHH are localized in the cytoplasm and the nucleus of Arabidopsis leaf cells, and possibly the chloroplast as well, even though no known targeting signals are detected in either protein. Thus, I hypothesized that the multiple localizations of these enzymes may be mediated by a specific protein interaction with other proteins. This interaction-mediated translocation is suggested by an earlier study of *Xenopus* SAHH

(xSAHH), which showed that xSAHH moves from the cytoplasm to the nucleus through an interaction with mRNA (guanine-7-) methyltransferase (xCMT) (Radomski et al., 2002). Further immunohistological localization studies demonstrated that the N-terminal 55 amino acids of xSAHH are sufficient for binding to xCMT (Radomski et al., 2002). However, the deletion analysis of Arabidopsis SAHH documented in Chapter 1 demonstrates that the insertion region (IR; Gly^{150}-Lys^{190}), located in the middle of the protein, is required for nuclear localization of Arabidopsis SAHH. Moreover, the study using a green fluorescent protein (GFP)-tagged IR polypeptide shows that the IR does not act as a nuclear localization signal (NLS) itself. These findings led to the hypothesis that SAHH is targeted to specific locations by a protein-protein interaction via its IR (Chapter 1).

Although SAHH is a highly conserved enzyme in all organisms (Stepkowski et al., 2005), the presence of the IR is a clear distinction between mammalian and plant SAHHs. Due to the additional 40-45 residues of IR in plant SAHHs (41-residues in Arabidopsis SAHH) their molecular weights are 53-57 kDa, as compared to the ~48 kDa in mammalian SAHHs. The IR is present in plants and parasitic protozoa of eukaryotes, and many prokaryotes including photosynthetic bacteria (Stepkowski et al., 2005). Crystal structures of SAHHs show that the IR does not affect the formation of the SAHH tetramer since it is lies on the outside of the monomer polypeptide (Reddy et al., 2008; Tanaka et al., 2004). Therefore, the surface-exposed IR could be accessible as a binding epitope, and thus may potentially provide interactions with other proteins.

The second part of the research outlined in this thesis focuses on the mechanism of the dynamic translocation of ADK and SAHH to maintain the MT activities within plant cells.

This involved examining if the movement of these methyl-recycling enzymes is mediated by a specific protein interaction.

2. Materials and Methods

2.1. Source of chemicals

Chemicals used in the research were obtained from Sigma-Aldrich, Bioshop, or BioBasic, unless otherwise noted. Taq polymerases were either generated from Moffatt lab (for colony PCR or T-DNA screening purpose) or purchased from New England Biolabs (Finnzyme's Phusion DNA Polymerase; for cloning purpose). Restriction enzymes and T4 DNA ligase were obtained from New England Biolabs or Fermentas. All primers were synthesized by Sigma-Aldrich.

2.2. Plant growth

Arabidopsis thaliana (accession Columbia) seeds were planted either on soil or agar media. Seeds were sterilized in a 1.5 mL microfuge tube containing approximately 0.1 ml of seeds for growth on media. The tube was placed in a glass desiccator jar beside a 400 mL beaker containing 100 mL of Javex (3-6% (v/v) sodium hypochlorite) and 3 mL of concentrated HCl. The seeds were sterilized by chlorine gas for 1 hour. Sterilized seeds were then sown onto a Petri plate containing 25 mL of half strength MS salt media (Murashige and Skoog, 1962) supplemented with B5 vitamins (2.5 mM MES, 30 g/L sucrose, 8g/L agar, pH 5.7-5.8). The plates were then incubated for 24-72 hours in the

dark and transferred to a TC7 tissue culture chamber (Conviron) with 24 hours light (150 ± 20 µmol m^{-2} s^{-1} PAR) at 20℃.

A 50:50 soil mixture of "Sunshine LC1 Mix" and "Sunshine LG3 Germination mix" (SunGro) filled in two different sizes of pots (25 or 530 cm^2) was used for growth on soil. After applying seeds, these pots placed in flats (1600 cm^2) were covered with plastic domes and incubated in the dark at 4℃ for 72 hours, and transferred to a growth chamber under long day conditions with 16 hours light (150 ± 20 µmol m^{-2} s^{-1} photosynthetically active radiation) at 22℃. Plants were watered and fertilized with 20:20:20 fertilizer (Plant-Prod) until they reached full maturity. Watering was then stopped and allowed to dry out and seeds were collected.

2.3. Stable transformation of Arabidopsis

Arabidopsis transformation was performed by the floral dip method (Clough and Bent, 1998). About 4-week-old healthy Arabidopsis plants containing immature floral buds were used for this method. *Agrobacterium tumefaciens* GV3101 strain carrying the transgene of interest was inoculated and grown overnight in 3 mL of Luria-Bertani (LB) media containing appropriate antibiotics at 28℃. This culture was then diluted 100-fold in 100 or 200 mL fresh LB media supplemented with the same antibiotics and incubated for another 16 hours. The culture was centrifuged at 5,500 rpm for 10 minutes, resuspended in 5 % sucrose and 0.05 % Silwet L-77 (Lehle Seeds, Cat. No. VIS-01) at an OD_{600} of 0.8, and placed in a tin foil baking dish (height 17 x, width 17, depth 5 cm). For dipping, the floral part of plants were submerged in the prepared solution and gently

agitated for 5-10 seconds. After floral dipping, the plants were covered with plastic dome to maintain high humidity and kept at room temperature away from direct sunlight for 24 hours. The plants were then transferred back to the growth chambers until seed set. T1 seeds of transformants carrying a *Bar* resistance gene were sown and selected on soil by spraying 40 mg/mL of the herbicide glufosinate ammonium (BASTA; "Wipeout" from Wilson Lab.) once a day for 1-2 weeks, when the plants are about 15-day-old with fully expanded first leaves. The successful T1 plants were transplanted into individual pots for further research.

2.4. Transient expressions in tobacco plants and Arabidopsis protoplasts

Transient expression in tobacco (*Nicotiana benthamiana*) plants was performed using the agro-infiltration method (Voinnet et al., 2003). For this, *A. tumefaciens* strains carrying the transgene of interest were inoculated in 3 mL of LB media containing appropriate antibiotics overnight at 28℃. This culture was diluted 100-fold in 6 mL fresh LB media supplemented with 10 mM MES (pH 5.6), 20 μM acetosyringone, and appropriate antibiotics, and incubated for another 16 hours. After centrifugation, the pellets were resuspended at an OD_{600} of 1.0 in 10 mM $MgCl_2$, 10 mM MES, 150 μM acetosyringone, and mixed with equal volumes of a culture of P19 strain (OD_{600} of 1.0). The mixtures were kept at room temperature for at 2-3 hours and then infiltrated onto the underside of 4-6 weeks old tobacco leaves with a gentle pressure. The infiltrated-plants were covered with plastic and kept at room temperature for 24 hours, and then transferred

back to the growth chambers. Fluorescent signals were monitored 3-5 days after infiltration under a confocal laser scanning microscopy.

Transient expression in Arabidopsis protoplasts was carried out using the PEG-mediated transfection method (Sheen, 1996). Protoplasts were isolated from 3-4 weeks old Arabidopsis leaves grown on a half strength MS salt media-containing plate. Leaf samples were collected using a r*azor blade* in a 50 ml tube containing 1% (w/v) cellulase R-10 (Yakult Honsha, Japan), 0.25% (w/v) macerozyme R-10 (Calbiochem, Germany), 20 mM MES, 0.1% (w/v) BSA, 400 mM mannitol, 10 mM $CaCl_2$, 20 mM KCl, and KOH (pH 5.6). After 3 hours incubation at 22℃ in darkness, the sample was filtered through a 100 μm nylon mesh and transferred to a new tube containing 21% (w/v) sucrose. After centrifugation at 98 x g for 10 minutes to remove the nonviable protoplasts and cellular debris, intact protoplasts floating on the upper layer were carefully washed twice with W5 solution (154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, 5 mM glucose, 1.5 mM MES, KOH, pH 5.6). The pellets were resuspended in 400 mM mannitol, 15 mM $MgCl_2$, 5 mM MES, KOH (pH 5.6) at a final concentration of 4×10^5 protoplasts/mL. For transfection, 100 μL of protoplasts (2×10^4) were gently mixed with 10 μL of plasmid DNA (20 μg) and incubated with 110 μL of PEG solution (40% w/v PEG, 400 mM mannitol, 100 mM $CaCl_2$) at room temperature for 30 minutes. The sample was washed twice with W5 solution, resuspended in 1 mL of W5 solution, and incubated in the dark or light condition for 24 hours. Fluorescent signals were then monitored under a confocal laser scanning microscopy.

2.5. Subcellular localization of CMT and ASDH

The coding sequence of *CMT* gene was isolated from wild-type Arabidopsis cDNA. For this, the full-length coding sequence for CMT (At3g20650) was amplified using a primer set (CMT-F Nco and CMT-R Bam; Table 1). The PCR product was then cloned into the *Nco*I and *Bam*HI sites of the pSAT6-EGFP-N1 vector (Tzfira et al., 2005) to generate the 35S:CMT-EGFP fusion vector. To construct pSAT6-EGFP-CMT, the *CMT1* gene was amplified using a primer set (CMT-F Hin and CMT-R Bam; Table 1) and cloned into the *Hin*dIII and *Bam*HI sites of the pSAT6-EGFP-C1 vector (Appendix I for the pSAT vector system).

To investigate the nuclear localization signal (NLS) sequence of CMT, several EGFP or DsRed2 fusion constructs were generated by PCR amplifications using specific primers for different regions of CMT. To create a construct pSAT6-CMTNLS-DsRed2, the predicted NLS sequence (RKRGESDGARRSGRR) was PCR amplified using a long forward primer (CMTNLS-F; Table 1) that contains nucleotide sequences responsible for the NLS sequence and a reverse primer (DsRed2R-Nco; Table 1) that is specific for the DsRed2. The amplified fragment was cloned into the *Nco*I site of the pSAT6-DsRed2-N1 vector.

For the construct pSAT6-CMTΔNLS-EGFP, the first 1,032 bp of *CMT* coding sequence was PCR amplified using a primer set (CMT-F Nco and CMT342-R; Table 1) and cloned into the *Nco*I and *Eco*RI sites of the pSAT6-EGFP-N1 vector, which resulted in the vector pSAT6-CMT(1-344)-EGFP. Using a primer set (CMT345-F and CMT-R Bam; Table 1), the last 84 bp of *CMT* coding sequence was PCR amplified and cloned

into the *Eco*RI site of the pSAT6-CMT(1-344)-EGFP. Since the two primers CMT342-R and CMT345-F (Table 1) contained nucleotide mutations to substitute the predicted NLS sequence (RKRGESDGARRSGRR) with an alternate sequence (SNSGEFYVPPTTVPR), these two-step cloning allowed the creation of the pSAT6-CMTΔNLS-EGFP vector.

For both constructs pSAT6-nCMT-EGFP and pSAT6-cCMT-EGFP, the vector pSAT6-CMT-EGFP was digested with either *Nco*I/*Pst*I or *Pst*I/*Bam*HI, which produces N-terminal CMT (177 aa) or C-terminal CMT (202 aa) fragment, respectively. Each DNA fragment was then ligated with the pSAT6-EGFP-N1 vector cut with *Nco*I/*Pst*I or *Pst*I/*Bam*HI.

The coding sequence for ASDH (At1g14810) was amplified using a primer set (ASDH-F Xho.Nco and ASDH-R Xma). The PCR product was then cloned into the *Nco*I/*Xma*I sites of pSAT6-EGFP-N1 vector to generate the 35:ASDH-EGFP vector. To test the *ASDH* native promoter, the 1.2-kb upstream region of the *ASDH* gene was amplified from wild-type Arabidopsis genomic DNA using a primer set (ASDHp-F Age and ASDHp-R Nco; Table 1). The PCR product was then cloned into the *Age*I/*Nco*I sites of the pSAT6-ASDH-EGFP vector to substitute the 35S promoter with the *ASDH* native promoter, which resulted in the pSAT6- ASDHp:ASDH-EGFP.

For control vectors for the nuclear and plastid localization, DsRed2 fusions with the NLS sequence of Simian virus 40 (SV40) large T antigen (pSAT4-SV40NLS-DsRed2) or with the Arabidopsis small subunit of Rubisco complex (RbcS; pSAT4-RbcS-DsRed2) were used, respectively (Chapter 1).

The expression cassette of all pSAT6 and pSAT4 vectors was transferred into PI-*Psp*I or I-*Sce*I unique recognition sites of the pPZP-RCS2 binary vector for plant transformation. To identify their subcellular localization, EGFP or DsRed2 fusions were transiently expressed in Arabidopsis protoplasts and tobacco plants, or stably transformed into Arabidopsis plants. Expression of the fusion proteins was visualized by confocal laser scanning microscopy (Carl Zeiss LSM 510; Jena, Germany) with different filer sets: GFP (excitation 488 nm, emission 505/530 nm), RFP (excitation 543 nm, emission 560/615 nm), chlorophyll autofluorescence (excitation 488 nm, emission 650 nm).

2.6. Bimolecular fluorescence complementation (BiFC)

The pSAT-BiFC vector system (Citovsky et al., 2008) was used to visualize protein-protein interactions in living cells. To do this, the target interacting proteins were cloned into either N-terminal (YN, amino acids 1–174) or C-terminal (YC, amino acids 175–238) fragments of the enhanced yellow fluorescent protein (EYFP). For ADK1-YN, a coding sequence for ADK1 was transferred from pSAT6-ADK1-EGFP (Chapter 1) to pSAT4-YN-N1 vector using *Nco*I and *Xma*I sites. For SAHH1-YC, the 1.4-Kb *Age*I/*Not*I fragment of the pSAT1-YC-N1 was replaced with the 1.9-Kb *Age*I/*Not*I fragment of the pSAT6-EGFP-N1 vector to create the pSAT6-YC-N1. The coding sequence for SAHH1 was transferred from the pSAT6-SAHH1-EGFP (Chapter 1) to the pSAT6-YC-N1 vector using *Nco*I and *Xma*I sites. For YC-SAHH1, the 1.4-Kb *Age*I/*Not*I fragment of the pSAT1-YC-C1 was replaced with the 1.9-Kb *Age*I/*Not*I fragment of the pSAT6-EGFP-N1 vector to create the pSAT6-YC-C1. The coding sequence for SAHH1 was transferred

from pSAT6-EGFP-SAHH1 (Chapter 1) to the pSAT6-YC-C1 vector using a *Kpn*I site. For SAHH1-YN, the coding sequence for SAHH1 was transferred from the pSAT6-SAHH1-EGFP (Chapter 1) to pSAT4-YN-N1 vector using *Nco*I and *Bam*HI sites. For SAHH1(Δ150-190)-YC, the SAHH1(Δ150-190) DNA fragment cut with *Nco*I/*Xma*I from the pSAT6-SAHH(Δ150-190)-EGFP (Chapter 1) was cloned into the pSAT6-YC-N1 vector. For atCMT fusion constructs, the coding sequence for atCMT was digested with *Nco*I and *Bam*HI from the atCMT-EGFP vector, and inserted into the pSAT4-YN-N1 vector and the pSAT6-YC-N1 vector to create the pSAT4-atCMT-YN and the pSAT6-atCMT-YC, respectively. For plant transformation, the expression cassette of various combinations of YN/YC fusion constructs was cloned into the pPZP-RCS2 binary vector using PI-*Psp*I and I-*Sce*I unique recognition sites for YN fusions and YC fusions, respectively. Their expressions were monitored by confocal laser scanning microscopy with a filer set (excitation 514 nm, emission 560/600 nm).

2.7. Strep tag purification

To identify protein complexes, StrepII affinity purifications were performed using a pXCS-HAStrep vector carrying the eight-amino acid StrepII epitope (WSHPQFEK) in plant cells (Witte et al., 2004). For construct pXCS-ADK1-Strep, the coding sequence for ADK1 was amplified from the pSAT6-ADK1-EGFP (Chapter 1) using a primer set (ADK1-F Xho and ADK1cR-Sma; Table 1), and cloned into the *Xho*I and *Xma*I sites of the pXCS-HAStrep vector. For pXCS-SAHH1-Strep, the 1.4-Kb *SAHH1* gene fragment cut with *Pst*I/*Xma*I from the pSAT6-YC-SAHH1 was cloned into the pXCS-HAStrep

vector. For pXCS-SAHH1(Δ150-190)-Strep, the SAHH1(Δ150-190) DNA fragment cut with *Nco*I/*Xma*I from the pSAT6-SAHH(Δ150-190)-YC was cloned into the pXCS-SAHH1-Strep vector. For ASDH-Strep, the coding sequence for ASDH was amplified from wild-type Arabidopsis cDNA using a primer set (ASDH-F Xho.Nco and ASDH-R Xma; Table 1), and cloned into the *Xho*I and *Xma*I sites of the pXCS-HAStrep vector. These Strep fusion vectors were then introduced into *A. tumefaciens* strain GV3101::pMP90RK for stable transformation of Arabidopsis plants. At least, twenty of T1 transgenic plants per each line were screened by spraying BASTA, and the expression of Strep fusion proteins in each T1 lines were confirmed by immunoblotting. For purification, 700 mg of leave tissues from 3-week old T2 plants were homogenized in 1 mL of extraction buffer [(100 mM Tris pH 8.0, 5 mM EDTA pH 8.0, 200 mM NaCl, 10 mM DTT, 0.5% (v/v) Triton X-100, 10% (v/v) glycerol, 1mM phenylmethylsulfonyl fluoride (PMSF), protease inhibitor cocktail (Roche)], and centrifuged at 20,000 x *g* for 15 minutes at 4℃. The supernatant was then transferred to a new microfuge tube, and incubated with 40 μL StrepTactin Macroprep (50% suspension; Novagen) in an end-over-end rotation wheel for 15 minutes at 4℃ to allow the binding of Strep fusion proteins and StrepTactin beads. After centrifugation at 700 x *g* for 30 seconds at 4℃, the beads were washed five times with 0.8 mL of wash buffer (100 mM Tris pH 8.0, 0.5 mM EDTA pH 8.0, 200 mM NaCl, 2 mM DTT, 0.5% (v/v) Triton X-100, 10% (v/v) glycerol, 1mM PMSF, protease inhibitor cocktail). Washed beads were then resuspended in one volume of 2 x SDS sample buffer [60 mM Tris-HCl, pH 6.8, 25% (v/v) glycerol, 2% (w/v) SDS, 0.1% (w/v) bromophenol blue, 5% (v/v) β-mercaptoethanol] and boiled for 6 minutes to

release bound proteins. The liquid phase containing the detached proteins was used for SDS-PAGE gel.

2.8. Protein identification by Mass spectrometry

Protein bands obtained from StrepII affinity purifications were excised from SDS-PAGE gels stained overnight with Colloidal Coomassie brilliant blue G-250. The gel pieces were washed with water and destained with 50 mM NH_4HCO_3/50% acetonitrile (ACN). Proteins were reduced by incubation with 10 mM DTT in 100 mM NH_4HCO_3 at 50°C for 30 minutes, alkylated by incubation with 55 mM iodoacetamide in 100 mM NH_4HCO_3 for 30 minutes, and dehydrated with 100% ACN. After being air-dried, the proteins were rehydrated and digested at 37°C overnight with a trypsin solution (Promega) and 50 μL of 50 mM NH_4HCO_3 (pH 8.0). The peptides were extracted by vortexing and then concentrated to 10 μL in a Savant SpeedVac. The tryptic peptides were cleaned using the C-18 ZipTip system (Millipore) and eluted with 5 μL of 50% ACN. The protein samples were either sent to the Proteomics and Mass Spectrometry Centre (Dalhousie University) for Liquid Chromatography-Mass Spectrometry (LC/MS/MS) analysis or analyzed using a Waters Micromass QTOF Ultima global mass spectrometer (Department of Chemistry, University of Waterloo) with nanospray injection as the sample delivery method. Proteins were identified using PEAKS software 3.1 (Bioinformatics Solutions Inc., Waterloo, ON) (Ma et al., 2003) against sequences in the Swiss-Prot database.

2.9. Yeast two-hybrid assay

Yeast two-hybrid analysis was performed by using the DY1 yeast strain harboring the *lacZ* reporter plasmid pSH18-34, the bait plasmid pEG-202, and the prey plasmid pJG4-6 as described previously (Varrin et al., 2005). To construct bait and prey vectors for SAHH1, the coding sequence for SAHH1 was amplified from the pSAT6-SAHH1-EGFP using a primer set (SAHH-F Bam and SAHH1sR; Table 1) or a primer set (SAHH1cDNcF and SAHH1sR; Table 1). The *Bam*HI/*Nco*I PCR fragment was cloned into the pEG-202 vector, and the *Nco*I PCR fragment was cloned into the pJG4-6 vector. For ADK1 prey construct, the coding sequence for ADK1 was amplified from the pSAT6-ADK1-EGFP using a primer set (AK1cDNcF and ADK-R Xho; Table 1), and cloned into *Nco*I and *Xho*I sites of the pJG4-6 vector. For an ADK1 bait construct, the ADK1 coding sequence was amplified from the pJG-ADK1 using a primer set (ADK-F Eco and ADK-R Xho; Table 1), and ligated with the pJG-ADK1 using an *Eco*RI site. For prey and bait constructs for CMT, the coding sequence for CMT was amplified from the pSAT6-CMT-EGFP using a primer set (CMT-F Nco and CMT-R Eco; Table 1) or using a primer set (CMT-F Eco and CMT-R Eco; Table 1), respectively. Each PCR products were then cloned into the *Nco*I and *Eco*RI sites of the pJG4-6 or the *Eco*RI site of the pEG-202 vector. For ASDH prey construct, the coding sequence for ASDH was amplified from the pSAT6-ASDH-EGFP using a primer set (ASDH-F Xho.Nco and ASDH-R Xho; Table 1), then cloned into *Nco*I and *Xho*I sites of the pJG4-6 vector.

Along with positive interacting partner Dbf4 and Rad53 (Varrin et al., 2005), various combinations of bait and prey vectors were transformed into the yeast strain DY-1, which already contains the reporter vector pSH18-34, using the DMSO-enhanced LiAc-method

(Hill et al., 1991). For liquid culture assay, single colonies of each transformants were grown overnight in synthetic complete (SC) medium (Amberg et al., 2005) lacking tryptophan, uracil, and histidine. The appropriate amount of cultures was then further incubated at 30℃ to reach $1x10^7$ cells/mL. After centrifugation at 2,000 x *g* for 5 minutes, the pellet was washed once with sterile water and resuspended in 20 mL of 2% galactose-1% raffinose medium (BD Bioscience) lacking tryptophan, uracil, and histidine. The resuspended cells were then incubated at 30℃ for 6 hours for induction. Then $5x10^6$ cells was transferred to a microfuge tube, washed once with Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, pH7.0), and resuspended in 0.5 mL Z buffer containing 28.8 mM β-mercaptoethanol. The activity of β-galactosidase (β-Gal) was assayed using o-nitrophenyl-β-d-galactopyranoside (ONPG) as a substrate. The reaction was stopped by adding 0.25 mL of 1 M Na_2CO_3 after development of a yellow color, and reaction tubes were centrifuged at 14,000 rpm for 10 minutes. The optical density (OD) of each supernatant was measured using a spectrophotometer at 420 nm. The β-Gal activity was calculated as below:

$$\beta\ \text{Gal activity} = \frac{1{,}000 \times \text{OD}_{420}}{\text{reaction time (min)} \times \text{culture volume (mL)} \times \text{OD}_{600}}$$

2.10. Generation of antibodies

Both rabbit anti-ADK and anti-SAHH antibodies were previously generated from Moffatt lab. Rabbit anti-CMT antibody was raised against His-tagged atCMT recombinant protein expressed from pET30a+ (Novagen) in BL21 (codon+). Rabbit anti-

ASDH antibody was raised against ASDH recombinant protein expressed from pPAL7 (BioRad) in BL21 (DE3) pLysS. Each crude serum was obtained after four booster injections given to a white New Zealand rabbit. Five hundred µg of recombinant protein mixed with an equal volume of complete Freund's adjuvant was given for the first injection, followed by repeated three boosts with 300 µg of recombinant protein mixed with an equal volume of incomplete Freund's adjuvant at intervals of 3 weeks. The rabbit was bled 3 weeks after the last injection. After centrifugation for 15 min at 1500 x g, each crude serum was collected and used for immunoblotting at a 1:300 (anti-CMT) or 1:5,000 (anti-ASDH) dilution.

2.11. Immunoprecipitation and other pull-down assays

For immunoprecipitation, 300 mg of leave tissues from 3-week old T2 Arabidopsis plants carrying GFP-tagged overexpression constructs were homogenized in 1.5 mL extraction buffer (50 mM Tris pH 7.5, 100 mM NaCl, 1 mM DTT, 0.2% (v/v) Triton X-100, 1 mM PMSF, protease inhibitor cocktail), and centrifuged at 20,000 x *g* for 10 minutes at 4℃. The supernatant was then transferred to a new microfuge tube, and incubated with 25 µL protein A-sepharose (50% suspension; Sigma-Aldrich) in an end-over-end rotation wheel for 15 minutes at 4℃. After centrifugation at 700 x *g* for 30 seconds at 4℃, the pre-cleared supernatant was incubated with the appropriate antibody for 3 hours at 4℃ and further incubated with 50 µL protein A-sepharose for 2 hours at 4℃. The immunoprecipitates were washed five times with extraction buffer and

resuspended in 25 µL of 2 x SDS sample buffer and boiled for 6 minutes to release bound proteins. The liquid phase was used for SDS-PAGE gel.

To create GST fusion constructs, the coding sequence for importin α (Impα; At3g06720) was PCR amplified from a cDNA prepared from wild-type Arabidopsis plants using a primer set (Impa-F Eco and Impa-R Xho; Table 1). The PCR products were then cloned into the *Eco*RI/*Xho*I site of the pGEX4T-1 (Amersham Pharmacia) to generate the GST-Impα. For GST-CMT vector, *CMT* was amplified using a primer set (CMT-F Bam CMT-R Bam; Table 1) and inserted into the *Bam*HI site of the pGEX4T-1. Both GST-Impα and GST-CMT, along with a GST alone control, were expressed in *E. coli* BL21 and purified with glutathione-Sepharose 4B beads (Amersham Pharmacia). To do this, *E. coli* strains were grown at 37℃ to an OD_{600} of 0.5–0.7. The expression of GST fusion proteins was induced by adding 0.2 mM isopropyl-thio-β-galactoside (IPTG) for 6 hours at 30℃. After centrifugation at 3,500 x *g* for 15 minutes at 4℃, the cell pellet was resuspended in ice-cold PBS buffer containing 1% (v/v) Triton X-100 and protease inhibitor cocktail, and disrupted by four cycles of sonication with 10 seconds bursts and 30 seconds resting on ice. The GST fusion proteins were purified using a glutathione-Sepharose affinity column by gravity flow.

For the GST pull-down assay, leave tissues from 3-week old wild-type Arabidopsis plants were homogenized in extraction buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.5% (v/v) Triton X-100, 0.2% (v/v) glycerol, 1 mM PMSF, protease inhibitor cocktail), and centrifuged at 10,000 x *g* for 10 minutes at 4℃. The supernatant was incubated with 20 µL GST•Mag agarose beads (50% suspension; Novagen) in an end-over-end rotation wheel for 15 minutes at 4℃ to pre-clear the supernatant. Three µg of purified GST fusion

proteins were incubated with 1 mg of the pre-cleared total crude plant protein extracts for 2 hours at 4°C, and then further incubated for 1 hour with 30 µL GST•Mag agarose beads. The beads were recovered by placing the tube in a magnetic stand (Dyanl Biotech) for 30 seconds and washed five times with wash buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.5% (v/v) Triton X-100, 1 mM PMSF, protease inhibitor cocktail). The beads were resuspended in 15 µL of 2 x SDS sample buffer and boiled for 6 minutes to release bound proteins, and the liquid phase was used for SDS-PAGE.

His-tag pull-down assays were performed using His-tagged recombinant proteins prepared from *E. coli* BL21 cells and plant crude extracts from wild-type Arabidopsis plant leaves. The proteins bound to His-tagged recombinant proteins were precipitated by His•Mag agarose beads (50% suspension; Novagen). The same buffers and conditions were used as in the GST pull-down assay.

2.12. SDS-PAGE and Western blotting

Protein samples were prepared by homogenizing plant leave tissues in three volumes of extraction buffer to the tissue (50 mM Tris pH 7.5, 1 mM EDTA, 8 mM $MgCl_2$, 1% (v/v) β-mercaptoethanol, 1 mM PMSF; unless otherwise stated) using an ice-chilled glass homogenizer. After centrifugation at 10,000 x *g* for 10 minutes at 4℃, the supernatant containing a total crude plant protein extracts was transferred to a new microfuge tube and protein concentrations were quantified by Bradford assay (Bradford, 1976). Prepared protein samples were boiled in SDS sample buffer for 5 minutes and then subjected to SDS-PAGE analyses. An SDS-PAGE gel (1 mm or 1.5 mm (T) × 8 cm (W) × 7.3 cm (H))

was prepared with a 10% or 12.5% separating gel and a 5% stacking gel according to standard methods (Laemmli, 1970) using the mini-protein 3 system (Bio-Rad). Protein samples were applied to each lane of the gel along with a prestained or an unstain protein marker (Fermentas, Canada), and the gels were run at 80–100 V in a running buffer containing 0.1% (w/v) SDS, 25 mM Tris base, pH 8.3, and 192 mM glycine.

For Western blotting, after electrophoresis the gel was pre-equilibrated twice for 15 minutes in a transfer buffer containing 48 mM Tris, 0.0375% (w/v) SDS, 20% (v/v) methanol, and 38 mM glycine, and the separated proteins were transferred onto nitrocellulose membranes at 20 V for 45 minutes in a Semi-dry electroblotting device (Bio-Rad). The membrane was then stained with Ponceau S (0.2 % (w/v) Ponceau S, 3 % (w/v) trichloroacetic acid (TCA)) to confirm equal protein loading and subsequently washed with PBS-T (1% (w/v) nonfat dry milk powder, 0.3% (v/v) Tween 20 in PBS). The membrane was then blocked with 0.1% (w/v) polyvinyl alcohol (PVA) in PBS for 30 seconds and incubated with the primary antibody diluted in PBS-T for overnight at 4℃. After three times of wash with PBS-T, the membrane was incubated with an alkaline phosphatase-conjugated secondary antibody (anti-mouse IgG or anti-rabbit IgG; Sigma-Aldrich) for 1 hour at room temperature. After three times of washing with PBS-T, the chemifluorescent detection was performed using enhanced chemifluorescence (ECF) substrate and the image acquired using the Typhoon 9400 laser scanning system (Amersham Bioscience).

For yeast proteins, cell pellets were collected by centrifugation at 2,000 x *g* for 5 minutes, and disrupted by using a bead beater with 0.5 mm glass beads (eight 20 seconds bursts and 30 seconds resting on ice; BioSpec Products) in a lysis buffer containing 10

mM Tris (pH 8.0), 1 mM EDTA, 1% (v/v) Triton X-100, 140 mM NaCl, and protease inhibitor cocktail. After centrifugation at 10,000 x *g* for 30 seconds, the supernatant was used for Western blotting. The expression of bait and prey proteins were detected using a rabbit polyclonal anti-LexA antibody (Invitrogen) and a mouse monoclonal anti-HA antibody (Sigma-Aldrich).

2.13. Screening T-DNA insertion lines

To identify homozygous T-DNA tagged lines, genomic DNA extracted from young leaves of T2 progeny were screened by PCR using primers specific to sequences flanking the insertion site (LP and RP primers) or in the left border (LB3 primer; Table 1) of the T-DNA insertion sequence as indicated in the figure of the result section. PCR reaction was performed with an initial 5 minute denaturation at 95 ℃, followed by 35 cycles (each cycle: 95 ℃, 30 seconds; 55 ℃, 30 seconds; and 72 ℃, 1 minute), then a final 10 minutes at 72 ℃. The primer sets used for CMT line are (CS846178-LP and CS846178-RP) and (LB3 and CS846178-RP), and the one for ASDH are (CS813591-LP and CS813591-RP) and (LB3 and CS813591-RP) (primers are listed in Table 1).

Table 1. Primers used for cloning and T-DNA screening

Name	Sequence (5' to 3')*	RE	Tm
CMT-F Nco	CTTccatggTTATGAAACGAGGATTCTCCG	*Nco*I	65
CMT-R Bam	CTAggatccAGCTATCGATATACAAAACATC	*Bam*HI	58
CMT-F Hin	CTTaagcttGTATGAAACGAGGATTCTCCG	*Hin*dIII	62
CMTNLS-F	ATccatggACAGAAAGAGGGGAGAATCAGATGGAGC TCGAAGGAGCGGGCGGAGGATGGCCTCCTCCGAG	*Nco*I	92
DsRED2R-Nco	AGCccatggTCTTCTTCTGCATC	*Nco*I	58
CMT342-R	CAGgaattcTCCGCTGTTACTCAAGACAAAAGAG	*Eco*RI	66
CMT345-F	GAgaattcTATGTACCTCCAACGACCGTGCCGTGGAA GAATGGG	*Eco*RI	78
ASDH-F Xho.Nco	CAActcgagccatggCGACGTTCACTCATC	*Nco*I/ *Xho*I	71
ASDH-R Xma	CAAcccgggGAGAAGCATCTCAGCGATC	*Xma*I	71
ASDHp-F Age	ACTTaccggtTAGTTGGTTGTGAG	*Age*I	53
ASDHp-R Nco	GTCGccatggTCGCCGCCGTGAG	*Nco*I	72
ADK1-F Xho	CTTctcgagATGGCCTTCCTCTGATTTCG	*Xho*I	66
ADK1cR-Sma	ATcccgggGTTGAAGTCTGGTTTCTC	*Xma*I	64
SAHH-F Bam	AAggatccGTATGGCGTTGCTCGTCGAG	*Bam*HI	69
SAHH1sR	TTTccatggCTCAGTACCTGTAGTGAGGAGG	*Nco*I	65
SAHH1cDNcF	TCAAccatggCGTTGCTCGTC	*Nco*I	57
AK1cDNcF	CATccatggCTTCCTCTGATTTC	*Nco*I	56
ADK-R Xho	AActcgagTTAGTTGAAGTCTGGTTTCTC	*Xho*I	57
ADK-F Eco	TTgaattcATGGCTTCCTCTGATTTCG	*Eco*RI	62
CMT-R Eco	CACgaattcAGCTATCGATATAC	*Eco*RI	48
CMT-F Eco	ATgaattcATGAAACGAGGATTCTCCG	*Eco*RI	61
ASDH-R Xho	GAActcgagCGGGTTCAAATTCAGAGAAG	*Xho*I	64
Impa-F Eco	AAgaattcATGTCACTGAGACCCAACGC	*Eco*RI	63
Impa-R Xho	AActcgagTCAGCTGAAGTTGAATCCTC	*Xho*I	60
CMT-F Bam	CTTggatccATGAAACGAGGATTCTC	*Bam*HI	60
LB3	TAGCATCTGAATTTCATAACCAATCTCGATACAC	N/A	62
CS846178-LP	TGTCTTTATCGGAACAATGCC	N/A	51
CS846178-RP	TCATTTTCCCATTCTTCCTCC	N/A	52
CS813591-LP	GCTAAAAACCAGAATCACGAGC	N/A	52
CS813591-RP	TGCCCGAAGATGTTACAAGTC	N/A	51

*Lowercased sequences represent the restriction enzyme (RE) sites as indicated.

2.14. Artificial microRNA constructs for ADK and SAHH

Artificial microRNA (amiRNA) constructs were designed using online amiRNA designing tool (WMD2 Web MicroRNA Designer; http://wmd2.weigelworld.org) as described in the web site (Schwab et al., 2005).

To design specific amiRNAs targeting both *ADK1* and *ADK2* transcripts (amiADK), both *ADK1* and *ADK2* genes were used as input target genes in the MicroRNA Designer. amiSAHH was designed in the same manner using both *SAHH1* and *SAHH2* genes. Among amiRNA candidates provided by the MicroRNA Designer, two amiRNAs in a different region of *ADK* (amiADK#1 and amiADK#2) and *SAHH* (amiSAHH#1 and amiSAHH#2) genes were selected for each. The amiADK#1 sequence (TAT GGG CAA ATG TAT GTC CTC) targets 5'-GAG GAC AAA CAT TTG CCC ATG-3' in *ADK1* and 5'-GAA GAC AAA CAT TTG CCC ATG-3' in *ADK2*; amiADK#2 sequence (TTA ACT GGT TGC CCC AGG GAC) targets 5'-GTT CCT GGG GCA ACC AGT TAC-3' in *ADK1* and 5'-ATT CCT GGG GCA ACC AGT TAC-3' in *ADK2*, respectively. amiSAHH#1 sequence (TAC CTT GTA TTC ACG GAC ACT) targets 5'-AGT GGC CGT GAA TAC AAG GTC-3' in both *SAHH1* and *SAHH2*; amiSAHH#2 sequence (TAC ACC ATC TGA CTA CAG CAC) targets 5'-GTG CTG AAG TCA GAT GGT GTT-3' in *SAHH1* and 5'-GCG CTG AAG TCA GAT GGT GTT-3' in *SAHH2*, respectively. These amiRNAs were generated by PCR using four overlapping primers (Table 2) and a pRS300-based vector as a template (Schwab et al., 2005). The PCR-amplified amiRNA fragments (~400 bp) were cloned into *Eco*RI/*Bam*HI (for amiADK) or *Hin*dIII/*Bam*HI (for amiSAHH) of the modified pSAT4 or the pSAT6 vector (EGFP between the CaMV 35S promoter and 35S terminator was removed from both vectors).

These constructs were transferred into the PI-*Psp*I site (for pSAT6) or I-*Sce*I site (for pSAT4) of pPZP-RCS2 binary vector for plant transformation (Appendix I for the pSAT vector system).

