University of Nevada, Reno

Genetic Engineering of Plant Seeds to Increase Thiamin (Vitamin B_1) Content

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Cell and Molecular Biology

by

Mohammad Yazdani

Dr. David Shintani/Dissertation Advisor

December, 2015

THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

MOHAMMAD YAZDANI

Entitled

Genetic Engineering of Plant Seeds to Increase Thiamin (Vitamin B1) Content

be accepted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

David Shintani, Advisor

Jeff Harper, Committee Member

Ian Wallace, Committee Member

Gary Blomquist, Committee Member

Stanley Omaye, Graduate School Representative

David W. Zeh, Ph. D., Dean, Graduate School

December, 2015

ABSTRACT

Thiamine (Vitamin B_1) in the form of thiamine pyrophosphate (TPP) is an essential cofactor for the function of numerous enzymes which are involved in central metabolism such as citric acid cycle, pentose phosphate pathway, Calvin cycle, isoprenoid biosynthesis, and branched-chain amino acid biosynthesis. All living organisms need thiamine. However, human and animals can synthesize TPP from thiamine, but they are not able to synthesize thiamine *de novo*. Therefore, human and animals must obtain thiamine from their diet to maintain a normal metabolism. Severe thiamine deficiency causes the lethal disease beriberi and Wernicke-Korsakoff syndrome in humans. The enzymes involved in thiamine de novo biosynthesis pathway are well known in microorganisms and plants, but little is known regarding the salvage pathways in plants. In order to have better insight about the thiamine salvage pathways in plants, the homologs of bacterial ThiM (thiazole kinase) were analyzed. It has been revealed that this protein in plants has thazole kinase activity which is important for thiamine salvage. In addition, analyzing the TenA_E proteins in plants shows that these proteins have amidohydrolase and aminohydrolase activity to form 4-amino-5-hydroxymethyl-2-methylpyrimidine (HMP) from the salvage of thiamine breakdown products. Thiamine plays a vital role in resistance against biotic and abiotic stresses in plants in addition to its role as a cofactor. It has been shown that elevated levels of thiamine content achieved by the seed overexpression of *Thi4*, *ThiC*, and *ThiE* genes can enhance the seed germination and seedlings viability under abiotic stress conditions. Additionally, thiamine and TPP over-producing lines shows altered seed carbon partitioning.

Dedication

I would like to dedicate my dissertation to my beloved parents, Esmaeil Yazdani and Banou Rokni Alaei, who believed in me and encouraged me throughout my education and for their never ending love; and to my beloved wife Maryam Rahmati Ishka, and sons, Danial and Aryan.

With lots of love,

Mohammad

12/16/2015

ACKNOWLEDGMENTS

I would never have been able to complete the following dissertation without support and guidance from several people who helped me tremendously during my Ph.D. study.

First, I would like to sincerely thank my advisor Dr. David Shintani for his excellent support, guidance, and patience during my experience in his lab. I would also like to thank my committee members Dr. Jeff Harper, Dr. Ian Wallace, Dr. Gary Blomquist, and Dr. Stanley Omaye for their advice and support that allowed me to pursue my research.

I wish to thank Dr. Christie Howard, Dr. Dylan Kosma, Dr. Upul Hathwaik and Dr. Rubi Figueroa Teran who as good friends and mentors. They were always helpful and willing to offer their time to me to better my insight on my research projects.

I would also like to thank to my good friends, Roy Sorensen, Shayne Urrutia, and Nathan Tatar who were very helpful not only in the lab but also out of the lab when I needed assistance. Additionally, I would like to thank other past and present Shintani lab members Sylvia Wines, Travis Garcia, Tharika Yoe, Quan Vong, Wendy Tan, Gareth Sexton, Mark Mendoza, and Samara Favela for their assistance in the lab.

In addition, I would also like to thank to Elizabeth Brown, James Davis, Richard Lohaus, and Sharon Young for their patience, the technical support, and instrumental assistance.

I would also like to thank to my parents, brothers, and sisters who were always supporting me and encouraging me with their best wishes.

Last but not least, I want to thank to my wife, Maryam Rahmati Ishka, for her unconditional love, encouragement, and support, as well as my sons Danial and Aryan for their patience and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Overview of thiamin metabolism in living organisms

INTRODUCTION

Thiamin (Vitamin B_1) in the form of cofactor thiamin pyrophosphate (TPP) is an essential metabolite for the function of numerous enzymes involved in carbohydrate, isoprenoid, lipid, and amino acid metabolism in living organisms (Krampitz, 1969; Jordan, 2003; LeClere et al., 2004; Nosaka, 2006; Goyer, 2010).

While thiamin can be synthesized by plants and microorganisms but animals are not able to synthesize this vitamin *de novo*. Hence, it must be taken up from their diet to drive their metabolism properly (Jurgenson et al., 2009; Yazdani et al., 2013). The enzymes in the human body can convert thiamin mono- and pyrophosphate into free thiamin which is the absorbable form of thiamin (Said and Mohammed, 2006). Vitamin B_1deficiency have been associated with human disorders including beriberi (Lonsdale, 2006), alcoholic ketoacidosis, Wernicke-Korsakoff syndrome and Alzheimer's disease (Mimori et al., 1996).

Thiamin biosynthesis in various organisms

In organisms with the capability of synthesizing Vitamin B_1, thiamin monophosphate (TMP) is produces by combining 4-amino-5-hydroxymethyl-2-methylpyrimidine pyrophosphate (HMP-PP) and 4-methyl-5-(β-hydroxyethyl) thiazole phosphate (HET-P) (Begley et al., 1999). TMP is then converted to the cofactor form, TPP, either by direct phosphorylation in enteric bacteria (rod-shaped Gram-negative

bacteria) (Cummings and MacFarlane, 1997) or by de-phosphorylation to thiamin followed by pyrophosphorylation in aerobic bacteria, yeast (Spenser e al., 1997) and plants (Roje, 2007). The two heterocyclic precursors of thiamin pyrophosphate, HMP-PP and HET-P, are biosynthesized through independent pathways.

Bacteria:

In enteric bacteria, (i.e. *E.coli*), HMP-PP is derived from 5-aminoimidazole ribonucleotide (AIR), a precursor shared between thiamin and *de novo* purine biosynthesis (Estramareix and Therisod, 1984; Estramareix and David, 1990). Hydroxymethylpyrimidine phosphate (HMP-P) synthase (ThiC) is essential for the conversion of AIR to HMP-P (Vanderhorn et al., 1993). HMP-P is subsequently phosphorylated to HMP-PP by ThiD (ThiE) which is a bifunctional enzyme with kinase activity (Roje, 2007).

Bacteria can synthesize HET-P using cysteine, 1-deoxy-D-xylulose-5-phosphate (DXP), and tyrosine (in *E. coli* and *Salmonella typhimurium*) or glycine (in *Bacillus subtilis)* in a complex reaction catalyzed by several enzymes (Settembre et al., 2003; Dorrestein et al., 2004; Begley, 2006). The enzymes ThiFSGH (Begley et al., 1999), ThiI (Palenchar et al., 2000), and IscS (Lauhon et al., 2000) are involved in thiazole formation (figure 2). Cys has been identified as the precursor of the thiazole sulfur atom. The sulfur atom is initially transferred to IscS, a cysteine desulfurase (Zheng et al., 1994), in the form of a persulfide of an active site Cys residue (Flint, 1996). The sulfur atom is then passed to ThiI (Webb et al., 1997) as a persulfide of residue Cys-456 (Palenchar et al., 2000). The sulfur atom is then transferred to the C terminus of ThiS, converting it to a thiocarboxylate. This final sulfur transfer reaction requires the activation of ThiS as its

acyladenylate. The adenylation reaction is catalyzed by ThiF, which has been observed to form a complex with ThiS (Taylor et al., 1998a). Finally, tyrosine is converted to dehydroglycine by ThiH (tyrosine lyase). The thiocarboxy C terminus of ThiS, DXP, and dehydroglycine are all combined together by ThiG (thiazole synthase) to form thiazole phosphate (Jurgenson et al., 2009).

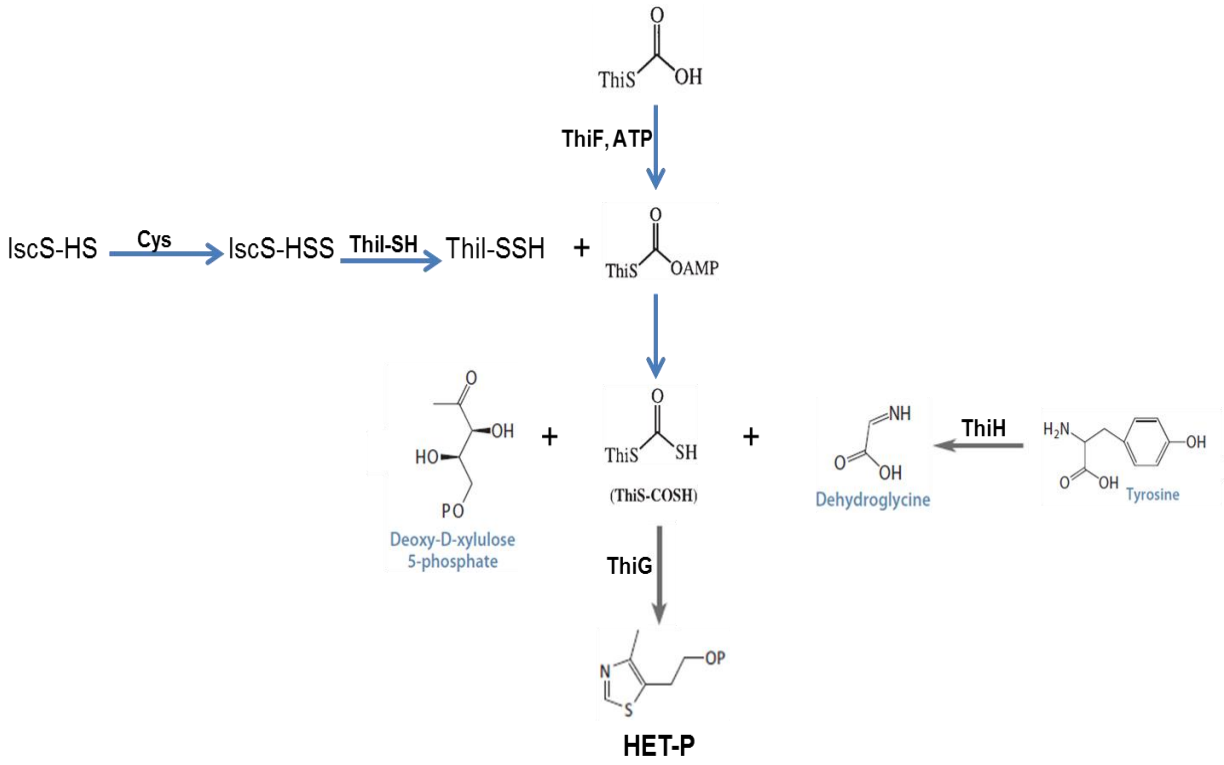


Figure 1: De novo HET-P biosynthesis pathway in bacteria. In contrast to yeast and plants, in bacteria six different enzymes are involved to synthesize thiazole phosphate.