Table 2. Primers used to construct amiRNAs

Target	Name	Sequence (5' to 3')
pRS300	miR319-A	CTGCAAGGCGATTAAGTTGGGTAAC
	miR319-B	GCGGATAACAATTTCACACAGGAAACAG
amiADK#1	miAK1-I	GATATGGGCAAATGTATGTCCTCTCTCTCTTTTGTATTCC
	miAK1-II	GAGAGGACATACATTTGCCCATATCAAAGAGAATCAATGA
	miAK1-III	GAGAAGACATACATTAGCCCATTTCACAGGTCGTGATATG
	miAK1-IV	GAAATGGGCTAATGTATGTCTTCTCTACATATATATTCCT
amiADK#2	miAK2-I	GATTAACTGGTTGCCCCAGGGACTCTCTCTTTTGTATTCC
	miAK2-II	GAGTCCCTGGGGCAACCAGTTAATCAAAGAGAATCAATGA
	miAK2-III	GAGTACCTGGGGCAAGCAGTTATTCACAGGTCGTGATATG
	miAK2-IV	GAATAACTGCTTGCCCCAGGTACTCTACATATATATTCCT
amiSAHH#1	amiSH1-I	GATACCTTGTATTCACGGACACTTCTCTCTTTTGTATTCC
	amiSH1-II	GAAGTGTCCGTGAATACAAGGTATCAAAGAGAATCAATGA
	amiSH1-III	GAAGCGTCCGTGAATTCAAGGTTTCACAGGTCGTGATATG
	amiSH1-IV	GAAACCTTGAATTCACGGACGCTTCTACATATATATTCCT
amiSAHH#2	amiSH2-I	GATACACCATCTGACTACAGCACTCTCTCTTTTGTATTCC
	amiSH2-II	GAGTGCTGTAGTCAGATGGTGTATCAAAGAGAATCAATGA
	amiSH2-III	GAGTACTGTAGTCAGTTGGTGTTTCACAGGTCGTGATATG
	amiSH2-IV	GAAACACCAACTGACTACAGTACTCTACATATATATTCCT

3. Results

3.1. Interaction between ADK1 and SAHH1

Given that ADK1 and SAHH1 provide related activities to SAM recycling and both enzymes are often detected in same subcellular compartments (Chapter 1), I investigated a possible protein-protein interaction between the two enzymes using several different approaches. In support of such an interaction, molecular docking (ZDOCK), based on the crystal structures of Plasmodium SAHH monomer (*Plasmodium falciparum*; PDB ID, 1V8B:A chain) and human ADK (PDB ID, 1BX4) that are highly homologous to those of Arabidopsis, predicted the best interaction between ADK1 and SAHH1 would occur via the IR. This is the same region that was essential for the nuclear localization of SAHH1 (Chapter 1) (Andrew Doxey, personal communication; Appendix II). The molecular docking analysis was done without the knowledge of the results presented in Chapter 1. As much as this modeling prediction was interesting, it required experimental testing.

To examine an ADK1-SAHH1 interaction, a pull-down assay was performed using wild-type Arabidopsis plant extracts and His-tagged ADK1 recombinant protein (henceforth referred to as His-ADK1) that was produced in *E. coli* and purified. His-ADK1 protein incubated with the plant crude extracts was precipitated by His•Mag agarose beads, and the precipitates (pellet) and supernatant were analyzed by immunoblotting with anti-SAHH antibody (Figure 1A). The result showed that His-ADK1 can pull-down SAHH1 from the wild-type plant extracts (Figure 1A, asterisk), suggesting that the two proteins may directly interact with each other or be part of a protein complex. To verify this observation, co-immunoprecipitation was further carried

out using wild-type Arabidopsis plant extracts and anti-SAHH antibody. To do this, wild-type ADK1 or His-ADK proteins were immunoprecipitated with anti-SAHH antibody and the precipitates were analyzed by immunoblotting with anti-ADK antibody (Figure 1B and C). These results revealed that both wild-type ADK1 (Figure 1B, asterisk and Figure 1C, lane 3) and His-tagged recombinant ADK1 (Figure 1C, lane 4) were successfully pulled-down by anti-SAHH. Control reactions carried out without adding anti-SAHH antibody failed to pull-down ADK1 proteins (Figure 1B, lane 2 and Figure 1C, lane 4).

The interaction was also confirmed by a pull-down assay using extracts of plants overexpressing ADK1 fused to a Strep epitope (Figure 2A and B). The strep tag was fused to the C-terminus of a full-length *ADK1* cDNA expressed under the control of CaMV 35S promoter. This construct was stably transformed into Arabidopsis plants. More than twenty T1 lines were screened by BASTA treatment and successful transgenic lines were verified by immunoblotting using monoclonal anti-HA (Influenza Hemagglutinin) antibody (Appendix IIIA). The T2 progeny of selected transgenic lines were used to purify the Strep tag protein and hopefully other interactors. For this, the purification conditions were optimized using different buffer compositions to determine the optimal stringency that allowed for removal of non-specific binding of proteins (Appendix IIIB). This optimized condition was applied to the follow-up experiments described in this thesis. The protein extracts prepared from transgenic lines were purified by StrepTactin-Macroprep resin and the precipitates were analyzed by SDS-PAGE and visualized with silver stain (Figure 2A). Several polypeptides were visualized on the stained SDS-PAGE along with the overexpressed ADK1-Strep proteins, but it was hard

to confirm that SAHH1 was present based on the staining. Thus, the same precipitates were analyzed by immunoblotting with anti-SAHH antibody (Figure 2B). The analyses indicated that ADK1-Strep can pull-down wild-type SAHH1 (Figure 2B).

Immunoprecipitation using various Arabidopsis transgenic plants expressing GFP fusions was performed in an attempt to locate regions of SAHH1 that are essential for its interaction with ADK1 (Figure 2C). The plant extracts were immunoprecipitated with anti-ADK serum using protein A Sepharose beads and analyzed by immunoblotting with anti-GFP antibody. The results showed that SAHH1-GFP was successfully pulled-down by anti-ADK (Figure 2C, lane 2), whereas the GFP-SAHH1 was hardly detected (Figure 2C, asterisk), and anti-ADK failed to pull-down SAHH1ΔIR-GFP as well as GLU3-GFP (β-1,3-glucanase 3)(Yaish et al., 2006) that was used as a negative control (Figure 2C). These results suggested that ADK1 may interact with SAHH1 through its N-terminal region or the IR.

Figure 1. Protein interaction between ADK1 and SAHH1.

(A) His-tag pull-down assay using ADK1 recombinant protein from *E. coli* BL21 cells and plant crude extracts prepared from wild-type Arabidopsis plant leaves. His-ADK1 proteins were precipitated by His•Mag agarose beads (Novagen). Both supernatant and pellet fractions were subjected to 10% SDS-PAGE and immunoblotted with anti-SAHH antibody (1:5,000). Asterisks indicate polypeptides indicative of an interaction between ADK1 and SAHH1. S, supernatant; P, pellet of beads. (B) Co-immunoprecipitation (IP) using extracts prepared from wild-type Arabidopsis plants. Extracts were immunoprecipitated with anti-SAHH antibody. The immunoprecipitates were resolved by 10% SDS-PAGE and immunoblotted with anti-ADK antibody (1:3,000). Asterisks indicate the polypeptides consistent with an interaction between ADK1 and SAHH1. (C) Co-immunoprecipitation using extracts prepared from wild-type Arabidopsis plants with/without His-ADK1 proteins. Extracts were immunoprecipitated with anti-SAHH antibody. The immunoprecipitates were resolved by 10% SDS-PAGE and immunoblotted with anti-ADK antibody (upper panel) and anti-SAHH antibody (lower panel). Asterisks indicate the polypeptides indicative of an interaction between ADK1 and SAHH1 (Lane 3) or His-ADK1 and SAHH1 (Lane4). Molecular masses are indicated on the left in kilodaltons (kDa).

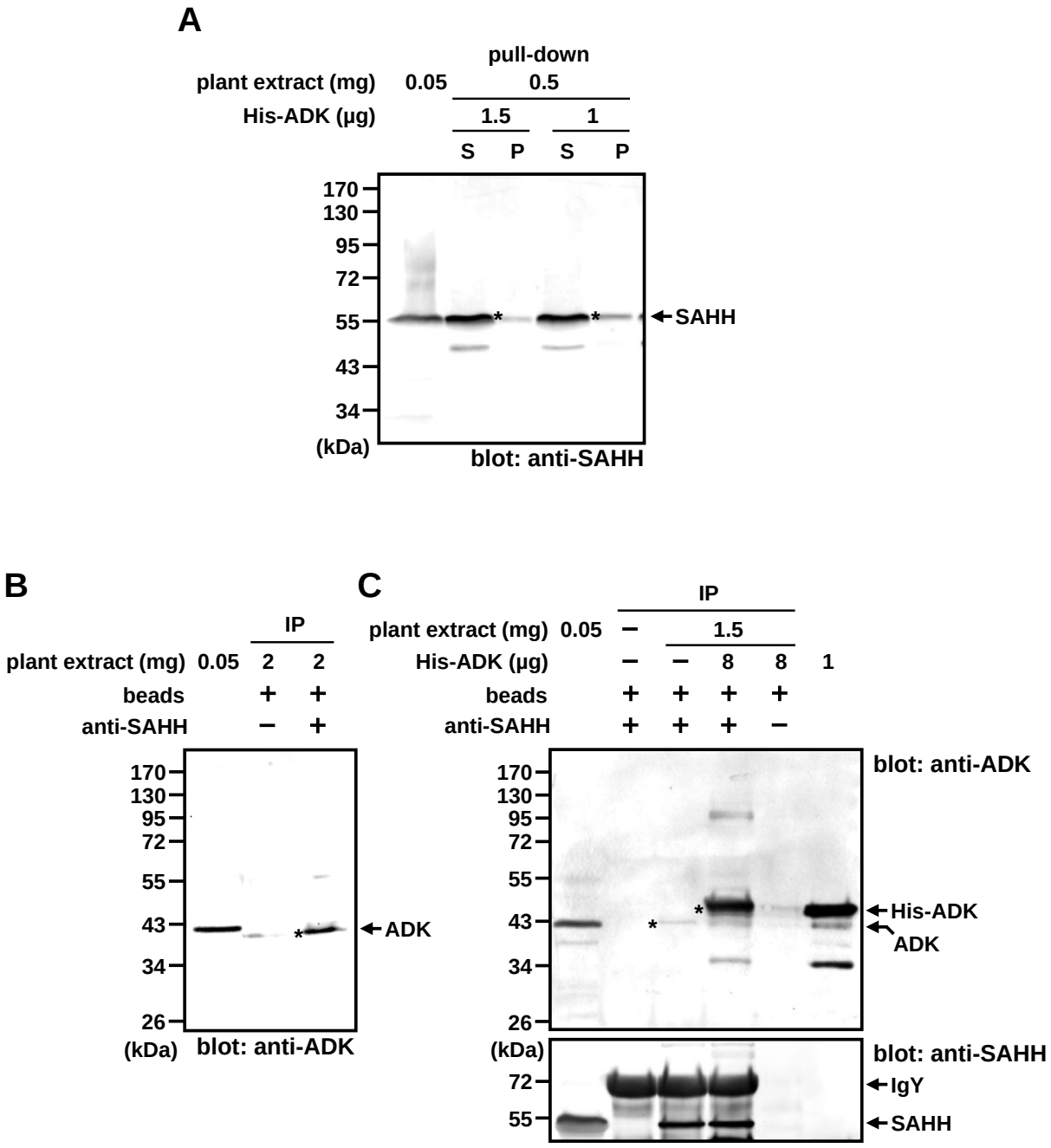
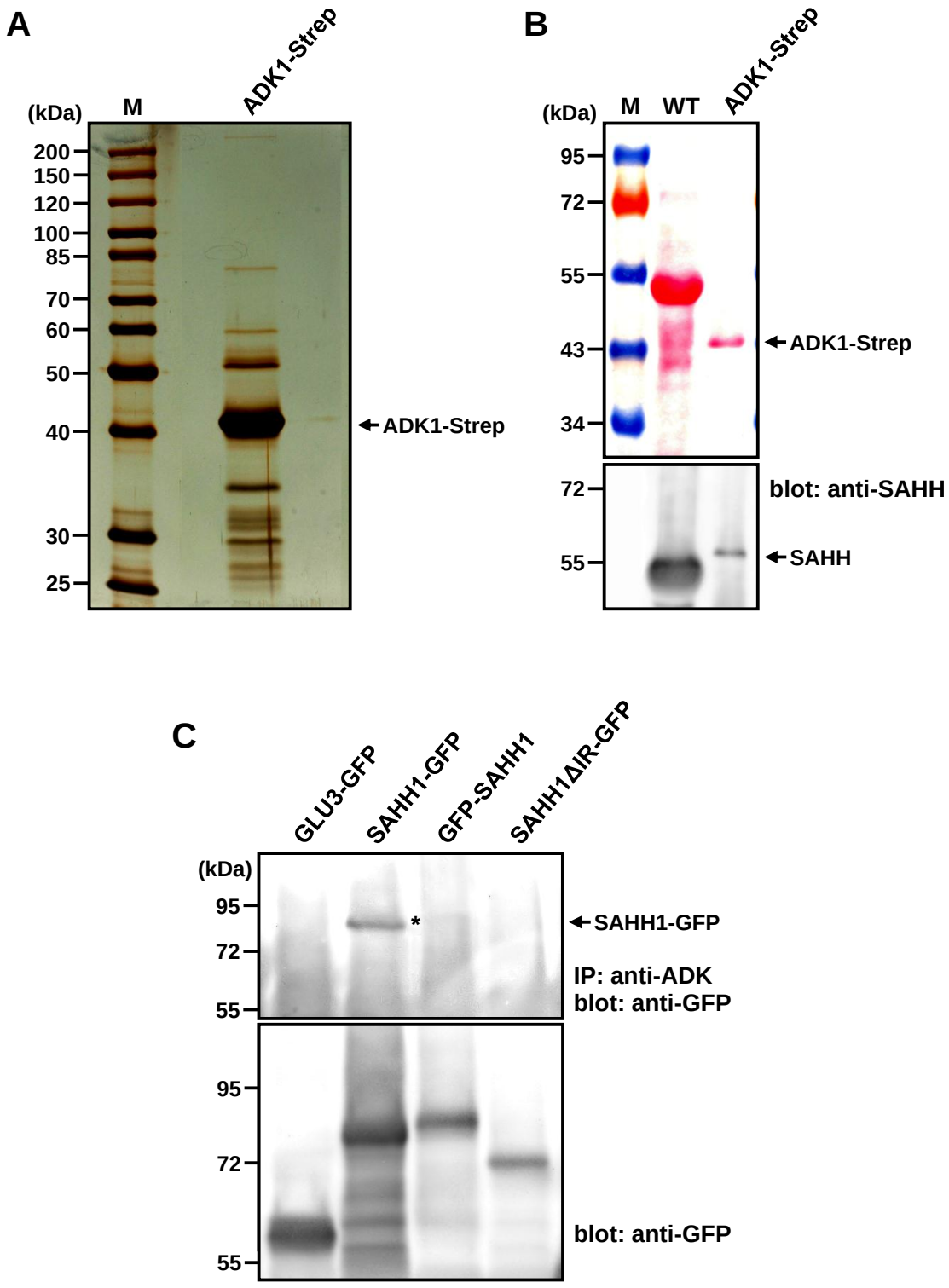

A
pull-down
plant extract (mg)
0.05
0.5
His-ADK (μg)
1.5
1
S
P
S
P
170
130
95
72
55
43
34
(kDa)
SAHH
blot: anti-SAHH
B
IP
plant extract (mg)
0.05
2
2
beads
+
+
anti-SAHH
−
+
170
130
95
72
55
43
34
26
(kDa)
ADK
blot: anti-ADK
C
IP
plant extract (mg)
0.05
−
1.5
His-ADK (μg)
−
−
8
8
1
beads
+
+
+
+
anti-SAHH
+
+
+
−
170
130
95
72
55
43
34
26
(kDa)
blot: anti-ADK
His-ADK
ADK
72
55
blot: anti-SAHH
IgY
SAHH

Figure 2. ADK1-Strep purification and immunoprecipitation to examine ADK1-SAHH1 interaction.

(A) Strep purification using extracts prepared from 700 mg of Arabidopsis plant leaves expressing an ADK1-StrepII fusion protein. Extracts were precipitated by incubation with StrepTactin beads. The precipitates were resolved by 10% SDS-PAGE and the polypeptides visualized by silver stain. (B) The precipitates were also analyzed by Ponceau S stain (upper panel) and immunoblotting with anti-SAHH antibody (lower panel). (C) Immunoprecipitation using extracts prepared from various Arabidopsis transgenic plants expressing GFP fusion proteins as indicated above the panel. Extracts were immunoprecipitated with anti-ADK antibody. The immunoprecipitates were resolved by 10% SDS-PAGE and immunoblotted with anti-GFP antibody (upper panel). Each extract was also analyzed by immunoblotting with anti-GFP antibody, in the absence of an IP reaction (lower panel). Asterisks indicate a weak protein band of GFP-SAHH1. Molecular masses are indicated on the left in kilodaltons (kDa). GLU3, β-1,3-glucanase 3; M, size marker; WT, twenty-five micrograms of protein extract obtained from wild-type Arabidopsis plants.


A
ADK1-Strep
(kDa)
M
200
150
120
100
85
70
60
50
40
30
25
ADK1-Strep
B
ADK1-Strep
(kDa)
M
WT
95
72
55
43
34
ADK1-Strep
72
55
blot: anti-SAHH
SAHH
C
GLU3-GFP
SAHH1-GFP
GFP-SAHH1
SAHH1ΔIR-GFP
(kDa)
95
72
55
*
SAHH1-GFP
IP: anti-ADK
blot: anti-GFP
95
72
55
blot: anti-GFP

3.2. Bimolecular fluorescence complementation for ADK1-SAHH1 interaction

Interactions between ADK1 and SAHH1 were further investigated using a bimolecular fluorescence complementation (BiFC) assay in tobacco (*Nicotiana benthamiana*) and Arabidopsis plants. To do this, ADK1 and SAHH1 coding sequences were subcloned upstream or downstream of N-terminal (YN; amino acids 1–174) and C-terminal fragments of YFP (YC; amino acids 175–238). The fusion proteins were located between the CaMV 35S promoter and 35S transcriptional terminator (Citovsky et al., 2008). The entire expression cassettes for both YN and YC were then subcloned into a plant binary vector, and transiently co-expressed in or stably transformed into plants to examine the reconstitution of YFP fluorescence.

Co-transfection of the ADK1-YN along with the SAHH1-YC resulted in strong YFP signals in the nucleus and cytoplasm of tobacco epidermal cells (Figure 3A), while YFP fluorescence was generally weaker when ADK1-YN was co-infiltrated with the YC-SAHH1 (Figure 3B). No or only very faint fluorescence was detected when ADK1-YN was co-transfected with the SAHH1ΔIR (Figure 3C). These results are consistent with those from the previous immunoprecipitation study (Figure 2C).

Native SAHH1 exists as a tetramer (Poulton, 1981), and it was thus of interest to determine whether the SAHH1-fusion proteins had formed functional complexes. To test this, two SAHH1 proteins were fused to either YN or YC. Strong fluorescence detected from the interaction between SAHH1-YN and SAHH1-YC/YC-SAHH1 indicative of tetramer formation (Figure 3D and E). Interestingly, similar fluorescent signals were

obtained from the ADK1-ADK1 interaction (Figure 3F), even though most ADKs are known to act as monomers with one exception of *Mycobacterium tuberculosis* ADK that exists as a dimer (Long et al., 2003).

As a control for the BiFC assay, ADK1-YN and free YC, SAHH1-YC and free YN, or free YN and YC were co-expressed in tobacco plants to test whether YC or YN fusion proteins could produce functional YFP fluorophores with the respective YC and YN protein fragments alone. The plants expressing these control constructs showed a faint YFP fluorescence, although the fluorescent signals were substantially weaker than those obtained from the interactions between ADK1 and SAHH1 or two SAHH1s that gave strong and clear YFP signals (Figure 3G-I). In addition, all these constructs were also transiently expressed in tobacco plants under the control of the actin promoter by replacing the CaMV 35S double enhanced promoter. Overall, these plants showed similar patterns of YFP fluorescence as those with 35S promoter, although the signals were slightly weaker (data not shown).

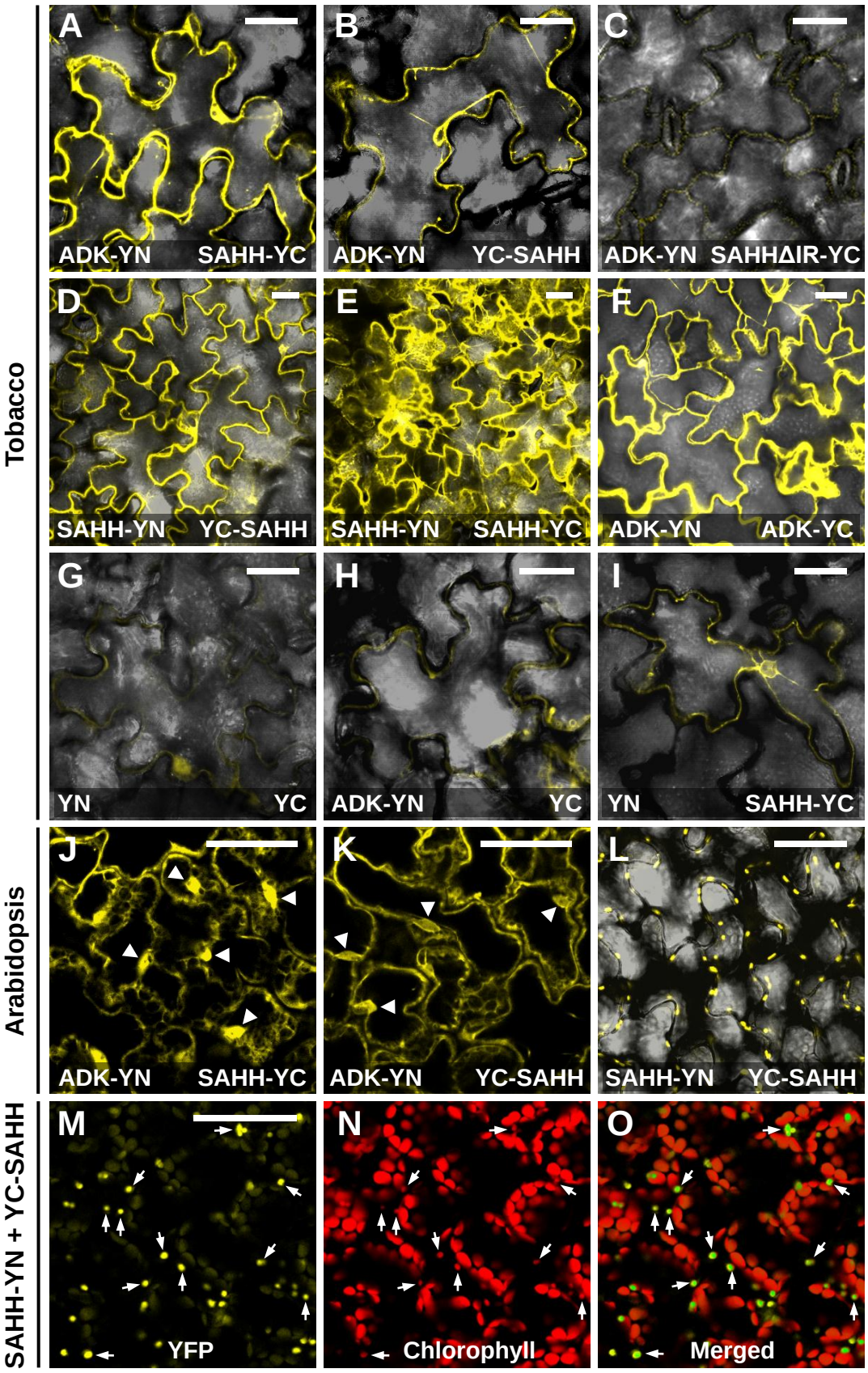
Some of the constructs for the BiFC were also examined in Arabidopsis stable transformants. Consistent with the result of transient expression in tobacco, the transgenic Arabidopsis leaves co-expressing ADK1-YN and SAHH1-YC or ADK1-YN and YC-SAHH1 showed strong YFP signals, while the signal obtained from ADK1-YN and YC-SAHH1 combination was slightly weaker than the other, especially in the nuclei (Figure 3J and K). Interestingly, co-expression of SAHH1-YN and YC-SAHH1 for the SAHH1 dimerization test in Arabidopsis plants resulted in strong punctate-like fluorescent signals surrounding epidermal cell walls and faint fluorescent signals in chloroplasts (Figure 3L), whereas the Arabidopsis leaves co-expressing of SAHH1-YN and SAHH1-YC showed a

similar fluorescence pattern to those obtained from transient expression in tobacco plants (data not shown). Moreover, the punctate-like YFP signals clearly overlapped with the autofluorescence of chlorophyll (Figure 3M-O; arrows). This result supports the previous data that suggested a possible localization of SAHH1 in plastids (Chapter 1). Taken together, these observations demonstrate that ADK1 and SAHH1 interact and this interaction may depend on the presence of the IR in SAHH1. In addition, the N-terminal region of SAHH1 may be also involved in the interaction with ADK.

To test if only ADK1 and SAHH1 were involved in the interaction, I performed yeast two-hybrid assays using each protein as prey/bait or bait/prey. Surprisingly, the interaction between ADK1 and SAHH1 could not be observed by this assay, whereas very strong interaction signals were detected when two SAHH1s were used for prey and bait (Appendix IV). These results suggested that additional plant-specific factors/proteins may be required for the ADK1 and SAHH1 interaction, since their interaction was detected only when plant extracts were added for the pull-down assay or the interaction was detected in living tissue by BiFC.

Figure 3. BiFC assay for ADK1-SAHH1 interaction.

The two fusion proteins, as indicated in the figure, were transiently expressed in tobacco (*Nicotiana benthamiana*) leaves (A-I) and stably transformed in Arabidopsis plants (J-O). Both tobacco and Arabidopsis leaves were imaged using confocal microscopy. Yellow colour represents fluorescent signals of reconstituted YFP complexes. The autofluorescent signal of chlorophyll is depicted in red (N). Arrows indicating the overlap of the YFP and chlorophyll autofluorescence are shown in green (O). Arrowheads indicate nuclei. Note the CaMV 35S promoter directs expression of all these constructs. YN, N-terminal fragment of YFP; YC, C-terminal fragment of YFP. Scale bars represent 50 μm.


Tobacco
Arabidopsis
SAHH-YN + YC-SAHH
A
ADK-YN SAHH-YC
B
ADK-YN YC-SAHH
C
ADK-YN SAHHΔIR-YC
D
SAHH-YN YC-SAHH
E
SAHH-YN SAHH-YC
F
ADK-YN ADK-YC
G
YN YC
H
ADK-YN YC
I
YN SAHH-YC
J
ADK-YN SAHH-YC
K
ADK-YN YC-SAHH
L
SAHH-YN YC-SAHH
M
YFP
N
Chlorophyll
O
Merged

3.3. Involvement of cap methyltransferase in ADK1-SAHH1 interactions

The nuclear localization of both ADK1 and SAHH1 was confirmed using EGFP fusions as documented in Chapter 1. Theoretically, ADK1 can just diffuse into the nucleus due to its small size (38 kDa) which is well below the diffusion limit of the nuclear pore (~70 kDa) (D'Angelo et al., 2009) and the tetramer of SAHH1 (~212 kDa) is too big for diffusion. However when each protein was fused to two GFPs that increase the size of each monomer to ~92 kDa (ADK1) and 107 kDa (SAHH1; 428 kDa for tetramer), they were still successfully targeted to the nucleus by active transport (Chapter 1).

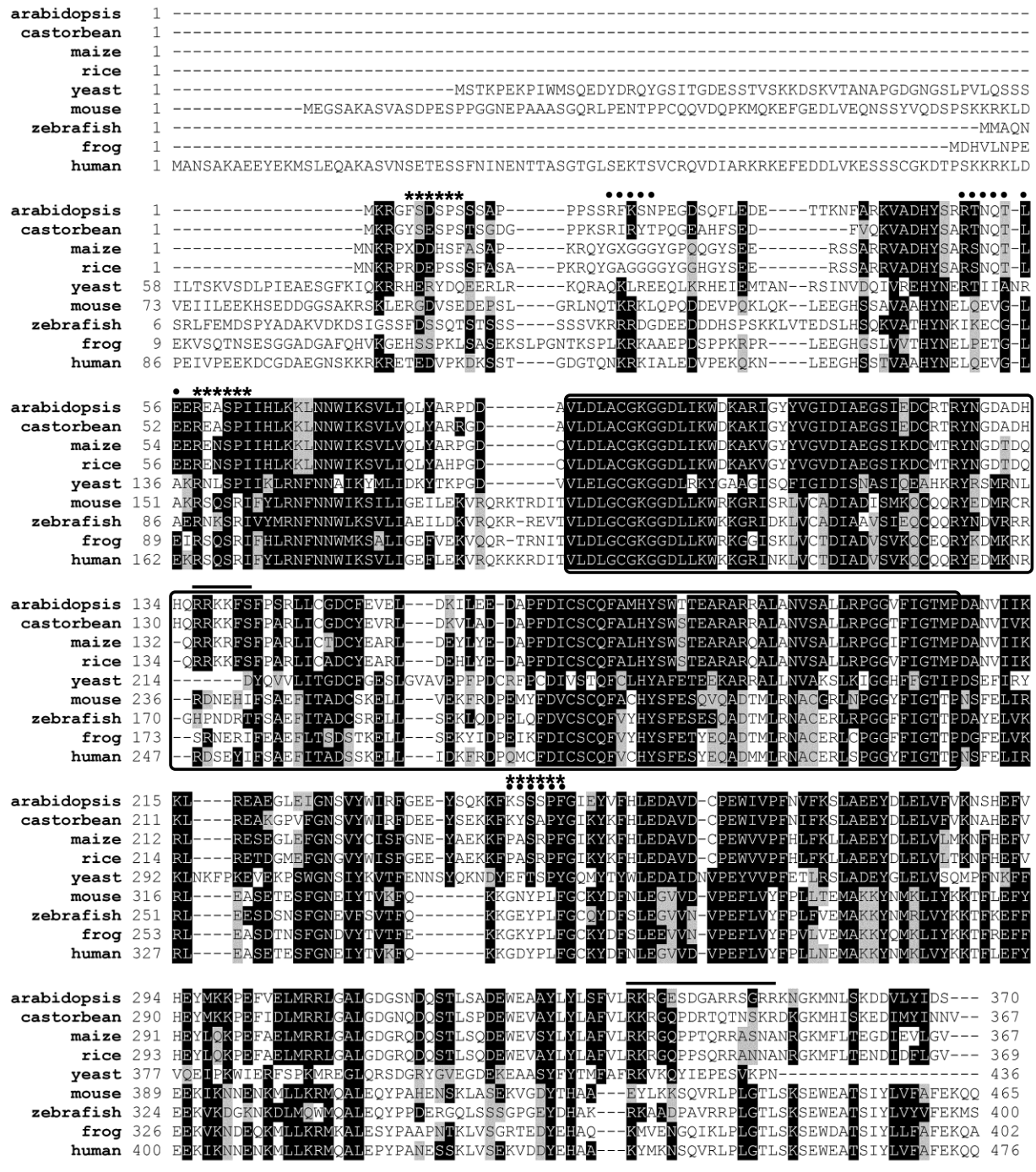
Since no conventional N-terminal signal sequence was located in either protein, the nuclear localization of both enzymes may rely on alternate pathways such as protein-protein interactions or trafficking mediated by other proteins that have a nuclear localization signal (NLS). The most likely candidate protein for this targeting is mRNA cap methyltransferase, as demonstrated by a previous study using *Xenopus laevis* oocytes showing that the translocation of *Xenopus* SAHH from the cytoplasm to the nucleus is mediated by a protein interaction with mRNA (guanine-7-) methyltransferase (Radomski et al., 2002). To test this, an Arabidopsis mRNA cap methyltransferase coding sequence was isolated from cDNA prepared from wild-type Arabidopsis plants. Among five mRNA capping enzymes (At3g09100, At3g20650, At3g52210, At5g01290, At5g28210) that are present in Arabidopsis (TAIR, 2010), At3g20650 was chosen based on its high sequence similarity with *Xenopus* cap MT. Thus, the full-length cDNA coding for Arabidopsis cap MT was isolated by PCR amplification using gene-specific primers based on the sequence of At3g20650. Sequence analysis of the full-length Arabidopsis

cap MT gene reveals that it has 13 exons spanning approximately 2,769 bp of genomic DNA and a 1,110 bp open reading frame (ORF) encoding 370 amino acids with a predicted molecular weight of 42.2 kDa; this was designated as CMT (Cap MethylTransferase).

A homology search of Arabidopsis CMT with the deduced amino acid (aa) sequence of other cap MT genes suggested that Arabidopsis CMT belongs to a family of SAM-dependent MTs, with a highly conserved SAM-dependent MTs domain (Figure 4). The amino acid sequence of Arabidopsis CMT shares 70-79% of aa identity with that of other plant CMTs used in the alignment, and 31-34% with proteins for other organisms including *Xenopus*, mouse, zebrafish, yeast, and human (Figure 4). Sequence analysis using bioinformatic tools provided by the ELM server (http://elm.eu.org/) showed that Arabidopsis CMT contains both potential 14-3-3 motifs and Class IV WW domains that mediate phosphorylation-dependent protein-protein interactions (Figure 4) (Gould et al., 2010). In addition, two potential NLS, a monopartite NLS (amino acid residues at 136-141) and a bipartite NLS (amino acid residues at 339-353), were predicted by the ELM server and ScanProsite (http://expasy.org/tools/scanprosite/), respectively (Figure 4).

Figure 4. Comparison of the deduced amino acid sequences of CMTs found in various organisms.

The alignment was performed using AlignX (Vector NTI, Invitrogen) and edited in Microsoft Word. Conserved residues are black-shaded, and similar residues are in gray. Numbers on the left of sequence represent the amino acid sequence. SAM-dependent methyltransferase domains are boxed and putative NLS sequences of Arabidopsis CMT are indicated by lines above the sequence. Black dots and asterisks above the sequence indicate 14-3-3 motif and WW domain, respectively. The amino acid sequences used in the alignment are as following: Arabidopsis (*Arabidopsis thaliana*; Q9LHQ7), castorbean (*Ricinus communis*; EEF33995), maize (*Zea mays*; ACG37354), rice (*Oryza sativa*; Q6Z9U7), yeast (*Saccharomyces cerevisiae*; P32783), mouse (*Mus musculus*; Q9D0L8), zebrafish (*Danio rerio*; Q1MTD3), frog (*Xenopus laevis*; Q9I8S2), human (*Homo sapiens*; BAA82447)**.**


arabidopsis
castorbean
maize
rice
yeast
mouse
zebrafish
frog
human

3.4. Subcellular localization of CMT

To determine its subcellular localization, CMT was fused to EGFP and examined in tobacco and Arabidopsis plants (Figure 5). Transient expression of the 35S::CMT-EGFP fusion protein in Arabidopsis protoplasts showed strong fluorescent signals in the nucleus (Figure 5G). The nuclear signal of CMT-EGFP fusion was further verified by its co-expression along with SV40NLS-DsRed2 that was successfully used as a nuclear localization control in Chapter 1 (Figure 5 A-F). In both transiently co-infiltrated tobacco (Figure 5 A-C) and stably co-transformed Arabidopsis leaves (Figure 5 D-F), the green fluorescent signals of CMT clearly overlapped with the red fluorescent signals of SV40NLS in the nuclei (Figure 5C and F). In addition, the tobacco leaves transiently expressing EGFP-CMT showed the similar fluorescent signals as obtained from the other orientation of fusion protein (Figure 5H).

To search for specific regions which are important for its nuclear targeting, several CMT deletion variants were fused to EGFP or DsRed2 and transiently expressed in tobacco leaves using an infiltration method. To test if one of the potential NLS sequences, $RKRGESDGARRSGRR^{339\text{-}353}$, acts as a nuclear localization signal, the 15-amino acid segment was fused to N-terminus of DsRed2 (Figure 5I) or deleted from CMT and fused to EGFP (Figure 5J). Tobacco leaves expressing CMTNLS-DsRed2 showed fluorescent signals in both cytoplasm and the nucleus no doubt because its small size allows it to diffuse between these two compartments. Therefore this result might not suggest that the 15-amino acid segment acts as a NLS (Figure 5I). This conclusion is supported by the observation that tobacco leaves expressing CMTΔNLS-EGFP displayed the nuclear specific localization, indicating that the deletion of the 15-amino acid segment did not

affect nuclear localization of CMT (Figure 5J). In addition, the CMT protein was split into two large fragments, N-terminal (1-170) and C-terminal (171-370), and each fragment was fused to N-terminus of EGFP to estimate the location of a NLS in CMT (Figure 5K and L). Tobacco leaves expressing nCMT-EGFP showed fluorescent signals only in the nucleus as observed with the full-length CMT-EGFP fusion (Figure 5K), whereas the leaves expressing the C-terminal CMT fused to EGFP displayed a diffuse staining pattern in the cytoplasm as well as in the nucleus (Figure 5L). These results suggested that the N-terminal deletion of CMT disrupted its nuclear localization, indicating the NLS sequence is perhaps located in N-terminal region of CMT (amino acids 1-170).

In addition to the nuclear localization of CMT, a punctate pattern of green fluorescent signals was also often observed in the epidermal cells of tobacco leaves transiently expressing CMT-EGFP fusion proteins (Figure 6A). These signals closely overlapped the red chlorophyll autofluorescence of the chloroplasts (Figure 6C; arrows). Interestingly, this enzyme is predicted to be targeted to chloroplasts by numerous bioinformatic sources including WoLF PSORT (Horton et al., 2007) and TargetP (Emanuelsson et al., 2000).

The subcellular localization of CMT was more carefully examined using the chloroplast-targeting sequence from the small subunit of Rubisco complex (RbcS) that was successfully used in Chapter 1, as a control (Figure 6). Arabidopsis transgenic plants co-expressing CMT-EGFP and RbcS-DsRed2 also displayed the punctate GFP patterns as well as nuclear localization signals (Figure 6D). While these punctate GFP signals closely overlapped with the red fluorescent signals of RbcS-DsRed2, no distinct GFP signals were detected in the chloroplasts of mesophyll cells (Figure 6D-F). The punctate patterns

of both green fluorescent signals of CMT-EGFP and red fluorescent signals of RbcS-DsRed2 were apparently located in the epidermal layers of the leaves, since they were visible only when the epidermal cell layers were in focus (Figure 6G-I).

Taken together, these results suggested that CMT is a nuclear-targeting enzyme and may also be targeted to the plastids in the epidermal cells. Although the nuclear localization signal peptide of CMT was not determined precisely in the present study, it seems to be positioned within the first 170 amino acids of the CMT polypeptide.

Figure 5. Nuclear localization of CMT.

(A-C) Tobacco leaves transiently co-expressing 35S::CMT-EGFP (A) and 35S::SV40NLS-DsRed2 (B) were imaged using confocal microscopy. (D-F) Leaves of Arabidopsis transgenic plants co-expressing 35S::CMT-EGFP (D) and 35S::SV40NLS-DsRed2 (E) were examined. (G) Transient expression of the 35S::CMT-EGFP fusion protein in protoplast. (H) Transient expression of the 35S::EGFP-CMT fusion protein in tobacco plants. (I) Transient expression of the 35S::CMTNLS-DsRed2 fusion protein in tobacco plants. (J) Transient expression of the 35S::CMTΔNLS-EGFP fusion protein in tobacco plants. (K-L) Transient expression of EGFP fused to N-terminal (K) or C-terminal (L) of CMT in tobacco plants. The green and red colours represent EGFP and DsRed2 signals, respectively. The autofluorescent signal of chlorophyll is depicted in red (G). Note the overlap of the EGFP and DsRed2 signals is shown in yellow (C and F). Arrows indicate the aggregates of GFP fusion proteins. Chl, chloroplast; N, Nuclei; nCMT, N-terminal fragment of CMT (amino acids 1-170); cCMT, C-terminal fragment of CMT (amino acids 171-370). Scale bars represent 20 μm.

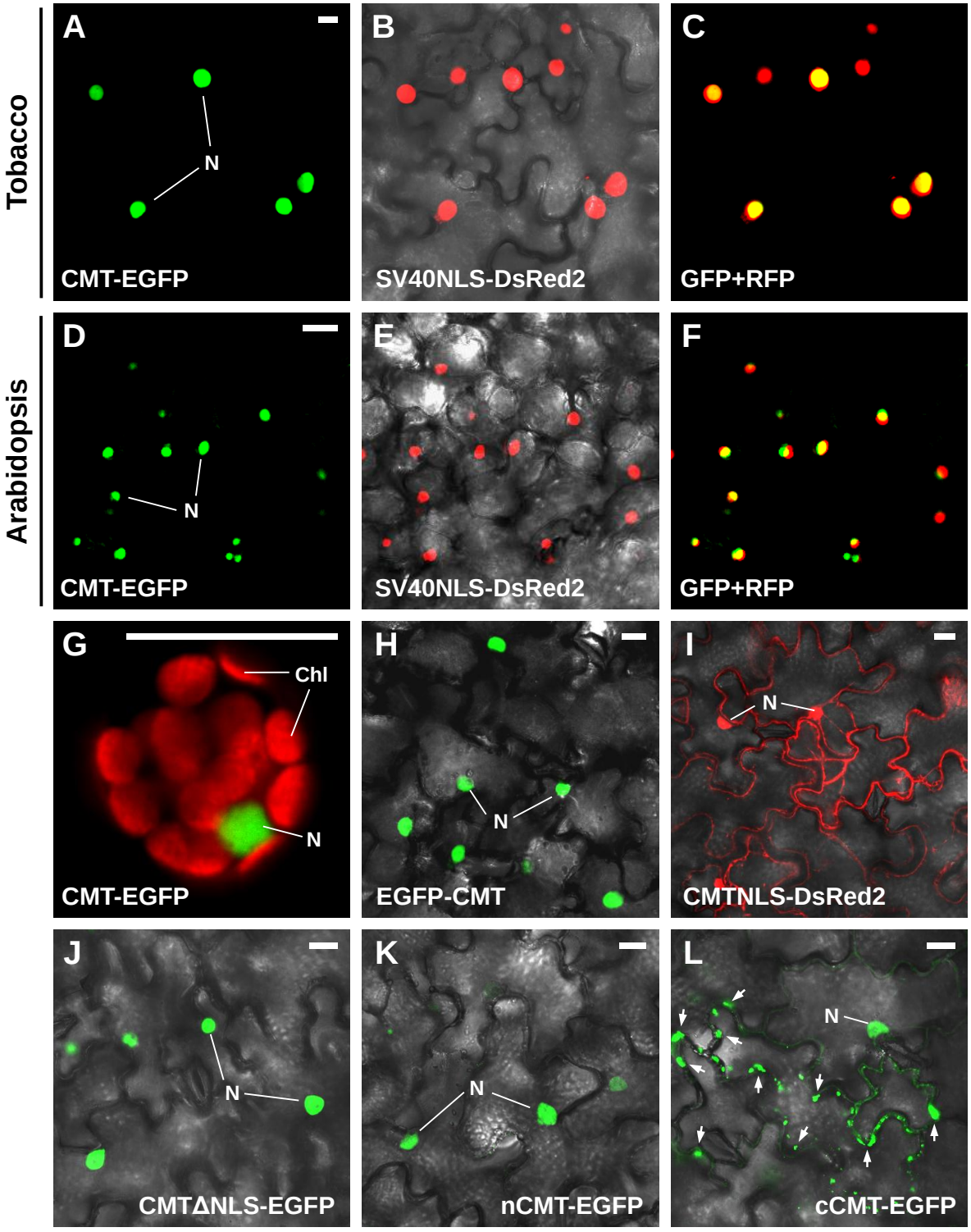
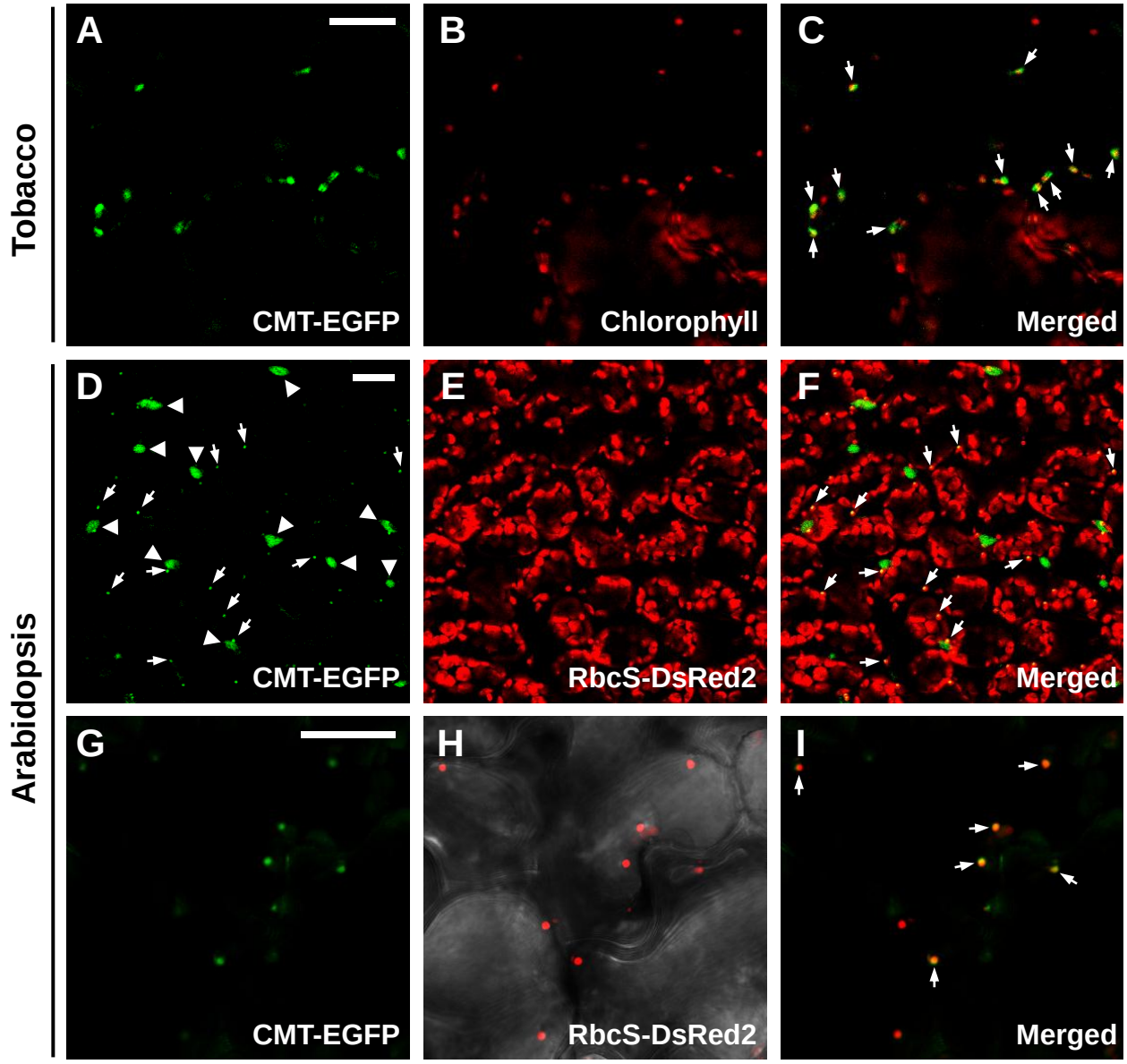

Tobacco
Arabidopsis
A
B
C
D
E
F
G
H
I
J
K
L
N
Chl
CMT-EGFP
SV40NLS-DsRed2
GFP+RFP
CMT-EGFP
SV40NLS-DsRed2
GFP+RFP
CMT-EGFP
EGFP-CMT
CMTNLS-DsRed2
CMTΔNLS-EGFP
nCMT-EGFP
cCMT-EGFP

Figure 6. Possible plastidic localization of CMT.

(A-C) Transient expression of the CMT-EGFP fusion protein in tobacco leaves. (D-I) Transgenic Arabidopsis leaves co-expressing CMT-EGFP (D and G) and RbcS-DsRed2 (E and H) were imaged using confocal microscopy. The green and red colours represent EGFP and DsRed2 signals, respectively. The autofluorescent signal of chlorophyll is depicted in red in B and C. Merged images of the EGFP and chlorophyll autofluorescence (C) and of the EGFP and DsRed2 (F and I) are shown in yellow. Arrowheads indicate nuclei. Arrows indicate colocalization of CMT-EGFP fusions with chlorophyll autofluorescence or RbcS-DsRed2. The 35S promoter was used to direct the expression of all constructs. RbcS, small subunit of Rubisco complex. Scale bars represent 20 μm.


Tobacco
Arabidopsis
A
CMT-EGFP
B
Chlorophyll
C
Merged
D
CMT-EGFP
E
RbcS-DsRed2
F
Merged
G
CMT-EGFP
H
RbcS-DsRed2
I
Merged

3.5. Protein-protein interaction of CMT

Nuclear import of proteins containing a NLS is initiated by the protein importin-α (Impα) that recognizes and binds to the NLS of the target protein in the cytoplasm (Gorlich and Kutay, 1999). Then, importin-β (Impβ) mediates translocation of the Impα-NLS containing protein complex from the cytoplasm to the nucleus through the nuclear pore complex (NPC) (Gorlich and Kutay, 1999). To examine CMT-Impα binding, glutathione *S*-transferase (GST) pull-down assays were performed using purified GST-Impα and wild-type Arabidopsis plant protein extracts. For this, 1,599 bp of full-length *Impα* gene (At3g06720) was cloned from a cDNA prepared from wild-type Arabidopsis plants using gene-specific primers. GST-Impα recombinant proteins were then expressed in *E. coli* and purified. The pull-down precipitates were analyzed by immunoblotting with a rabbit anti-CMT polyclonal antibody that was raised against His-CMT recombinant protein. The result showed that GST-Impα pulled-down CMT in the wild-type plant extracts but GST did not, suggesting that CMT bound to Impα through its NLS (Figure 7A).

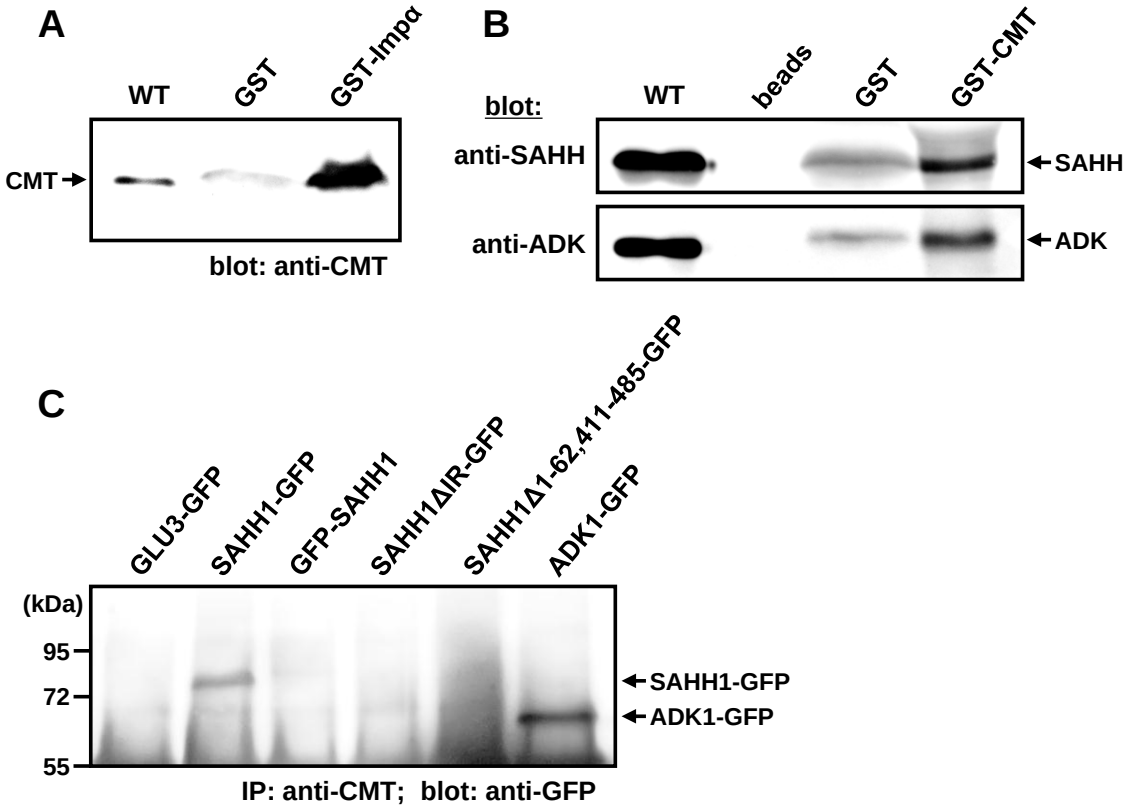
To investigate whether CMT interacts with ADK and SAHH, GST pull-down assays were performed using purified GST alone or GST-CMT recombinant proteins. The pulled-down products were detected by anti-SAHH and anti-ADK antibodies. Both SAHH and ADK were pulled-down by GST-CMT, but not by the GST alone or the beads alone (Figure 7B). Similar results were obtained with immunoprecipitation assays using an anti-CMT antibody to precipitate ADK or SAHH-GFP fusion proteins from leaf extracts of Arabidopsis transgenic plants. Both SAHH1-GFP and ADK1-GFP were successfully pulled-down by anti-CMT, whereas a negative control, GLU3-GFP was not

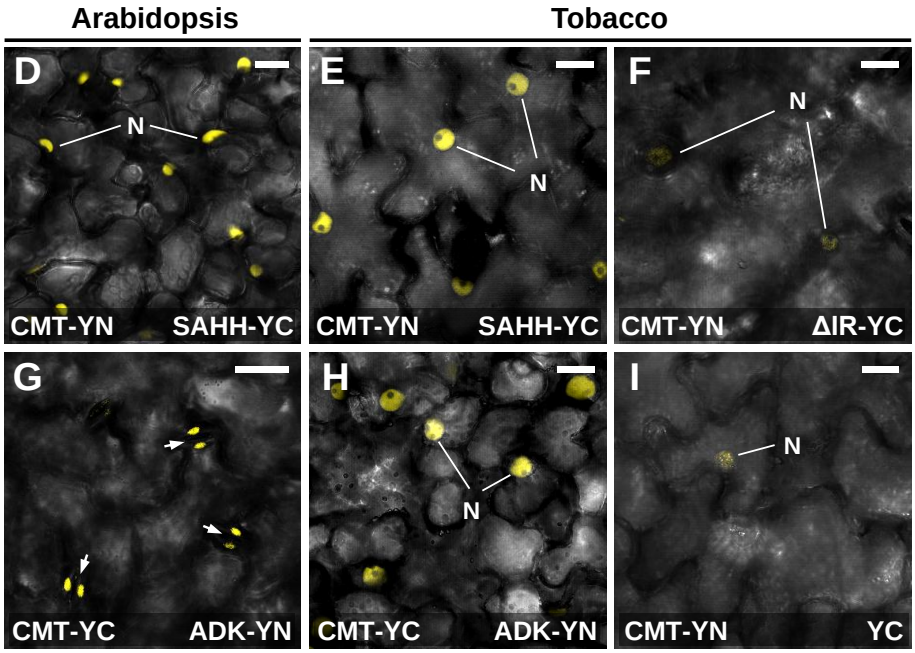
detected by immunoblotting (Figure 7C). In addition, anti-CMT antibody failed to pull-down other variants of SAHH1 fusions including GFP-SAHH1, SAHH1ΔIR-GFP, and SAHH1Δ1-62,411-485-GFP, suggesting that the interaction between CMT and SAHH1 may rely on specific regions of SAHH1 (Figure 7C).

To assess these interactions *in vivo*, BiFC assays were further applied in both tobacco and Arabidopsis plants (Figure 7D-I). For this, the full-length CMT coding sequence was fused to the N-terminus of YN or YC to produce CMT-YN and CMT-YC, with their expression driven by the 35S promoter. Consistent with the results obtained with GST pull-down and immunoprecipitation assays, co-expression of CMT-YN along with SAHH-YC resulted in strong YFP signals in the nucleus of both stably transformed Arabidopsis (Figure 7D) and transiently expressed tobacco plants (Figure 7E). Similar expression patterns were observed when CMT-YC was co-expressed with the ADK-YN (Figure 7G and H). However, no or very faint fluorescent signals were detected in the nucleus when CMT-YN was co-expressed with the SAHH1ΔIR, suggesting the deletion of IR disrupted the interaction between CMT and SAHH (Figure 7F). As a negative control, faint YFP signals were barely detected in the nucleus when CMT-YN was co-expressed with YC (Figure 7I). Taken together with the results obtained from GST pull-down and immunoprecipitation assays, BiFC results suggest that CMT interacts with both ADK and SAHH, and the nuclear targeting of both enzymes can be mediated by CMT.

Figure 7. Protein interaction of CMT with other proteins.

(A) GST pull-down assay to examine the interaction between CMT and importin α. GST fusion proteins were overexpressed and purified from *E. coli* BL21 cells. GST fusion proteins incubated with plant extracts prepared from wild-type Arabidopsis leaves were precipitated by GST•Mag agarose beads. The precipitates were resolved by SDS-PAGE and immunoblotted with anti-CMT antibody (1:300); the bound antibody was detected with alkaline phosphatase-conjugated secondary antibody and enhanced chemifluorescence (ECF) using the Typhoon 9400 laser scanning system. (B) GST pull-down assay to examine the interaction between CMT and ADK/SAHH. Immunoblotting analyses were performed using anti-SAHH (1:5,000) or anti-ADK (1:3,000) antibody. (C) Immunoprecipitation using the extracts prepared from various Arabidopsis transgenic plants expressing a GFP fusion as indicated above the panel. Extracts were immunoprecipitated with anti-CMT. The immunoprecipitates were resolved by SDS-PAGE and immunoblotted with anti-GFP. (D-I) BiFC assay for CMT interactions. Two YN/YC fusion proteins indicated in each figures were transiently expressed in tobacco (*Nicotiana benthamiana*) leaves (E, F, H, and I) and stably transformed in Arabidopsis plants (D and G). Both tobacco and Arabidopsis leaves were imaged using confocal microscopy. Yellow colour represents fluorescent signals of reconstituted YFP complexes. Arrows indicate the nuclei of guard cells of leaf epidermis. Note the CaMV 35S promoter directed the expression of each fusion protein. Molecular masses are indicated on the left in kilodaltons (kDa). Impα, importin α; M, size marker; WT, twenty-five micrograms of protein extracts obtained from wild-type Arabidopsis plants; YN, N-terminal fragment of YFP; YC, C-terminal fragment of YFP; ΔIR, SAHH1ΔIR. Scale bars represent 20 μm.


A
WT
GST
GST-Impα
CMT
blot: anti-CMT
B
blot:
WT
beads
GST
GST-CMT
anti-SAHH
SAHH
anti-ADK
ADK
C
GLU3-GFP
SAHH1-GFP
GFP-SAHH1
SAHH1ΔIR-GFP
SAHH1Δ1-62,411-485-GFP
ADK1-GFP
(kDa)
95
72
55
SAHH1-GFP
ADK1-GFP
IP: anti-CMT; blot: anti-GFP


Arabidopsis
Tobacco
D
N
CMT-YN
SAHH-YC
E
N
CMT-YN
SAHH-YC
F
N
CMT-YN
ΔIR-YC
G
CMT-YC
ADK-YN
H
N
CMT-YC
ADK-YN
I
N
CMT-YN
YC

3.6. Identification of other SAHH protein interactors: Isolation of aspartate semialdehyde dehydrogenase

To recover other protein interactors of SAHH, an affinity purification was performed using Arabidopsis plants expressing the SAHH1-HA.Strep fusion protein. A HA-StrepII epitope tag was fused to the C-terminus of SAHH1 and the expression of the fusion gene was directed by the CaMV 35S promoter. The SAHH1-HA.Strep fusion construct was transformed into Arabidopsis and the successful transgenic lines were selected by immunoblotting analysis using anti-HA antibody (Figure 8A). A large-scale purification of the Strep-tagged SAHH was performed using StrepTactin-Macroprep resin and the precipitates were analyzed by Coomassie Brilliant Blue staining (Figure 8B). As a result of the binding to SAHH1 fusion protein, several protein bands were visualized on the Coomassie-stained SDS-PAGE along with the overexpressed SAHH1 protein. Among these, seven gel pieces were excised and the proteins were treated with trypsin and subjected to MS spectrometry analysis (LC-MS/MS). Numerous polypeptides were identified to reside in these gel pieces as documented in Table 3. Most of these polypeptides were considered to be due to non-specific interactions apart from aspartate-semialdehyde dehydrogenase which showed a strong interaction with SAHH1 (Figure 8B; #5). Thus it was further studied to verify its protein interaction with SAHH and its subcellular localization.

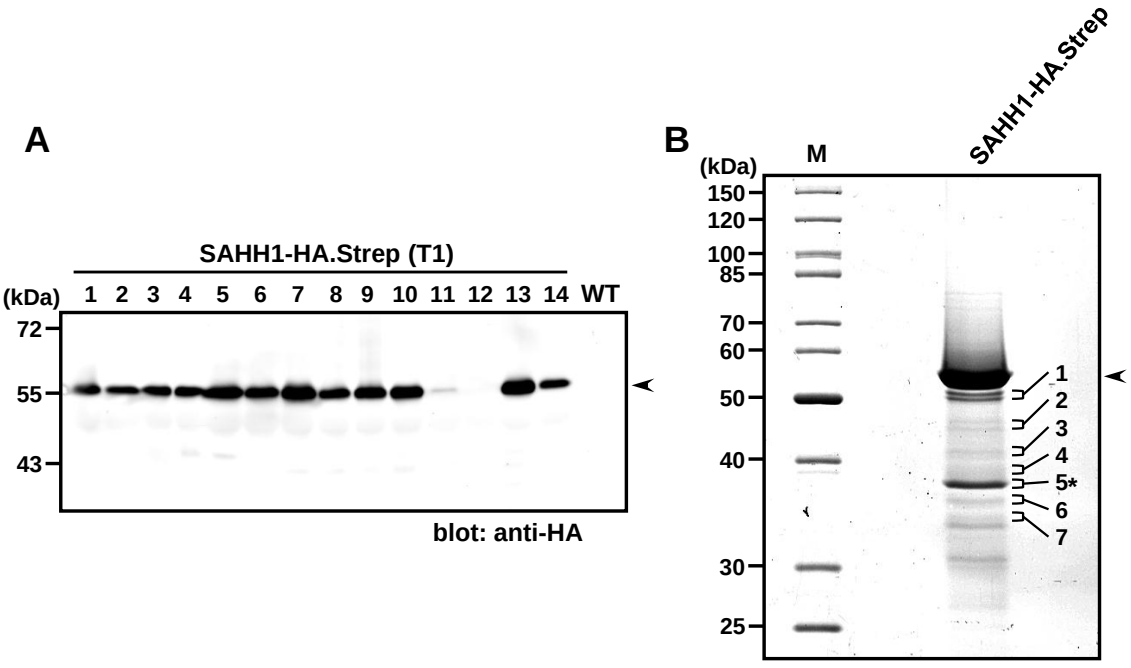
A full-length sequence encoding aspartate-semialdehyde dehydrogenase (ASDH; EC 1.2.1.11) was isolated from cDNA prepared from wild-type Arabidopsis plants by PCR amplification using gene-specific primers based on the sequence of At1g14810. The *ASDH* gene is composed of 7 exons spanning approximately 1,950 bp of genomic DNA

and 1,128 bp of open reading frame (ORF) encoding 375 amino acids with a predicted molecular weight of 40.7 kDa (Figure 8C). Sequence analysis using the deduced amino acid sequences of ASDH revealed that it contains a putative transit peptide (amino acid residues 1-35) and the active site residue Cys-170 (Figure 8C) identified by site-specific mutagenesis of a bacterial ASDH (Karsten and Viola, 1992).

ASDH, which catalyzes the second step of the pathway for the biosynthesis of essential amino acids derived from aspartate, is a common intermediate for threonine, lysine, and methionine biosynthesis (Azevedo et al., 1997). In this reaction, the β-aspartyl phosphate produced from aspartate by aspartate kinase activity, the first step of the pathway, is converted to aspartate semialdehyde by ASDH in an NADPH-dependent reaction. The product, aspartate semialdehyde is the first branch point intermediate between the lysine and threonine/methionine pathways.

Figure 8. SAHH1-Strep purification and ASDH identification.

(A) Immunoblot analysis of StrepII-tagged SAHH1 proteins in transgenic Arabidopsis. Total leaf protein extracts (15 μg) prepared from wild-type and individual T1-plants expressing SAHH1-HA.StrepII were analyzed by SDS-PAGE and immunoblotting using anti-HA antibody. (B) Strep purification using extracts prepared from the Arabidopsis plants expressing SAHH1-StrepII fusion. The extracts were precipitated by StrepTactin beads. The precipitates were resolved by SDS-PAGE and stained with Coomassie Brilliant Blue. The polypeptides within gel pieces #1 to #7 were analyzed by mass spectrometry. Arrowheads indicate overexpressed SAHH1-HA.StrepII fusion proteins in Arabidopsis. Molecular masses are indicated on the left in kilodaltons (kDa). M, size marker. (C) Nucleotide and deduced amino acid sequences of ASDH. Boxed amino acid residues indicate transit peptide of ASDH. An active site (Cys-170) of ASDH is circled. Asterisk indicates the stop codon of ASDH. The numbers on the right represent the nucleotide sequence and amino acid sequence, respectively.



C

```
atggcgacgttcactcatcaaaccccacaaacccatttcctctctcgccttcctctcagagctaaacctagacacttctcc   81
 M  A  T  F  T  H  Q  T  P  Q  T  H  F  L  S  R  L  P  L  R  A  K  P  R  H  F  S    27
gcaagagtcaaaatgtctctccaagaatctgcgccttcactcgccgtcgttggcgttaccggagccgtcggacaagaattc  162
 A  R  V  K  M  S  L  Q  E  S  A  P  S  L  A  V  V  G  V  T  G  A  V  G  Q  E  F    54
ctctccgtcttatccgatcgagacttcccttatagctccatcaagatgcttgcatcgaaacgctccgcagggaagcgtgtt  243
 L  S  V  L  S  D  R  D  F  P  Y  S  S  I  K  M  L  A  S  K  R  S  A  G  K  R  V    81
gccttcgatggccatgaatacacggtggaggagctcacggcggatagcttcaacggtgtagacatagctttgttcagtgct  324
 A  F  D  G  H  E  Y  T  V  E  E  L  T  A  D  S  F  N  G  V  D  I  A  L  F  S  A   108
ggtggatccataagcaaagagtttggtcctctggcggcggagaaaggaaccattgttgttgataatagctcggcttttaga  405
 G  G  S  I  S  K  E  F  G  P  L  A  A  E  K  G  T  I  V  V  D  N  S  S  A  F  R   135
atggttgatggagttccgcttgtgattcctgaagtgaatccagaagcaatgaaagggatcaaagttggaatgggcaaaggg  486
 M  V  D  G  V  P  L  V  I  P  E  V  N  P  E  A  M  K  G  I  K  V  G  M  G  K  G   162
gctttgattgcaaaccctaattgctccaccattatctgtttgatggctgttacgcctcttcatcatcacgcaaaggtgaag  567
 A  L  I  A  N  P  N  C  S  T  I  I  C  L  M  A  V  T  P  L  H  H  H  A  K  V  K   189
agaatggtggttagtacatatcaagcagctagtggtgctggtgctgcagctatggaagagcttgttcagcagactcgtgag  648
 R  M  V  V  S  T  Y  Q  A  A  S  G  A  G  A  A  A  M  E  E  L  V  Q  Q  T  R  E   216
gttttagagggtaaacctccgacttgtaacatcttcgggcagcagtatgcatttaacttgttttcacataatgctcctatt  729
 V  L  E  G  K  P  P  T  C  N  I  F  G  Q  Q  Y  A  F  N  L  F  S  H  N  A  P  I   243
cttgacaacggttacaacgaagaggaaatgaaacttgtgaaggagacaaggaagatctggaatgacacagaagtaaaagta  810
 L  D  N  G  Y  N  E  E  E  M  K  L  V  K  E  T  R  K  I  W  N  D  T  E  V  K  V   270
acagcaacgtgtatacgtgttccagttatgcgtgctcatgcagagagtgtgaatcttcagtttgagaatcctcttgatgag  891
 T  A  T  C  I  R  V  P  V  M  R  A  H  A  E  S  V  N  L  Q  F  E  N  P  L  D  E   297
aacacagcaagggagatactgaagaaagcacctggagtttatataatagatgaccgtgcttcaaatactttccctactcca  972
 N  T  A  R  E  I  L  K  K  A  P  G  V  Y  I  I  D  D  R  A  S  N  T  F  P  T  P   324
ctcgatgtctctaacaaagacgatgtagcggttggtaggataaggcgagacgtgtctcaagatggcaatttcgggctggac 1053
 L  D  V  S  N  K  D  D  V  A  V  G  R  I  R  R  D  V  S  Q  D  G  N  F  G  L  D   351
atattcgtttgtggagaccaaatacgcaaaggagctgctctgaatgctgttcagatcgctgagatgcttctctga       1128
 I  F  V  C  G  D  Q  I  R  K  G  A  A  L  N  A  V  Q  I  A  E  M  L  L  *         375
```

Table 3. Protein identification by Mass spectrometry

Bands	Protein ID	ATG number
#1	acetyl-CoA carboxylase	At5g35360
	tubulin beta chain	At5g62690
	ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit	AtCg00490
#2	ribulose-1,5-bisphosphate carboxylase/oxygenase activator (rubisco activase)	At2g39730
	putative chloroplast translation elongation factor EF-Tu	At4g20360
#3	rubisco activase	At2g39730
#4	polygalacturonase inhibitor-like protein	At3g20820
	glyceraldehyde-3-phosphate dehydrogenase C subunit (GapC)	At3g04120
	glyceraldehyde-3-phosphate dehydrogenase (NADP)	At3g26650
#5	aspartate-semialdehyde dehydrogenase (ASDH)	At1g14810
#6	SAHH (degraded protein fragments)	At4g13940
#7	acetyl-CoA carboxylase biotin-containing subunit	At5g16390
	ABC transporter-like protein	At5g02270

3.7. Subcellular localization of ASDH

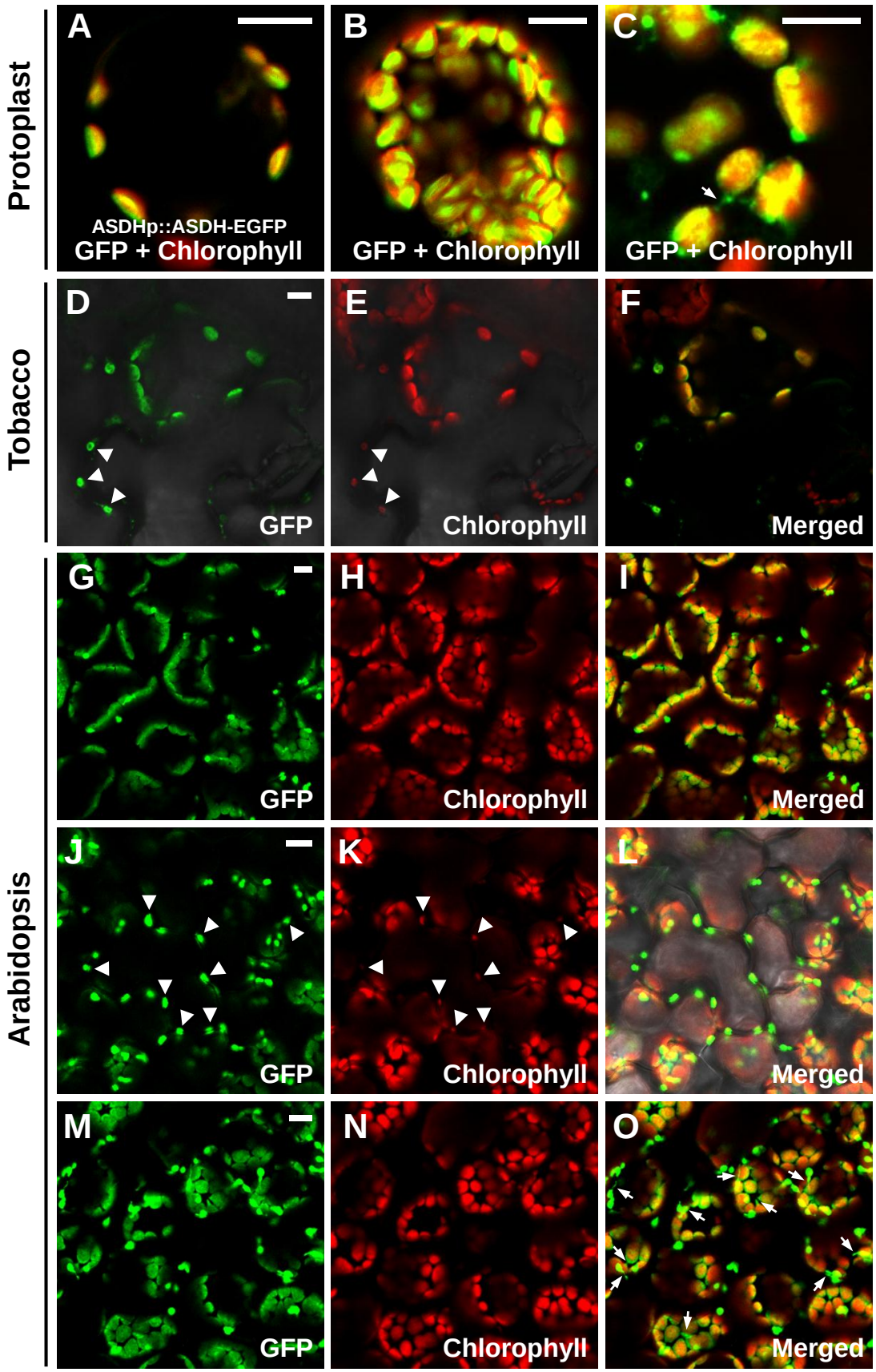
To study its subcellular localization, the full-length *ASDH* coding sequence was fused to EGFP under the control of CaMV 35S promoter or its native promoter and examined in tobacco and Arabidopsis plants (Figure 9). To test its endogenous promoter, the 1.2 kb upstream region of the *ASDH* gene was PCR amplified and the fusion construct was transiently expressed in Arabidopsis protoplasts. The green fluorescent signals of ASDHp::ASDH-EGFP clearly overlapped with the autofluorescence of chlorophyll, indicating that ASDH is targeted to the chloroplasts (Figure 9A). Transiently transformed protoplasts expressing ASDH under the control of CaMV 35S promoter displayed similar localization patterns as those with the endogenous promoter, suggesting that the use of the constitutive 35S promoter did not affect the movement of the ASDH-GFP fusion protein (Figure 9B). In addition, the fluorescent signals of ASDH-EGFP were often detected in stromule-like structures in protoplasts. Figure 9C showed the fluorescence in stromule-like structures between two chloroplasts, indicating the possible movement of GFP between chloroplasts that are connected by a stromule (Figure 9C; arrow).

Tobacco leaves transiently expressing ASDH-GFP under the control of 35S promoter showed fluorescence in the chloroplasts similar to that obtained with protoplasts, and the signals overlapped with those of the autofluorescence of chlorophyll (Figure 9D-F). Moreover, punctate-like fluorescent signals surrounding epidermal cell walls were detected in the tobacco leaves, and the signals closely overlapped with the autofluorescence of chlorophyll (Figure 9D-F). However, unlike the result with protoplasts, no fluorescent signals in stromule-like structures were visualized in transiently expressed tobacco leaves.