Yeast:

Like all thiamin-synthesizing organisms, yeasts first separately synthesize two precursors, HET-P and HMP-PP, which are then condensed into TMP. This organism can use cysteine as a sulfur donor, glycine, and D-pentulose-5-phosphate as substrates for thiazole synthesis enzyme to synthesize HET-P. The D-pentulose-5-phosphate can be substitute with D-ribulose-5-phosphate or DXP, which show the link between thiamin biosynthesis and the pentose phosphate pathway (Hohmann and Meacock, 1998). A mechanism of thiazole synthesis was also proposed by Chatterjee et al. (2007). In this mechanism NAD^+ is a source for a five-carbon carbohydrate and the intermediate is an ADP adduct of 5-(2-hydroxyethyl)-4-methylthiazole-2-carboxylic acid.

HMP-P is synthesized in yeast cells by the enzyme Thi5 (Hohmann and Meacock, 1998) from histidine and pyridoxal-5-phosphate. The thiamin and vitamin B6 biosynthesis pathways are linked by Pyridoxal-5-phosphate (Zeidler et al., 2003). The next step is the phosphorylation by Thi20 enzyme to produce HMP-PP which is condensed with HET-P by a bifunctional enzyme Thi6 to form TMP (Hohmann and Meacock, 1998).

Yeasts are not able to do direct phosphorylation of TMP to obtain the TPP. Hence, TMP first is dephosphorylated to thiamin. The thiamin is then phosphorylated to form TPP (Nosaka et al., 1993). Yeast can also uptake the free thiamin from the environment to produce TPP (Enjo et al., 1997).

Plants:

In plants, the pyrimidine moiety of thiamin is synthesized via a pathway which is identical to bacterial pathway. The first step is the conversion of AIR to HMP-P which is

catalyzed by ThiC and requires S-adenosylmethionine (SAM) and reduced nicotinamide. The HET-P biosynthesis pathway in plants and yeasts is similar (Chatterjee et al., 2007, 2008b). Thi4 is the only enzyme involved in thiazole biosynthesis which has been identified so far in eukaryotes. Sequence homologs of the yeast Thi4 were also found in some plants such as *Zea mays* (Belanger et al., 1995), *Alnus glutinosa* (Ribeiro et al., 1996), *Arabidopsis thaliana* (Machado et al., 1996), and *Oryza sativa* (Wang et al., 2006). Thi4 cDNAs from *Z. mays*, *A. glutinosa*, and *A. thaliana* could complement the Thi4 yeast mutant deficient in Thi4. In addition, *Arabidopsis tz* (thiazole requirement) mutants are chlorotic and die early during development unless supplemented with thiazole or thiamin (Feenstra, 1964; Redei, 1965). Experimental evidence supports localization of HET-P biosynthesis in plastids via the yeast pathway (Julliard and Douce, 1991), as putative Thi4 from *Z. mays* (Belanger et al., 1995) and *A. thaliana* (Chabregas et al., 2001) were detected in these organelles using immunogold labeling. Thi4 was also detected in mitochondria in *A. thaliana* (Chabregas et al., 2001, 2003), suggesting these organelles as another site of HET-P biosynthesis in some plants.

In addition to their role in thiamin biosynthesis, Thi4 from *Arabidopsis* and yeast were shown to be involved in mitochondrial DNA damage tolerance (Machado et al., 1996, 1997). Therefore, the dual targeting of Thi4 enables this enzyme to function in protection against DNA damage when targeted to mitochondria (Chabregas et al., 2001, 2003) and to function in thiamin biosynthesis when targeted to chloroplasts (Julliard and Douce, 1991; Belanger et al., 1995; Chabregas et al., 2001).

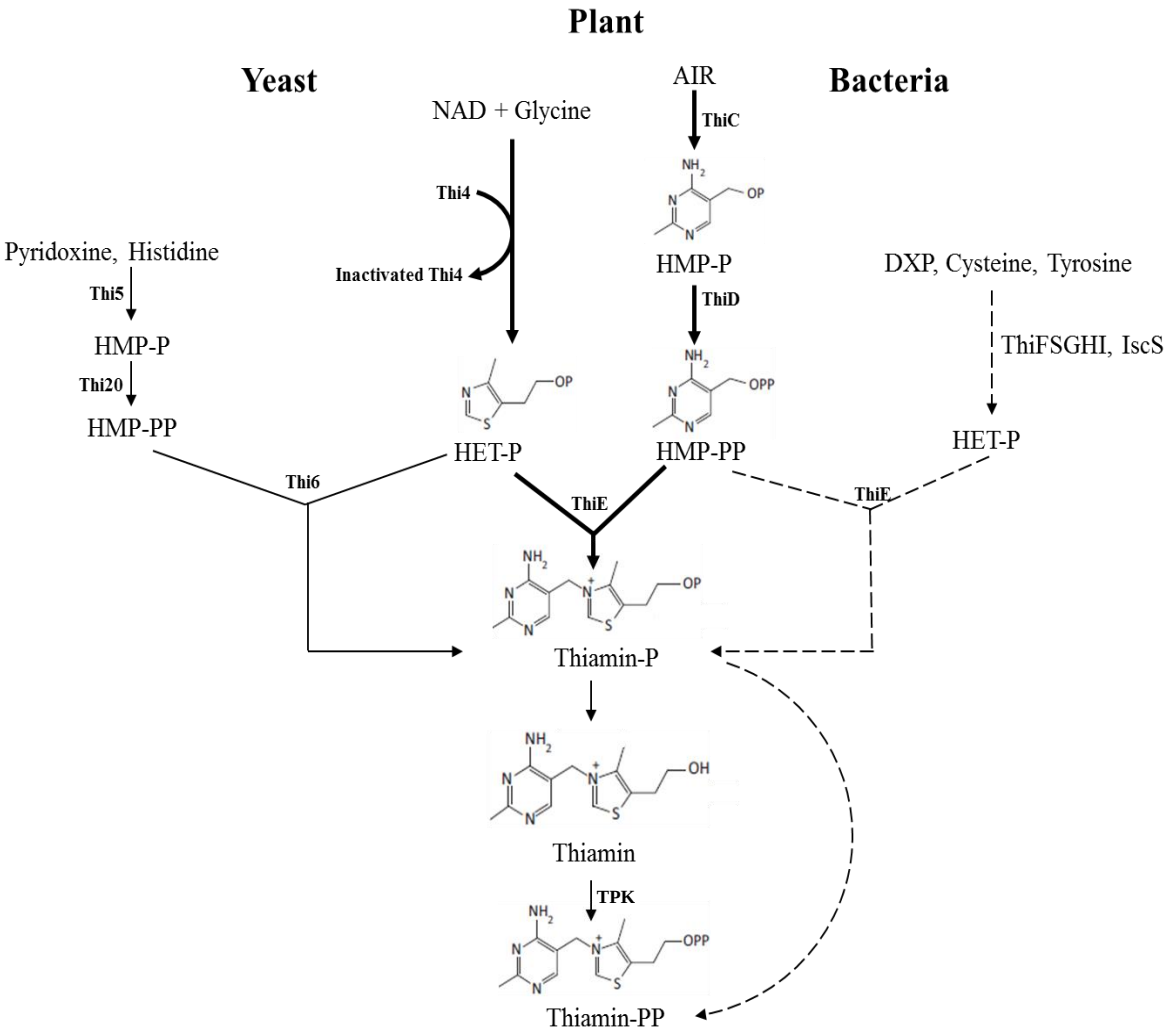


Figure 2: Thiamin biosynthesis pathways in different organisms. Plants synthesize HET-P and HMP-PP via the yeast and bacterial pathway, respectively. The noticeable feature of the pathway is that the Thi4 uses itself as sulphur donor for thiazole formation (Chatterjee et al., 2011).

THIAMIN DEGRADATION AND SALVAGE

Two classes of enzymes responsible for the thiamin degradation were identified in bacteria. These two enzymes are different in sequence and structure and comprise

thiaminase I and II (TenA) which are able to break the thiamin down to its thiazole and pyrimidine moieties (Jurgenson et al., 2009). In addition, thiaminases have different substrates for cleavage of the C-N bond between thiazole and pyrimidine heterocyclic rings in thiamin structure (Jurgenson et al., 2009). While thiaminase II only uses H_2O for thiamin degradation, thiaminase I degrades thiamin using the exchange of the thiazole ring with some molecules such as pyridine, cysteine, aniline, quinolone and dithiothreitol (Lienhard, 1970; Costello et al., 1996; Jurgenson et al., 2009). In bacteria such as *Bacillus subtilis* it has been shown that thiaminase II functions in the regeneration of the pyrimidine moiety of thiamin rather than in thiamin degradation (Jenkins et al., 2007). Although the significance of thiaminase II in salvaging of pyrimidine has been shown in bacteria (Jenkins et al., 2007) and yeast (Onozuka et al., 2008), the physiological significance of thiamin degradation has not been fully identified in bacteria (Fitzpatrick and Thore, 2014) and plants (Goyer, 2010).

Because of thiamin and TPP instability (McCourt et al., 2006, Fitzpatrick et al., 2012), microorganisms and plants have developed salvage mechanism to re-use the thiamin degradation products (thiazole and pyrimidine) for thiamin biosynthesis (Jurgenson et al., 2009)to save the metabolic cost of HET-P and HMP-PP re-synthetizing. Resent studies in fungi have shown that the Thi4 (HET-P synthase) is a single turnover (suicidal) enzyme which donates a sulfur atom from an active site Cys residue to the synthesis of HET-P (Chatterjee et al., 2011) therefore the production of a single thiazole molecule irreversibly inactivates HET-P synthase. Consequently, this makes *de novo* thiazole biosynthesis though Thi4 costly in metabolic terms (Praekelt et al., 1994). As

such, thiazole salvage in fungi and plants is highly energetically beneficial (Chatterjee et al., 2011; Gerdes et al., 2012; Yazdani et al., 2013).

Toward this end, the enzyme, thiazole kinase (ThiM) has been shown to convert salvaged thiazole to thiazole phosphate which can be condensed with HMP-PP to reform thiamin monophosphate (Jurgenson et al., 2009). HET kinase has been thoroughly biochemically characterized in *Bacillus subtilis* (Zhang et al., 1997; Campobasso et al, 2000) and *Salmonella typhimurium* (Petersen and Downs, 1997). The gene encoding the plant HET kinase has recently been identified and the corresponding enzyme has been functionally characterized (Yazdani et al, 2013).

Enzymes involved in HMP-PP recycling have also just recently identified in plants. These enzymes share a high degree of amino acid sequence similarity with the TenA protein family. In *Arabidopsis*, TenA_E has been shown catalyze the convertion of the thiamin degradation products *N*-formyl-4-amino-5-aminomethyl-2-methylpyrimidine (formylamino-HMP) and 4-amino-5-aminomethyl-2-methylpyrimidine (amino-HMP) to HMP (Zallot et al., 2014).

ROLE OF THIAMIN IN STRESS TOLERANCE

Abiotic stresses such as drought, salinity, extreme temperatures, chemical toxicity and oxidative stress are serious threats to agriculture. Abiotic stress is the main cause of crop loss and is attributed to reducing crop yields globally by more than 50% (Boyer, 1982; Bray et al., 2000). Furthermore, these stresses can cause morphological, physiological, biochemical and molecular changes which adversely affect plant growth and productivity (Wang et al. 2001a).

In addition to its role as a cofactor, thiamin has been demonstrated to play a key role in biotic and abiotic stresses in plants (Ahn et al., 2005; Dong et al., 2015; Sayed and Gadallah, 2002; Tunc-Ozdemir et al., 2009; Conrath et al., 2002). In relashonship to biotic stresss, Ahn and co-workers (2005) defined a novel function for thiamin in plants as a defense activator against biotic stresses. They reported that application of thiamin to rice, tobacco, tomato, cucumber, and *Arabidopsis* triggers the expression of the genes involved in systemic acquired resistance and could inhibit diseases caused by semibiotrophic and biotrophic pathogens. Additionally, it has been recently reported that overexpression of *Thi4* and *ThiC* genes using constitutive promoters in *Arabidopsis* plants could increase the leaves total thiamin content up to 3.4-fold and after inoculation of these plants with *Pseudomonas syringae*, the plants had lower populations of this pathogen compared to the wild type plants (Dong et al., 2015).

Regarding abiotic stresses, several studies have shown that plants respond to stress by increasing thiamin biosynthesis. Tunc-Ozdemir et al, (2009) demonstrated that various abiotic stress treatments including high salt, drought, high_light and, paraquat treatment; include plants to synthesize elevated levels of thiamin. This stress induced increase thiamin was shown to correlate with increased gene expression of thiamin biosynthetic genes. Additionally, Rapala-Kozik et al, (2008) showed that the thiamin content of *Zea mays* seedlings increased under drought, high salt, and oxidative stress conditions and this increase was accompanied by an increase in the activity of some thiamin biosynthetic enzymes.

These observations suggest a rolr for thiamin in abiotic stress tolerance. In support this idea, Sayed and Gadallah, (2002) showed that the application of thiamin on

shoot and root of salt-stressed sunflower plants could alleviate the detrimental effects of salt stress on plant growth by improving of cell membrane integrity and K^+ uptake. Tunc-Ozdemir et al. (2009) later showed that the exogenous application of thiamin could protect *Arabidopsis* from paraquat induced oxidative damage. The authors discuss the possibility that thiamin pyrophosphate (TPP) could be a critical factor in plant stress tolerance as its levels were higher than the free thiamin and TMP. This could be related to the role of TPP in the production of the reductants NADH and NADPH to combat oxidative stress damage (Tunc-Ozdemir et al., 2009; Rapala-Kozik et al., 2012).

Taken together, the data regarding the effects of thiamin on biotic and abiotic stress tolerance is mostly based on the results obtained from feeding studies using the exogenous thiamin in bacteria, yeasts, and plants growth medium and also based on the complementation assays using thiamin biosynthetic genes in the thiamin mutants of various living organisms.

REFERENCES

Ahn, I.P., Kim, S., Lee, Y.H. (2005). Vitamin B1 functions as an activator of plant disease resistance. Plant Physiol. 138: 1505-1515.

Begley, T.P. (2006). Cofactor biosynthesis: an organic chemist's treasure trove. Nat. Prod. Rep. 23: 15-25.

Begley, T.P., Downs, D.M., Ealick, S.E., McLafferty, F.W., Van Loon, A., Taylor, S., Campobasso, N., Chiu, H.J., Kinsland, C., Reddick, J.J., Xi, J. (1999). Thiamin biosynthesis in prokaryotes. Arch. Microbiol. 171: 293-300.

Belanger, F., Leustek, T., Chu, B., Kirz, A. (1995). Evidence for the thiamine biosynthetic pathway in higher plant plastids and its developmental regulation. Plant Mol. Biol. 29: 809-821.

Boyer, J.S. (1982). Plant productivity and environment. Science, 218: 443-448.

Bray, E.A., Bailey-Serres, J., Weretilnyk, E. (2000). Responses to abiotic stresses. In: Gruissem W, Buchannan B, Jones R (eds) Biochemistry and molecular biology of plants. American Society of Plant Physiologists, Rockville, MD, pp 1158-1249.

Campobasso, N., Mathews I.I., Begley, T.P., Ealick, S.E. (2000). Crystal structure of 4-methyl-5-β-hydroxyethylthiazole kinase from *Bacillus subtilis* at 1.5 ˚A resolution. Biochemistry, 39: 7868-7877.

Chabregas, S.M., Luche, D.D., Farias, L.P., Ribeiro, A.F., Van Sluys, M.A., Menck, C.F., Silva-Filho, M.C. (2001). Dual targeting properties of the N-terminal signal sequence of *Arabidopsis thaliana* THI1 protein to mitochondria and chloroplasts. Plant Mol. Biol. 46: 639-650.

Chabregas, S.M., Luche, D.D., Van Sluys, M.A., Menck, C.F.M., Silva-Filho, M.C. (2003). Differential usage of two in-frame translational start codons regulates subcellular localization of *Arabidopsis thaliana* THI1. J. Cell Sci. 116: 285-291.

Chatterjee, A., Abeydeera, N.D., Bale, S., Pai, P.J., Dorrestein, P.C., Russell, D.H., Ealick, S.E., Begley, T.P. (2011). *Saccharomyces cerevisiae* THI4p isasuicide thiamine thiazole synthase. Nature, 478: 542-546.

Chatterjee, A., Jurgenson, C.T., Schroeder, F.C., Ealick, S.E., Begley, T.P. (2007). Biosynthesis of thiamin thiazole in eukaryotes: conversion of NAD to an advanced intermediate. J. Am. Chem. Soc. 129: 2914-2922.

Chatterjee, A., Schroeder, F.C., Jurgenson, C.T., Ealick, S. E., Begley, T.P. (2008b). Biosynthesis of the thiamin-thiazole in eukaryotes: identification of a thiazole tautomer intermediate. J. Am. Chem. Soc. 130: 11394-11398.

Conrath, U., Pieterse, C.M., Mauch-Mani, B. (2002). Priming in plant–pathogen interactions. Trends Plant Sci. 7: 210-216.

Costello, C.A, Kelleher, N.L., Abe, M., McLafferty, F.W., Begley, T.P. (1996). Mechanistic studies on thiaminase I. Overexpression and identification of the active site nucleophile. J. Biol. Chem. 271: 3445-3452.

Cummings, J.H., MacFarlane, G.T. (1997). "Role of intestinal bacteria in nutrient metabolism". Clin. Nutr. 16: 3-9.

Dong, W., Stockwell, V.O., Goyer, A. (2015). Enhancement of thiamin content in *Arabidopssis thaliana* by metabolic engineering. Plant Cell Physiol. doi: 10.1093/pcp/pcv148.

Dorrestein, P.C., Zhai, H., McLafferty, F.W., Begley, T.P. (2004). The Biosynthesis of the Thiazole Phosphate Moiety of Thiamin. Chem. Biol. 11: 1373-1381.

Enjo, F., Nosaka, K., Ogata, M., Iwashima, A. Nishimura, H. (1997). Isolation and characterization of a thiamin transport gene, *THI10*, from *Saccharomyces cerevisiae*. J. Biol. Chem. 272: 19165-19170.

Estramareix, B., David, S. (1990). Conversion of 5-aminoimidazole ribotide to the pyrimidine of thiamin in enterobacteria: study of the pathway with specifically labeled samples of riboside. Biochim. Biophys. Acta. 1035: 154-60.

Estramareix, B., Therisod, M. (1984). Biosynthesis of thiamin: 5-aminoimidazole ribotide as the precursor of all the carbon atoms of the pyrimidine moiety. J. Am. Chem. Soc. 106: 3857-3860.

Feenstra, W. (1964). Isolation of nutritional mutants in *Arabidopsis thaliana*. Genetica, 35: 259-269.

Fitzpatrick, T.B., Basset, G.J., Borel, P., Carrari, F., DellaPenna, D., Fraser, P.D., Hellmann, H., Osorio, S., Rothan, C., Valpuesta, V., Caris-Veyrat, C., Fernie, A.R. (2012). Vitamin deficiencies in humans: can plant science help? Plant Cell. 24: 395-414.

Fitzpatrick, T.B., Thore, S. (2014). Complex behavior: from cannibalism to suicide in the vitamin B1 biosynthesis world. Curr. Opin. Struc. Biol. 29: 34-43.

Flint, D.H. (1996). *Escherichia coli* contains a protein that is homologous in function and N-terminal sequence to the protein encoded by the *nifS* gene of *Azotobacter vinelandii* and that can participate in the synthesis of the Fe-S cluster of dihydroxy-acid dehydratase. J. Biol. Chem. 271: 16068-16074.

Gerdes, S., Lerma-Ortiz, C., Frelin, O., Seaver, S.M., Henry, C.S., de Crécy-Lagard, V., Hanson, A.D. (2012). Plant B vitamin pathways and their compartmentation: a guide for the perplexed. J. Exp. Bot. 63: 5379-5395.

Goyer, A. (2010). Thiamine in plants: Aspects of its metabolism and functions. Phytochemistry, 71: 1615-1624.

Hohmann, S., Meacock, P.A. (1998). Thiamin metabolism and thiamin diphosphate-dependent enzymes in the yeast *Saccharomyces cerevisiae*: genetic regulation. Biochim. Biophys. Acta. 1385: 201-219.

Jenkins, A.H., Schyns, G., Potot, S., Sun, G., Begley, T.P. (2007). A new thiamin salvage pathway. Nat. Chem. Biol. 3: 492-497.