The subcellular localization of ASDH was further examined in transgenic Arabidopsis plants generated by *Agrobacterium*-mediated transformation. Consistent with the result of transient expressions in both Arabidopsis protoplasts and tobacco leaves, ASDH-EGFP fusion proteins under the control of 35S promoter were successfully targeted to the chloroplasts and their signals clearly overlapped with chlorophyll autofluorescence (Figure 9G-I). The punctate GFP signals observed in tobacco leaves were also detected in transformed Arabidopsis plants (Figure 9J-L; arrowheads). Occasionally the punctate GFP signals were even stronger than that of chloroplasts. In order to assess if these punctate signals were an artifact caused by the overexpression of ASDH by the constitutive 35S promoter, the Arabidopsis plants expressing ASDH-EGFP under the control of its native promoter were generated. These plants also displayed similar punctate-GFP signals surrounding epidermal cell walls and faint fluorescent signals in chloroplasts of mesophyll cells (data not shown). Moreover, the stromule-like GFP signals observed in protoplasts (Figure 9C) were also visualized in Arabidopsis leaves (Figure 9M-O; arrows). However, no GFP signals were detected in either the nucleus or cytoplasm. Together, these results suggested that ASDH is a chloroplast-specific enzyme which also resides in stromules.

Figure 9. Subcellular localization of ASDH.

(A) Transient expression of ASDH-EGFP under the control of its native promoter in Arabidopsis protoplasts. (B-C) Transient expression of 35S::ASDH-EGFP in Arabidopsis protoplasts. (D-F) Transient expression of 35S::ASDH-EGFP in tobacco plants. (G-I) Transgenic Arabidopsis leaves expressing 35S::ASDH-EGFP. (J-L) Expression of 35S::ASDH-EGFP in Arabidopsis epidermis showing punctate patterns. (M-O) Expression of 35S::ASDH-EGFP in Arabidopsis leaves showing stromule-like structures. The green and red colours represent EGFP signals and chlorophyll autofluorescence, respectively. Note the overlap of the EGFP and chlorophyll autofluorescent signals is shown in yellow. Arrowheads indicate punctate GFP signals. Arrows indicate the stromule-like structures. Scale bars represent 10 μm.


Protoplast
Tobacco
Arabidopsis
A
ASDHp::ASDH-EGFP
GFP + Chlorophyll
B
GFP + Chlorophyll
C
GFP + Chlorophyll
D
GFP
E
Chlorophyll
F
Merged
G
GFP
H
Chlorophyll
I
Merged
J
GFP
K
Chlorophyll
L
Merged
M
GFP
N
Chlorophyll
O
Merged

3.8. Protein interaction of ASDH and SAHH

To assess the protein interaction between SAHH and ASDH, a pull-down assay was performed using wild-type Arabidopsis plant extracts and His-tagged full-length ASDH recombinant proteins that were expressed in *E. coli* and purified. His-tagged proteins incubated with the plant crude extracts were precipitated by His•Mag agarose beads, and the precipitates were analyzed by SDS-PAGE and Coomassie Brilliant Blue staining (Figure 10A, left panel). However, no visible SAHH protein band was detected on the Coomassie-stained SDS-PAGE, although the His-ASDH proteins were successfully overexpressed (Figure 10A, left panel). To confirm whether SAHH proteins were pulled-down by His-ASDH, the precipitates were further analyzed by immunoblotting with anti-SAHH antibody (Figure 10A, right panel). The result showed that His-ASDH can pull-down SAHH from the total protein extracts of wild-type Arabidopsis plant (Figure 10A, right panel).

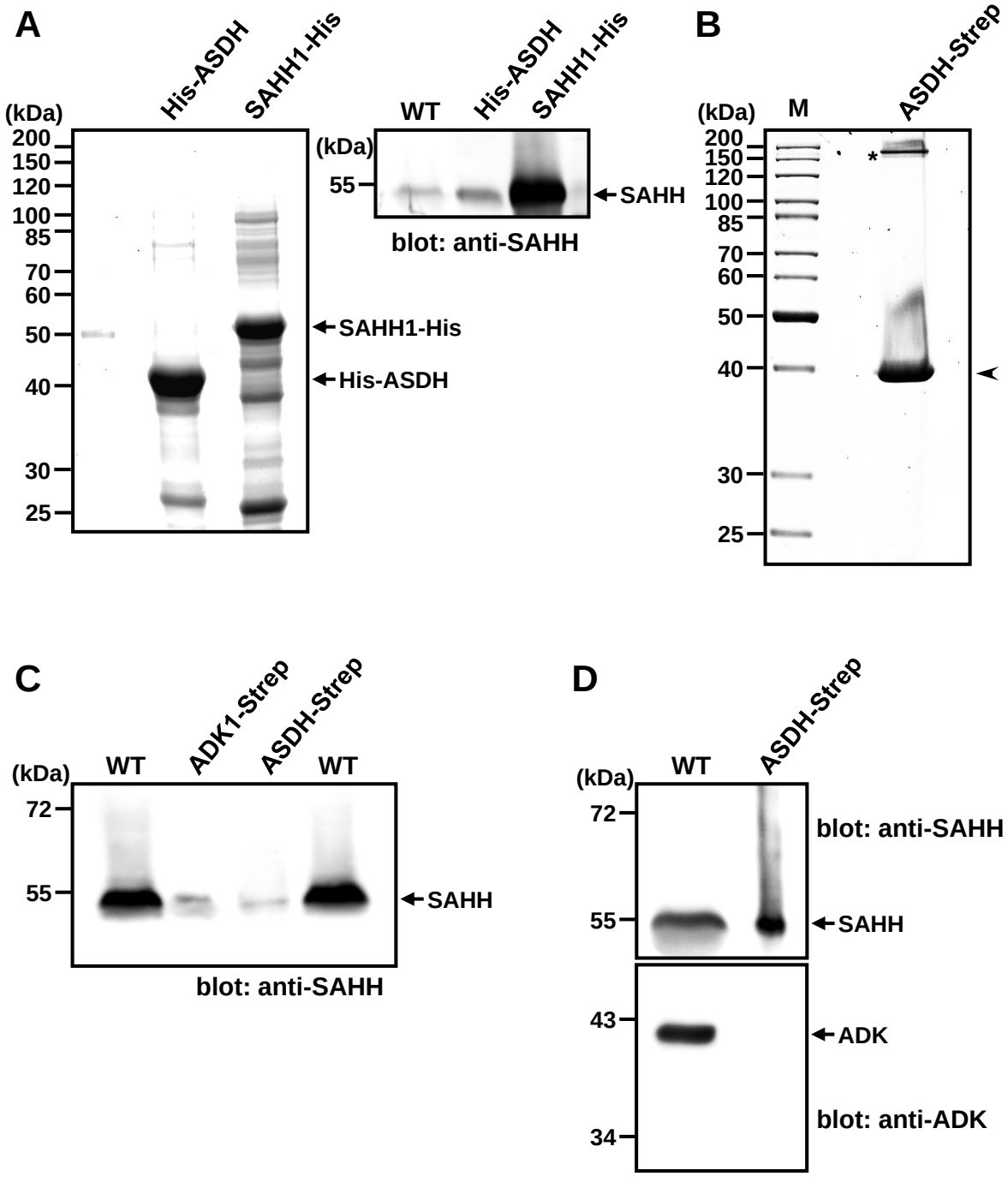
In addition, a Strep purification was performed using Arabidopsis plants expressing an ASDH-Strep fusion to test whether ASDH-Strep can pull-down SAHH. For this, a Strep tag was fused to C-terminus of full-length ASDH and transformed in Arabidopsis. The protein extracts prepared from the successful transgenic lines were precipitated by StrepTactin-Macroprep resin and the precipitates were analyzed by Coomassie Brilliant Blue staining (Figure 10B). However, there was no detectable SAHH protein band on the Coomassie-stained SDS-PAGE, while the strong protein band of the overexprssed ASDH-Strep proteins was visualized (Figure 10B; arrowhead). Interestingly, one clear protein band was detected approximately at 170 kDa in the ASDH-Strep purification (Figure 10B; asterisk). The identity of the protein band was confirmed as ASDH by mass

spectrometric analysis. Since SAHH proteins were not visible with the Coomassie staining, the precipitates prepared from ASDH-Strep and ADK1-Strep expressing plants were analyzed by immunoblotting with anti-SAHH antibody (Figure 10C). The results indicated that SAHH proteins were pulled-down by ASDH-Strep as well as by ADK1-Strep (Figure 10C).

To investigate whether ASDH interacts with ADK, the precipitates prepared from the Arabidopsis plants expressing ASDH-Strep were analyzed by immunoblotting with anti-ADK antibody (Figure 10D). While the strong SAHH signals were detected by anti-SAHH antibody as expected (Figure 10D, upper panel), no ADK proteins were detected by the immunoblotting (Figure 10D, lower panel). To verify this result, the interaction between ASDH and ADK was further examined by a co-immunoprecipitation and a pull-down assay using purified His-tagged proteins, but no detectable interaction signals were confirmed (data not shown). Together, all these results suggested that ASDH interacts with SAHH but not with ADK.

Figure 10. Interaction between SAHH and ASDH.

(A) His-tag pull-down assay using ASDH recombinant protein and plant crude extracts prepared from *E. coli* BL21 cells and wild-type Arabidopsis plant leaves, respectively. His-ASDH proteins incubated with the plant extracts were precipitated by His•Mag agarose beads. The precipitates were subjected to SDS-PAGE and Coomassie Brilliant Blue staining (left panel) or immunoblotted with anti-SAHH antibody (1:5,000; right panel). (B) Strep purification using extracts prepared from the Arabidopsis plants expressing ASDH-Strep fusion protein using StrepTactin beads. The precipitates were resolved by SDS-PAGE and Coomassie Brilliant Blue stained. An arrowhead indicates the overexpressed ASDH-Strep fusion proteins in Arabidopsis. (C) Immunoblot analysis using the products of ADK1-Strep and ASDH-Strep purifications. Strep purification was performed using extracts prepared from the Arabidopsis plants expressing each Strep fusion protein. The precipitates were resolved by SDS-PAGE and analyzed by immunoblotting with anti-SAHH antibody (1:5,000). (D) Immunoblot analysis using the product of ASDH-Strep purification to examine the interaction between ASDH and ADK. The precipitates were analyzed by immunoblotting with anti-SAHH (1:5,000; upper panel) and anti-ADK (1:3,000; lower panel) antibodies. Note the CaMV 35S promoter directed the expression of each Strep fusion protein. Molecular masses are indicated on the left in kilodaltons (kDa). M, size marker; WT, twenty-five micrograms of protein extracts obtained from wild-type Arabidopsis plants.


A
His-ASDH
SAHH1-His
(kDa)
200
150
120
100
85
70
60
50
40
30
25
SAHH1-His
His-ASDH
WT
His-ASDH
SAHH1-His
(kDa)
55
SAHH
blot: anti-SAHH
B
(kDa)
M
ASDH-Strep
200
150
120
100
85
70
60
50
40
30
25
*
C
ADK1-Strep
ASDH-Strep
(kDa)
WT
WT
72
55
SAHH
blot: anti-SAHH
D
ASDH-Strep
(kDa)
WT
72
55
blot: anti-SAHH
SAHH
43
34
ADK
blot: anti-ADK

3.9. Protein interaction of ASDH and SAHH variants

Additional Strep-tagged constructs fused to SAHH1 variants were created and used for purifications in an attempt to define specific regions in SAHH which are important for its protein interaction with ASDH (Figure 11). For this, a Strep-tag was fused to the N-terminus of full-length SAHH1, the C-terminus of SAHH1ΔIR, the C-terminus of yeast SAHH, the C-terminus of SAHH1ΔPDP ($PDP^{164-166}$ to AAA; Chapter 1) or the C-terminus of SAHH1ΔTST ($TST^{167-169}$ to AAA; Chapter 1). All these fusion constructs were transformed into Arabidopsis plants, and the successful transgenic lines were selected by BASTA treatment and verified by immunoblotting using anti-HA or anti-Strep antibodies (data not shown). I failed to obtain transgenic plants expressing either SAHH1ΔIR-Strep or yeast SAHH-Strep despite screening more than 50 T1-transgenic plants for each line (data not shown). However, the other fusion proteins were expressed well in Arabidopsis and successfully used for each purification.

First, the protein extracts prepared from the plants expressing Strep-SAHH1 were used for the purification to test if the N-terminal fusion on SAHH affected its interaction with ASDH (Figure 11A). The precipitates of Strep-SAHH1 purified extracts were analyzed by SDS-PAGE and Coomassie Brilliant Blue staining along with the one from SAHH1-Strep. No visible ASDH was detected from the Strep-SAHH1 purification, whereas both fusion proteins, SAHH1-Strep and Strep-SAHH1 were successfully expressed and the ASDH protein band from the SAHH1 purification was clearly detected on the Coomassie-stained SDS-PAGE as expected (Figure 11A). This result indicated that the N-terminal SAHH1 fusion may interrupt its interaction with ASDH.

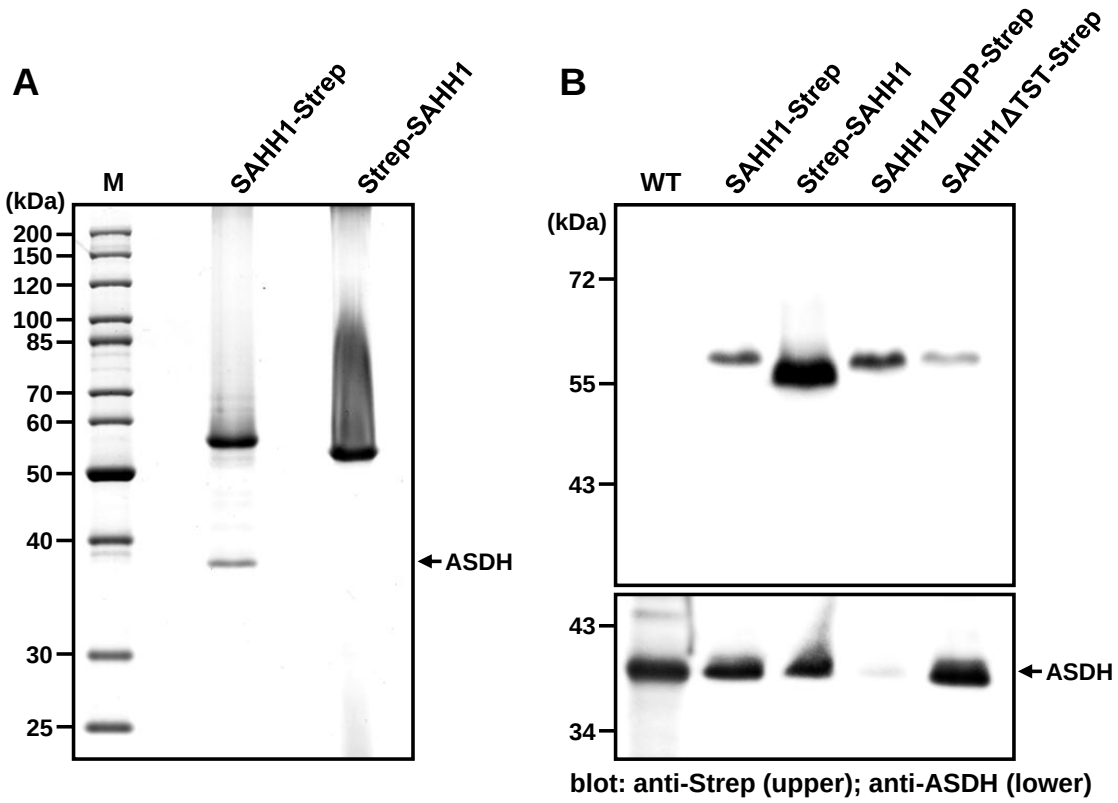
The interaction between Strep-SAHH1 and ASDH was further examined by immunoblotting along with two SAHH1 mutant fusions, SAHH1ΔPDP and SAHH1ΔTST (Figure 11B). For this, a rabbit anti-ASDH polyclonal antibody was raised against His-ASDH recombinant protein (data not shown). The protein extracts prepared from the plants expressing these Strep fusion proteins were used for the purification, and the precipitates were analyzed by SDS-PAGE and immunoblotting with anti-ASDH antibody (Figure 11B). The successful expression of the Strep fusion proteins was confirmed by immunoblotting with anti-Strep antibody (Figure 11B, upper panel); binding of ASDH was detected by anti-ASDH antibody (Figure 11B, lower panel). Interestingly, ASDH was detected in the Strep-SAHH1 purification, although it was not visible with Coomassie staining. However, this happened only when Strep-SAHH1 fusion proteins were highly expressed in Arabidopsis; the SAHH1-Strep fusion proteins were much less effective, suggesting that N-terminal region of SAHH1 is somehow involved in the interaction with ASDH. More surprisingly, the purification using the extracts prepared from the plants expressing SAHH1ΔPDP-Strep fusion showed almost no or very weak ASDH binding, whereas the other fusion proteins including SAHH1-Strep and SAHH1ΔTST-Strep successfully pulled down ASDH (Figure 11B). This result indicated that the substitution of PDP to AAA at positions 164-166 of SAHH1 disrupted its interaction with ASDH.

Two different sizes of ASDH were often detected from the SAHH1ΔTST purification by both Coomassie-stained SDS-PAGE (Figure 11C) and immunoblotting with anti-ASDH antibody (Figure 11D); no ASDH proteins were detected from the SAHH1ΔPDP purification as expected (Figure 11C and D). Based on the size difference, the smaller

ASDH polypeptide (~36.6 kDa) is the one normally obtained from other SAHH Strep fusion purifications. This protein lacks the transit sequence corresponds to that predicted by the Genbank record AAG33078. To confirm if the larger polypeptide (~38 to 39 kDa) is the precursor protein (~40.7 kDa, predicted) that contains the transit peptide, the identity of both protein bands were analyzed by mass spectrometry. No tryptic peptides belonging to the transit peptides of ASDH (amino acid sequence 1–35) were found in either polypeptide, suggesting that both may represent the mature protein.

Figure 11. Interaction between ASDH and SAHH variants.

(A) Strep purification using extracts prepared from 700 mg of Arabidopsis plant leaves expressing SAHH1-Strep or Strep-SAHH1 fusion proteins. The extracts were precipitated by StrepTactin beads. The precipitates were resolved by SDS-PAGE and stained with Coomassie Brilliant Blue. (B) Immunoblot analysis using the products of Strep purifications. Strep purifications were performed using extracts prepared from the Arabidopsis plants expressing each Strep fusion protein. The precipitates were resolved by SDS-PAGE and analyzed by immunoblotting with anti-Strep (1:1,000; upper panel) and anti-ASDH (1:2,000; lower panel) antibodies. (C) Strep purification using extracts prepared from the Arabidopsis plants expressing SAHH1ΔPDP-Strep or SAHH1ΔTST-Strep fusion protein. The precipitates were resolved by SDS-PAGE and stained with Coomassie Brilliant Blue. (D) Immunoblot analysis using Strep purification products. Strep purification was performed using extracts prepared from the Arabidopsis plants expressing SAHH1ΔPDP-Strep or SAHH1ΔTST-Strep fusion protein. The precipitates were resolved by SDS-PAGE and analyzed by immunoblotting with anti-ASDH. Asterisks indicate the larger size of ASDH protein band. Note the 35S promoter was used to direct the expression of all Strep fusion constructs. Molecular masses are indicated on the left in kilodaltons (kDa). M, size marker; WT, twenty-five micrograms of protein extracts obtained from wild-type Arabidopsis plants.


A
M
SAHH1-Strep
Strep-SAHH1
(kDa)
200
150
120
100
85
70
60
50
40
30
25
ASDH
B
WT
SAHH1-Strep
Strep-SAHH1
SAHH1ΔPDP-Strep
SAHH1ΔTST-Strep
(kDa)
72
55
43
43
34
ASDH
blot: anti-Strep (upper); anti-ASDH (lower)
C
SAHH1ΔPDP-Strep
SAHH1ΔTST-Strep
(kDa)
M
200
150
120
100
85
70
60
50
40
30
25
ΔPDP-Strep
ΔTST-Strep
ASDH
*
D
WT
SAHH1ΔPDP-Strep
SAHH1ΔTST-Strep
(kDa)
43
34
*
ASDH
blot: anti-ASDH

3.10. Protein interaction with APT

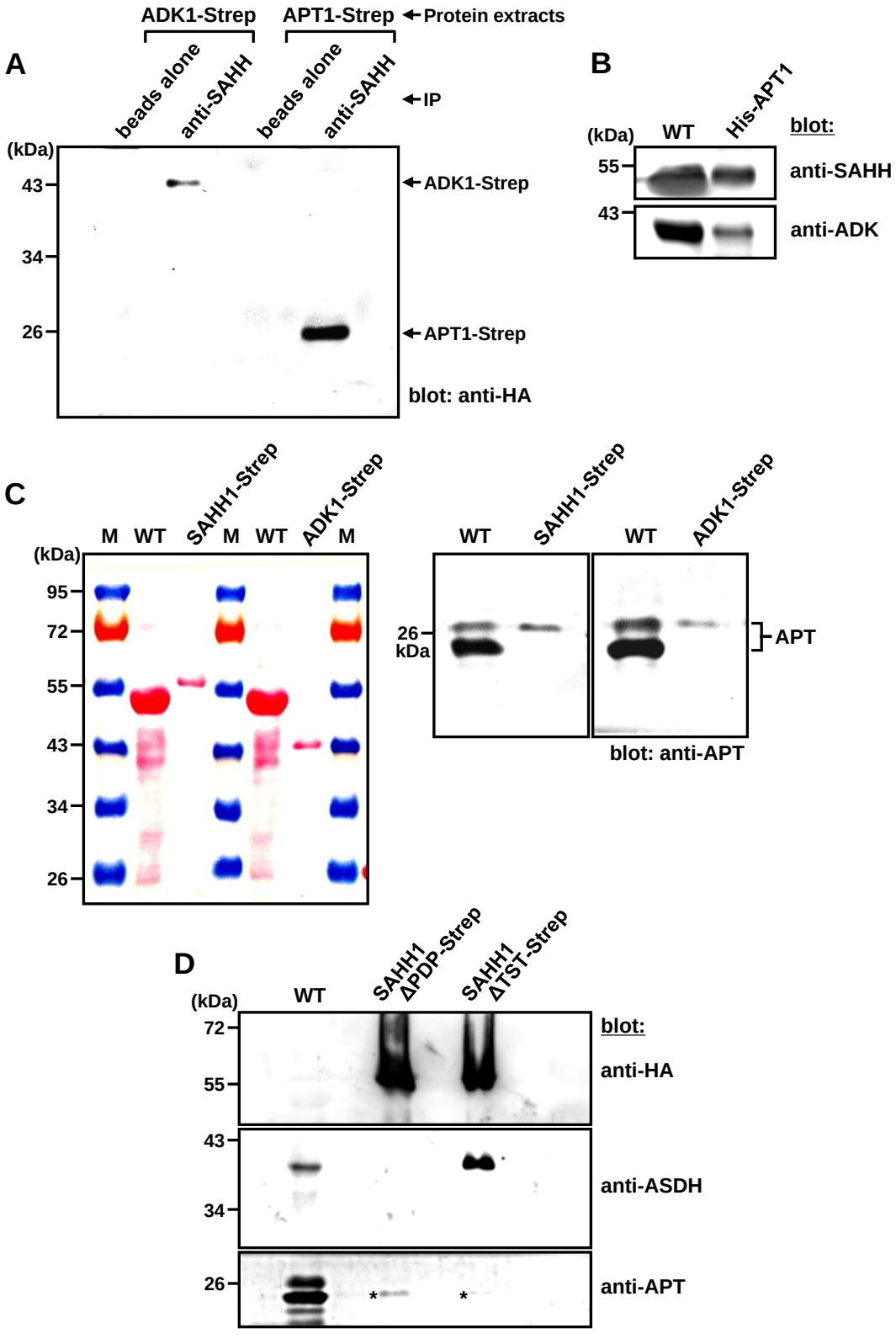
In addition to the protein interaction between SAHH and ASDH, I found evidence that other proteins may also be associated with the ADK-SAHH complex. These unexpected results were obtained while attempting to assess the protein interaction between ADK and SAHH. Co-immunoprecipitation was performed using total protein extracts prepared from the Arabidopsis transgenic plants expressing 35S::ADK1-Strep and anti-SAHH antibody to pull-down the Strep fusion protein. The protein extracts containing overexpressed ADK1-Strep proteins were immunoprecipitated with anti-SAHH antibody, and the precipitates were analyzed by immunoblotting with anti-HA antibody. ADK1-Strep was successfully pulled-down by anti-SAHH, whereas beads alone without adding anti-SAHH that was used as a negative control failed to pull-down ADK1-Strep proteins (Figure 12A). In addition to this, one more negative control experiment was performed using the transgenic Arabidopsis plants overexpressing adenine phosphoribosyltransferase (APT) fused to a Strep tag (Facciuolo, 2009). APT is the purine salvage enzyme which catalyzes the conversion of adenine to adenosine monophosphate (Moffatt and Somerville, 1988). EST analysis reveals that there are two forms of APT1, one differs from the other by the presence of an additional 5´ terminal exon, and the short one has been shown to reside in plastids (Facciuolo, 2009). Interestingly, the control experiment revealed that APT1-Strep proteins were also immunoprecipitated by anti-SAHH serum, indicating the possible interaction between SAHH and APT (Figure 12A). A similar result was obtained with a pull-down assay using wild-type Arabidopsis plant extracts and His-tagged APT1 recombinant proteins. Purified His-APT1 proteins incubated with plant extracts were precipitated by His•Mag

agarose beads and the precipitates were immunoblotted with anti-SAHH and anti-ADK antibodies. The result showed that His-APT1 proteins can pull-down both SAHH and ADK from the total plant extracts (Figure 12B).

Moreover, their interactions were further examined by Strep purifications using the plant lines expressing SAHH1-Strep or ADK1-Strep (Figure 12C). The purification products were stained with Ponceau S to check the expression of fusion proteins (Figure 12C, left panel) and analyzed by immunoblotting with an anti-APT antibody (Figure 12C, right panel). The immunoblot results showed that both SAHH1-Strep and ADK1-Strep pulled-down a long form of APT1 (Figure 12C, right panel). To test whether the SAHH1 mutations with substitution of 3 amino acid residues affect its interaction with APT, both plant lines expressing SAHH1ΔPDP-strep or SAHH1ΔTST-Strep were used for purification (Figure 12D). The overexpressed fusion proteins were confirmed with anti-HA and the interaction between ASDH and the two fusion proteins were detected by anti-ASDH. Consistent with the earlier results, no interaction signal was detected from the SAHH1ΔPDP-Strep purification, whereas strong signals were obtained with the SAHH1ΔTST-Strep purification. Unlike the result with the full-length SAHH1-Strep purification showing the interaction with a long form of APT (Figure 12C, right panel), both SAHH1ΔPDP-Strep and SAHH1ΔTST-Strep bound to a short form of APT but not with a long form (Figure 12D; asterisks). The signals obtained with the SAHH1ΔPDP-Strep purification were generally stronger than those of SAHH1ΔTST-Strep purification.

Figure 12. Interaction of APT with ADK and SAHH.

(A) Co-immunoprecipitation using the extracts prepared from 700 mg of Arabidopsis transgenic plant leaves expressing ADK1-Strep or APT1-Strep. The extracts were immunoprecipitated with anti-SAHH serum using protein A sepharose beads. The immunoprecipitates were resolved by SDS-PAGE and immunoblotted with anti-HA antibody (1:4,000). (B) His-tag pull-down assay using His-APT1 recombinant protein and plant crude extracts prepared from *E. coli* BL21 cells and wild-type Arabidopsis plants, respectively. His-APT1 proteins incubated with the plant extracts were precipitated by His•Mag agarose beads. The precipitates were subjected to SDS-PAGE and immunoblotted with anti-SAHH (1:5,000; upper panel) and anti-ADK (1:3,000; lower panel) antibodies. (C) Strep purification using extracts prepared from the Arabidopsis plant expressing SAHH1-Strep or ADK1-Strep fusion protein to examine the interaction between APT and SAHH/ADK. The extracts were precipitated by StrepTactin beads. The precipitates were resolved by SDS-PAGE and Ponceau S stained (left panel), and analyzed by immunoblotting with anti-APT antibody (1:10,000; right panel). (D) Immunoblot analysis using Strep purification products. Strep purifications were performed using extracts prepared from the Arabidopsis plants expressing SAHH1ΔPDP-Strep or SAHH1ΔTST-Strep fusion protein. The precipitates were resolved by SDS-PAGE and analyzed by immunoblotting with anti-HA (1:4,000), anti-ASDH (1:2,000), and anti-APT (1:10,000) antibodies as indicated. Asterisks indicate the weak protein bands of APT. Note the 35S promoter directs expression of all Strep fusion constructs. Molecular masses are indicated on the left in kilodaltons (kDa). M, size marker; WT, Twenty-five micrograms of protein extracts obtained from wild-type Arabidopsis plants.


A
ADK1-Strep
APT1-Strep
Protein extracts
beads alone
anti-SAHH
beads alone
anti-SAHH
IP
(kDa)
43
34
26
ADK1-Strep
APT1-Strep
blot: anti-HA
B
(kDa)
WT
His-APT1
blot:
55
43
anti-SAHH
anti-ADK
C
M
WT
SAHH1-Strep
M
WT
ADK1-Strep
M
(kDa)
95
72
55
43
34
26
WT
SAHH1-Strep
WT
ADK1-Strep
26
kDa
APT
blot: anti-APT
D
(kDa)
WT
SAHH1 ΔPDP-Strep
SAHH1 ΔTST-Strep
72
55
43
34
26
blot:
anti-HA
anti-ASDH
anti-APT
*
*

3.11. T-DNA insertion mutants of CMT and ASDH

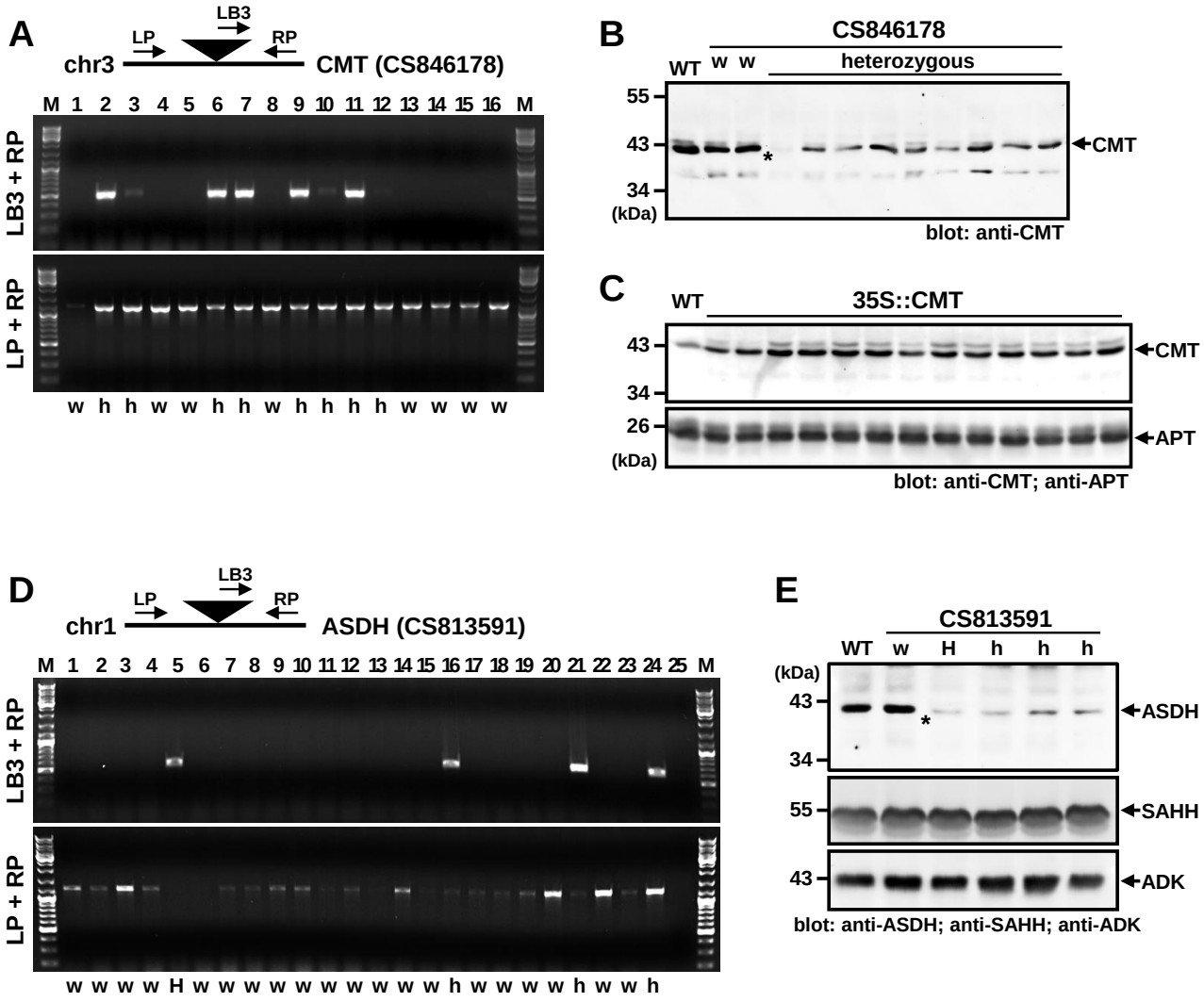
The previous results document interactions between ADK/SAHH and CMT and between SAHH and ASDH. I next investigated to see whether CMT knock-out lines could be recovered and if so, did the loss of CMT affect SAHH or ADK functionality/phenotypic changes. For this, the *cmt* (SAIL_1252_D05; CS846178) and *asdh* (SAIL_293_E04; CS813591) T-DNA insertion lines (both are inserted in exons of the indicated gene) were obtained from the Arabidopsis SALK collection at the Arabidopsis Biological Resource Center (ABRC) (Alonso et al., 2003). To monitor the genetic segregation of T-DNA and to screen homozygous lines, the T-DNA tagged lines were genotyped using genomic DNAs isolated from Arabidopsis seedlings by PCR reactions using the primers, LP, RP, and LB3 (Figure 13A and D). However, I failed to isolate homozygous *cmt* plants from more than 60 individual seedlings (SAIL_1252_D05; CS846178) examined. The segregation ratio of wild-type, heterozygous, and homozygous plants was about 1:1:0, indicating a gametophytic defect (Figure 13A). This abnormal segregation ratio suggested that CMT is essential for plant growth and development. The protein expression level of CMT in heterozygous mutant lines was generally lower than that of wild-type plants (Figure 13B), but no distinct phenotypes were observed. To create CMT overexpression lines, the construct of the full-length CMT under the control of CaMV 35S promoter was transformed in wild-type Arabidopsis (Figure 13C). However, I could not obtain the plants exhibiting any phenotypic changes, while CMT protein expression in transgenic plants was increased approximately 2-fold compared to that of wild-type plants (Figure 13C). In addition, two more T-DNA insertion lines (SALK_148957 and SALK_148949) for the other potential cap methyltransferase

(At3g52210), which was not used in this study, were screened by PCR. The result for both T-DNA tagged lines showed a segregation ratio close to 1:2:1 (data not shown). However, homozygous mutant lines for both SALK_148957 and SALK_148949 exhibited no distinct phenotypes (data not shown).

A total of 24 seedlings of an *asdh* T-DNA insertion line were tested to determine their homo- or heterozygosity by PCR (Figure 13D). The results revealed that one (lane #5) and three (lane #16, #21, #24) plants were found as homozygous and heterozygous, respectively, and the remaining plants (twenty) were wild type (Figure 13D). The protein expression level of ASDH was reduced in all homozygous and heterozygous mutant plants compared to that of wild-type plants, whereas no changes in protein expression were observed for either ADK or SAHH (Figure 13E). However, no phenotypes were observed in homozygous and heterozygous mutant plants when grown under normal conditions (16h day length; data not shown). This may due to residual ASDH activity, indicating that an ASDH homozygote line (Figure 13D, lane#5) as identified by the PCR was not stable homozygote mutant. This line needs to be re-examined in order to determine the precise genotype of this line.

Figure 13. T-DNA insertion lines for CMT and ASDH.

(A) Genotyping of T-DNA insertion line for CMT (CS846178; SAIL_1252_D05) with PCR using two different primer sets. LP, forward primer generated from the 7th exon of *CMT* gene; RP, reverse primer generated from the 13th exon of *CMT*; LB3, forward primer located in T-DNA region. The LP and RP primers produce the 1.2 kb PCR fragment only when there is no T-DNA insertion in the genome, while the LB3 and RP primers produce the ~0.5 kb fragment only when T-DNA is present in the genome. (B) Immunoblot analysis using T-DNA insertional mutant lines to check their protein expression level. The protein extracts prepared from wild-type and heterozygous Arabidopsis plants were immunoblotted with anti-CMT antibody. (C) Immunoblot analysis using CMT overexpression lines. The protein extracts prepared from wild-type and transformed Arabidopsis plants expressing 35S::CMT were immunoblotted with anti-CMT (upper panel) and anti-APT antibodies (lower panel). (D) Genotyping of T-DNA insertion line for ASDH (CS813591; SAIL_293_E04) with PCR using two different primer sets. LP, forward primer generated from the 5' upstream region of *CMT* gene; RP, reverse primer generated from the 3rd exon of *CMT* gene; LB3, forward primer located in T-DNA region. The LP and RP primers produce the 1.2 kb PCR fragment only when there is no T-DNA insertion in the genome, while the LB3 and RP primers produce the ~0.6 kb fragment only when T-DNA is present in the genome. (E) Immunoblot analysis using T-DNA insertional mutant lines to check their protein expression level. The protein extracts prepared from wild-type, homozygous and heterozygous Arabidopsis plants were immunoblotted with anti-ASDH, anti-SAHH, anti-ADK antibodies as indicated. Asterisks indicate the weak protein bands of CMT (B) and ASDH (E). Twenty micrograms of protein extracts from wild type and T-DNA insertion lines were used for immunoblot analyses. Molecular masses are indicated on the left in kilodaltons (kDa). Chr1, Arabidopsis chromosome 1; chr3, Arabidopsis chromosome 3; H, homozygote; h, heterozygote; W, wild-type plants; M, size marker; WT, Protein extracts obtained from wild-type Arabidopsis plants.