Jordan, F. (2003). Current mechanistic understanding of thiamin diphosphate-dependent enzymatic reactions. Nat. Prod. Rep. 20: 184-201.

Julliard, J.H., Douce, R. (1991). Biosynthesis of the thiazole moiety of thiamin (vitamin B1) in higher plant chloroplasts. PNAS. 88: 2042-2045.

Jurgenson, C.T., Begley, T.P., Ealick, S.E. (2009). The structural and biochemical foundations of thiamin biosynthesis. Annu. Rev. Biochem. 78: 569-603.

Krampitz, L.O. (1969). Catalytic functions of thiamin diphosphate. Annu. Rev. Biochem. 38: 213-240.

Lauhon, C.T., Kambampati, R. (2000). The *iscS* gene in *Escherichia coli* is required for the biosynthesis of 4-thiouridine, thiamin, and NAD. J. Biol. Chem. 275: 20096-20103.

LeClere, S., Rampey, R.A., Bartel, B. (2004). IAR4, a gene required for auxin conjugate sensitivity in *Arabidopsis*, encodes a pyruvate dehydrogenase E1alpha homolog. Plant Physiol. 135: 989-999.

Lienhard, G.E. (1970). Kinetic evidence for a (4-amino-2-methyl-5-pyrimidinyl)methyl-enzyme intermediate in the thiaminase I reaction. Biochemistry, 9: 3011-3020.

Lonsdale, D. (2006). A review of the biochemistry, metabolism and clinical benefits of thiamine and its derivatives. eCAM. 3: 49-59.

Machado, C.R., Costa de Oliveira., R., Boiteux, S., Praekelt, U.M., Meacock P.A., Menck, C.F. M. (1996). *thi1*, a thiamine biosynthetic gene in *Arabidopsis thaliana*, complements bacterial defects in DNA repair. Plant Mol. Biol. 31: 585-593.

Machado, C.R., Praekelt, U.M., Costa de Oliveira, R., Barbosa, A.C., Byrne, K.L., Meacock, P. A., Menck, C.F.M. (1997). Dual role for the yeast *THI4* gene in thiamine biosynthesis and DNA damage tolerance. J. Mol. Biol. 273:114-121.

McCourt, J.A., Nixon, P.F., Duggleby, R.G. (2006). Thiamin nutrition and catalysis-induced instability of thiamin diphosphate. Br. J. Nutr. 96: 636-638.

Mimori, Y., Katsuoka, H., Nakamura, S. (1996). Thiamine therapy in Alzheimer's disease. Metab. Brain Dis. 11: 89-94.

Nosaka, K. (2006). Recent progress in understanding thiamin biosynthesis and its genetic regulation in *Saccharomyces cerevisiae*. Appl. Microbiol. Biotechnol. 72: 30-40.

Nosaka, K., Kaneko, Y., Nishimura, H., Iwashima, A. (1993). Isolation and characterization of a thiamin pyrophosphokinase gene, *THI80*, from *Saccharomyces cerevisiae*. J. Biol. Chem. 268:17440-17447.

Onozuka, M., Konno, H., Kawasaki, Y., Akaji, K., Nosaka, K. (2008). Involvement of thiaminase II encoded by the THI20 gene in thiamin salvage of *Saccharomyces cerevisiae*. FEMS Yeast Res. 8: 266-275.

Palenchar, P.M., Buck, C.J., Cheng, H., Larson, T.J., Mueller, E.G. (2000). Evidence that ThiI, an enzyme shared between thiamine and 4-thiouridine biosynthesis, may be a sulfurtransferase that proceeds through a persulfide intermediate. J. Biol. Chem. 275: 8283-8286.

Petersen, L.A., Downs, D.M. (1997). Identification and characterization of an operon in *Salmonella typhimurium* involved in thiamine biosynthesis. J. Bacteriol. 179:4894-4900.

Praekelt, U.M., Byrne, K.L., Meacock, P.A. (1994). Regulation of *THI4* (*MOL1*), a thiamine-biosynthetic gene of *Saccharomyces cerevisiae*. Yeast, 10: 481-490.

Rapala-Kozik, M., Kowalska, E., Ostrowska, K. (2008). Modulation of thiamine metabolism in *Zea mays* seedlings under conditions of abiotic stress. J. Exp. Bot. 59: 4133-4143.

Rapala-Kozik, M., Wolak, N., Kujda, M., Banas, A.K. (2012). The upregulation of thiamine (vitamin B1) biosynthesis in *Arabidopsis thaliana* seedlings under salt and osmotic stress conditions is mediated by abscisic acid at the early stages of this stress response. BMC Plant Biol. 12: 1-14.

Redei, G. (1965). Genetic blocks in the thiamine synthesis of the angiosperm *Arabidopsis*. Am. J. Bot. 52: 834-841.

Ribeiro, A., Praekelt, U., Akkermans, A.D., Meacock, P.A., van Kammen, A., Bisseling, T., Pawlowski, K. (1996). Identification of agthi1, whose product is involved in biosynthesis of the thiamine precursor thiazole, in actinorhizal nodules of *Alnus glutinosa*. Plant J. 10: 361-368.

Roje, S. (2007). Vitamin B biosynthesis in plants. Phytochemistry, 68: 1904-1921.

Said, H.M., Mohammed, Z.M. (2006). Intestinal absorption of water soluble vitamins: an update. Curr. Opin. Gastroenterol. 22: 140-146.

Sayed, S.A., Gadallah, M.A.A. (2002). Effects of shoot and root application of thiamin on salt-stressed sunflower plants. Plant Growth Regul. 36: 71-80.

Settembre, E., Begley, T.P., Ealick, S.E. (2003). Structural biology of enzymes of the thiamin biosynthesis pathway. Curr. Opin. Struct. Biol. 13: 739-747.

Spenser, I.D., White, R.L. (1997). Biosynthesis of Vitamin B_1 (Thiamin): An instance of biochemical diversity. Angew. Chem. Int. Ed. Engl. 36:1032-1046.

Taylor, S.V., Kelleher, N.L., Kinsland, C., Chiu, H.J., Costello, C.A., Backstrom, A.D., McLafferty, F.W., Begley, T. (1998 a). Thiamin biosynthesis in *Escherichia coli*: identification of ThiS thiocarboxylate as the immediate sulfur donor in the thiazole formation. J Biol Chem. 273: 16555-16560.

Tunc-Ozdemir, M., Miller, G., Song, L., Kim, J., Sodek, A., Koussevitzky, S., Misra, A.N., Mittler, R., Shintani, D. (2009). Thiamin confers enhanced tolerance to oxidative stress in *Arabidopsis*. Plant Physiol. 151: 421-432.

Vanderhorn, P.B., Backstrom, A.D., Stewart, V., Begley, T.P. (1993). Structural genes for thiamine biosynthetic enzymes (thiCEFGH) in *Escherichia coli* K-12. J. Bacteriol. 175: 982-992.

Wang, G., Ding, X., Yuan, M., Qiu, D., Li, X., Xu, C., Wang, S. (2006). Dual function of rice *OsDR8* gene in disease resistance and thiamine accumulation. Plant Mol. Biol. 60: 437-449.

Wang, W.X., Vinocur, B., Shoseyov, O., Altman, A. (2001a). Biotechnology of plant osmotic stress tolerance: physiological and molecular considerations. Acta Hort. 560:285-292.

Webb, E.A., Claas, K., Downs, D.M. (1997). Characterization of *thiI*, a new gene involved in thiazole biosynthesis in *Salmonella typhimurium*. J Bacteriol. 179: 4399-4402.

Yazdani, M., Zallot, R., Tunc-Ozdemir M., de Crécy-Lagard, V., Shintani D.K., Hanson, A.D. (2013). Identification of the thiamin salvage enzyme thiazole kinase in Arabidopsis and maize. Phytochemistry, 94: 68-73.

Zallot, R., Yazdani, M., Goyer, A.**,** Ziemak, M.J., Guan, J.C., McCarty, D.R., De Crécy-Lagard, V., Gerdes, S., Garrett, T.J., Benach, J., Hunt, J.F., Shintani, D.K., Hanson, A.D. (2014). Salvage of the thiamin pyrimidine moiety by plant TenA proteins lacking an active-site cysteine. Biochem. J. 463: 145-155.

Zeidler, J., Sayer, B.G., Spenser, I.D. (2003). Biosynthesis of vitamin B1 in yeast. Derivation of the pyrimidine unit from pyridoxine and histidine. Intermediacy of urocanic acid. J. Am. Chem. Soc. 125: 13094-13105.

Zhang, Y., Taylor, S.V., Chiu, H.J., Begley, T.P. (1997). Characterization of the *Bacillus subtilis* thiC operon involved in thiamine biosynthesis. J. Bacteriol. 179:3030-3035.

Zheng, L., White, R.H., Cash, V.L., Dean, D.R. (1994). Mechanism for the Desulfurization of L-Cysteine Catalyzed by the nifS Gene Product. Biochemistry, 33: 4714-4720.

CHAPTER 2

Identification of the thiamin salvage enzyme thiazole kinase in *Arabidopsis* and maize

Mohammad Yazdani, Rémi Zallot, Meral Tunc-Ozdemir, Valérie de Crécy-Lagard, David K. Shintani , Andrew D. Hanson

This chapter was published in Phytochemistry, (2013) 94:68-73.

Author contribution: Bioinformatics and *E.coli* mutant studies were performed by our collaborators in the University of Florida, Gainesville, USA.

ABSTRACT

The breakdown of thiamin (vitamin B_1) and its phosphates releases a thiazole moiety, 4-methyl-5-(2-hydroxyethyl)thiazole (THZ), that microorganisms and plants are able to salvage for re-use in thiamin synthesis. The salvage process starts with the ATP-dependent phosphorylation of THZ, which in bacteria is mediated by ThiM. The *Arabidopsis* and maize genomes encode homologs of ThiM (At3g24030 and GRMZM2G094558, respectively). Plasmid-driven expression of either plant homolog restored the ability of THZ to rescue *Escherichia coli thiM* deletant strains, showing that the plant proteins have ThiM activity in vivo. Enzymatic assays with purified recombinant proteins confirmed the presence of THZ kinase activity. Furthermore, ablating the *Arabidopsis* At3g24030 gene in a thiazole synthesis mutant severely impaired rescue by THZ. Collectively, these results show that ThiM homologs are the main source of THZ kinase activity in plants and are consequently crucial for thiamin salvage.

INTRODUCTION

Thiamin (vitamin B_1), as its active diphosphate form, is an essential cofactor for various enzymes that make or break C–C bonds (Müller et al., 2009b). Plants and most microorganisms can synthesize thiamin *de novo*, but animals cannot and thus require it in the diet (Jurgenson et al., 2009). The *de novo* biosynthesis pathway and most of the enzymes involved are known in microorganisms and plants (Jurgenson et al., 2009 and Goyer, 2010). In this pathway, the thiazole and pyrimidine moieties of thiamin are made separately and coupled together to form thiamin phosphate, which is then converted to the diphosphate. The plant thiamin biosynthesis pathway is shown in Figure 1.

Thiamin and thiamin diphosphate are chemically and enzymatically labile (McCourt et al., 2006, Goyer, 2010 and Fitzpatrick et al., 2012), and microorganisms and plants have the capacity to re-use the thiazole and pyrimidine fragments from thiamin breakdown for thiamin synthesis (Li and Rédei, 1969 and Jurgenson et al., 2009). For the thiazole moiety, 4-methyl-5-(2-hydroxyethyl)thiazole (THZ), the key salvage step is phosphorylation to give 4-methyl-5-(2-phosphonooxyethyl)thiazole (THZ-P) (Figure. 1).

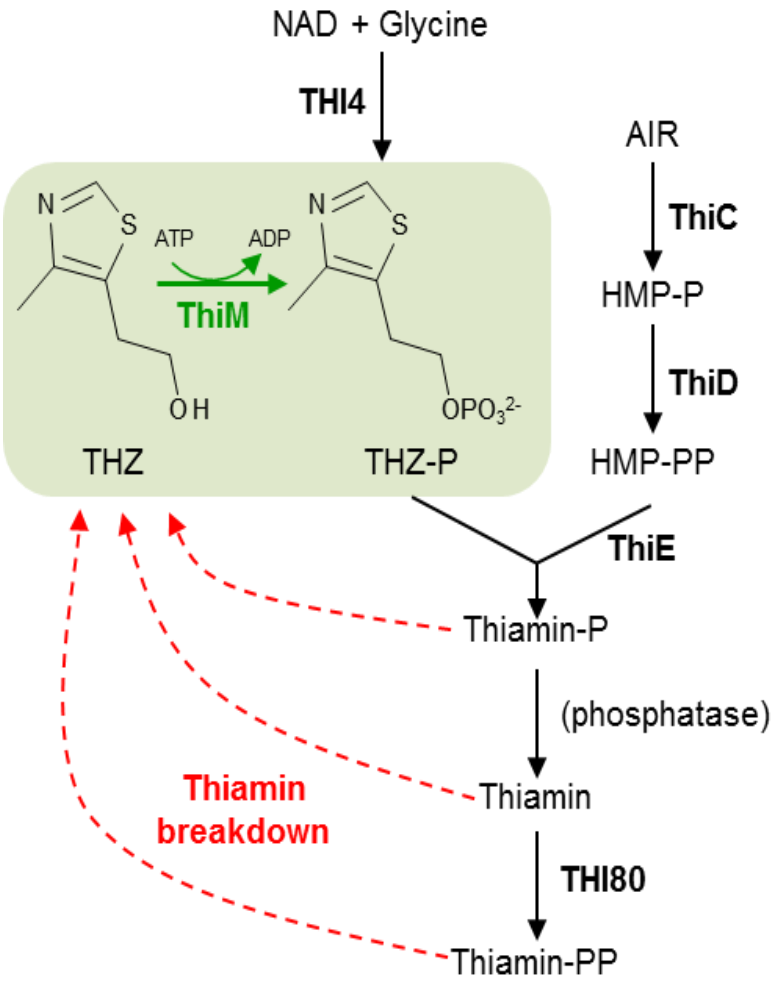


Figure 1: Thiamin biosynthesis and salvage pathways that occur in plants. Enzymes are identified by their gene names in *E. coli* or yeast, as follows: ThiC, phosphomethylpyrimidine synthase; ThiD, phosphomethylpyrimidine kinase; ThiE, thiamin-phosphate diphosphorylase; THI4, single turnover thiazole biosynthesis enzyme; ThiM, thiazole kinase; THI80, thiamin diphosphokinase. Compounds: AIR, 5-aminoimidazole ribotide; HMP-P, 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate; HMP-PP, 4-amino-5-hydroxymethyl-2-methylpyrimidine diphosphate; THZ, 4-methyl-5-(2-hydroxyethyl)-thiazole; THZ-P, 4-methyl-5-(2-phosphonooxyethyl)thiazole; Thiamin-P, thiamin monophosphate; Thiamin-PP, thiamin diphosphate (9). Note that in *E. coli* THZ-P is made by a different route (not shown) that involves ThiG, ThiH, and ThiS.

The THZ kinase (EC 2.7.1.50) responsible for this step is encoded by *thiM* in *Escherichia coli* and other bacteria, and by the 3′ region of *THI6* in *Saccharomyces cerevisiae* (Mizote and Nakayama, 1989, Nosaka et al., 1994, Jurgenson et al., 2009 and Paul et al., 2010). Nothing is yet known, however, about the THZ kinase enzyme in plants and the plant THZ kinase gene has not been identified (Goyer, 2010). Identifying this gene has become particularly worthwhile in light of the recent realization that THZ salvage in fungi and plants is highly energetically beneficial (Chatterjee et al., 2011 and Gerdes et al., 2012). The THZ synthesis protein THI4 (Figure 1) is a single-turnover enzyme from which a cysteine residue provides the THZ sulfur atom (Chatterjee et al., 2011), so that producing a single THZ molecule irreversibly inactivates a THI4 polypeptide, comprised of ~350 amino acids. Therefore, each THZ molecule salvaged in effect saves the energy cost of re-synthesizing a whole 350-residue protein.

Here, the *Arabidopsis* and maize (*Zea mays*) THZ kinase genes were identified by demonstrating that ThiM homologs from these plants can functionally replace *E. coli* ThiM, that the recombinant plant proteins have THZ kinase activity, and that ablating the *Arabidopsis* gene results in severe loss of ability to salvage THZ from the medium.

EXPERIMENTAL

Bioinformatics

The sequence of maize ThiM was from Maizesequence.org (http://maizesequence.org/index.html); other sequences were from NCBI or the Joint Genome Institute (http://www.jgi.doe.gov/). Sequence alignments were made with ClustalW and phylogenetic trees were constructed using MEGA5 (Tamura et al., 2011). Comparative analysis of bacterial genomes was made using the SEED database and its

tools (http://pubseed.theseed.org/) (Aziz et al., 2012). *Arabidopsis* transcript data were from CSB.DB (http://csbdb.mpimp-golm.mpg.de/csbdb/dbxp/ath/int/ath_xpsgq.html); maize transcript data were from qTeller (http://qteller.com/qteller3/).

E. coli strains and growth conditions

E. coli strains were grown in Luria–Bertani medium (LB) at 37 °C. Media were solidified with 15 g/l of agar. Kanamycin (Kan, 50 μg/ml) and ampicillin (Amp, 100 μg/ml) were added as required. M9 medium containing 0.4% glucose, prepared as described (Sambrook and Russell, 2001), was supplemented with 3.5 μM thiamin. HCl (Sigma–Aldrich) or 3.5 μM THZ (Sigma–Aldrich). Strains Δ*thiM* JW2091-2, Δ*thiG* JW5549-1, and Δ*thiH* JW3953-2 and the corresponding wild type strain BW25113 were from the Keio collection (Baba et al., 2006). These strains were checked by PCR (primers are listed in Table S1), and for growth on M9 medium with and without thiamin. The Δ*thiM* strain was used to make the double mutants Δ*thiG* Δ*thiM* and Δ*thiH* Δ*thiM* by recombineering (Datsenko and Wanner, 2000). Briefly, the Δ*thiM* strain was first transformed with pCP20 to eliminate the Kan^R cassette by Flp recombinase (Cherepanov and Wackernagel, 1995). Transformants were selected on LB plus Amp at 30 °C, restreaked twice on LB, grown at 42 °C, and then checked for absence of Kan and Amp resistance. A Δ*thiM* strain that had lost Kan resistance was transformed with pKD46 and grown at 30 °C on LB plus Amp. After selection and restreaking of transformants, expression of the Red genes (γ, β, and *exo*) was induced in liquid culture (LB plus Amp) with 1 mM arabinose. Cells were then transformed separately with two PCR products obtained by amplifying the Δ*thiG* and Δ*thiH* Kan deletion cassette plus flanking regions, allowing recombination of these amplicons, and generating the double deletants Δ*thiG*

Δ*thiM* and Δ*thiH* Δ*thiM* (primers are listed in Table S1). Transformants were restreaked and selected on LB plus Kan at 42 °C and then checked for their phenotype. The double deletant strains were checked for growth on M9 medium alone or plus 3.5 μM THZ or 3.5 μM thiamin.

Construction of complementation vectors

E. coli (strain MG1655) *thiM* (EcThiM) was amplified from genomic DNA. The *Arabidopsis* At3g24030 cDNA (AtThiM) was amplified from clone U61784 obtained from ABRC. The maize GRMZM2G094558 cDNA (ZmThiM) was amplified from a tassel primordium cDNA library obtained from R.J. Schmidt (University of California San Diego). Primers used for amplifications are listed in Table S1. For functional complementation assays, coding sequences were cloned into pBAD24 (Guzman et al., 1995). EcThiM was cloned into the *Nco*I and *Hin*dIII restriction sites and used as a positive control; cloning into the *Nco*I site changed the second codon from CAA (Gln) to GAA (Glu). AtThiM and ZmThiM were cloned into *Nco*I and *Hin*dIII sites. The resulting constructs (pBAD24::EcThiM, pBAD24::AtThiM, and pBAD24::ZmThiM) were confirmed by sequencing.

Functional complementation experiments

The Δ*thiG* Δ*thiM* and Δ*thiH* Δ*thiM* double deletants were transformed with pBAD24 alone (negative control) or with pBAD24 containing EcThiM (positive control), AtThiM, or ZmThiM. Complementation tests were made by streaking selected transformants on M9 plates containing 0.4% glucose, minus or plus 3.5 μM THZ or 3.5 μM thiamin. Arabinose was not added to induce gene expression as the basal

expression level sufficed. To avoid THZ or thiamin carry-over, two successive restreaks were made. Images are from the second restreak, after incubation for 32 h at 37 °C.