A
LB3
LP
RP
chr3
CMT (CS846178)
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 M
LB3 + RP
LP + RP
w h h w w h h w h h h h w w w w
B
CS846178
WT
w w
heterozygous
55
43
34
(kDa)
*
CMT
blot: anti-CMT
C
WT
35S::CMT
43
34
26
(kDa)
CMT
APT
blot: anti-CMT; anti-APT
D
LB3
LP
RP
chr1
ASDH (CS813591)
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 M
LB3 + RP
LP + RP
w w w w H w w w w w w w w w w h w w w w h w w h
E
CS813591
WT w H h h h
(kDa)
43
34
55
43
*
ASDH
SAHH
ADK
blot: anti-ASDH; anti-SAHH; anti-ADK

3.12. Creation of artificial microRNA silenced lines for ADK and SAHH

Lines deficient in either ADK or SAHH were generated using an artificial microRNA (amiRNA)-based gene silencing method to evaluate their functional significance in plant development. Although the sADK lines, created by overexpression of the ADK1 cDNA in the sense orientation, have been well characterized, and their phenotype is well documented (Moffatt et al., 2002), it was important to verify that off-target silencing was not contributing to this phenotype. For this, amiRNA should be more effective since it provides target-specific gene silencing (Schwab et al., 2005).

AmiRNAs have become a widely used tool for silencing genes of interest in plants (Schwab et al., 2006). The small miRNAs (21-24 nucleotide) target a complementary single-stranded mRNA, ultimately resulting in degradation of target mRNA or inhibition of translation (Ossowski et al., 2008).

Specific amiRNAs targeting both *ADK1* and *ADK2* transcripts (amiADK) or both *SAHH1* and *SAHH2* transcripts (amiSAHH) were designed using an online artificial microRNA designing tool (WMD2 Web MicroRNA Designer; http://wmd2.weigelworld.org). Currently, WMD version 3 is released (http://wmd3.weigelworld.org). To design specific 21-mer amiRNA sequences for amiADK and amiSAHH, either *ADK1*/*ADK2* or *SAHH1*/*SAHH2* genes were used as input target genes in the MicroRNA Designer. The output of this online tool provided a list of possible amiRNAs sequences that would potentially be effective in silencing *ADK* or *SAHH* genes. However, since there is no guarantee that any of the amiRNA candidate

sequences will work properly, two amiRNAs in a different region of *ADK* (amiADK#1 and amiADK#2) and *SAHH* (amiSAHH#1 and amiSAHH#2) were chosen for each (see Materials and Methods for details). These amiRNAs were constructed by PCR using a pRS300-based vector (Schwab et al., 2005) containing a miR319a precursor sequence as the PCR template and four primers predicted by WMD2. These PCR reactions allow to replace the endogenous miRNA and miRNA* sequences by the 21mers of amiADK and amiADK*. Approximately 400 bp of PCR-amplified amiRNAs were expressed under the control of 35S promoter in Arabidopsis plants. A total of seven different transgenic lines expressing amiRNAs as followings were generated: amiADK#1, amiADK#2, amiADK#1 + amiADK#2, amiSAHH#1, amiSAHH#2, amiSAHH#1 + amiSAHH#2, and amiADK#1 + amiSAHH#2.

No plants displaying a distinct phenotype were obtained from transgenic lines expressing amiADK#2 or amiSAHH#1, whereas the lines expressing amiADK#1, amiADK#1 + amiADK#2, amiSAHH#2, or amiSAHH#1 + amiSAHH#2 showed clear mutant phenotypes (Figure 14 and 15). These results suggested that at least the amiRNAs amiADK#1 and amiSAHH#2 are effective in silencing of ADK and SAHH, respectively, although it is not clear whether amiADK#2 and amiSAHH#1 are not functional for gene silencing. Therefore, the lines expressing amiADK#1 or amiSAHH#2 were followed for further analysis.

To establish the plant phenotype of *adk*, amiADK plants expressing 35S::amiADK#1 along with wild-type plants were monitored for their growth and development from germination to senescence (Figure 14). A total of fifteen individual T1-transgenic plants were selected by BASTA spraying. Although they showed a narrow range of phenotypes,

they exhibited similar patterns of abnormal morphology compared to wild-type plants. Overall, amiADK plants showed delayed development throughout the whole growth and development period. Wrinkled leaves and delayed primary shoot development were observed in 3 to 4-week old amiADK lines compared to wild-type plants (Figure 14A and F, upper box). The delayed primary shoot development of amiADK plants ultimately makes the mature plants small (approximately 5 to 10-fold smaller than wild type) (Figure 14B). Floral bud and silique formation were also delayed (Figure 14B and C), and both were clustered as shown in Figure 14C (a, b) for clustered inflorescences and Figure 14F (lower box) for the clustered siliques. The mature siliques of amiADK were much smaller than those of wild type and produced fewer seed than wild-type plants (Figure 14F). Furthermore, the lines showed delayed senescence as they still retain the green rosette and cauline leaves and flowers at 13 weeks, long after the wild type had died (Figure 14D and E). In this late growth stage, they were still small but displayed bushy phenotypes with a number of secondary branches (Figure 14D and E).

Plant phenotype of *sahh* mutants was also examined using amiSAHH plants expressing 35S::amiSAHH#2 compared to wild-type plants throughout the plant development (Figure 15). A range of phenotypes were observed amongst twenty individual T1-transgenic plants. They can be divided into four groups based on their different phenotypic traits: five plants had a defect in primary shoot formation and died by 12-13 weeks (data not shown); one plant with single primary shoot which failed to form siliques (Figure 15A-D); four plants had clustered buds and siliques (Figure 15E and F) and 10 normal-looking plants (data not shown). Overall, the plants displaying

abnormal morphologies showed delayed primary shoot development and leaf senescence compared to those of wild-type plants.

The phenotype of the first group of five plants showing the defect in primary shoot formation was almost identical to the second group of 1 mutant plant. The only difference is whether they form a primary shoot or not. Both first and second groups of plants exhibited abnormal leaf development with a number of rosette leaves without the primary shoot formation until they were 9-week-old (Figure 15A), the colour of rosette leaves started to change from green to purple by 7-8 weeks, and they started to senesce by 8-9 weeks. The second group of plants formed the primary shoot in 9 weeks, and few bud-like structures formed on top of the shoot by 10 weeks (Figure 15B and C); this is in comparison to wild-type plants that started shoot development at 2.5-3 weeks, under these growth conditions. By week 12-13 the bud-like structures started to dry and the plant ultimately died with the failure of seed set at 19 weeks (Figure 15D). Interestingly, the third group of amiSAHH plants displayed similar phenotypes to those previously observed in amiADK plants including clustered buds and siliques (Figure 15E and F). These plants were also smaller than wild-type plants and displayed delayed senescence (Figure 15E and F).

To analyze protein expression levels in these lines, immunoblot analyses were performed using total protein extracts prepared from leaf tissue (Figure 16). Figure 16A shows the substantially reduced ADK abundance in all homozygous amiADK plants examined compared to wild-type plants; there was no change in SAHH protein levels in both wild-type and amiADK plants (Figure 16A). The SAHH protein levels of the plants with the defect in primary shoot formation (first group) were not detectable or very low,

and those of normal-looking plants (fourth group) were slightly lower than those in wild-type plants (Figure 16B). However, ADK protein levels were not affected by reduced SAHH levels in any of the amiSAHH plants examined (Figure 16B). The SAHH protein levels of 5-week-old amiSAHH plant (Figure 15A, second group) were significantly decreased and almost no protein was detected in 10 weeks (Figure 15B and C, Figure 16C). The other amiSAHH lines displaying clustered buds and siliques (Figure 15E and F, third group) also showed reduced SAHH protein levels with no change in ADK and APT protein levels (data not shown).

In addition, amiRNA mutant plants co-expressing 35S::amiADK#1 and 35S::amiSAHH#2 were generated to examine *adk*/*sahh* double knock-outs. Some of these T1-transgenic plants showed decreased protein levels of ADK, SAHH, or both ADK and SAHH, but none of the plants had an abnormal phenotype (Figure 16D).

Moreover, in an attempt to see whether overexpression of both SAHH isoforms SAHH1/SAHH2 causes any phenotypic changes, both Mas promoter::SAHH1 and Actin promoter::SAHH2 constructs were transformed into Arabidopsis and the SAHH protein levels in these plants were examined by immunoblotting (Figure 16E). Interestingly, these lines showed two different patterns of protein expression levels, one with no (lane 2) or lower (lanes 4, 6, 8, 12) and the other with slightly higher (lanes 3, 5, 6, 11) than those in wild-type plants, but ADK protein levels were not affected by the changes of SAHH levels (Figure 16E). However, none of these plants displayed a distinct phenotype, despite the decrease or increase in SAHH abundance (data not shown).

Finally, a homozygous amiADK plant was crossed with a heterozygous SAHH1-EGFP plant to examine whether the movement of SAHH is affected by suppression of ADK protein levels (Figure 16F). The ADK and EGFP fusion protein levels in the two cross lines (#1 and #2) were confirmed by immunblotting with anti-ADK (Figure 16F, left panel) and anti-GFP (Figure 16F, right panel) antibodies. The ADK protein levels were significantly decreased in both F1 lines, similar to those in amiADK plants (Figure 16F, left panel), whereas SAHH1-EGFP fusion protein expression was detected only in the #2 F1 line with the similar protein levels to those in SAHH1-EGFP plants (Figure 16F, right panel), indicating the successful expression of both amiADK and SAHH1-EGFP in the #2 F1 line. However, no distinct change in SAHH-EGFP movement was observed in the amiADK x SAHH1-EGFP line compared to that in the wild-type background, based on the confocal microscopic examination (data not shown). However this was not quantitatively evaluated.

Figure 14. Phenotype of amiADK plants.

(A) 4-week-old wild-type plant and a transgenic Arabidopsis plant expressing 35S::amiADK. The boxed image shows an amiADK plant with different angle (B) 6-week-old wild-type plant and a transgenic Arabidopsis plant expressing 35S::amiADK. (C) Enlarged image of the amiADK plant shown in B. a and b show the enlarged images of boxed regions. (D) 9-week-old amiADK plant. The enlarged image represents the boxed region. (E) 13-week-old amiADK plant. (F) The size difference in silique of wild-type and amiADK plants. The boxed images show the wrinkled leaf shape of 3-week-old and clustered siliques of 7-week-old amiADK plants, respectively. amiADK, Artificial microRNA-ADK; WT, wild-type Arabidopsis plants.

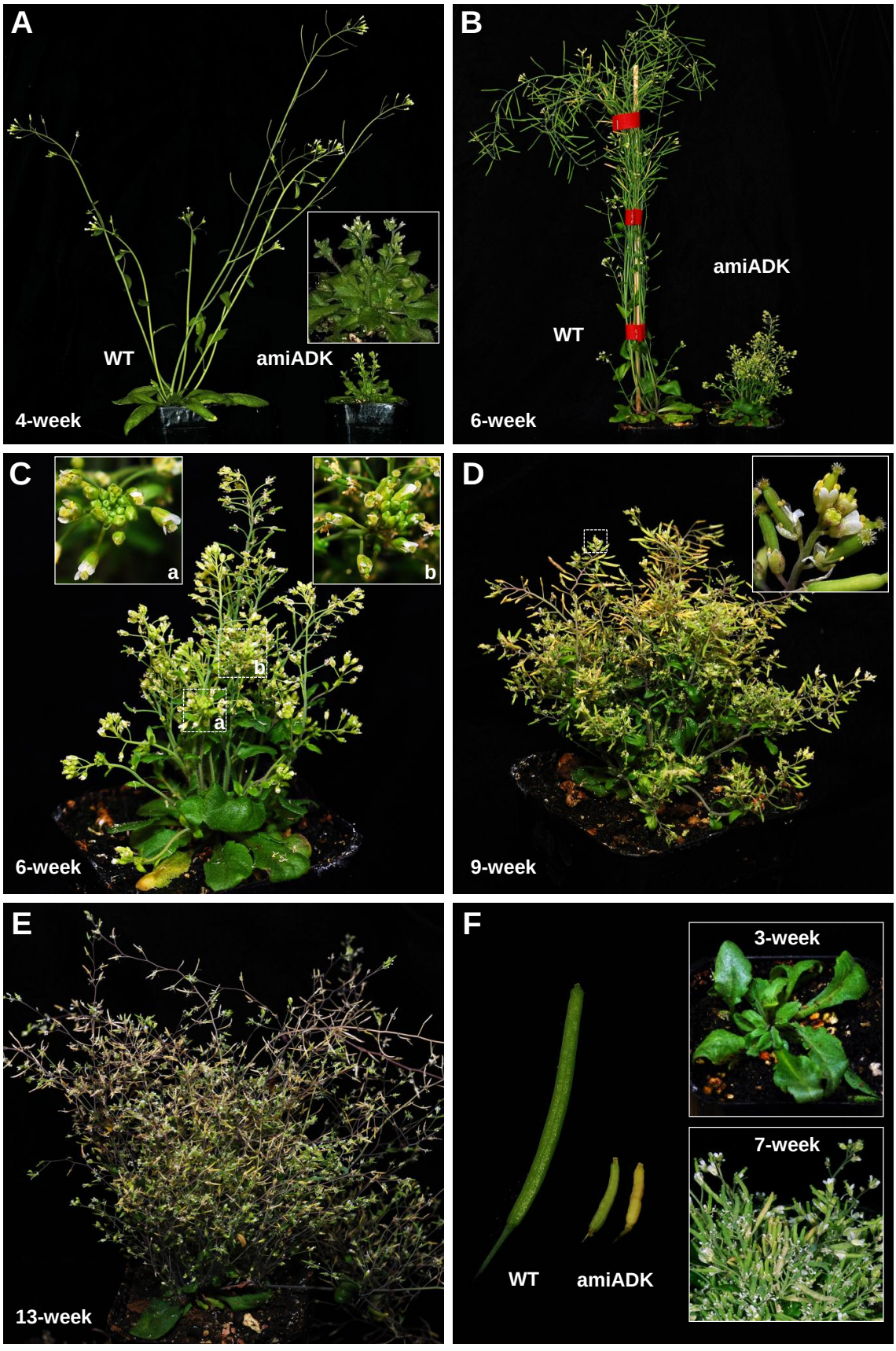

A
WT
amiADK
4-week
B
amiADK
WT
6-week
C
a
b
b
a
6-week
D
9-week
E
13-week
F
WT
amiADK
3-week
7-week

Figure 15. Phenotype of amiSAHH plants.

(A) 5-week-old wild-type plant and a transgenic Arabidopsis plant expressing 35S::amiSAHH (first group). The boxed and enlarged image shows an amiSAHH plant with different angle (B) 10-week-old wild-type plant and a transgenic Arabidopsis plant expressing 35S::amiSAHH (second group). (C) Enlarged image of the amiSAHH plant shown in B. a and b show the enlarged images of boxed regions. (D) 13-week-old amiSAHH plant. The enlarged image represents the boxed region. (E) 10-week-old amiSAHH plant (third group) showing similar phenotype to that of amiADK plants. (F) Enlarged image of the boxed region shown in E. The boxed images show the clustered inflorescences (a) and the similar phenotype to that observed in the plants generated by the overexpression of 35S::SAHH2-EGFP fusion (b). amiSAHH, Artificial microRNA-SAHH; WT, wild-type Arabidopsis plants.

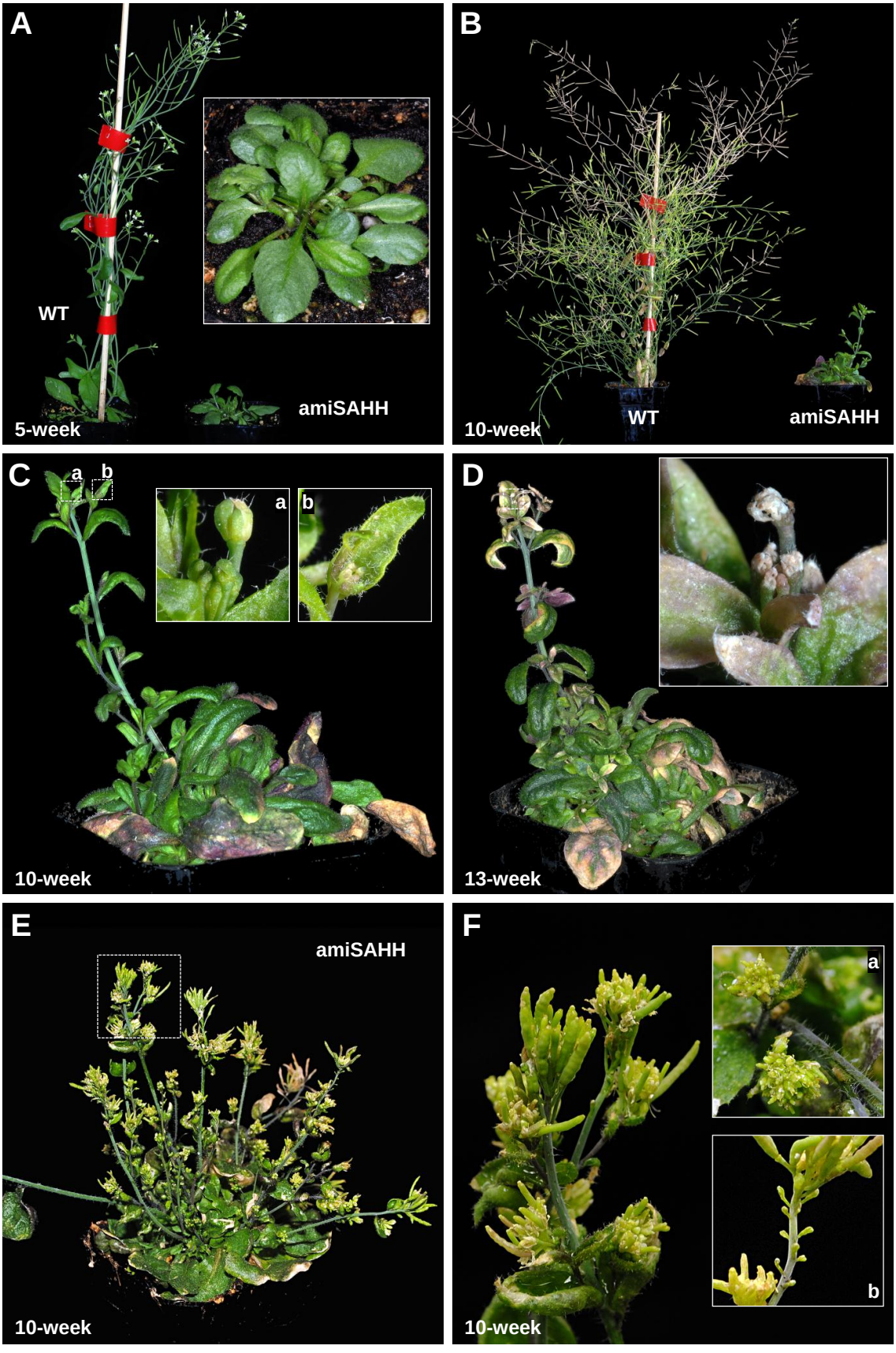
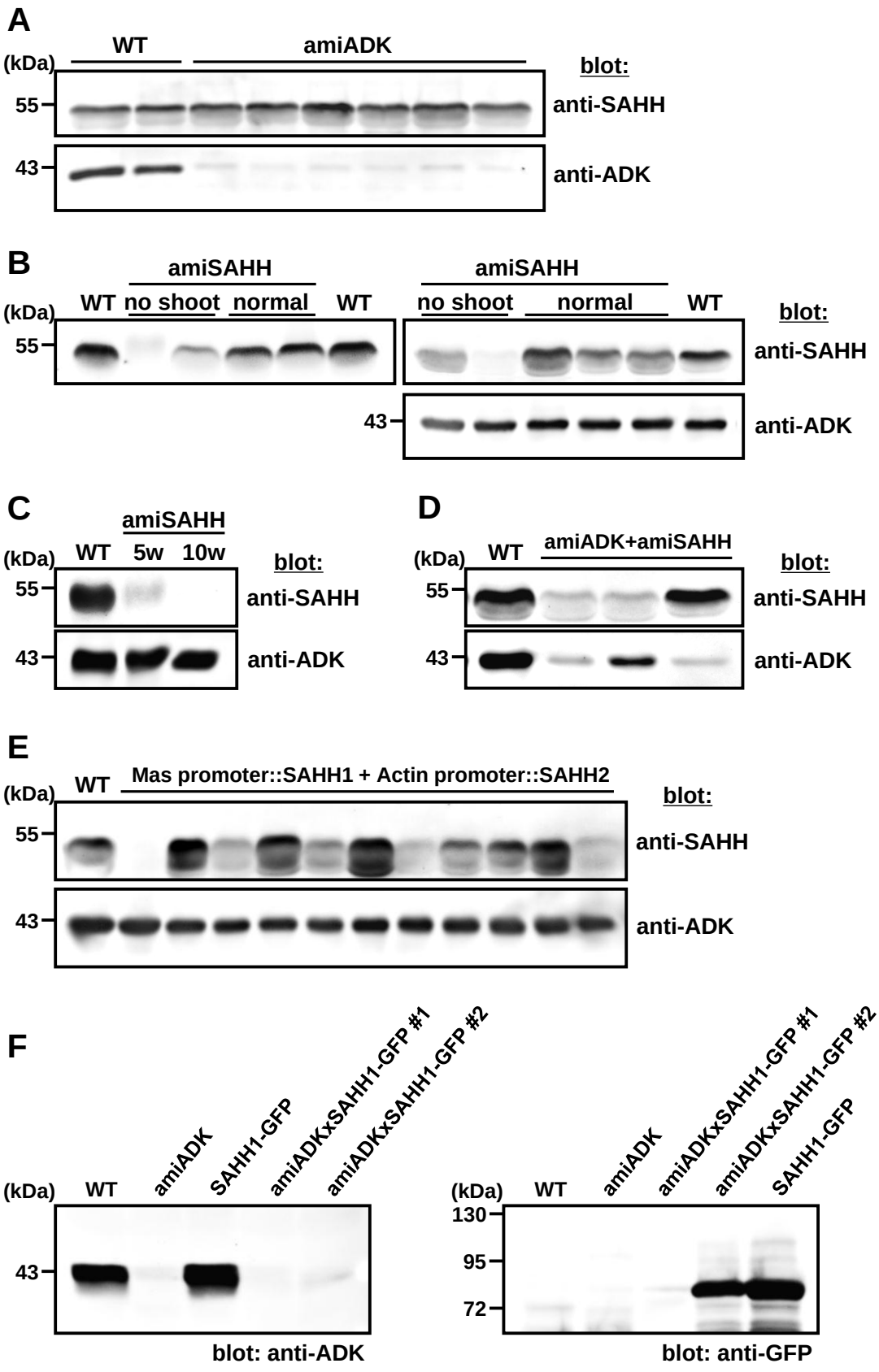

A
WT
amiSAHH
5-week
B
10-week
WT
amiSAHH
C
a
b
a
b
10-week
D
13-week
E
amiSAHH
10-week
F
a
b
10-week

Figure 16. Immunoblot analyses of amiRNA lines.

(A) The extracts prepared from wild-type and homozygous amiADK plants were immunoblotted with antiSAHH and antiADK. (B) The extracts prepared from wild-type and amiSAHH plants were immunoblotted with anti-SAHH (1:5,000) and anti-ADK (1:3,000) antibodies. The 'no shoot' and 'normal' indicate amiSAHH plants showing defects in primary shoot formation (first group) and a normal looking phenotype like wild-type plants (fourth group), respectively. (C) The extracts prepared from wild-type Arabidopsis plants, 5-week-old and 10-week-old amiSAHH plants were immunoblotted with antiSAHH and antiADK antibodies. This plant is identical to that shown in Figure 15A-D (first and second group). (D) The extracts prepared from wild-type and transgenic Arabidopsis plants expressing both amiADK and amiSAHH were immunoblotted with anti-SAHH and anti-ADK antibodies. (E) Immunoblot analysis of transgenic plants expressing both SAHH1 and SAHH2 driven by the Mas and Actin promoters, respectively. The extracts prepared from wild-type and transgenic Arabidopsis plants expressing both Mas promoter::SAHH1 and Actin promoter::SAHH2 were immunoblotted with anti-SAHH and anti-ADK antibodies. (F) Immunoblot analysis to verify a cross line (amiADK x SAHH1-EGFP). The extracts prepared from Arabidopsis plants expressing amiADK or SAHH1-EGFP, and two cross lines (amiADK x SAHH1-EGFP) along with wild-type Arabidopsis plants were immunoblotted with anti-ADK and anti-GFP (1:1,000) antibodies. Twenty micrograms of protein extracts prepared from wild-type and transgenic plants were used for immunoblot analyses. Molecular masses are indicated on the left in kilodaltons (kDa). WT, protein extracts prepared from wild-type Arabidopsis plants.


A
WT
amiADK
(kDa)
blot:
55
anti-SAHH
43
anti-ADK
B
amiSAHH
amiSAHH
WT
no shoot
normal
WT
no shoot
normal
WT
(kDa)
blot:
55
anti-SAHH
43
anti-ADK
C
amiSAHH
WT
5w
10w
(kDa)
blot:
55
anti-SAHH
43
anti-ADK
D
WT
amiADK+amiSAHH
(kDa)
blot:
55
anti-SAHH
43
anti-ADK
E
WT
Mas promoter::SAHH1 + Actin promoter::SAHH2
(kDa)
blot:
55
anti-SAHH
43
anti-ADK
F
WT
amiADK
SAHH1-GFP
amiADKxSAHH1-GFP #1
amiADKxSAHH1-GFP #2
(kDa)
43
blot: anti-ADK
WT
amiADK
amiADKxSAHH1-GFP #1
amiADKxSAHH1-GFP #2
SAHH1-GFP
(kDa)
130
95
72
blot: anti-GFP

4. Discussion and conclusions

4.1. ADK interacts with SAHH in planta

I have previously demonstrated that in Arabidopsis and tobacco, ADK and SAHH are localized to the same subcellular compartments including the nucleus, cytoplasm, and possibly plastids too (Chapter 1). These observations led me to test the hypothesis that this targeting is mediated by a specific protein-protein interaction. In this study, several assays including co-immunoprecipication, BiFC, and pull-down using His- or GST-tagged proteins demonstrated an interaction between ADK and SAHH in planta. Moreover, the results of co-immunoprecipitation using transgenic lines expressing GFP fusions and BiFC using various combinations revealed that the interaction of ADK and SAHH is mediated by specific regions of SAHH. The deletion of 41 amino acids in the middle of SAHH1 (IR; Gly^{150}-Lys^{190}), a region that was previously identified as a crucial region for its nuclear targeting in the Chapter 1, disrupted SAHH's interaction with ADK (Figure 2C and 3C). Furthermore, the addition of EGFP (~27 kDa) or C-terminal fragment of YFP (~9 kDa including linker residues) at the N-terminus of SAHH1 also interferes with this interaction but does not completely eliminate it, as shown in Figure 2C and 3B. In contrast, the addition of a protein tag to the N-terminus of SAHH1 did not interrupt its tetramer formation based on the yeast two-hybrid and BiFC results. Both prey and bait SAHHs, which were fused with a DNA-binding domain (~22.3 kDa) or a transcription activation domain (~11.4 kDa) at the N-terminus of each SAHH1, produced strong interaction signals (Appendix IV). These results indicated that ADK and SAHH could form a complex and this is mainly mediated by the surface-exposed IR of SAHH, indicating another role of IR as an interacting domain. This interaction result is consistent

with the docking model between Plasmodium SAHH and human ADK which shows the best interaction through the IR of SAHH (Appendix II).

The close association between ADK and SAHH has been evidenced by several previous studies. Overall phenotypic traits of ADK-deficient plants (Moffatt et al., 2002) are similar to those observed in *SAHH*-silenced tobacco plants (Tanaka et al., 1997), suggesting that ADK and SAHH deficiencies affect similar physiological and developmental processes. Moreover, sADK plants have decreased SAHH activity due to an accumulation of Ado which causes the reversible SAHH reaction to flow in the synthesis direction (Moffatt et al., 2002). Earlier work performed in the Moffatt lab revealed that both SAHH and ADK are often detected in the same tissues of Arabidopsis plants, with their transcript abundance and enzyme activities being co-ordinately regulated during plant development (Pereira et al., 2007; Schoor, 2007). Together, all these previous findings support a collaborative relationship between these two key enzymes in plants to sustain SAM-dependent transmethylation reactions. However, the functional importance and exact roles of their interaction remain to be elucidated.

4.2. Dynamic interaction of ADK and SAHH

The interaction between ADK and SAHH documented here could be either dynamic or indirect due to following reasons. This interaction was obtained only when concentrated plant extracts (1.5-2.0 mg/ml) was used for the immunoprecipitation assay (Figure 1B and C), whereas it was barely detected when lower amounts of plant protein (<1.5 mg/ml) were used (data not shown). Moreover, no interaction was found in the yeast two-hybrid

assay (Appendix IV) and in the pull-down assays in which only the two recombinant proteins are present (data not shown). These results suggest that either the ADK/SAHH complex requires other protein components that are not present with recombinant proteins or in yeast, or perhaps specific post-translational modifications are required for this interaction.

There have been several reported or predicted post-translational modifications of ADK and SAHH, which regulate many different processes during plant development. For instance, SAHH is identified as one of candidates for protein S-nitrosylation, the covalent attachment of nitric oxide (NO; a small, reactive signaling molecule) to the thiol group of cysteine residues. This modification was detected in SAHH recovered from Arabidopsis cell cultures and leaves, by the biotin switch method in combination with nano liquid chromatography and tandem mass spectrometry (nanoLC/MS/MS) analysis (Lindermayr et al., 2005). Interestingly SAHH activity is inhibited by NO *in vitro*, suggesting the possibility that S-nitrosylation may regulate transmethylation reactions by inhibiting the activity of SAHH (Christian Lindermayr; personal communication). Recently, another nitric oxide mediated post-translational modification was reported in sunflower (*Helianthus annuus* L.) hypocotyls (Chaki et al., 2009). In this research, the *in vitro* analyses using a combination of immunological and proteomic methods demonstrate that SAHH is a potential target for the protein tyrosine nitration, that occurs under oxidative stress conditions by the addition of a nitro ($-NO_2$) group to an aromatic ring carbon of tyrosine residues. The result of nitration reaction test by exposing the hypocotyl to a nitrating agent (SIN-1; a peroxynitrite donor) shows that 49-94% of SAHH activity was inhibited by various concentrations (0.5-5 mM) of SIN-1 (Chaki et al., 2009). *In vitro*

phosphorylation assays demonstrate that ADK1 is phosphorylated by the Arabidopsis sucrose nonfermenting-related kinase (SnRK2.8) that is down-regulated when plants are deprived of nutrients and growth is reduced (Shin et al., 2007). In addition, the segment VKAE$^{151\text{-}154}$ within the IR of SAHH is predicted as a potential sumoylation site by the ELM server (http://elm.eu.org/) (Chapter 1). Sumoylation plays a crucial role in various cellular processes by altering protein activities, changing subcellular localization, or providing a platform for protein-protein interactions (Geiss-Friedlander and Melchior, 2007). The segment EKTGQVPDPTSTDN$^{158\text{-}171}$ is predicted to be phosphopeptide binding domain by the ELM server. Within this segment, both Ser^{168} and Thr^{169} residues are potential phosphorylation sites (http://www.cbs.dtu.dk/services/NetPhos/) (Chapter 1). All these data suggest that ADK/SAHH complex could possibly be a target for specific post-translational modifications such as S-nitrosylation, tyrosine nitration, sumoylation and phosphorylation. The validity of these predictions and their functional significance remain to be elucidated.

4.3. CMT may direct ADK and SAHH to the nucleus via protein-protein interactions

In the previous chapter of this thesis, it is proposed that the nuclear localization of ADK and SAHH, each which has no conventional NLS sequence, is mediated by a protein-protein interaction with other nuclear targeting proteins (Chapter 1). I tested this hypothesis using one MT that is specifically targeted to the nucleus. I isolated a cDNA encoding Arabidopsis mRNA cap (guanine-7-) methyltransferase, CMT, and confirmed it

is a highly conserved member of the SAM-dependent MT family (Figure 4). EGFP fusions with this coding sequence showed that CMT is predominantly localized to the nucleus (Figure 5A-H), and its NLS sequence may be located in the N-terminal region of CMT (Figure 5K). The binding assay between CMT and GST-Impα demonstrated that the nuclear localization of CMT is an NLS-dependent nuclear protein import (Figure 7A). Interaction of CMT with both ADK and SAHH was confirmed by GST pull-down (Figure 7B), immunoprecipitation (Figure 7C) and BiFC (Figure 7D-I).

These results supported the hypothesis that ADK and SAHH could be directed to the nucleus via the interaction with CMT through an NLS-mediated mechanism. This is consistent with the findings of Radomski et al. (2002), who showed that the translocation of SAHH from the cytoplasm to the nucleus is mediated by a protein-protein interaction with CMT (xCMT) in *Xenopus laevis* (Radomski et al., 2002). However, in *Xenopus* the 55 amino acids of N-terminal part of xSAHH are sufficient for binding to xCMT (Radomski et al., 2002), whereas in Arabidopsis, the IR of SAHH was required for this interaction (Figure 7C and F). This observation is correlated with earlier findings of the deletion analysis showing that IR-deleted SAHH failed to localize to the nucleus (Chapter 1). Taken together, these results indicate that in Arabidopsis CMT is capable of directing cytoplasmic ADK and SAHH to the nucleus by a protein interaction, and moreover its interaction with SAHH is mediated by the IR.

Interaction of MTs with another enzyme(s) involved in SAM cycle was suggested by Shen et al. (2002) based on their study of Arabidopsis *S*-adenosyl-L-methionine synthetase (SAMS; EC 2.5.1.6) mutant plants. These mutants have decreased SAMS activity and SAM concentration, as well as 22% decreased lignin content (Shen et al.,

2002). These authors proposed that SAMS may associate with a MT involved in lignin biosynthesis that requires large amounts of SAM, and that this association for metabolite channeling purposes could allow more efficient methylation of lignin (Shen et al., 2002). Another study using an Arabidopsis *SAHH* missense mutant shows that reduced SAHH activity resulted in a decrease in genomic cytosine (non-CG) methylation that is controlled by CMT3 DNA MT (DNA (cytosine-5-)-methyltransferase; EC 2.1.1.37) (Mull et al., 2006). More recently, a study examining the same Arabidopsis SAHH mutant and a specific SAHH inhibitor, DHPA, demonstrated that SAHH is an important regulator of chromatin remodeling by DNA/histone methylation (Baubec et al., 2010). In addition, both pectin methylation and DNA methylation are reduced in sADK lines, suggesting that the wavy leaf morphology of the sADK may be due to altered pectin methyl-esterification (Pereira et al., 2006). Taken together, these observations suggest possible interactions between SAM recycling enzymes and various MTs for efficient transmethylation reactions throughout plant development. In addition, it is possible that other enzymes involved in SAM cycle may also be associated in the ADK:SAHH complex, since APT1 is shown to interact with both ADK and SAHH (Figure 12).

4.4. The interaction with ASDH is required for plastid localization of SAHH

I used an affinity purification strategy to identity other interactors of SAHH. As a result of this experiment I focused on aspartate-semialdehyde dehydrogenase (ASDH) as a novel interactor of SAHH (Figure 8B). ASDH catalyzes the second step of the

biosynthesis pathway of four essential amino acids (isoleucine, lysine, methionine, and threonine) derived from aspartate (Appendix V). This aspartate-derived amino acid is essential for animal diets, since this pathway does not exist in mammals (Jander and Joshi, 2010).

However, despite of the importance of this pathway, ASDH has received little attention from researchers and has not been well studied in plants. Three decades ago, the activity of ASDH was assessed in the callus and suspension culture extracts of maize: it is inhibited by methionine and slightly reduced by lysine and threonine (Gengenbach et al., 1978). The only detailed study of plant ASDH is a report on the cloning and purification of an Arabidopsis ASDH based on its similarity to microbial enzymes (Paris et al., 2002). The Arabidopsis ASDH that was overproduced and purified from *E. coli* was shown to be a homodimer composed of 36 kDa subunits, and exhibit ASDH enzyme activity (Paris et al., 2002).

Many enzymes involved in lysine and threonine synthesis are known to be localized in plastids with N-terminal transit peptides, including aspartate kinase (EC 2.7.2.4)(Wallsgrove et al., 1983), homoserine kinase (EC 2.7.1.39)(Lee and Leustek, 1999) and threonine synthase (EC 4.2.99.2)(Curien et al., 1996), whereas methionine and SAM appear to be synthesized in the cytosol (Galili, 1995). Sequence analysis of ASDH predicted that it contains a putative transit peptide (amino acid residues 1-35), suggesting its possible localization to plastids (Figure 8C). The localization study using EGFP fusion proteins showed that ASDH is targeted to the chloroplasts as expected (Figure 9). In addition, surprisingly the stromule-like GFP signals emanating from the chloroplast were often visualized in these chloroplasts or in the connection between chloroplasts (Figure

9M-O; arrows). Although the exact mechanisms of their formation and dynamics are yet unclear, stromules, stroma-filled tubular extensions of the plastid envelope membrane, are known to have several different roles (Hanson and Sattarzadeh, 2008). Since stromules increase the surface area of the plastid envelope and allow close proximity between plastids and other organelles, they may facilitate the exchange or transfer of molecules between plastids or plastids and the cytoplasm/other organelles (Hanson and Sattarzadeh, 2008). In addition, it was proposed that stromules may also facilitate plastid-to-nucleus signaling based on the observation that stromules lie within grooves and invaginations of the nuclear envelope (Kwok and Hanson, 2004). Together, the subcellular localization data suggested that chloroplast-specific enzyme ASDH can probably move from one plastid to another, to the cytoplasm, or other organelles. Based on the data reported here, I propose that some of the plastidic ASDH is delivered to the cytosol through stromules and via its strong interaction it brings SAHH in vicinity of the plastid surface. This localization of SAHH would be ideally suited to hydrolyze the SAH released by the SAM/SAH exchanger located in the outer membrane of plastids (Bouvier et al., 2006; Ravanel et al., 2004) (Figure 17). The presence of a SAM/SAH exchanger (or a SAM transporter) in Arabidopsis mitochondria (Palmieri et al., 2006) and plastids (Bouvier et al., 2006) was recently reported. In both cases, the transporters are capable of exchanging SAH produced in plastids or mitochondria for cytosolic SAM thereby maintaining the SAM/SAH ratio within the organelles. Thus, the SAH pumped out via the transporter could possibly be hydrolyzed by SAHH residing near the plastid surface (Figure 17).