Production of recombinant ThiM proteins

The *Arabidopsis* and maize ThiM sequences were amplified by PCR using primers given in Table S1 and cloned between the *Nco*I and *Xho*I sites of pET28b (+) for *Arabidopsis* and *Eco*RI and *Xho*I sites of pET28b (+) for maize, which adds a C-terminal hexahistidine tag. Sequence-verified constructs were electroporated into *E. coli* strain BL21 (DE3). For protein isolation, transformants were grown at 37 °C in LB medium until OD_{600} reached about 0.6, at which point isopropyl β-d-1-thiogalactopyranoside was added to give a final concentration of 1 mM. After a further 6 h at 30 °C, cells were harvested by centrifugation and disrupted using a sonicator (60 Sonic Dismembrator, Fisher Scientific) in lysis buffer (50 mM sodium phosphate, pH 8.0, 0.5 M NaCl). After centrifugation to clear, proteins were purified using ProBond™ nickel-chelating resin (Invitrogen) according to the manufacturer's protocol. Eluted fractions were desalted using PD-10 columns (GE Healthcare) equilibrated in 50 mM Tris–HCl, pH 8.0; freshly desalted proteins were used for activity assays.

Thiazole kinase assays

Activity was assayed at 25 °C either spectrophotometrically using a pyruvate kinase-lactate dehydrogenase-coupled system or radiometrically using labeled THZ. The coupled spectrophotometric assay (Reddick et al., 1998) uses ADP and NADH as substrates, and THZ-dependent consumption of NADH is measured by monitoring the decrease in absorbance at 340 nm. The assay buffer was 100 mM Tris–HCl, pH 8.0. The radiometric assay uses [^{14}C]THZ as substrate and measures the phosphate ester product,

THZ-P, after ion exchange separation (Atkins and Canvin, 1971). Assays (100 µl) contained 50 mM Tris–HCl, pH 8.0, 10 mM ATP, 10 mM $MgCl_2$, 0.2 mM THZ containing 0.1 µCi hydroxyethyl[2-^{14}C]thiazole (Moravek Biochemicals, Inc.), and 100 µg of *Arabidopsis* ThiM or 30 µg of maize ThiM. Reactions were run for 1 h, then held on ice until they were applied to 0.5 cm (diameter) × 1 cm (height) columns of AG1-X8, 200–400 mesh, in the formate form. After washing the columns with H_2O 3 ml, THZ-P was eluted with 1 M ammonium formate (3 ml) and quantified by scintillation counting.

***Arabidopsis* mutants and growth conditions**

The SALK_123358C homozygous T-DNA knockout line for ThiM (At3g24030) was obtained from the *Arabidopsis* Biological Resource Center (ABRC) and verified by semiquantitative RT-PCR. The *tz-1* (Li and Rédei, 1969) THI4 mutant was obtained from ABRC (stock number CS3375) and crossed with the ThiM knockout; the resulting F1 plants were then selfed to produce the double homozygous mutant. Seed of the single and double mutants, and wild type Columbia, were cultured on MS medium (Murashige and Skoog, 1962) supplemented with 2% sucrose, 100 mg/l inositol, 0.5 mg/l nicotinic acid and 0.5 mg/l pyridoxine, with or without 100 µM thiamin or various concentrations of THZ. The plated seeds were vernalized at 4 °C for two days and germinated at 25 °C under constant light (150 µE m^{-2} s^{-1}). Plants had grown for 15 days when pictures were taken.

ThiM transcript analysis

Wild type Columbia plants were grown for two weeks as above with thiamin or THZ supplements at the indicated levels. RNA was extracted and ThiM transcript levels

were quantified as described (Tunc-Ozdemir et al., 2009). In brief, mRNA was extracted using the RNeasy plant mini kit (Qiagen). First-strand cDNA was synthesized using (2 μg) of total RNA, GeneRacer oligo dT Primer, and SuperScript II RNase H−reverse transcriptase (Invitrogen). Amplification and detection of the transcript levels were performed with an ABI7000 real-time PCR system (PE Biosystems) using the primers given in Table S1. The reaction conditions were 55 °C for 2 min, followed by 95 °C for 15 min and then 50 cycles of 76 °C for 45 s, 94 °C for 45 s, and 56 °C for 45 s. The *eEF-1α* gene tagged with VIC dye (Qiagen) was included in every well as an internal control. Relative-fold changes in expression levels were calculated using the $2^{-\Delta\Delta CT}$ method as described in Applied Biosystems User Bulletin No. 2 (P/N 4303859). PCR for each sample was repeated at least in duplicate. The significance of differences between data sets was evaluated by Student's *t* test.

RESULTS and DISCUSION

Identification of plant ThiM homologs

BlastP searches of *Arabidopsis* and maize protein databases using the *E. coli* ThiM sequence detected single homologs (At3g24030 and GRMZM2G094558) that are 42% identical to the *E. coli* protein and 61% identical to each other. Neither protein has apparent targeting signals or has so far been detected in organelles (Sun et al., 2009 and Tanz et al., 2013). Similar ThiM homologs were found in other angiosperms, gymnosperms, and lower plants (Figure 2A). The phylogeny of the plant sequences generally tracks organismal phylogeny (Figure 2A). These observations indicate that a ThiM-like protein, probably located in the cytosol, has been present in plants since their origin and has persisted throughout their subsequent radiation.

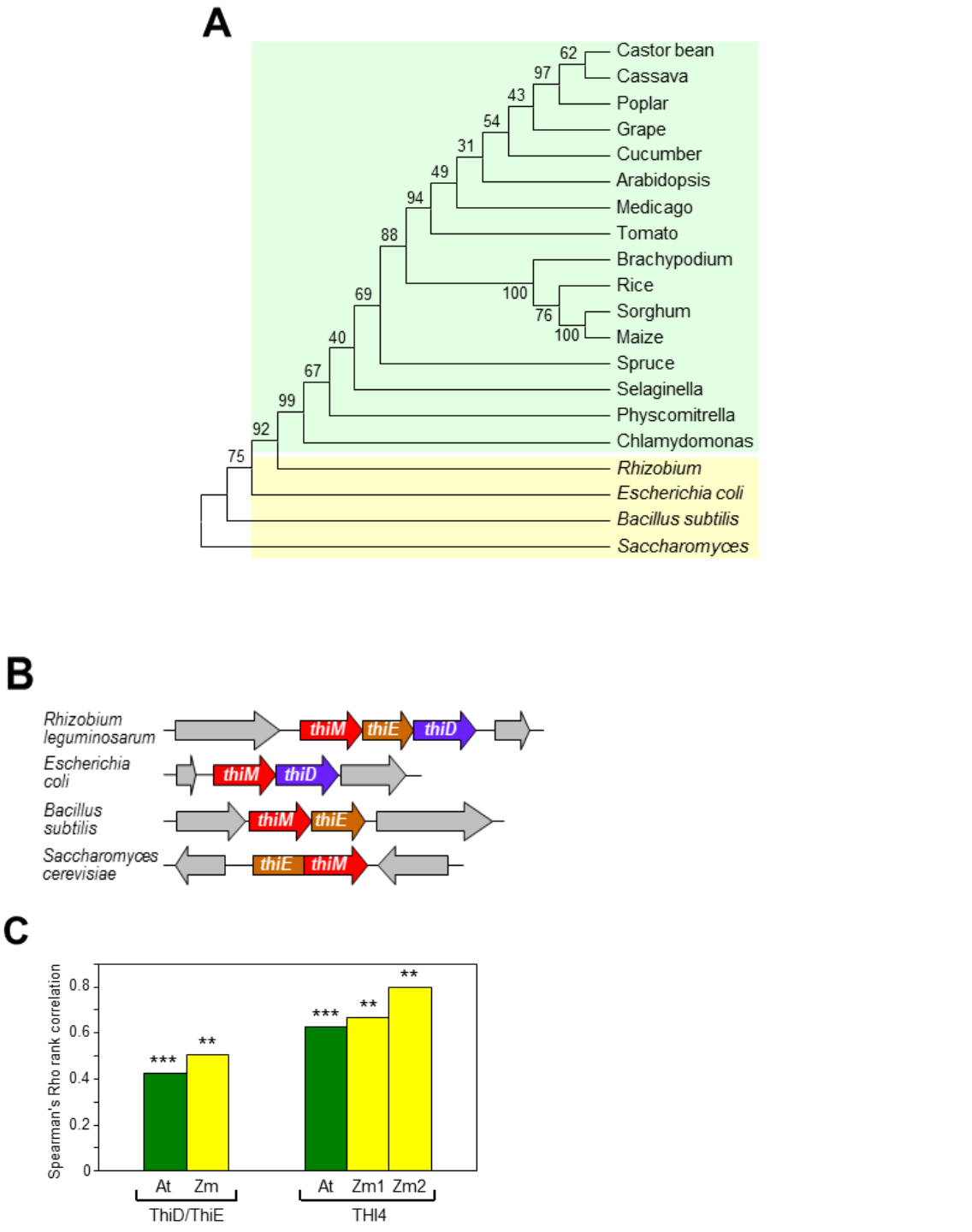


Figure 2: Phylogenetic analysis of plant and microbial ThiM proteins and comparative genomic evidence connecting them with the thiamin pathway. (A) Phylogenetic tree of plant ThiM homologs (highlighted in green) and four microbial ThiM proteins with experimentally proven THZ kinase activity. Sequences were aligned with ClustalW; the tree was constructed by the neighbor joining method with MEGA5. Bootstrap values (%)

for 1,000 replicates are shown next to nodes. Only tree topology is shown. *Rhizobium*, *Rhizobium leguminosarum* bv. viciae 3841; *Saccharomyces*, *Saccharomyces cerevisiae*. (B) Clustering in operonic structures of bacterial *thiM* genes with *thiE* (thiamin-phosphate diphosphorylase) and/or *thiD* (phosphomethylpyrimidine kinase) genes, and the *thiE-thiM* (*THI6*) fusion gene in yeast. (C) Correlations between expression in various organs of the *Arabidopsis* (At) and maize (Zm) ThiM homolog genes (At3g24030, GRMZM2G094558) and the genes encoding the ThiD/ThiE fusion enzyme (At1g22940, GRMZM2G401934) or the THI4 thiazole synthesis enzyme, of which *Arabidopsis* has one (At5g54770) and maize has two (Zm1, GRMZM2G018375; Zm2, GRMZM2G074097). Asterisks indicate rank correlation coefficient values that are significant at $^{**}P < 0.01$ or $^{***}P < 0.001$.