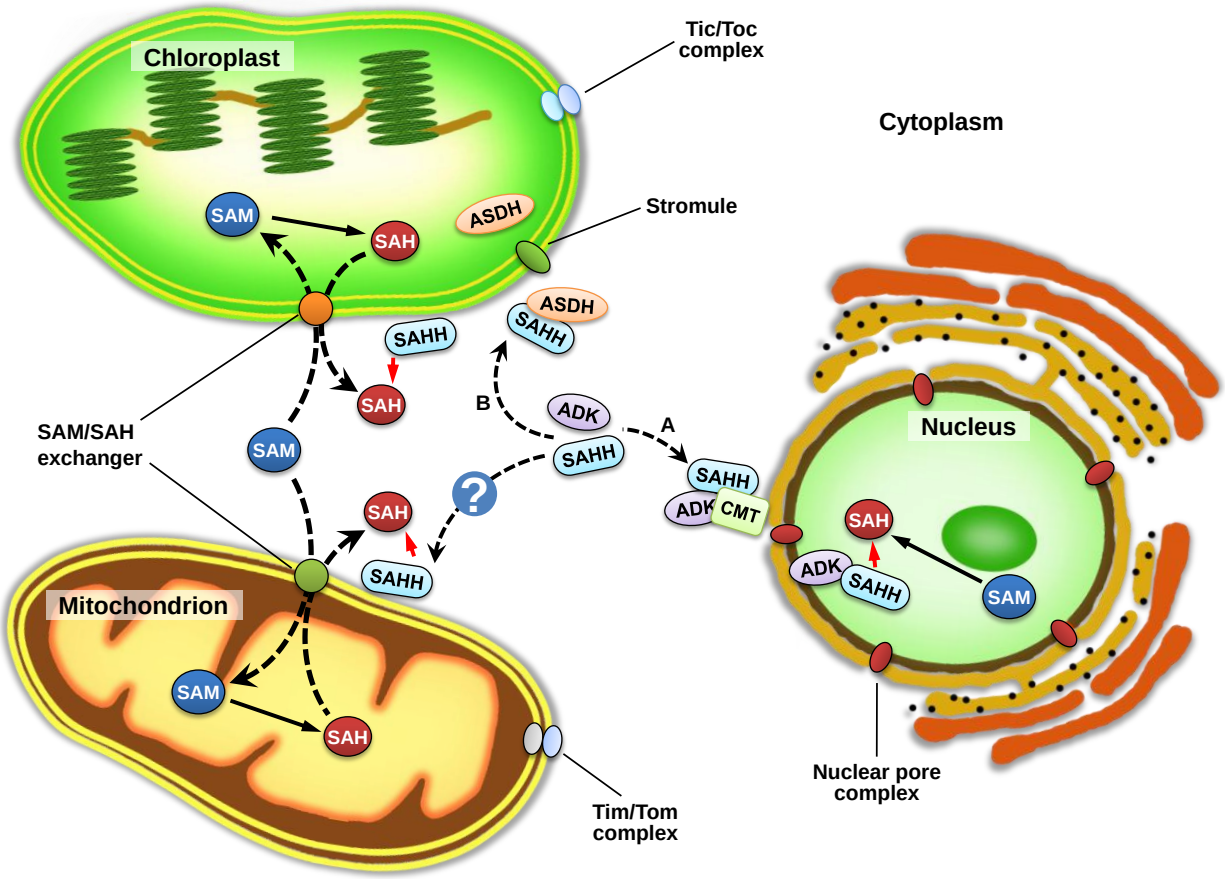
Further affinity purifications using Arabidopsis lines expressing mutated SAHH1-Strep fusion proteins suggested that the interaction between ASDH and SAHH1 is dependent upon the presence of the PxP motif of SAHH1 that is located within the IR and is important for protein channel gating and protein interaction (Labro et al., 2003). However, mutation in potential phosphorylation site TST^{167} of SAHH1 substituted with three alanines, resulted in strong interaction with ASDH, whereas two distinct ASDH were detected by both Coomassie blue staining and immunoblotting (Figure 11C and D). The larger ASDH band might be the product of a post-translational modification of ASDH, since both bands were identified as a mature form of ASDH by LC-MS/MS analysis. However, in this study it is not possible to ascertain why the larger band only binds to the TST^{167} mutated version of SAHH1, whereas it does not bind to intact SAHH1. Moreover, a docking model also predicted that the surface-exposed IR loop which contains PDPTSTDN is to be located at the SAHH:ASDH interface (Andrew Doxey, personal communication; Appendix VI). Although the mechanism of the PDP and TST mediated-interaction could be quite complex, these results at least indicated that the mutation in PDP^{164} of SAHH1 disrupts the IR loop of SAHH significantly enough to disrupt binding, whereas mutating the TST region has a lesser effect. In addition, these results are correlated with earlier findings that Arabidopsis protoplasts transiently expressing the PDP164AAA-EGFP displayed a punctate pattern of GFP fluorescence (Chapter 1). It is possible that the punctate pattern may arise from aggregates of PDP164AAA-EGFP that fail to localize to chloroplasts due to the disruption of binding to ASDH, suggesting that the interaction with ASDH is required for plastid localization of SAHH.

In addition, another possible role of SAHH/ASDH interaction may exist in maintaining the proper balance between methionine and threonine biosynthesis pathways by indirectly regulating SAM. This could have important consequences since SAM can increase threonine synthase by up to 85-fold in Arabidopsis (Laber et al., 1999). This hypothesis is supported by the two facts: first, SAHH is known to be important for maintaining SAM levels or at least the ratio of SAM:SAH (Moffatt and Weretilnyk, 2001); second, ASDH is only one enzyme responsible for the conversion of aspartyl phosphate to aspartate-semialdehyde that is the branch point intermediate between the lysine and threonine/methionine biosynthesis pathways (Jander and Joshi, 2010). Moreover, it is also proposed that a chloroplastic SAM/SAH exchanger not only allows efficient one-carbon metabolism between the cytoplasm and chloroplasts by exchanging metabolites, but also serves as a key element in the regulation of the synthesis of the aspartate-derived amino acids in the chloroplast (Ravanel et al., 2004).

Although it is probably too ambitious to give a conclusion, beside the possible role of the SAHH/ASDH interaction for the plastid localization of SAHH, it may also play another role in the aspartate-derived amino acid biosynthesis pathway.

Figure 17. Working model for ADK and SAHH localization.

The arrows (A and B) represent two possible solutions to maintain methylation reactions throughout the cell. (A) Both ADK and SAHH are localizing to the nucleus that is one of the locations where the SAH is produced. (B) ASDH located in the plastid is perhaps delivered to the cytoplasm through the stromules and it brings sufficient SAHH near the plastid via a strong interaction. Thus, the plastidic SAH released into the cytoplasm via a SAM/SAH exchanger could possibly be metabolized in vicinity of the plastid surface. Red arrows indicate the hydrolysis of SAH released from a SAM transporter of each organelle. ADK, adenosine kinase; ASDH, ; CMT, mRNA cap methyltransferase; SAH, *S*-adenosylhomocysteine; SAHH, *S*-adenosylhomocysteine hydrolase; SAM, *S*-adenosylmethionine; Tic, translocon at the inner envelope membrane of chloroplasts; Tim, translocator at the inner membrane of mitochondria; Toc, translocon at the outer envelope membrane of chloroplasts; Tom, translocator at the outer membrane of mitochondria.


Chloroplast
Tic/Toc complex
Cytoplasm
Stromule
SAM
SAH
ASDH
SAHH
SAM/SAH exchanger
B
ADK
A
Nucleus
CMT
?
Mitochondrion
Nuclear pore complex
Tim/Tom complex

SUMMARY

The study outlined in this thesis focuses on the subcellular localization of methyl-recycling enzymes and the mechanism of their movement between compartments. In order to determine the location of ADK and SAHH in a plant cell, various fluorescent protein fusion constructs were created and expressed in tobacco and Arabidopsis, and imaged by laser scanning confocal microscopy. This work demonstrates that both ADK and SAHH are capable of localizing at least to the cytoplasm and the nucleus, and possibly to the chloroplast as well. Moreover, the deletion analysis of SAHH reveals that the 41 amino acids of the IR (Gly^{150}-Lys^{190}), which is only present in plants and parasitic protozoan SAHHs among eukaryotes, is crucial for nuclear targeting of SAHH. Further mutational analysis of SAHH within the IR indicates that Pro^{164}-Asp^{165}-Pro^{166} (potential PxP motif) may be the key residues for SAHH localization in a plant cell. However, the IR fused to a heterologous protein (2xGFP) was localized to almost everywhere in the cell, indicating that the IR is not an autonomous localization signal, but it may rather serve as an interaction domain to be associated with other proteins that can direct SAHH to the destination. The crystal structures of SAHH for *Plasmodium falciparum* and *Mycobacterium tuberculosis* show that the IR appears to be exposed outside of each monomer and separated from the main body of the tetramer, suggesting that the IR can quite possibly be accessible as a binding epitope for interactions with other proteins.

The second part of my research involved examining protein interactions using various methods including co-immunoprecipitation, BiFC, yeast two-hybrids, and pull-down assays using purified recombinant His- and GST-tagged proteins or Arabidopsis plant extracts expressing Strep fusion proteins. These experiments demonstrate that ADK and

SAHH possibly interact in Arabidopsis, suggesting a collaborative relationship of the two enzymes to maintain MT activities throughout plant development. However, this interaction could only be detected in living plants cells, or in concentrated plant extracts *in vitro*; no interaction is observed in the yeast two-hybrid assay. These findings suggest that the interaction is either dynamic or indirect, requiring a cofactor/another protein(s) or specific post-translational modifications.

A putative Arabidopsis mRNA cap MT, CMT, required for adding a methyl group to every mRNA synthesized in the nucleus, is a potential candidate to guide SAHH or ADK to the nucleus. The results of EGFP fusion proteins and the pull-down assay using GST-Impα indicate that CMT is a nuclear-targeting protein through the NLS-dependent nuclear import pathway. Moreover, the possible interaction of CMT with both ADK and SAHH is confirmed by several methods, supporting the hypothesis that the nuclear targeting of both enzymes can be mediated by CMT. In addition, affinity purification using Arabidopsis plants expressing SAHH1-Strep reveals a strong interaction of SAHH with ASDH involved in the aspartate-derived amino acid biosynthesis pathway, which is only present in plants and prokaryotes. Expression of the EGFP fusions demonstrates that ASDH is localized to the chloroplasts and the stromule-like structures that emanate from the chloroplast, suggesting that ASDH can possibly move from plastids to other compartments through the stromule. The strong interaction between SAHH and ASDH leads me to speculate that their association may be required for plastid-targeting of SAHH. Moreover, this interaction may also contribute to the balance between methionine and threonine biosynthesis by regulating SAM concentration that is known to trigger threonine synthase activity.

In addition, the surface-exposed IR of SAHH appears to be a key element for binding to all three proteins presented in this study. Moreover, the mutation in Pro^{164}-Asp^{165}-Pro^{166} of SAHH1 disrupts the binding to ASDH, indicating that the PxP motif is possibly crucial for protein interactions. This IR-mediated interaction might also involve phosphorylation at the Ser^{168} and Thr^{169} residues. The IR loop seems to be evolutionary important residues and is most likely to be a plant specific segment among eukaryotic SAHHs. While the IR is found in many prokaryotes including cyanobacteria and proteobacteria, which are endosymbiotic origins of eukaryotic organelles, most other eukaryotes appear to have lost the segment during the divergence of their lineages. This leads to the hypothesis that the IR may play an important role in interacting with plant specific enzymes such as ASDH. It remains to be investigated whether these interactions allow the formation of a multiple-protein complex including other enzymes involved in SAM cycle.

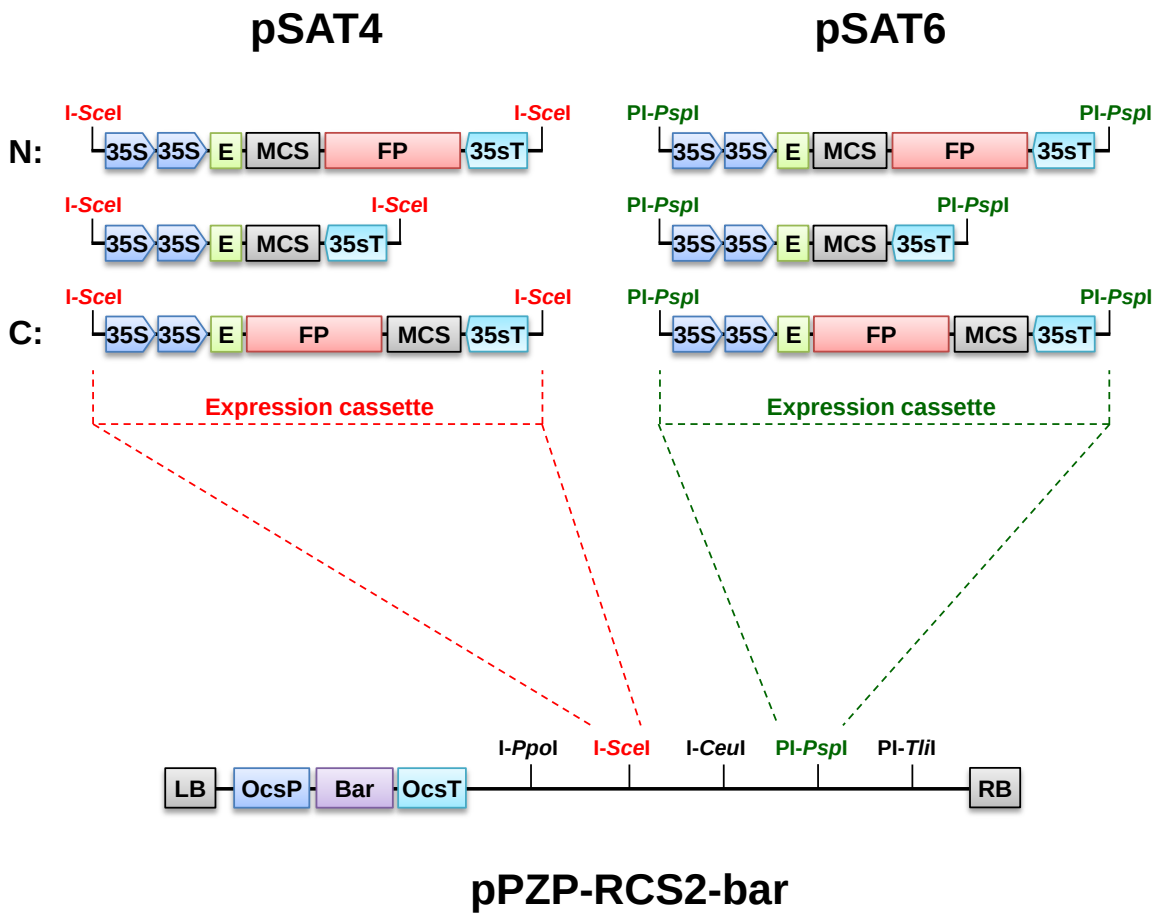
All the results presented in this thesis lead to the following conclusions concerning the two scenarios for both enzymes ADK and SAHH to accomplish SAH metabolism throughout the cell to maintain the MT activities. First, it is clearly shown that ADK and SAHH reside in the nucleus, one of the locations where the SAH is produced. Secondly, the localization of ADK and SAHH in the cytoplasm and the surface of chloroplasts supports the hypothesis that the plastidic SAH released into the cytoplasm via a SAM/SAH exchanger could possibly be metabolized in vicinity of the plastid surface where ADK and SAHH may reside.

To sum up, this thesis first reports a novel interaction between a constitutive and fundamental enzyme, SAHH and the essential enzyme in plants, ASDH that does not

exist in other eukaryotes. This interaction provides further insight into how proteins are targeted to multiple locations to meet plant needs for growth and development. Future research will have to evaluate the functional significance of CMT and ASDH in plant development, and demonstrate whether these proteins really effect the subcellular localization of ADK and SAHH. This can be done by using CMT and ASDH mutant plants or using specific inhibitors for ADK or SAHH. Moreover, it remains to be determined what other factor(s) are involved in these interactions.

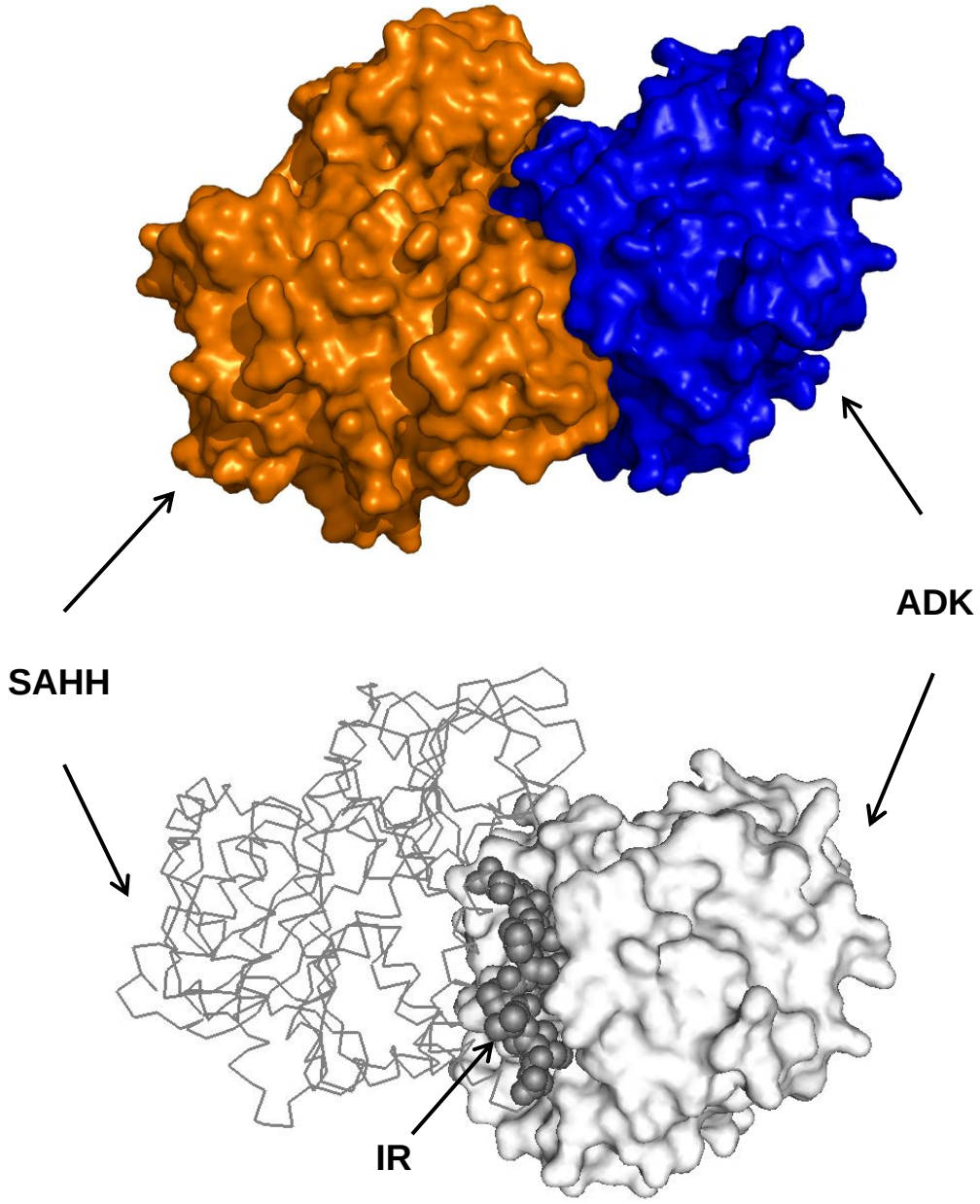
APPENDICES

Appendix I. The pSAT vector system



35S, the cauliflower mosaic virus 35S promoter; 35sT, the cauliflower mosaic virus 35S terminator; Bar, the herbicide bialaphos resistance gene (*bar*); E, translational enhancer (5' UTR from tobacco etch virus); FP, fluorescent proteins; LB, left border; MCS, multiple cloning site; OcsP, the octopine synthase promoter; OcsT, the octopine synthase terminator; RB, right border. MCS of pSAT and pSAT-N vectors contains following restriction enzyme cutting sites: *Nco*I, *Bgl*II, *Xho*I, *Sac*I, *Hin*dIII, *Eco*RI, *Sal*I, *Kpn*I, *Sac*I, *Apa*I, *Xma*I, *Bam*HI. MCS of pSAT-C vectors contains same cutting sites as in the pSAT-N vectors, except for *Nco*I site.

Appendix II. Docking model for ADK:SAHH interaction



Computational docking was performed by Andrew Doxey (Stanford University) using ZDOCK (Chen et al., 2003) with the closest homologous structures available in the Protein Data Bank, Plasmodium SAHH monomer (PDB ID 1V8B:A chain) and Human ADK (PDB ID 1BX4).

Appendix III. Purification of ADK1-Strep

(A) Immunoblotting with anti-HA antibody to screen successful transformants expressing 35S::ADK1-HA.Strep.

(B) Optimization of Strep-purification conditions

Left, Protein extracts were incubated with StrepTactin Macroprep beads for 30 minutes and washed 5 times with 0.5 mL wash buffer.

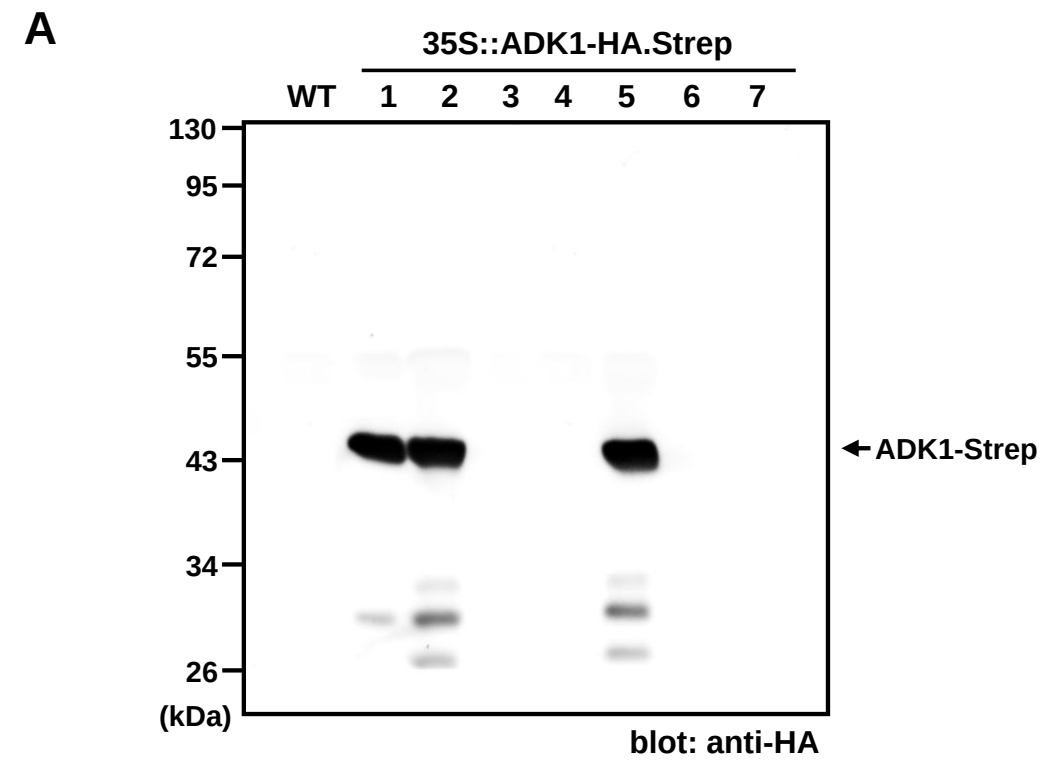
(Extraction buffer) 100mM Tris, 5mM EDTA, 100mM NaCl, 15mM DTT, 0.5% Triton X-100, 1mM PMSF, protease inhibitor cocktail

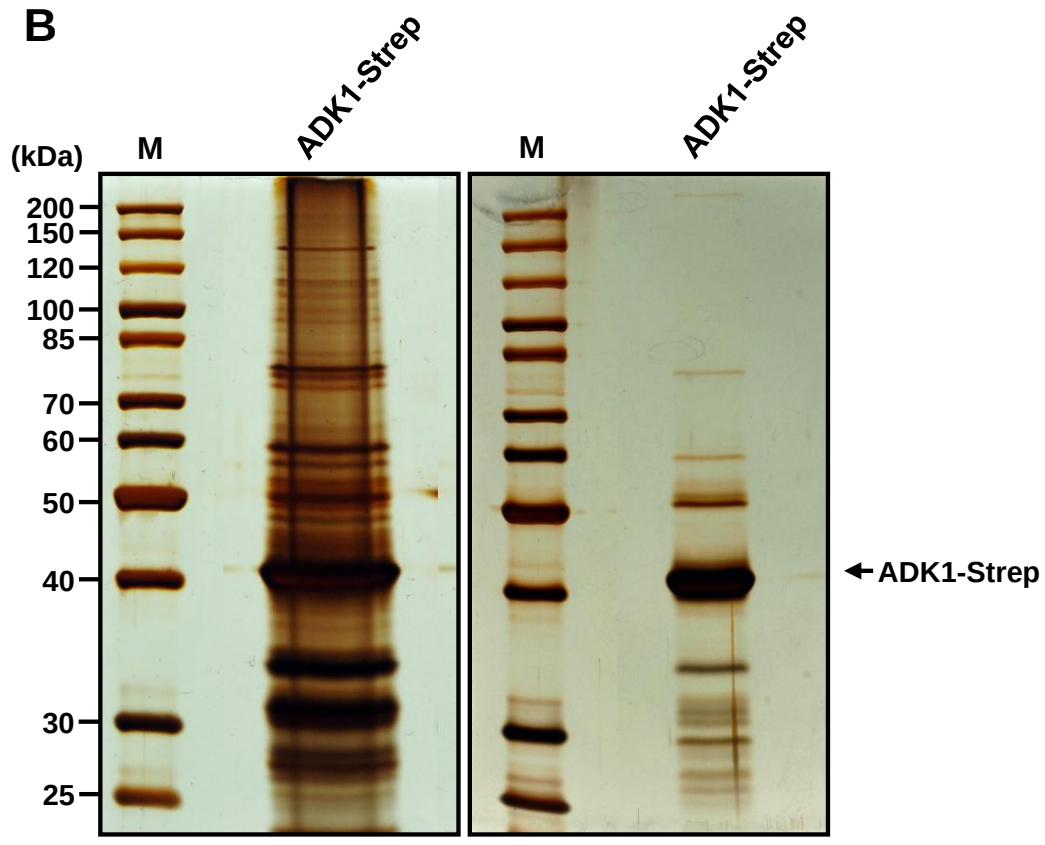
(Wash buffer) 100mM Tris, 0.5mM EDTA pH8.0, 100mM NaCl, 2mM DTT, 0.005% Triton X-100, 1mM PMSF, protease inhibitor cocktail

Right, Protein extracts were incubated with StrepTactin Macroprep beads for 10 minutes and washed 5 times with 0.8 mL wash buffer.

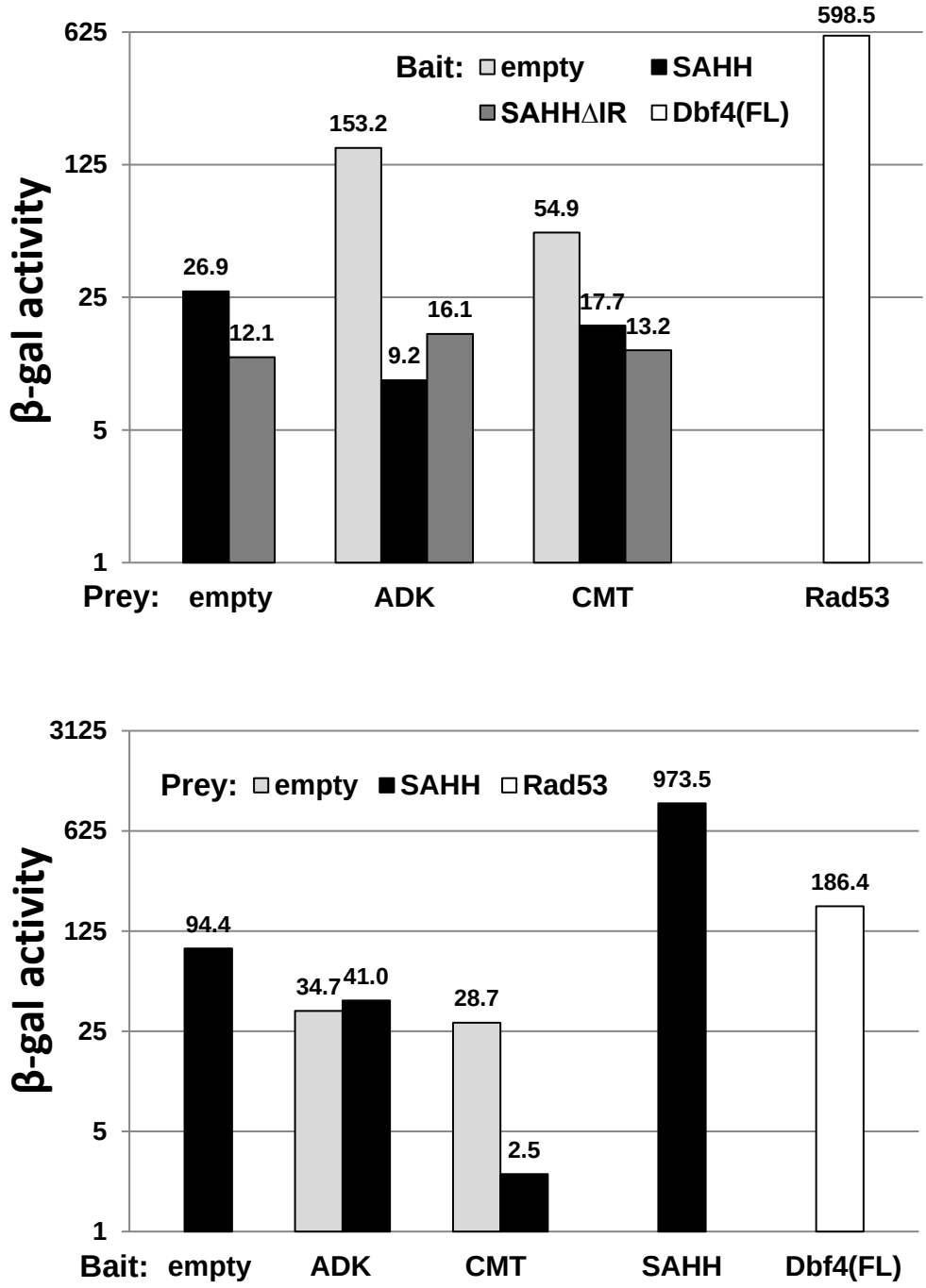
(Extraction buffer) 100 mM Tris, 5 mM EDTA, 200 mM NaCl, 10 mM DTT, 0.5% Triton X-100, 10% glycerol, 1mM PMSF, protease inhibitor cocktail

(Wash buffer) 100 mM Tris, 0.5 mM EDTA, 200 mM NaCl, 2 mM DTT, 0.5% Triton X-100, 10% glycerol, 1mM PMSF, protease inhibitor cocktail


A
35S::ADK1-HA.Strep
WT
1
2
3
4
5
6
7
130
95
72
55
43
34
26
(kDa)
ADK1-Strep
blot: anti-HA

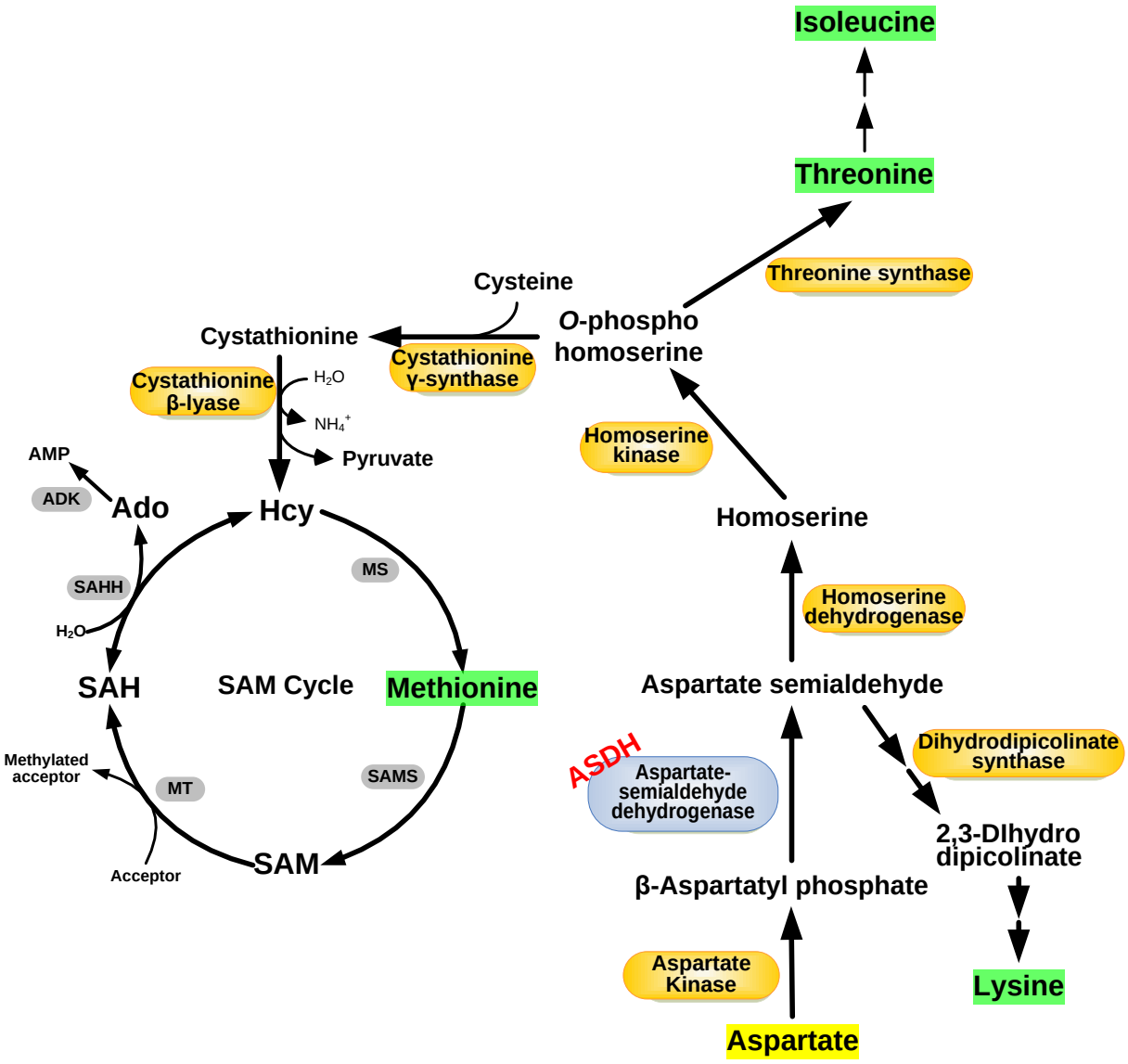

B
(kDa)
M
ADK1-Strep
M
ADK1-Strep
200
150
120
100
85
70
60
50
40
30
25
ADK1-Strep

Appendix IV. Yeast two-hybrid assay



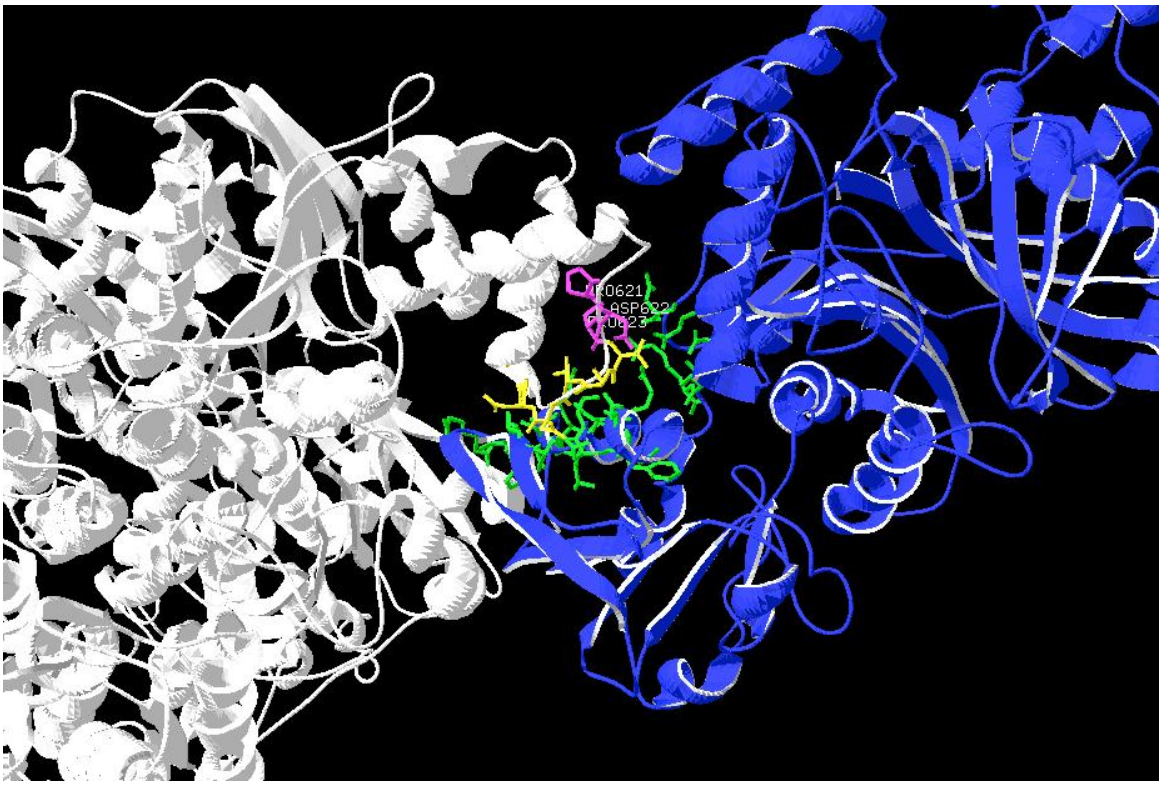
Yeast two hybrid results using various combinations of bait and prey proteins. Dbf4 and Rad53 were tested along with other samples as positive control.