Apart from experimental evidence (Mizote and Nakayama, 1989, Nosaka et al., 1994, Zhang et al., 1997 and Karunakaran et al., 2006), comparative genomic evidence robustly links microbial *thiM* genes with thiamin, based on chromosomal clustering, or fusion, of *thiM* with the thiamin synthesis/salvage genes *thiD* or *thiE* (Figure 2B). The clustering of these genes in operonic structures, or their fusion, results in co-expression. There is analogous evidence for co-expression of the plant ThiM homologs with the same thiamin genes. Thus, expression of the *Arabidopsis* homolog is strongly positively correlated ($P < 0.001$) with that of the bifunctional ThiD-ThiE gene, and there is a similar correlation in maize (Figure 2C). Moreover, the expression of the *Arabidopsis* ThiM homolog is highly correlated with that of the de novo THZ synthesis gene THI4 (Figure 2C), and there are comparable correlations in maize between ThiM and both of its THI4 genes (Figure 2C). These comparative transcriptomic data reinforce the evidence from sequence homology that plant ThiM homologs are good candidates for the missing THZ kinase genes.

Plant ThiM proteins can functionally replace ThiM in *E. coli*

A functional complementation assay for ThiM (thiazole kinase) activity was developed in *E. coli* by deleting either of two de novo THZ synthesis genes, *thiG* or *thiH* (Leonardi et al., 2003), as well as *thiM*. The resulting double deletant strains can neither synthesize nor salvage THZ and consequently can be rescued by thiamin but not by THZ (Figure 3).

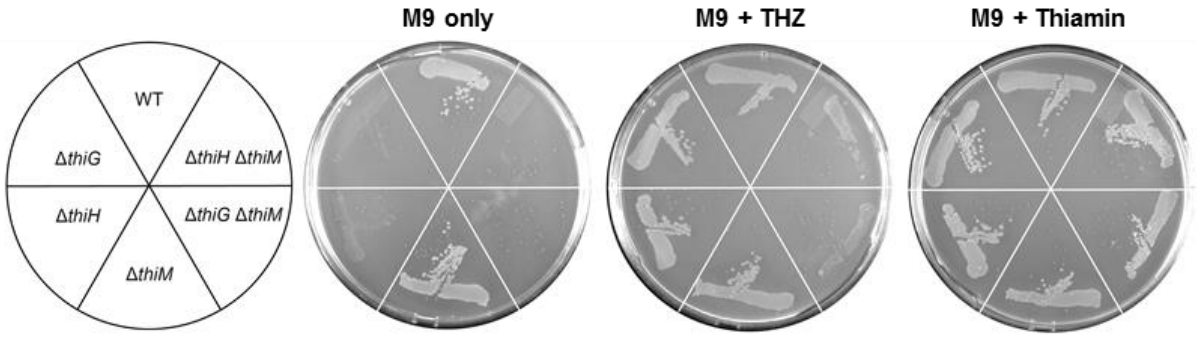


Figure 3: Growth phenotypes of the double deletant strains used for complementation assays**.** Wild type (WT) *E. coli* BW25113, Δ*thiG*, Δ*thiH*, and Δ*thiM* single deletant strains, and Δ*thiG* Δ*thiM* and Δ*thiH* Δ*thiM* double deletant strains were streaked twice in succession (to avoid THZ or thiamin carry-over) on M9 minimal medium alone or supplemented with 3.5 µM THZ or 3.5 µM thiamin. Images are for the second streak after incubation for 32 h at 37ºC.

As expected, expression of the native *E. coli thiM* gene from a plasmid restored the capacity of both strains for rescue by THZ (Figure 4), establishing that the Δ*thiG* Δ*thiM* and Δ*thiH ΔthiM* strains are suitable for complementation tests of ThiM activity. Both strains were used to test *Arabidopsis* At3g24030 and maize GRMZM2G094558 expression constructs. Each construct restored growth on THZ as effectively as the *E. coli thiM* positive control in both the Δ*thiG* Δ*thiM* strain (Figure 4A) and the Δ*thiH*

ΔthiM strain (Figure 4B). These results establish that either plant gene can functionally replace *thiM* in *E. coli*.

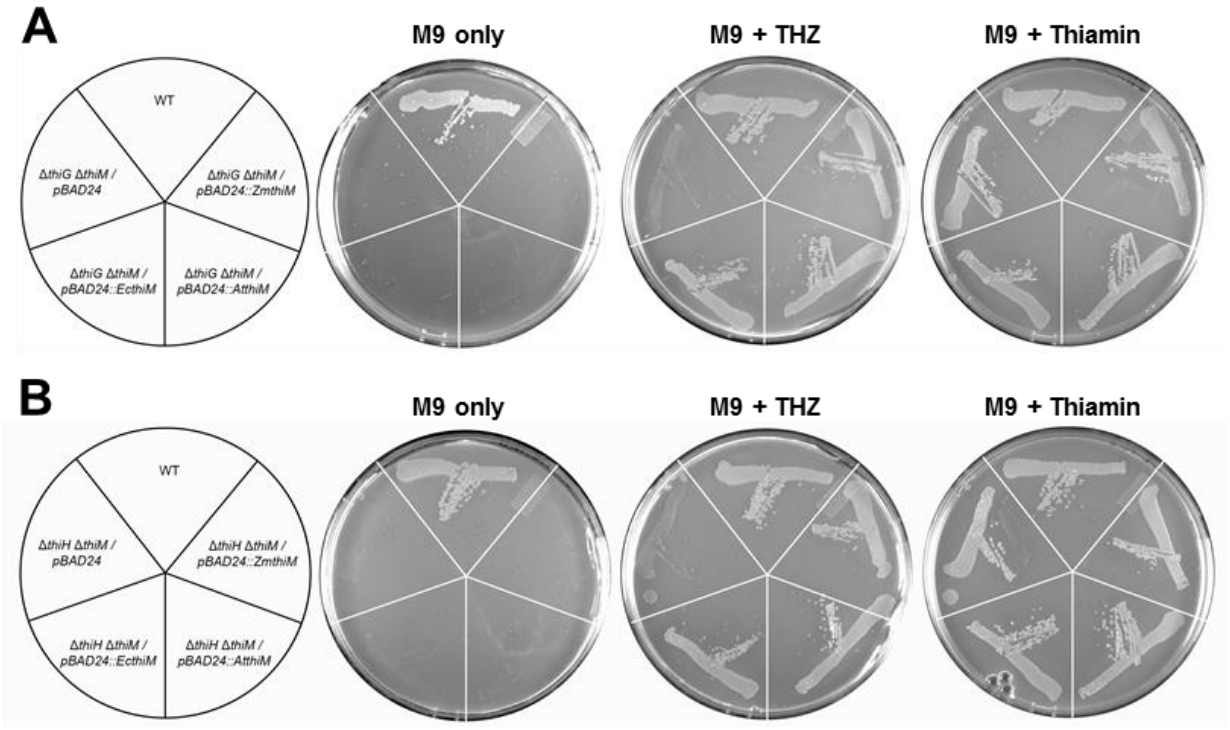


Figure 4: Functional complementation of the *E. coli ΔthiG ΔthiM* and *ΔthiH ΔthiM* double deletant strains by ThiM homologs from *Arabidopsis* and maize. The wild type (WT) strain BW25113 and the Δ*thiG* Δ*thiM* strain (A) or the Δ*thiH* Δ*thiM* strain (B) harboring pBAD24 alone (vector, negative control) or encoding *E. coli ThiM* (*ThiM, positive control*), At3g24030 (At) or GRMZM2G094558 (Zm) were streaked on M9 minimal medium alone or plus 3.5 μM THZ or 3.5 μM thiamin.

Recombinant plant ThiM proteins have THZ kinase activity

To directly test whether *Arabidopsis* and maize ThiM homologs have THZ kinase activity, both proteins, with a C-terminal hexahistidine tag, were expressed in *E. coli*, purified by Ni^{2+}-affinity chromatography (Figure 5), and assayed for THZ kinase activity.

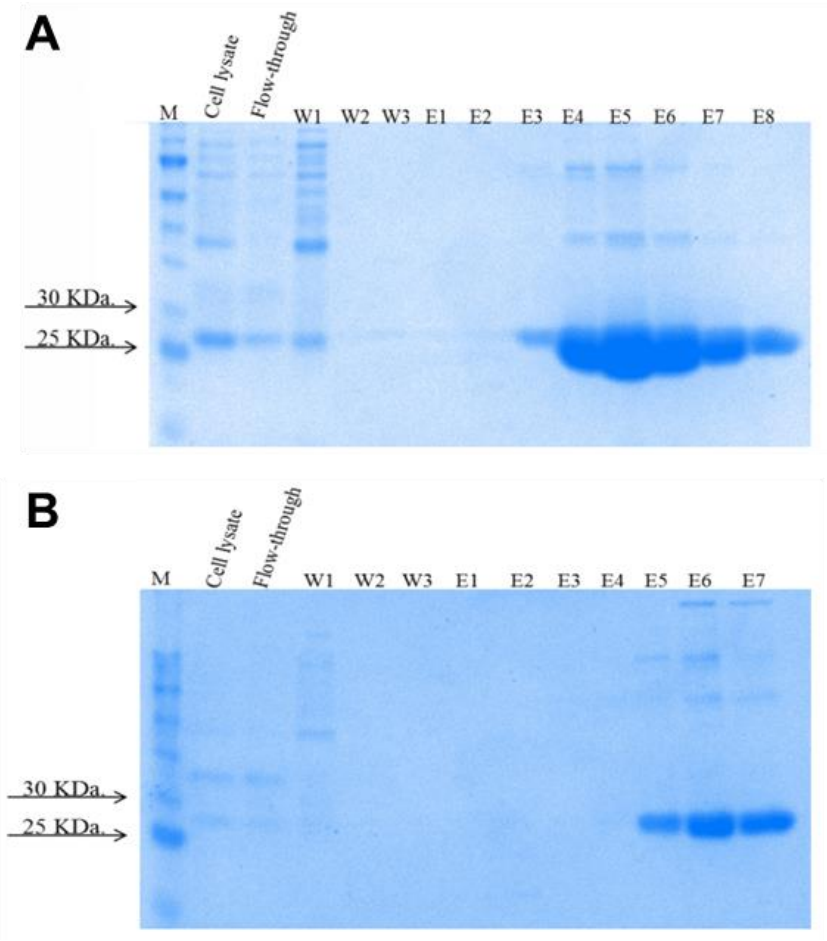


Figure 5: Ni2+-affinity purification of recombinant *Arabidopsis* ThiM (A) and maize ThiM (B)**.** Whole cell lysate, the column flow-through, and sucessive wash (W) and elution (E) fractions were separated by SDS-PAGE. Gels were stained with Coomassie blue. M, molecular mass markers.

Activity was measured using a coupled spectrophotometric assay in which thiazole-dependent ADP formation is coupled via pyruvate kinase and lactate dehydrogenase to NADH consumption (Reddick et al., 1998), or using a radiometric assay in which the THZ-P product is quantified after ion-exchange separation (Atkins and

Canvin, 1971). Both assays detected activity with both enzymes; the activity of the maize enzyme was much higher than that of the *Arabidopsis* enzyme (Figure 6).