Appendix V. The aspartate pathway



In this pathway, four essential amino acids (lysine, threonine, isoleucine and methionine) are produced in a series of reactions. ASDH converts aspartyl phosphate to aspartate semialdehyde that is the first branch point intermediate between the lysine and threonine/methionine pathways.

Appendix VI. Docking model for SAHH:ASDH interaction



Computational docking was performed by Andrew Doxey (Stanford University) using ZDOCK (Chen et al., 2003) with the closest homologous structures available in the Protein Data Bank, Plasmodium SAHH monomer (PDB ID 1V8B:A chain) and *Thermus thermophilus* ASDH (PDB ID 2YV3). The IR loop that contains PDPTSTDN is predicted to be located at the SAHH:ASDH interface. White, SAHH; purple, PDP; yellow, TSTDN; blue, ASDH; green, ASDH residues contacting the IR loop.

REFERENCES

Abrahams, S., Hayes, C.M., and Watson, J.M. (1995). Expression patterns of three genes in the stem of lucerne (*Medicago sativa*). Plant Mol Biol *27*, 513-528.

Agrimi, G., Di Noia, M.A., Marobbio, C.M., Fiermonte, G., Lasorsa, F.M., and Palmieri, F. (2004). Identification of the human mitochondrial S-adenosylmethionine transporter: bacterial expression, reconstitution, functional characterization and tissue distribution. Biochem J *379*, 183-190.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., *et al.* (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science *301*, 653-657.

Amberg, D.C., Burke, D.J., and Strathern, J.N. (2005). Methods in yeast genetics: a cold spring harbor laboratory course manual, Vol 2005 Edition (New York, Cold Spring Harbor Laboratory Press).

Andres, C.M., and Fox, I.H. (1979). Purification and properties of human placental adenosine kinase. Journal of Biological Chemistry *254*, 11388-11393.

Angrand, P.-O., Segura, I., Völkel, P., Ghidelli, S., Terry, R., Brajenovic, M., Vintersten, K., Klein, R., Superti-Furga, G., Drewes, G., *et al.* (2006). Transgenic Mouse Proteomics Identifies New 14-3-3-associated Proteins Involved in Cytoskeletal Rearrangements and Cell Signaling. Molecular & Cellular Proteomics *5*, 2211-2227.

Arnau, J., Lauritzen, C., Petersen, G.E., and Pedersen, J. (2006). Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. Protein Expression and Purification *48*, 1-13.

Auer, C. (1999). The Arabidopsis mutation cym changes cytokinin metabolism, adenosine nucleosidase activity and plant phenotype. International Symposium on Auxins and Cytokinins in Plant Development *42*, S3.

Ault-Riché, D.B., Yuan, C.S., and Borchardt, R.T. (1994). A single mutation at lysine 426 of human placental S-adenosylhomocysteine hydrolase inactivates the enzyme. Journal of Biological Chemistry *269*, 31472-31478.

Azevedo, R.A., Arruda, P., Turner, W.L., and Lea, P.J. (1997). The biosynthesis and metabolism of the aspartate derived amino acids in higher plants. Phytochemistry *46*, 395-419.

Azpiroz-Leehan, R., and Feldmann, K.A. (1997). T-DNA insertion mutagenesis in Arabidopsis: going back and forth. Trends Genet *13*, 152-156.

Bae, W., Lee, Y.J., Kim, D.H., Lee, J., Kim, S., Sohn, E.J., and Hwang, I. (2008). AKR2A-mediated import of chloroplast outer membrane proteins is essential for chloroplast biogenesis. Nat Cell Biol *10*, 220-227.

Barić, I., Fumić, K., Glenn, B., Ćuk, M., Schulze, A., Finkelstein, J.D., James, S.J., Mejaški-Bošnjak, V., Pažanin, L., Pogribny, I.P., *et al.* (2004). S-adenosylhomocysteine hydrolase deficiency in a human: A genetic disorder of methionine metabolism. Proceedings of the National Academy of Sciences of the United States of America *101*, 4234-4239.

Bartel, P.L., Roecklein, J.A., SenGupta, D., and Fields, S. (1996). A protein linkage map of *Escherichia coli* bacteriophage T7. Nature Genetics *12*, 72-77.

Baubec, T., Dinh, H.Q., Pecinka, A., Rakic, B., Rozhon, W., Wohlrab, B., von Haeseler, A., and Mittelsten Scheid, O. (2010). Cooperation of Multiple Chromatin Modifications Can Generate Unanticipated Stability of Epigenetic States in Arabidopsis. Plant Cell, tpc.109.072819.

Bauch, A., and Superti-Furga, G. (2006). Charting protein complexes, signaling pathways, and networks in the immune system. Immunological Reviews *210*, 187-207.

Baulcombe, D.C. (2007). Molecular biology: amplified silencing. Science *315*, 199-200.

Bedouelle, H., and Duplay, P. (1988). Production in *Escherichia coli* and one-step purification of bifunctional hybrid proteins which bind maltose - export of the Klenow polymerase into the periplasmic space. European Journal of Biochemistry *171*, 541-549.

Bent, A.F. (2000). Arabidopsis in planta transformation. Uses, mechanisms, and prospects for transformation of other species. Plant Physiol *124*, 1540-1547.

Bethin, K.E., Petrovic, N., and Ettinger, M.J. (1995). Identification of a major hepatic copper binding protein as S-adenosylhomocysteine hydrolase. J Biol Chem *270*, 20698-20702.

Blackburn, M.R., Wakamiya, M., Caskey, C.T., and Kellems, R.E. (1995). Tissue-specific Rescue Suggests That Placental Adenosine Deaminase Is Important for Fetal Development in Mice. Journal of Biological Chemistry *270*, 23891-23894.

Blum, T., Briesemeister, S., and Kohlbacher, O. (2009). MultiLoc2: integrating phylogeny and Gene Ontology terms improves subcellular protein localization prediction. BMC Bioinformatics *10*, 274.

Bohnsack, M.T., Regener, K., Schwappach, B., Saffrich, R., Paraskeva, E., Hartmann, E., and Gorlich, D. (2002). Exp5 exports eEF1A via tRNA from nuclei and synergizes with other transport pathways to confine translation to the cytoplasm. EMBO J *21*, 6205-6215.

Boison, D., Scheurer, L., Zumsteg, V., Rulicke, T., Litynski, P., Fowler, B., Brandner, S., and Mohler, H. (2002). Neonatal hepatic steatosis by disruption of the adenosine kinase gene. Proc Natl Acad Sci U S A *99*, 6985-6990.

Bouche, N., and Bouchez, D. (2001). Arabidopsis gene knockout: phenotypes wanted. Curr Opin Plant Biol *4*, 111-117.

Boudet, A.M. (1998). A new view of lignification. Trends Plant Sci *3*, 67-71.

Bourgis, F., Roje, S., Nuccio, M.L., Fisher, D.B., Tarczynski, M.C., Li, C.J., Herschbach, C., Rennenberg, H., Pimenta, M.J., Shen, T.L., *et al.* (1999). S-methylmethionine plays a major role in phloem sulfur transport and is synthesized by a novel type of methyltransferase. Plant Cell *11*, 1485-1497.

Bourlard, T., SchaumannGaudinet, A., BruyantVannier, M.P., and Morvan, C. (1997). Various pectin methyltransferase activities with affinity for low and highly methylated pectins. Plant and Cell Physiology *38*, 259-267.

Bouvier, F., Linka, N., Isner, J.C., Mutterer, J., Weber, A.P., and Camara, B. (2006). Arabidopsis SAMT1 defines a plastid transporter regulating plastid biogenesis and plant development. Plant Cell *18*, 3088-3105.

Bracha-Drori, K., Shichrur, K., Katz, A., Oliva, M., Angelovici, R., Yalovsky, S., and Ohad, N. (2004). Detection of protein-protein interactions in plants using bimolecular fluorescence complementation. Plant Journal *40*, 419-427.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem *72*, 248-254.

Brady, T.G., and Hegarty, V.J. (1966). An investigation of plant seeds for adenosine deaminase. Nature *209*, 1027-1028.

Buist, N., Glenn, B., Vugrek, O., Wagner, C., Stabler, S., Allen, R., Pogribny, I., Schulze, A., Zeisel, S., Barić, I., *et al.* (2006). S -Adenosylhomocysteine hydrolase deficiency in a 26-year-old man. Journal of Inherited Metabolic Disease *29*, 538-545.

Capecchi, M.R. (1989). Altering the genome by homologous recombination. Science *244*, 1288-1292.

Caplan, J.L., Mamillapalli, P., Burch-Smith, T.M., Czymmek, K., and Dinesh-Kumar, S.P. (2008). Chloroplastic Protein NRIP1 Mediates Innate Immune Receptor Recognition of a Viral Effector. Cell *132*, 449-462.

Caputto, R. (1951). The enzymatic synthesis of adenylic acid; adenosinekinase. J Biol Chem *189*, 801-814.

Carrie, C., Giraud, E., and Whelan, J. (2009). Protein transport in organelles: Dual targeting of proteins to mitochondria and chloroplasts. FEBS J *276*, 1187-1195.

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell *136*, 642-655.

Chaki, M., Valderrama, R., Fernandez-Ocana, A.M., Carreras, A., Lopez-Jaramillo, J., Luque, F., Palma, J.M., Pedrajas, J.R., Begara-Morales, J.C., Sanchez-Calvo, B., *et al.* (2009). Protein targets of tyrosine nitration in sunflower (*Helianthus annuus* L.) hypocotyls. J Exp Bot *60*, 4221-4234.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., and Prasher, D.C. (1994). Green Fluorescent Protein as a Marker for Gene-Expression. Science *263*, 802-805.

Chen, C.M., and Eckert, R.L. (1977). Phosphorylation of Cytokinin by Adenosine Kinase from Wheat Germ. Plant Physiol *59*, 443-447.

Chen, C.M., and Kristopeit, S.M. (1981). Metabolism of cytokinin: Deribosylation of cytokinin ribonucleoside by adenosine nucleosidase from wheat germ cells. Plant Physiol *68*, 1020-1023.

Chen, R., Li, L., and Weng, Z. (2003). ZDOCK: an initial-stage protein-docking algorithm. Proteins *52*, 80-87.

Chen, S.B., Tao, L.Z., Zeng, L.R., Vega-Sanchez, M.E., Umemura, K., and Wang, G.L. (2006). A highly efficient transient protoplast system for analyzing defence gene expression and protein-protein interactions in rice. Molecular Plant Pathology *7*, 417-427.

Chinsky, J.M., Ramamurthy, V., Fanslow, W.C., Ingolia, D.E., Blackburn, M.R., Shaffer, K.T., Higley, H.R., Trentin, J.J., Rudolph, F.B., Knudsen, T.B., *et al.* (1990). Developmental expression of adenosine deaminase in the upper alimentary tract of mice. Differentiation *42*, 172-183.

Chou, T., and Talalay, P. (1972). The mechanism of S-adenosyl-L-methionine synthesis by purified preparations of bakers' yeast. Biochemistry *11*, 1065-1073.

Christensen, A.C., Lyznik, A., Mohammed, S., Elowsky, C.G., Elo, A., Yule, R., and Mackenzie, S.A. (2005). Dual-Domain, Dual-Targeting Organellar Protein Presequences in Arabidopsis Can Use Non-AUG Start Codons. Plant Cell *17*, 2805-2816.

Chudakov, D.M., Lukyanov, S., and Lukyanov, K.A. (2005). Fluorescent proteins as a toolkit for in vivo imaging. Trends Biotechnol *23*, 605-613.

Citovsky, V., Gafni, Y., and Tzfira, T. (2008). Localizing protein-protein interactions by bimolecular fluorescence complementation in planta. Methods *45*, 196-206.

Citovsky, V., Lee, L.Y., Vyas, S., Glick, E., Chen, M.H., Vainstein, A., Gafni, Y., Gelvin, S.B., and Tzfira, T. (2006). Subcellular localization of interacting proteins by bimolecular fluorescence complementation in planta. Journal of Molecular Biology *362*, 1120-1131.

Clarke, R., Lewington, S., and Landray, M. (2003). Homocysteine, renal function, and risk of cardiovascular disease. Kidney Int Suppl, S131-133.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J *16*, 735-743.

Cook, W.J., DeLucas, L.J., and Chattopadhyay, D. (2000). Crystal structure of adenosine kinase from Toxoplasma gondii at 1.8 A resolution. Protein Sci *9*, 704-712.

Corbesier, L., Prinsen, E., Jacqmard, A., Lejeune, P., Van Onckelen, H., Perilleux, C., and Bernier, G. (2003). Cytokinin levels in leaves, leaf exudate and shoot apical meristem of *Arabidopsis thaliana* during floral transition. J Exp Bot *54*, 2511-2517.

Cormack, B.P., Valdivia, R.H., and Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). Gene *173*, 33-38.

Cristalli, G., Costanzi, S., Lambertucci, C., Lupidi, G., Vittori, S., Volpini, R., and Camaioni, E. (2001). Adenosine deaminase: functional implications and different classes of inhibitors. Med Res Rev *21*, 105-128.

Cui, J., Li, P., Li, G., Xu, F., Zhao, C., Li, Y., Yang, Z., Wang, G., Yu, Q., Li, Y., *et al.* (2008). AtPID: Arabidopsis thaliana protein interactome database—an integrative platform for plant systems biology. Nucleic Acids Research *36*, D999-D1008.

Cui, X.A., Singh, B., Park, J., and Gupta, R.S. (2009). Subcellular localization of adenosine kinase in mammalian cells: The long isoform of AdK is localized in the nucleus. Biochem Biophys Res Commun *388*, 46-50.

Curien, G., Dumas, R., Ravanel, S., and Douce, R. (1996). Characterization of an *Arabidopsis thaliana* cDNA encoding an S-adenosylmethionine-sensitive threonine synthase. Threonine synthase from higher plants. FEBS Lett *390*, 85-90.

Curien, G., Job, D., Douce, R., and Dumas, R. (1998). Allosteric activation of Arabidopsis threonine synthase by S-adenosylmethionine. Biochemistry *37*, 13212-13221.

D'Angelo, M.A., Raices, M., Panowski, S.H., and Hetzer, M.W. (2009). Age-dependent deterioration of nuclear pore complexes causes a loss of nuclear integrity in postmitotic cells. Cell *136*, 284-295.

Dancer, J.E., Hughes, R.G., and Lindell, S.D. (1997). Adenosine-5'-phosphate deaminase. A novel herbicide target. Plant Physiol *114*, 119-129.

Dang, C.V., Barrett, J., Villagarcia, M., Resar, L.M.S., Kato, G.J., and Fearon, E.R. (1991). Intracellular leucine zipper interactions suggest c-myc hetero-oligomerization. Molecular and Cellular Biology *11*, 954-962.

Darling, J.A., Sullivan, W.J., Jr., Carter, D., Ullman, B., and Roos, D.S. (1999). Recombinant expression, purification, and characterization of *Toxoplasma gondii* adenosine kinase. Mol Biochem Parasitol *103*, 15-23.

Datko, A.H., and Mudd, S.H. (1988). Enzymes of Phosphatidylcholine Synthesis in Lemna, Soybean, and Carrot. Plant Physiol *88*, 1338-1348.

Datta, A.K., Bhaumik, D., and Chatterjee, R. (1987). Isolation and characterization of adenosine kinase from *Leishmania donovani*. J Biol Chem *262*, 5515-5521.

De La Haba, G., and Cantoni, G.L. (1959). The enzymatic synthesis of S-adenosyl-L-homocysteine from adenosine and homocysteine. J Biol Chem *234*, 603-608.

Diella, F., Haslam, N., Chica, C., Budd, A., Michael, S., Brown, N.P., Trave, G., and Gibson, T.J. (2008). Understanding eukaryotic linear motifs and their role in cell signaling and regulation. Front Biosci *13*, 6580-6603.

Dinkins, R.D., Majee, S.M., Nayak, N.R., Martin, D., Xu, Q., Belcastro, M.P., Houtz, R.L., Beach, C.M., and Downie, A.B. (2008). Changing transcriptional initiation sites and alternative 5'- and 3'-splice site selection of the first intron deploys Arabidopsis protein isoaspartyl methyltransferase2 variants to different subcellular compartments. Plant J *55*, 1-13.

Dixon, R.A., and Paiva, N.L. (1995). Stress-induced phenylpropanoid metabolism. Plant Cell *7*, 1085-1097.

Dolzhanskaya, N., Merz, G., Aletta, J.M., and Denman, R.B. (2006). Methylation regulates the intracellular protein-protein and protein-RNA interactions of FMRP. J Cell Sci *119*, 1933-1946.

Drakas, R., Prisco, M., and Baserga, R. (2005). A modified tandem affinity purification tag technique for the purification of protein complexes in mammalian cells. Proteomics *5*, 132-137.

Dunnwald, M., Varshavsky, A., and Johnsson, N. (1999). Detection of transient *in vivo* interactions between substrate and transporter during protein translocation into the endoplasmic reticulum. Molecular Biology of the Cell *10*, 329-344.

Durfee, T., Becherer, K., Chen, P.L., Yeh, S.H., Yang, Y.Z., Kilburn, A.E., Lee, W.H., and Elledge, S.J. (1993). The retinoblastoma protein associates with the protein phosphatase TYPE-1 catalytic subunit. Genes & Development *7*, 555-569.

Edwards, R. (1996). S-adenosyl-L-methionine metabolism in alfalfa cell cultures following treatment with fungal elicitors. Phytochemistry *43*, 1163-1169.

Edwards, R., and Dixon, R.A. (1991). Purification and characterization of S-adenosyl-L-methionine: caffeic acid 3-O-methyltransferase from suspension cultures of alfalfa (*Medicago sativa L.*). Arch Biochem Biophys *287*, 372-379.

Ehrhard, K.N., Jacoby, J.J., Fu, X.Y., Jahn, R., and Dohlman, H.G. (2000). Use of G-protein fusions to monitor integral membrane protein-protein interactions in yeast. Nature Biotechnology *18*, 1075-1079.

Eisenberg, D., Marcotte, E.M., Xenarios, I., and Yeates, T.O. (2000). Protein function in the post-genomic era. Nature *405*, 823-826.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature *411*, 494-498.

Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol *300*, 1005-1016.

Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci *8*, 978-984.

Engel, K. (2009). Functional analysis of putative adenosine recycling enzymes in Arabidopsis thaliana. In Department of Biology (Waterloo, Ontario, Canada, University of Waterloo).

Facchini, P.J. (2001). Alkaloid biosynthesis in plants: biochemistry, cell biology, molecular regulation, and metabolic engineering applications. Annu Rev Plant Physiol Plant Mol Biol *52*, 29-66.

Facciuolo, A. (2009). The functional significance of the alternative first exons of the *Arabidopsis thaliana APT1* gene. In Department of Biology (Waterloo, Ontario, Canada, University of Waterloo).

Fahlgren, N., Howell, M.D., Kasschau, K.D., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., Law, T.F., Grant, S.R., Dangl, J.L., *et al.* (2007). High-throughput sequencing of Arabidopsis microRNAs: Evidence for frequent birth and death of miRNA genes. PLoS ONE *2*, e219.

Faye, F., and LeFloch, F. (1997). Adenosine kinase of peach tree flower buds: Purification and properties. Plant Physiol Biochem *35*, 15-22.

Fields, S., and Song, O.K. (1989). A novel genetic system to detect protein-protein interactions. Nature *340*, 245-246.

Finnegan, E.J., Peacock, W.J., and Dennis, E.S. (2000). DNA methylation, a key regulator of plant development and other processes. Curr Opin Genet Dev *10*, 217-223.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Fisher, M.N., and Newsholme, E.A. (1984). Properties of rat heart adenosine kinase. Biochem J *221*, 521-528.

Flajolet, M., Rotondo, G., Daviet, L., Bergametti, F., Inchauspe, G., Tiollais, P., Transy, C., and Legrain, P. (2000). A genomic approach of the hepatitis C virus generates a protein interaction map. Gene *242*, 369-379.

Fox, I.H., and Kelley, W.N. (1978). The Role of Adenosine and 2'-Deoxyadenosine in Mammalian Cells. Annual Review of Biochemistry *47*, 655-686.

Fricker, M., Runions, J., and Moore, I. (2006). Quantitative fluorescence microscopy: from art to science. Annual Review of Plant Biology *57*, 79-107.

Fried, H., and Kutay, U. (2003). Nucleocytoplasmic transport: taking an inventory. Cellular and Molecular Life Sciences *60*, 1659-1688.

Fu, X.D., and Maniatis, T. (1990). Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. Nature *343*, 437-441.

Fujita, E., Node, M., and Hori, H. (1977). Terpenoids. Part 39. Total synthesis of gebberellins A15 and A37. J Chem Soc Perkin 1, 611-621.

Fuxreiter, M., Tompa, P., and Simon, I. (2007). Local structural disorder imparts plasticity on linear motifs. Bioinformatics *23*, 950-956.

Galili, G. (1995). Regulation of Lysine and Threonine Synthesis. Plant Cell *7*, 899-906.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res *31*, 3784-3788.

Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., *et al.* (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature *415*, 141-147.

Gehl, C., Waadt, R., Kudla, J., Mendel, R.-R., and Hansch, R. (2009). New GATEWAY vectors for high throughput analyses of protein-protein interactions by Bimolecular Fluorescence Complementation. Mol Plant *2*, 1051-1058.

Geisler-Lee, J., O'Toole, N., Ammar, R., Provart, N.J., Millar, A.H., and Geisler, M. (2007). A Predicted Interactome for Arabidopsis. Plant Physiol *145*, 317-329.

Geiss-Friedlander, R., and Melchior, F. (2007). Concepts in sumoylation: a decade on. Nat Rev Mol Cell Biol *8*, 947-956.

Gengenbach, B.G., Walter, T.J., Green, C.E., and Hibberd, K.A. (1978). Feedback Regulation of Lysine, Threonine, and Methionine Biosynthetic Enzymes in Corn1. Crop Sci *18*, 472-476.

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. Nat Rev Genet *10*, 94-108.

Gibson, K.D., Wilson, J.D., and Udenfriend, S. (1961). The enzymatic conversion of phospholipid ethanolamine to phospholipid choline in rat liver. J Biol Chem *236*, 673-679.

Giepmans, B.N., Adams, S.R., Ellisman, M.H., and Tsien, R.Y. (2006). The fluorescent toolbox for assessing protein location and function. Science *312*, 217-224.

Gingras, A.-C., Gstaiger, M., Raught, B., and Aebersold, R. (2007). Analysis of protein complexes using mass spectrometry. Nat Rev Mol Cell Biol *8*, 645-654.

Godge, M.R., Kumar, D., and Kumar, P.P. (2008). Arabidopsis *HOG1* gene and its petunia homolog PETCBP act as key regulators of yield parameters. Plant Cell Reports *27*, 1497-1507.

Goldfarb, D.S., Gariepy, J., Schoolnik, G., and Kornberg, R.D. (1986). Synthetic peptides as nuclear localization signals. Nature *322*, 641-644.

Gorlich, D., and Kutay, U. (1999). Transport between the cell nucleus and the cytoplasm. Annu Rev Cell Dev Biol *15*, 607-660.

Gould, C.M., Diella, F., Via, A., Puntervoll, P., Gemund, C., Chabanis-Davidson, S., Michael, S., Sayadi, A., Bryne, J.C., Chica, C., *et al.* (2010). ELM: the status of the 2010 eukaryotic linear motif resource. Nucleic Acids Res *38*, D167-180.

Gould, S.B., Waller, R.F., and McFadden, G.I. (2008). Plastid Evolution. Annual Review of Plant Biology *59*, 491-517.

Grigoriev, A. (2003). On the number of protein–protein interactions in the yeast proteome. Nucleic Acids Research *31*, 4157-4161.

Grinberg, A.V., Hu, C.D., and Kerppola, T.K. (2004). Visualization of Myc/Max/Mad family dimers and the competition for dimerization in living cells. Molecular and Cellular Biology *24*, 4294-4308.

Gully, D., and Bouveret, E. (2006). A protein network for phospholipid synthesis uncovered by a variant of the tandem affinity purification method in *Escherichia coli*. Proteomics *6*, 282-293.

Guranowski, A. (1979). Plant adenosine kinase: purification and some properties of the enzyme from *Lupinus luteus* seeds. Arch Biochem Biophys *196*, 220-226.

Guranowski, A., Montgomery, J.A., Cantoni, G.L., and Chiang, P.K. (1981). Adenosine analogs as substrates and inhibitors of S-adenosylhomocysteine hydrolase. Biochemistry *20*, 110-115.

Guranowski, A., and Pawelkiewicz, J. (1977). Adenosylhomocysteinase from yellow lupin seeds. Purification and properties. Eur J Biochem *80*, 517-523.

Guranowski, A., and Schneider, Z. (1977). Purification and characterization of adenosine nucleosidase from barley leaves. Biochim Biophys Acta *482*, 145-158.

Hanson, A.D., Gage, D.A., and Shachar-Hill, Y. (2000). Plant one-carbon metabolism and its engineering. Trends Plant Sci *5*, 206-213.

Hanson, A.D., Rivoal, J., Paquet, L., and Gage, D.A. (1994). Biosynthesis of 3-dimethylsulfoniopropionate in *Wollastonia biflora* (L.) DC. Evidence that S-methylmethionine is an intermediate. Plant Physiol *105*, 103-110.

Hanson, A.D., and Roje, S. (2001). One-Carbon Metabolism in Higher Plants. Annu Rev Plant Physiol Plant Mol Biol *52*, 119-137.

Hanson, M.R., and Sattarzadeh, A. (2008). Dynamic morphology of plastids and stromules in angiosperm plants. Plant Cell Environ *31*, 646-657.

Harel, A., and Forbes, D.J. (2004). Importin beta: conducting a much larger cellular symphony. Mol Cell *16*, 319-330.

He, F., Zhou, Y., and Zhang, Z. (2010). Deciphering the Arabidopsis Floral Transition Process by Integrating a Protein-Protein Interaction Network and Gene Expression Data. Plant Physiol *153*, 1492-1505.

Heazlewood, J.L., Tonti-Filippini, J., Verboom, R.E., and Millar, A.H. (2005). Combining experimental and predicted datasets for determination of the subcellular location of proteins in Arabidopsis. Plant Physiol *139*, 598-609.

Henikoff, S., Till, B.J., and Comai, L. (2004). TILLING. Traditional mutagenesis meets functional genomics. Plant Physiol *135*, 630-636.

Hill, J., Donald, K.A., and Griffiths, D.E. (1991). DMSO-enhanced whole cell yeast transformation. Nucleic Acids Res *19*, 5791.

Hochuli, E., Dobeli, H., and Schacher, A. (1987). New metal chelate adsorbent selective for proteins and peptides containing neighboring histidine-residues. Journal of Chromatography *411*, 177-184.

Hopp, T.P., Prickett, K.S., Price, V.L., Libby, R.T., March, C.J., Cerretti, D.P., Urdal, D.L., and Conlon, P.J. (1988). A short polypeptide marker sequence useful for recombinant protein identification and purification. Bio-Technology *6*, 1204-1210.

Horne, D.W., Holloway, R.S., and Wagner, C. (1997). Transport of S-adenosylmethionine in isolated rat liver mitochondria. Arch Biochem Biophys *343*, 201-206.

Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., and Nakai, K. (2007). WoLF PSORT: protein localization predictor. Nucleic Acids Res *35*, W585-587.

Hoth, S., Ikeda, Y., Morgante, M., Wang, X., Zuo, J., Hanafey, M.K., Gaasterland, T., Tingey, S.V., and Chua, N.H. (2003). Monitoring genome-wide changes in gene expression in response to endogenous cytokinin reveals targets in *Arabidopsis thaliana*. FEBS Lett *554*, 373-380.

Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W., and Zimmermann, P. (2008). Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. Adv Bioinformatics *2008*, 420747.

Hu, C.D., Chinenov, Y., and Kerppola, T.K. (2002). Visualization of interactions among bZip and Rel family proteins in living cells using bimolecular fluorescence complementation. Molecular Cell *9*, 789-798.

Hu, C.D., and Kerppola, T.K. (2003). Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. Nat Biotechnol *21*, 539-545.

Hu, Y., Komoto, J., Huang, Y., Gomi, T., Ogawa, H., Takata, Y., Fujioka, M., and Takusagawa, F. (1999). Crystal structure of S-adenosylhomocysteine hydrolase from rat liver. Biochemistry *38*, 8323-8333.

Ibrahim, R.K., Bruneau, A., and Bantignies, B. (1998). Plant O-methyltransferases: molecular analysis, common signature and classification. Plant Molecular Biology *36*, 1-10.

Ideker, T., Thorsson, V., Ranish, J.A., Christmas, R., Buhler, J., Eng, J.K., Bumgarner, R., Goodlett, D.R., Aebersold, R., and Hood, L. (2001). Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. Science *292*, 929-934.

Inoue, T., Higuchi, M., Hashimoto, Y., Seki, M., Kobayashi, M., Kato, T., Tabata, S., Shinozaki, K., and Kakimoto, T. (2001). Identification of CRE1 as a cytokinin receptor from Arabidopsis. Nature *409*, 1060-1063.

International_Human_Genome_Sequencing_Consortium (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

International_Rice_Genome_Sequencing_Project (2005). The map-based sequence of the rice genome. Nature *436*, 793-800.

Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001). A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proceedings of the National Academy of Sciences of the United States of America *98*, 4569-4574.

Ito, T., Ota, K., Kubota, H., Yamaguchi, Y., Chiba, T., Sakuraba, K., and Yoshida, M. (2002). Roles for the two-hybrid system in exploration of the yeast protein interactome. Molecular & Cellular Proteomics *1*, 561-566.

Jander, G., and Joshi, V. (2010). Recent progress in deciphering the biosynthesis of aspartate-derived amino acids in plants. Mol Plant *3*, 54-65.

Jarvis, P., and Robinson, C. (2004). Mechanisms of protein import and routing in chloroplasts. Curr Biol *14*, R1064-1077.

Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. (1987). GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J *6*, 3901-3907.

Jin, J.B., Kim, Y.A., Kim, S.J., Lee, S.H., Kim, D.H., Cheong, G.W., and Hwang, I. (2001). A new dynamin-like protein, ADL6, is involved in trafficking from the trans-Golgi network to the central vacuole in Arabidopsis. Plant Cell *13*, 1511-1526.

Jung, B., Florchinger, M., Kunz, H.-H., Traub, M., Wartenberg, R., Jeblick, W., Neuhaus, H.E., and Mohlmann, T. (2009). Uridine-Ribohydrolase Is a Key Regulator in the Uridine Degradation Pathway of Arabidopsis. Plant Cell *21*, 876-891.

Kaake, R.M., Wang, X., and Huang, L. (2010). Profiling of Protein Interaction Networks of Protein Complexes Using Affinity Purification and Quantitative Mass Spectrometry. Molecular & Cellular Proteomics *9*, 1650-1665.

Karsten, W.E., and Viola, R.E. (1992). Identification of an essential cysteine in the reaction catalyzed by aspartate-beta-semialdehyde dehydrogenase from *Escherichia coli*. Biochim Biophys Acta *1121*, 234-238.

Kawalleck, P., Plesch, G., Hahlbrock, K., and Somssich, I.E. (1992). Induction by fungal elicitor of S-adenosyl-L-methionine synthetase and S-adenosyl-L-homocysteine hydrolase mRNAs in cultured cells and leaves of *Petroselinum crispum*. Proc Natl Acad Sci U S A *89*, 4713-4717.

Keeling, P.J., and Palmer, J.D. (2008). Horizontal gene transfer in eukaryotic evolution. Nat Rev Genet *9*, 605-618.

Kerppola, T.K. (2006). Visualization of molecular interactions by fluorescence complementation. Nature Reviews Molecular Cell Biology *7*, 449-456.

Kessler, F., and Blobel, G. (1996). Interaction of the protein import and folding machineries in the chloroplast. Proceedings of the National Academy of Sciences of the United States of America *93*, 7684-7689.

Kim, I., Kobayashi, K., Cho, E., and Zambryski, P.C. (2005). Subdomains for transport via plasmodesmata corresponding to the apical-basal axis are established during Arabidopsis embryogenesis. Proc Natl Acad Sci U S A *102*, 11945-11950.

Kim, S., Park, G.H., and Paik, W.K. (1998). Recent advances in protein methylation: enzymatic methylation of nucleic acid binding proteins. Amino Acids *15*, 291-306.

Kleffmann, T., Russenberger, D., von Zychlinski, A., Christopher, W., Sjolander, K., Gruissem, W., and Baginsky, S. (2004). The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. Curr Biol *14*, 354-362.

Kocsis, M.G., Ranocha, P., Gage, D.A., Simon, E.S., Rhodes, D., Peel, G.J., Mellema, S., Saito, K., Awazuhara, M., Li, C., *et al.* (2003). Insertional Inactivation of the Methionine S-Methyltransferase Gene Eliminates the S-Methylmethionine Cycle and Increases the Methylation Ratio. Plant Physiol *131*, 1808-1815.

Koncz, C., and Schell, J. (1986). The Promoter of Tl-DNA Gene 5 Controls the Tissue-Specific Expression of Chimeric Genes Carried by a Novel Type of Agrobacterium Binary Vector. Mol Gen Genet *204*, 383-396.

Kornberg, A., and Pricer, W.E., Jr. (1951). Enzymatic phosphorylation of adenosine and 2,6-diaminopurine riboside. J Biol Chem *193*, 481-495.

Korndorfer, I.P., and Skerra, A. (2002). Improved affinity of engineered streptavidin for the Strep-tag II peptide is due to a fixed open conformation of the lid-like loop at the binding site. Protein Sci *11*, 883-893.

Koshiishi, C., Kato, A., Yama, S., Crozier, A., and Ashihara, H. (2001). A new caffeine biosynthetic pathway in tea leaves: utilisation of adenosine released from the S-adenosyl-L-methionine cycle. FEBS Lett *499*, 50-54.

Kramer, D.L., Porter, C.W., Borchardt, R.T., and Sufrin, J.R. (1990). Combined modulation of S-adenosylmethionine biosynthesis and S-adenosylhomocysteine metabolism enhances inhibition of nucleic acid methylation and L1210 cell growth. Cancer Res *50*, 3838-3842.

Krysan, P.J., Young, J.C., and Sussman, M.R. (1999). T-DNA as an Insertional Mutagen in Arabidopsis. Plant Cell *11*, 2283-2290.

Kudo, T., Kiba, T., and Sakakibara, H. (2010). Metabolism and long-distance translocation of cytokinins. J Integr Plant Biol *52*, 53-60.

Kwade, Z., Swiatek, A., Azmi, A., Goossens, A., Inze, D., Van Onckelen, H., and Roef, L. (2005). Identification of four adenosine kinase isoforms in tobacco By-2 cells and their putative role in the cell cycle-regulated cytokinin metabolism. J Biol Chem *280*, 17512-17519.

Kwok, E.Y., and Hanson, M.R. (2004). Plastids and stromules interact with the nucleus and cell membrane in vascular plants. Plant Cell Rep *23*, 188-195.

Kwon, Y.T., Kashina, A.S., and Varshavsky, A. (1999). Alternative splicing results in differential expression, activity, and localization of the two forms of arginyl-tRNA-protein transferase, a component of the N-end rule pathway. Mol Cell Biol *19*, 182-193.

Laber, B., Maurer, W., Hanke, C., Grafe, S., Ehlert, S., Messerschmidt, A., and Clausen, T. (1999). Characterization of recombinant *Arabidopsis thaliana* threonine synthase. Eur J Biochem *263*, 212-221.

Labro, A.J., Raes, A.L., Bellens, I., Ottschytsch, N., and Snyders, D.J. (2003). Gating of Shaker-type Channels Requires the Flexibility of S6 Caused by Prolines. Journal of Biological Chemistry *278*, 50724-50731.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680-685.

Lakey, J.H., and Raggett, E.M. (1998). Measuring protein-protein interactions. Current Opinion in Structural Biology *8*, 119-123.

Laukens, K., Lenobel, R., Strnad, M., Van Onckelen, H., and Witters, E. (2003). Cytokinin affinity purification and identification of a tobacco BY-2 adenosine kinase. FEBS Lett *533*, 63-66.

Lee, M., and Leustek, T. (1999). Identification of the gene encoding homoserine kinase from *Arabidopsis thaliana* and characterization of the recombinant enzyme derived from the gene. Arch Biochem Biophys *372*, 135-142.

Lee, Y.J., Kim, D.H., Kim, Y.W., and Hwang, I. (2001). Identification of a signal that distinguishes between the chloroplast outer envelope membrane and the endomembrane system in vivo. Plant Cell *13*, 2175-2190.

Levi, V., Jacobson, E.L., and Jacobson, M.K. (1978). Inhibition of poly(ADP-ribose) polymerase by methylated xanthines and cytokinins. FEBS Lett *88*, 144-146.

Levitan, A., Trebitsh, T., Kiss, V., Pereg, Y., Dangoor, I., and Danon, A. (2005). Dual targeting of the protein disulfide isomerase RB60 to the chloroplast and the endoplasmic reticulum. Proc Natl Acad Sci U S A *102*, 6225-6230.

Li, C.H., Yu, N., Jiang, S.M., Shangguan, X.X., Wang, L.J., and Chen, X.Y. (2008). Down-regulation of S-adenosyl-L-homocysteine hydrolase reveals a role of cytokinin in promoting transmethylation reactions. Planta *228*, 125-136.

Li, H.-m., and Chiu, C.-C. (2010). Protein Transport into Chloroplasts. Annual Review of Plant Biology *61*, 157-180.

Li, Y., and Horwitz, M.S. (1997). Use of green fluorescent protein in studies of apoptosis of transfected cells. Biotechniques *23*, 1026-1029.

Lin, M., Hu, B., Chen, L., Sun, P., Fan, Y., Wu, P., and Chen, X. (2009). Computational Identification of Potential Molecular Interactions in Arabidopsis. Plant Physiol *151*, 34-46.

Lindermayr, C., Saalbach, G., and Durner, J. (2005). Proteomic identification of S-nitrosylated proteins in Arabidopsis. Plant Physiol *137*, 921-930.

Lisby, M., Mortensen, U.H., and Rothstein, R. (2003). Colocalization of multiple DNA double-strand breaks at a single Rad52 repair centre. Nat Cell Biol *5*, 572-577.

Liska, A.J., Shevchenko, A., Pick, U., and Katz, A. (2004). Enhanced Photosynthesis and Redox Energy Production Contribute to Salinity Tolerance in Dunaliella as Revealed by Homology-Based Proteomics. Plant Physiol *136*, 2806-2817.

Lloyd, A., Plaisier, C.L., Carroll, D., and Drews, G.N. (2005). Targeted mutagenesis using zinc-finger nucleases in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America *102*, 2232-2237.

Loehrer, F., Angst, C., Brunner, F., Haefeli, W., and Fowler, B. (1998). Evidence for disturbed S-adenosylmethionine: S-adenosylhomocysteine ratio in patients with end-stage renal failure: a cause for disturbed methylation reactions? Nephrol Dial Transplant *13*, 656-661.

Logemann, E., Tavernaro, A., Schulz, W., Somssich, I.E., and Hahlbrock, K. (2000). UV light selectively coinduces supply pathways from primary metabolism and flavonoid secondary product formation in parsley. Proc Natl Acad Sci U S A *97*, 1903-1907.

Long, M.C., Escuyer, V., and Parker, W.B. (2003). Identification and characterization of a unique adenosine kinase from *Mycobacterium tuberculosis*. J Bacteriol *185*, 6548-6555.

Lu, G.T., Tang, Y.Q., Li, C.Y., Li, R.F., An, S.Q., Feng, J.X., He, Y.Q., Jiang, B.L., Tang, D.J., and Tang, J.L. (2009). An adenosine kinase exists in *Xanthomonas campestris* pathovar campestris and is involved in extracellular polysaccharide production, cell motility, and virulence. J Bacteriol *191*, 3639-3648.

Lu, S.C. (2000). S-Adenosylmethionine. Int J Biochem Cell Biol *32*, 391-395.

Lu, Z., Szafron, D., Greiner, R., Lu, P., Wishart, D.S., Poulin, B., Anvik, J., Macdonell, C., and Eisner, R. (2004). Predicting subcellular localization of proteins using machine-learned classifiers. Bioinformatics *20*, 547-556.

Lukyanov, K.A., Chudakov, D.M., Lukyanov, S., and Verkhusha, V.V. (2005). Innovation: Photoactivatable fluorescent proteins. Nat Rev Mol Cell Biol *6*, 885-891.

Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A., and Lajoie, G. (2003). PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. Rapid Commun Mass Spectrom *17*, 2337-2342.

Mackenzie, S.A. (2005). Plant organellar protein targeting: a traffic plan still under construction. Trends Cell Biol *15*, 548-554.

Mansour, S.L., Thomas, K.R., and Capecchi, M.R. (1988). Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature *336*, 348-352.

Maple, J., and Moller, S.G. (2007). Mutagenesis in Arabidopsis. Methods Mol Biol *362*, 197-206.

Marcus, A.I., Moore, R.C., and Cyr, R.J. (2001). The role of microtubules in guard cell function. Plant Physiol *125*, 387-395.

Marobbio, C.M., Agrimi, G., Lasorsa, F.M., and Palmieri, F. (2003). Identification and functional reconstitution of yeast mitochondrial carrier for S-adenosylmethionine. EMBO J *22*, 5975-5982.

Martienssen, R.A. (1998). Functional genomics: probing plant gene function and expression with transposons. Proc Natl Acad Sci U S A *95*, 2021-2026.

Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M., and Penny, D. (2002). Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proceedings of the National Academy of Sciences of the United States of America *99*, 12246-12251.

Masuta, C., Tanaka, H., Uehara, K., Kuwata, S., Koiwai, A., and Noma, M. (1995). Broad resistance to plant viruses in transgenic plants conferred by antisense inhibition of a host gene essential in S-adenosylmethionine-dependent transmethylation reactions. Proc Natl Acad Sci U S A *92*, 6117-6121.

Mathews, II, Erion, M.D., and Ealick, S.E. (1998). Structure of human adenosine kinase at 1.5 A resolution. Biochemistry *37*, 15607-15620.

May, T., and Soll, J. (2000). 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. Plant Cell *12*, 53-64.

McCallum, C.M., Comai, L., Greene, E.A., and Henikoff, S. (2000). Targeting Induced Local Lesions IN Genomes (TILLING) for Plant Functional Genomics. Plant Physiol *123*, 439-442.

McNally, T., Helfrich, R.J., Cowart, M., Dorwin, S.A., Meuth, J.L., Idler, K.B., Klute, K.A., Simmer, R.L., Kowaluk, E.A., and Halbert, D.N. (1997). Cloning and Expression

of the Adenosine Kinase Gene from Rat and Human Tissues. Biochemical and Biophysical Research Communications *231*, 645-650.

Merkle, T. (2003). Nucleo-cytoplasmic partitioning of proteins in plants: implications for the regulation of environmental and developmental signalling. Current Genetics *44*, 231-260.

Migchielsen, A.A.J., Breuer, M.L., Vanroon, M.A., Riele, H.T., Zurcher, C., Ossendorp, F., Toutain, S., Hershfield, M.S., Berns, A., and Valerio, D. (1995). Adenosine-deaminase-deficient mice die perinatally and exhibit liver-cell degeneration, atelectasis and small-intestinal cell-death. Nature Genetics *10*, 279-287.

Millar, A.H., Whelan, J., and Small, I. (2006). Recent surprises in protein targeting to mitochondria and plastids. Curr Opin Plant Biol *9*, 610-615.

Miller, M.W., Duhl, D.M., Winkes, B.M., Arredondo-Vega, F., Saxon, P.J., Wolff, G.L., Epstein, C.J., Hershfield, M.S., and Barsh, G.S. (1994). The mouse lethal nonagouti (a(x)) mutation deletes the S-adenosylhomocysteine hydrolase (Ahcy) gene. EMBO J *13*, 1806-1816.

Mitsui, S., Wakasugi, T., and Sugiura, M. (1993). A cDNA encoding the 57 kDa subunit of a cytokinin-binding protein complex from tobacco: the subunit has high homology to S-adenosyl-L-homocysteine hydrolase. Plant Cell Physiol *34*, 1089-1096.

Mlejnek, P., and Prochazka, S. (2002). Activation of caspase-like proteases and induction of apoptosis by isopentenyladenosine in tobacco BY-2 cells. Planta *215*, 158-166.

Moffatt, B., and Somerville, C. (1988). Positive selection for male-sterile mutants of Arabidopsis lacking adenine phosphoribosyl transferase activity. Plant Physiol *86*, 1150-1154.

Moffatt, B.A., Stevens, Y.Y., Allen, M.S., Snider, J.D., Pereira, L.A., Todorova, M.I., Summers, P.S., Weretilnyk, E.A., Martin-McCaffrey, L., and Wagner, C. (2002). Adenosine kinase deficiency is associated with developmental abnormalities and reduced transmethylation. Plant Physiol *128*, 812-821.

Moffatt, B.A., Wang, L., Allen, M.S., Stevens, Y.Y., Qin, W., Snider, J., and von Schwartzenberg, K. (2000). Adenosine kinase of Arabidopsis. Kinetic properties and gene expression. Plant Physiol *124*, 1775-1785.

Moffatt, B.A., and Weretilnyk, E.A. (2001). Sustaining S-adenosyl-L-methionine-dependent methyltransferase activity in plant cells. Physiologia Plantarum *113*, 435-442.

Morsy, M., Gouthu, S., Orchard, S., Thorneycroft, D., Harper, J.F., Mittler, R., and Cushman, J.C. (2008). Charting plant interactomes: possibilities and challenges. Trends Plant Sci *13*, 183-191.

Mowen, K.A., Schurter, B.T., Fathman, J.W., David, M., and Glimcher, L.H. (2004). Arginine methylation of NIP45 modulates cytokine gene expression in effector T lymphocytes. Mol Cell *15*, 559-571.

Mudd, S.H., and Datko, A.H. (1990). The *S*-methylmethionine cycle in *Lemna paucicostata*. Plant Physiology *93*, 623-630.

Mull, L., Ebbs, M.L., and Bender, J. (2006). A histone methylation-dependent DNA methylation pathway is uniquely impaired by deficiency in Arabidopsis S-adenosylhomocysteine hydrolase. Genetics *174*, 1161-1171.

Murashige, T., and Skoog, F. (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. Physiologia Plantarum *15*, 473-&.

Nair, R., and Rost, B. (2005). Mimicking cellular sorting improves prediction of subcellular localization. J Mol Biol *348*, 85-100.

Nakayasu, H., and Berezney, R. (1989). Mapping replicational sites in the eucaryotic cell nucleus. J Cell Biol *108*, 1-11.

Neduva, V., Linding, R., Su-Angrand, I., Stark, A., de Masi, F., Gibson, T.J., Lewis, J., Serrano, L., and Russell, R.B. (2005). Systematic discovery of new recognition peptides mediating protein interaction networks. PLoS Biol *3*, e405.

Nielsen, E., Akita, M., DavilaAponte, J., and Keegstra, K. (1997). Stable association of chloroplastic precursors with protein translocation complexes that contain proteins from both envelope membranes and a stromal Hsp 100 molecular chaperone. Embo Journal *16*, 935-946.

Nigg, E.A. (1997). Nucleocytoplasmic transport: signals, mechanisms and regulation. Nature *386*, 779-787.

Nooren, I.M., and Thornton, J.M. (2003a). Structural characterisation and functional significance of transient protein-protein interactions. J Mol Biol *325*, 991-1018.

Nooren, I.M.A., and Thornton, J.M. (2003b). Diversity of protein-protein interactions. Embo Journal *22*, 3486-3492.

Ossowski, S., Schwab, R., and Weigel, D. (2008). Gene silencing in plants using artificial microRNAs and other small RNAs. The Plant Journal *53*, 674-690.

Ozawa, T., Kaihara, A., Sato, M., Tachihara, K., and Umezawa, Y. (2001). Split luciferase as an optical probe for detecting protein-protein interactions in mammalian cells based on protein splicing. Analytical Chemistry *73*, 2516-2521.

Palmer, J.L., and Abeles, R.H. (1979). The mechanism of action of S-adenosylhomocysteinase. J Biol Chem *254*, 1217-1226.

Palmieri, L., Arrigoni, R., Blanco, E., Carrari, F., Zanor, M.I., Studart-Guimaraes, C., Fernie, A.R., and Palmieri, F. (2006). Molecular identification of an Arabidopsis S-adenosylmethionine transporter. Analysis of organ distribution, bacterial expression, reconstitution into liposomes, and functional characterization. Plant Physiol *142*, 855-865.

Paris, S., Wessel, P.M., and Dumas, R. (2002). Overproduction, purification, and characterization of recombinant aspartate semialdehyde dehydrogenase from *Arabidopsis thaliana*. Protein Expression and Purification *24*, 99-104.

Parker, N.B., Yang, X., Hanke, J., Mason, K.A., Schowen, R.L., Borchardt, R.T., and Yin, D.H. (2003). Trypanosoma cruzi: molecular cloning and characterization of the S-adenosylhomocysteine hydrolase. Exp Parasitol *105*, 149-158.

Pedelacq, J.D., Cabantous, S., Tran, T., Terwilliger, T.C., and Waldo, G.S. (2006). Engineering and characterization of a superfolder green fluorescent protein. Nat Biotechnol *24*, 79-88.

Pegg, A.E. (1986). Recent advances in the biochemistry of polyamines in eukaryotes. Biochem J *234*, 249-262.

Peiter, E., Montanini, B., Gobert, A., Pedas, P., Husted, S., Maathuis, F.J., Blaudez, D., Chalot, M., and Sanders, D. (2007). A secretory pathway-localized cation diffusion facilitator confers plant manganese tolerance. Proc Natl Acad Sci U S A *104*, 8532-8537.

Pereira, L., Schoor, S., Goubet, F., Dupree, P., and Moffatt, B. (2006). Deficiency of adenosine kinase activity affects the degree of pectin methyl-esterification in cell walls of *Arabidopsis thaliana*. Planta *224*, 1401-1414.

Pereira, L., Todorova, M., Cai, X., Makaroff, C., Emery, R., and Moffatt, B. (2007). Methyl recycling activities are co-ordinately regulated during plant development. J Exp Bot, erl275.

Pereira, L.A.R. (2004). Developmental expression of two methyl-recycling enzymes in *Arabidopsis thaliana*. In Department of Biology (Waterloo, ON, Canada, University of Waterloo).

Poulton, J.E. (1981). Transmethylation and demethylation reactions in the metabolism of secondary plant products, Vol 7 (New York, Academic Press Inc.).

Poulton, J.E., and Butt, V.S. (1975). Purification and properties of S-adenosyl-L-methionine: caffeic acid O-methyltransferase from leaves of spinach beet (Beta vulgaris L). Biochim Biophys Acta *403*, 301-314.

Poulton, J.E., and Butt, V.S. (1976). Purification and properties of S-adenosyl-L-homocysteine hydrolase from leaves of spinach beet. Arch Biochem Biophys *172*, 135-142.

Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G., and Cormier, M.J. (1992). Primary structure of the *Aequorea victoria* green-fluorescent protein. Gene *111*, 229-233.

Radomski, N., Barreto, G., Kaufmann, C., Yokoska, J., Mizumoto, K., and Dreyer, C. (2002). Interaction of S-adenosylhomocysteine hydrolase of *Xenopus laevis* with mRNA(guanine-7-)methyltransferase: implication on its nuclear compartmentalisation and on cap methylation of hnRNA. Biochim Biophys Acta *1590*, 93-102.

Radomski, N., Kaufmann, C., and Dreyer, C. (1999). Nuclear accumulation of S-adenosylhomocysteine hydrolase in transcriptionally active cells during development of *Xenopus laevis*. Mol Biol Cell *10*, 4283-4298.

Rain, J.C., Selig, L., De Reuse, H., Battaglia, V., Reverdy, C., Simon, S., Lenzen, G., Petel, F., Wojcik, J., Schachter, V., *et al.* (2001). The protein-protein interaction map of *Helicobacter pylori*. Nature *409*, 211-215.

Rajkarnikar, A., Kwon, H.J., and Suh, J.W. (2007). Role of adenosine kinase in the control of Streptomyces differentiations: Loss of adenosine kinase suppresses sporulation and actinorhodin biosynthesis while inducing hyperproduction of undecylprodigiosin in *Streptomyces lividans*. Biochem Biophys Res Commun *363*, 322-328.

Ravanel, S., Block, M.A., Rippert, P., Jabrin, S., Curien, G., Rebeille, F., and Douce, R. (2004). Methionine metabolism in plants: chloroplasts are autonomous for de novo methionine synthesis and can import S-adenosylmethionine from the cytosol. J Biol Chem *279*, 22548-22557.

Ravanel, S., Gakiere, B., Job, D., and Douce, R. (1998). The specific features of methionine biosynthesis and metabolism in plants. Proc Natl Acad Sci U S A *95*, 7805-7812.

Reddy, M.C., Kuppan, G., Shetty, N.D., Owen, J.L., Ioerger, T.R., and Sacchettini, J.C. (2008). Crystal structures of *Mycobacterium tuberculosis* S-adenosyl-L-homocysteine hydrolase in ternary complex with substrate and inhibitors. Protein Sci *17*, 2134-2144.

Reddy, M.C., Palaninathan, S.K., Shetty, N.D., Owen, J.L., Watson, M.D., and Sacchettini, J.C. (2007). High resolution crystal structures of *Mycobacterium tuberculosis* adenosine kinase: insights into the mechanism and specificity of this novel prokaryotic enzyme. J Biol Chem *282*, 27334-27342.

Regev-Rudzki, N., and Pines, O. (2007). Eclipsed distribution: a phenomenon of dual targeting of protein and its significance. Bioessays *29*, 772-782.

Remy, I., Montmarquette, A., and Michnick, S.W. (2004). PKB/Akt modulates TGF-beta signalling through a direct interaction with Smad3. Nature Cell Biology *6*, 358-365.

Reumann, S., Babujee, L., Ma, C., Wienkoop, S., Siemsen, T., Antonicelli, G.E., Rasche, N., Luder, F., Weckwerth, W., and Jahn, O. (2007). Proteome analysis of Arabidopsis

leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. Plant Cell *19*, 3170-3193.

Reyes-Prieto, A., Weber, A.P.M., and Bhattacharya, D. (2007). The Origin and Establishment of the Plastid in Algae and Plants. Annual Review of Genetics *41*, 147-168.

Richards, E.J., and Elgin, S.C. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. Cell *108*, 489-500.

Richter, S., and Lamppa, G.K. (1998). A chloroplast processing enzyme functions as the general stromal processing peptidase. Proceedings of the National Academy of Sciences of the United States of America *95*, 7463-7468.

Riewe, D., Grosman, L., Fernie, A.R., Zauber, H., Wucke, C., and Geigenberger, P. (2008). A cell wall-bound adenosine nucleosidase is involved in the salvage of extracellular ATP in *Solanum tuberosum*. Plant Cell Physiol *49*, 1572-1579.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nature Biotechnology *17*, 1030-1032.

Robins, M.J., Wnuk, S.F., Yang, X., Yuan, C.S., Borchardt, R.T., Balzarini, J., and De Clercq, E. (1998). Inactivation of S-adenosyl-L-homocysteine hydrolase and antiviral activity with 5',5',6',6'-tetradehydro-6'-deoxy-6'-halohomoadenosine analogues (4'-haloacetylene analogues derived from adenosine). J Med Chem *41*, 3857-3864.

Rocha, P.S.C.F., Sheikh, M., Melchiorre, R., Fagard, M., Boutet, S., Loach, R., Moffatt, B., Wagner, C., Vaucheret, H., and Furner, I. (2005). The Arabidopsis HOMOLOGY-DEPENDENT GENE SILENCING1 Gene Codes for an S-Adenosyl-L-Homocysteine Hydrolase Required for DNA Methylation-Dependent Gene Silencing. Plant Cell *17*, 404-417.

Rohila, J.S., Chen, M., Cerny, R., and Fromm, M.E. (2004). Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. Plant Journal *38*, 172-181.

Rohila, J.S., Chen, M., Chen, S., Chen, J., Cerny, R., Dardick, C., Canlas, P., Xu, X., Gribskov, M., Kanrar, S., *et al.* (2006). Protein-protein interactions of tandem affinity purification-tagged protein kinases in rice. Plant Journal *46*, 1-13.

Roje, S. (2006). S-Adenosyl-l-methionine: Beyond the universal methyl group donor. Phytochemistry *67*, 1686-1698.

Rossi, F., Charlton, C.A., and Blau, H.M. (1997). Monitoring protein-protein interactions in intact eukaryotic cells by beta-galactosidase complementation. Proceedings of the National Academy of Sciences of the United States of America *94*, 8405-8410.

Rubio, V., Shen, Y.P., Saijo, Y., Liu, Y.L., Gusmaroli, G., Dinesh-Kumar, S.P., and Deng, X.W. (2005). An alternative tandem affinity purification strategy applied to Arabidopsis protein complex isolation. Plant Journal *41*, 767-778.

Sakowicz, M., Grden, M., and Pawelczyk, T. (2001). Expression level of adenosine kinase in rat tissues. Lack of phosphate effect on the enzyme activity. Acta Biochim Pol *48*, 745-754.

Sapir-Mir, M., Mett, A., Belausov, E., Tal-Meshulam, S., Frydman, A., Gidoni, D., and Eyal, Y. (2008). Peroxisomal localization of Arabidopsis isopentenyl diphosphate isomerases suggests that part of the plant isoprenoid mevalonic acid pathway is compartmentalized to peroxisomes. Plant Physiol *148*, 1219-1228.

Schnell, D.J., Blobel, G., Keegstra, K., Kessler, F., Ko, K., and Soll, J. (1997). A consensus nomenclature for the protein-import components of the chloroplast envelope. Trends Cell Biol *7*, 303-304.

Schomburg, I., Chang, A., Ebeling, C., Gremse, M., Heldt, C., Huhn, G., and Schomburg, D. (2004). BRENDA, the enzyme database: updates and major new developments. Nucleic Acids Res *32*, D431-433.

Schoor, S.M. (2007). Determining the subcellular localization of adenosine kinase and SAH hydrolase and their roles in adenosine metabolism. In Department of Biology (Waterloo, Ontario, Canada, University of Waterloo).

Schröder, G., Eichel, J., Breinig, S., and Schröder, J. (1997). Three differentially expressed S-adenosylmethionine synthetases from *Catharanthus roseus*: molecular and functional characterization. Plant Molecular Biology *33*, 211-222.

Schwab, R., Ossowski, S., Riester, M., Warthmann, N., and Weigel, D. (2006). Highly specific gene silencing by artificial MicroRNAs in Arabidopsis. Plant Cell *18*, 1121-1133.

Schwab, R., Palatnik, J.F., Riester, M., Schommer, C., Schmid, M., and Weigel, D. (2005). Specific effects of microRNAs on the plant transcriptome. Dev Cell *8*, 517-527.

Schwacke, R., Fischer, K., Ketelsen, B., Krupinska, K., and Krause, K. (2007). Comparative survey of plastid and mitochondrial targeting properties of transcription factors in Arabidopsis and rice. Mol Genet Genomics *277*, 631-646.

Sebestova, L., Votruba, I., and Holy, A. (1984). Studies on S-adenyl-L-homocysteine hydrolase .11. S-adenosine-L-homocysteine hydrolase from *Nicotiana tabacum* L. - isolation and properties. Collect Czech Chem Commun *49*, 1543-1551.

Sedbrook, J.C. (2004). MAPs in plant cells: delineating microtubule growth dynamics and organization. Curr Opin Plant Biol *7*, 632-640.

Sganga, M.W., Aksamit, R.R., Cantoni, G.L., and Bauer, C.E. (1992). Mutational and nucleotide sequence analysis of S-adenosyl-L-homocysteine hydrolase from *Rhodobacter capsulatus*. Proc Natl Acad Sci U S A *89*, 6328-6332.

Shafer, B., Chu, C., and Shatkin, A.J. (2005). Human mRNA cap methyltransferase: alternative nuclear localization signal motifs ensure nuclear localization required for viability. Mol Cell Biol *25*, 2644-2649.

Shah, M. (2006). Searching for the subcellular targeting signals of adenosine kinase. In 499 Senior Honours Project (Waterloo, University of Waterloo).

Sheen, J. (1996). Ca2+-dependent protein kinases and stress signal transduction in plants. Science *274*, 1900-1902.

Shen, B., Li, C., and Tarczynski, M.C. (2002). High free-methionine and decreased lignin content result from a mutation in the Arabidopsis S-adenosyl-L-methionine synthetase 3 gene. Plant J *29*, 371-380.

Shimomura, O., Johnson, F.H., and Saiga, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the *Luminous hydromedusan*, Aequorea. J Cell Comp Physiol *59*, 223-239.

Shin, R., Alvarez, S., Burch, A.Y., Jez, J.M., and Schachtman, D.P. (2007). Phosphoproteomic identification of targets of the Arabidopsis sucrose nonfermenting-like kinase SnRK2.8 reveals a connection to metabolic processes. Proc Natl Acad Sci U S A *104*, 6460-6465.

Silva-Filho, M.C. (2003). One ticket for multiple destinations: dual targeting of proteins to distinct subcellular locations. Curr Opin Plant Biol *6*, 589-595.

Slade, A.J., Fuerstenberg, S.I., Loeffler, D., Steine, M.N., and Facciotti, D. (2005). A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. Nat Biotech *23*, 75-81.

Smith, D.B., and Johnson, K.S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. Gene *67*, 31-40.

Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A., and Kucherlapati, R.S. (1985). Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. Nature *317*, 230-234.

Snow, E.T., Foote, R.S., and Mitra, S. (1984). Base-pairing properties of O6-methylguanine in template DNA during in vitro DNA replication. J Biol Chem *259*, 8095-8100.

Spychala, J., Datta, N.S., Takabayashi, K., Datta, M., Fox, I.H., Gribbin, T., and Mitchell, B.S. (1996). Cloning of human adenosine kinase cDNA: sequence similarity to microbial ribokinases and fructokinases. Proc Natl Acad Sci U S A *93*, 1232-1237.

Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F.H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S., *et al.* (2005). A human protein-protein interaction network: a resource for annotating the proteome. *122*, 957-968.

Stepkowski, T., Brzezinski, K., Legocki, A.B., Jaskólski, M., and Béna, G. (2005). Bayesian phylogenetic analysis reveals two-domain topology of S-adenosylhomocysteine hydrolase protein sequences. Molecular Phylogenetics and Evolution *34*, 15-28.

Sterck, L., Rombauts, S., Vandepoele, K., Rouze, P., and Van de Peer, Y. (2007). How many genes are there in plants (... and why are they there)? Curr Opin Plant Biol *10*, 199-203.

Stewart, M. (2007). Molecular mechanism of the nuclear protein import cycle. Nat Rev Mol Cell Biol *8*, 195-208.

Summers, P.S., and Weretilnyk, E.A. (1993). Choline synthesis in spinach in relation to salt stress. Plant Physiol *103*, 1269-1276.

Suzuki, M., Yasumoto, E., Baba, S., and Ashihara, H. (2003). Effect of salt stress on the metabolism of ethanolamine and choline in leaves of the betaine-producing mangrove species *Avicennia marina*. Phytochemistry *64*, 941-948.

TAIR (2010). The Arabidopsis Information Resource. http://www.arabidopsis.org/.

Takei, K., Yamaya, T., and Sakakibara, H. (2004). Arabidopsis CYP735A1 and CYP735A2 encode cytokinin hydroxylases that catalyze the biosynthesis of trans-Zeatin. Journal of Biological Chemistry *279*, 41866-41872.

Tanaka, H., Masuta, C., Uehara, K., Kataoka, J., Koiwai, A., and Noma, M. (1997). Morphological changes and hypomethylation of DNA in transgenic tobacco expressing antisense RNA of the S-adenosyl-L-homocysteine hydrolase gene. Plant Mol Biol *35*, 981-986.

Tanaka, N., Nakanishi, M., Kusakabe, Y., Shiraiwa, K., Yabe, S., Ito, Y., Kitade, Y., and Nakamura, K.T. (2004). Crystal structure of S-adenosyl-L-homocysteine hydrolase from the human malaria parasite *Plasmodium falciparum*. Journal of Molecular Biology *343*, 1007-1017.

The_Arabidopsis_Genome_Initiative (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature *408*, 796-815.

Todorova, M. (2002). Characterization of the expression and enzyme activity of S-adenosyl-L-homocysteine hydrolase in major organs of *Arabidopsis thaliana*. In Department of Biology (Waterloo, Ontario, Canada, University of Waterloo).

Tourriere, H., Chebli, K., and Tazi, J. (2002). mRNA degradation machines in eukaryotic cells. Biochimie *84*, 821-837.

Townsend, J.A., Wright, D.A., Winfrey, R.J., Fu, F., Maeder, M.L., Joung, J.K., and Voytas, D.F. (2009). High-frequency modification of plant genes using engineered zinc-finger nucleases. Nature *459*, 442-445.

Tran, E.J., and Wente, S.R. (2006). Dynamic nuclear pore complexes: life on the edge. Cell *125*, 1041-1053.

Turker, M.S. (2002). Gene silencing in mammalian cells and the spread of DNA methylation. Oncogene *21*, 5388-5393.

Turner, M.A., Yang, X., Yin, D., Kuczera, K., Borchardt, R.T., and Howell, P.L. (2000). Structure and function of S-adenosylhomocysteine hydrolase. Cell Biochem Biophys *33*, 101-125.

Turner, M.A., Yuan, C.S., Borchardt, R.T., Hershfield, M.S., Smith, G.D., and Howell, P.L. (1998). Structure determination of selenomethionyl S-adenosylhomocysteine hydrolase using data at a single wavelength. Nat Struct Biol *5*, 369-376.

Tzfira, T., Tian, G.W., Lacroix, B., Vyas, S., Li, J., Leitner-Dagan, Y., Krichevsky, A., Taylor, T., Vainstein, A., and Citovsky, V. (2005). pSAT vectors: a modular series of plasmids for autofluorescent protein tagging and expression of multiple genes in plants. Plant Mol Biol *57*, 503-516.

Uetz, P. (2002). Two-hybrid arrays. Current Opinion in Chemical Biology *6*, 57-62.

Varrin, A.E., Prasad, A.A., Scholz, R.P., Ramer, M.D., and Duncker, B.P. (2005). A mutation in Dbf4 motif M impairs interactions with DNA replication factors and confers increased resistance to genotoxic agents. Mol Cell Biol *25*, 7494-7504.

Vaucheret, H., Vazquez, F., Crete, P., and Bartel, D.P. (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. Genes Dev *18*, 1187-1197.

Veraksa, A., Bauer, A., and Artavanis-Tsakonas, S. (2005). Analyzing protein complexes in Drosophila with tandem affinity purification-mass spectrometry. Developmental Dynamics *232*, 827-834.

Verdecia, M.A., Bowman, M.E., Lu, K.P., Hunter, T., and Noel, J.P. (2000). Structural basis for phosphoserine-proline recognition by group IV WW domains. Nat Struct Biol *7*, 639-643.

Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D. (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J *33*, 949-956.

von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S.G., Fields, S., and Bork, P. (2002). Comparative assessment of large-scale data sets of protein-protein interactions. Nature *417*, 399-403.

von Schwartzenberg, K., Kruse, S., Reski, R., Moffatt, B., and Laloue, M. (1998). Cloning and characterization of an adenosine kinase from Physcomitrella involved in cytokinin metabolism. Plant J *13*, 249-257.

Von Wettstein, D., Gough, S., and Kannangara, C.G. (1995). Chlorophyll biosynthesis. Plant Cell *7*, 1039-1057.

Waadt, R., Schmidt, L.K., Lohse, M., Hashimoto, K., Bock, R., and Kudla, J. (2008). Multicolor bimolecular fluorescence complementation reveals simultaneous formation of alternative CBL/CIPK complexes in planta. The Plant Journal *56*, 505-516.

Wakamiya, M., Blackburn, M.R., Jurecic, R., McArthur, M.J., Geske, R.S., Cartwright, J., Mitani, K., Vaishnav, S., Belmont, J.W., and Kellems, R.E. (1995). Disruption of the adenosine deaminase gene causes hepatocellular impairment and perinatal lethality in mice. Proceedings of the National Academy of Sciences of the United States of America *92*, 3673-3677.

Waldo, G.S., Standish, B.M., Berendzen, J., and Terwilliger, T.C. (1999). Rapid protein-folding assay using green fluorescent protein. Nat Biotechnol *17*, 691-695.

Walhout, A.J.M., Boulton, S.J., and Vidal, M. (2000). Yeast two-hybrid systems and protein interaction mapping projects for yeast and worm. Yeast *17*, 88-94.

Wallsgrove, R.M., Lea, P.J., and Miflin, B.J. (1983). Intracellular localization of aspartate kinase and the enzymes of threonine and methionine biosynthesis in green leaves. Plant Physiol *71*, 780-784.

Walter, M., Chaban, C., Schutze, K., Batistic, O., Weckermann, K., Nake, C., Blazevic, D., Grefen, C., Schumacher, K., Oecking, C., *et al.* (2004). Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. Plant Journal *40*, 428-438.

Wang, M., Borchardt, R.T., Schowen, R.L., and Kuczera, K. (2005). Domain motions and the open-to-closed conformational transition of an enzyme: a normal mode analysis of S-adenosyl-L-homocysteine hydrolase. Biochemistry *44*, 7228-7239.

Wang, Y.X., Shyy, J.Y.J., and Chien, S. (2008). Fluorescence proteins, live-cell imaging, and mechanobiology: Seeing is believing. Annu Rev Biomed Eng *10*, 1-38.

Wansink, D.G., Schul, W., van der Kraan, I., van Steensel, B., van Driel, R., and de Jong, L. (1993). Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. J Cell Biol *122*, 283-293.

Warthmann, N., Chen, H., Ossowski, S., Weigel, D., and Hervé, P. (2008). Highly specific gene silencing by artificial miRNAs in Rice. PLoS ONE *3*, e1829.

Weis, K. (2003). Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle. Cell *112*, 441-451.

Weretilnyk, E.A., Alexander, K.J., Drebenstedt, M., Snider, J.D., Summers, P.S., and Moffatt, B.A. (2001). Maintaining methylation activities during salt stress. The involvement of adenosine kinase. Plant Physiol *125*, 856-865.

Werner, T., Motyka, V., Laucou, V., Smets, R., Van Onckelen, H., and Schmulling, T. (2003). Cytokinin-deficient transgenic Arabidopsis plants show multiple developmental alterations indicating opposite functions of cytokinins in the regulation of shoot and root meristem activity. Plant Cell *15*, 2532-2550.

Williams, N.E. (2000). Immunoprecipitation procedures. In Methods in Cell Biology, Vol 62, pp. 449-453.

Winkler, S., Schwabedissen, A., Backasch, D., Bökel, C., Seidel, C., Bönisch, S., Fürthauer, M., Kuhrs, A., Cobreros, L., Brand, M., *et al.* (2005). Target-selected mutant screen by TILLING in Drosophila. Genome Research *15*, 718-723.

Witte, C.P., Noel, L.D., Gielbert, J., Parker, J.E., and Romeis, T. (2004). Rapid one-step protein purification from plant material using the eight-amino acid StrepII epitope. Plant Mol Biol *55*, 135-147.

Wortman, J.R., Haas, B.J., Hannick, L.I., Smith, R.K., Jr., Maiti, R., Ronning, C.M., Chan, A.P., Yu, C., Ayele, M., Whitelaw, C.A., *et al.* (2003). Annotation of the Arabidopsis genome. Plant Physiol *132*, 461-468.

Wright, D.A., Townsend, J.A., Winfrey, R.J., Irwin, P.A., Rajagopal, J., Lonosky, P.M., Hall, B.D., Jondle, M.D., and Voytas, D.F. (2005). High-frequency homologous recombination in plants mediated by zinc-finger nucleases. The Plant Journal *44*, 693-705.

Wu, X.Z., Li, F.L., Kolenovsky, A., Caplan, A., Cui, Y.H., Cutler, A., and Tsang, E.W.T. (2009). A mutant deficient in S-adenosylhomocysteine hydrolase in Arabidopsis shows defects in root-hair development. Botany *87*, 571-584.

Xu, Y., Piston, D.W., and Johnson, C.H. (1999). A bioluminescence resonance energy transfer (BRET) system: Application to interacting circadian clock proteins. Proceedings of the National Academy of Sciences of the United States of America *96*, 151-156.

Xylourgidis, N., and Fornerod, M. (2009). Acting out of character: regulatory roles of nuclear pore complex proteins. Dev Cell *17*, 617-625.

Yaish, M.W., Doxey, A.C., McConkey, B.J., Moffatt, B.A., and Griffith, M. (2006). Cold-active winter rye glucanases with ice-binding capacity. Plant Physiol *141*, 1459-1472.

Yamada, H., Suzuki, T., Terada, K., Takei, K., Ishikawa, K., Miwa, K., Yamashino, T., and Mizuno, T. (2001). The Arabidopsis AHK4 histidine kinase is a cytokinin-binding receptor that transduces cytokinin signals across the membrane. Plant and Cell Physiology *42*, 1017-1023.

Yang, X., and Borchardt, R.T. (2000). Overexpression, purification, and characterization of S-adenosylhomocysteine hydrolase from *Leishmania donovani*. Arch Biochem Biophys *383*, 272-280.

Yang, X., Hu, Y., Yin, D.H., Turner, M.A., Wang, M., Borchardt, R.T., Howell, P.L., Kuczera, K., and Schowen, R.L. (2003). Catalytic strategy of S-adenosyl-L-homocysteine hydrolase: transition-state stabilization and the avoidance of abortive reactions. Biochemistry *42*, 1900-1909.

Yin, D., Yang, X., Hu, Y., Kuczera, K., Schowen, R.L., Borchardt, R.T., and Squier, T.C. (2000). Substrate binding stabilizes S-adenosylhomocysteine hydrolase in a closed conformation. Biochemistry *39*, 9811-9818.

Yogev, O., and Pines, O. (2010). Dual targeting of mitochondrial proteins: Mechanism, regulation and function. Biochim Biophys Acta.

Yoshida, S., Ito, M., Callis, J., Nishida, I., and Watanabe, A. (2002). A delayed leaf senescence mutant is defective in arginyl-tRNA:protein arginyltransferase, a component of the N-end rule pathway in Arabidopsis. Plant J *32*, 129-137.

Yuan, C.S., Saso, Y., Lazarides, E., Borchardt, R.T., and Robins, M.J. (1999). Recent advances in S-adenosyl-L-homocysteine hydrolase inhibitors and their potential clinical applications. Expert Opin Ther Pat *9*, 1197-1206.

Zamyatnin, A.A., Solovyev, A.G., Bozhkov, P.V., Valkonen, J.P.T., Morozov, S.Y., and Savenkov, E.I. (2006). Assessment of the integral membrane protein topology in living cells. Plant Journal *46*, 145-154.

Zervos, A.S., Gyuris, J., and Brent, R. (1993). MXI1, a protein that specifically interacts with max to bind MYC-MAX recognition sites. Cell *72*, 223-232.

Zhang, X., Henriques, R., Lin, S.S., Niu, Q.W., and Chua, N.H. (2006). Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. Nat Protoc *1*, 641-646.

Zhong, S., Lin, Z., and Grierson, D. (2008). Tomato ethylene receptor-CTR interactions: visualization of NEVER-RIPE interactions with multiple CTRs at the endoplasmic reticulum. J Exp Bot *59*, 965-972.