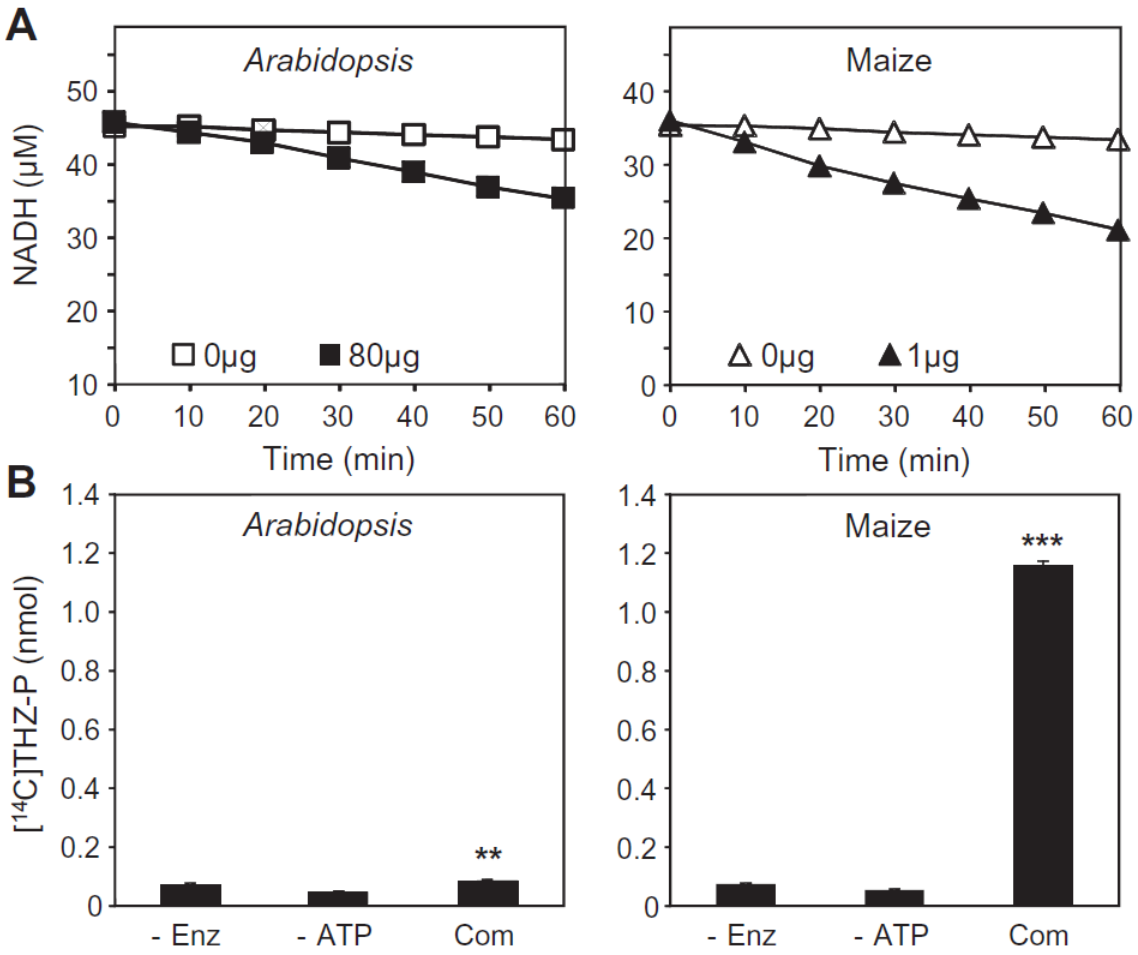


Figure 6: Detection of thiazole kinase activity of *Arabidopsis* and maize ThiM proteins. (A) Spectrophotometric assay in which ADP production is coupled to NADH consumption. Assays contained 80 μg of *Arabidopsis* ThiM, 1 μg of maize ThiM, or no enzyme as a control. Note the far higher activity of the maize enzyme. (B) Radiometric assay in which THZ-P formation is measured after separating it from THZ using an ion exchange column. The fraction containing THZ-P also contains traces of labeled THZ breakdown products, which result in a low level of radioactivity in control assays without enzyme (−Enz) or without ATP (−ATP). Complete assays (Com) contained [^{14}C]thiazole, ATP, and enzyme (100 μg for *Arabidopsis*, 30 μg for maize). Data are mean values of triplicate determinations ± SE. [^{14}C]THZ-P formation significantly exceeded the ^{14}C activity detected in the blank assays at $P < 0.01$ (**) for *Arabidopsis* and $P < 0.001$ (***) for maize (Student's *t*-test).

The coupled assay was used for kinetic characterization of both enzymes (Table 1). Their K_m values for THZ fell within the range (6-68 μM) found for other THZ kinases; the k_{cat} value for maize was in the middle of the reported range (0.14-2.2 s^{-1}) and that for *Arabidopsis* was just below it (Kawasaki, 1993, Campobasso et al., 2000, Wrenger et al., 2006 and Müller et al., 2009a). In connection with the low k_{cat} value for *Arabidopsis* ThiM, it may be noted that the value for another thiamin salvage kinase (hydroxymethylpyrimidine kinase) is equally low (Reddick et al., 1998). Overall, the plant ThiM enzymes can be categorized as typical THZ kinases.

Table 1
Kinetic parameters of thiazole kinase activity[a] of *Arabidopsis* and maize ThiM.

Plant	Substrate	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} μM^{-1})
Arabidopsis	Thiazole	30.5 ± 4.7	0.068 ± 0.01	22×10^{-4}
	ATP	44.8 ± 13.6	0.069 ± 0.02	15×10^{-4}
Maize	Thiazole	21.8 ± 3.4	1.79 ± 0.30	82×10^{-3}
	ATP	17.5 ± 5.0	2.39 ± 0.60	140×10^{-3}

[a] Data are the means of three experiments ± SE.

***Arabidopsis* ThiM is important for THZ salvage in planta**

In order to determine the role of ThiM in vivo, we first confirmed by RT-PCR analysis that a Salk homozygous T-DNA knockout line for At3g24030 lacked detectable ThiM mRNA (Figure 7), and crossed this line with a THI4 (TZ, At5g54770) mutant known to require THZ (Li and Rédei, 1969) in order to obtain a double homozygous mutant.

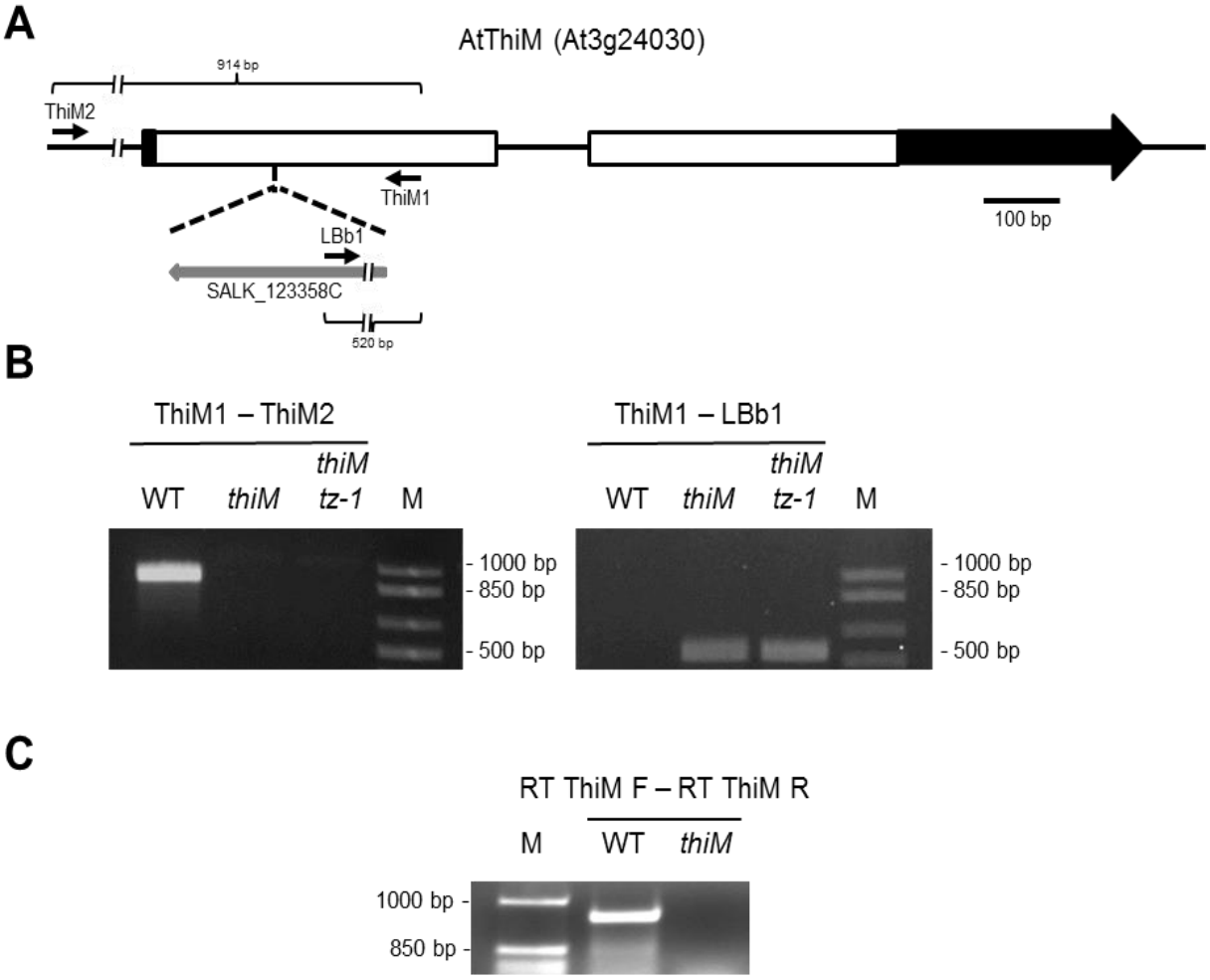


Figure 7: Characterization of *Arabidopsis* ThiM knockout lines. (A) Scheme showing the position of the T-DNA insertion (gray arrow) in the ThiM gene (At3g24030) in the SALK_123358C line. Introns are represented by lines and exons by boxes; 5′- and 3′-untranslated regions are in black. The positions of primers used for genomic PCR are shown by black arrows. (B) PCR analysis of genomic DNA of representative wild type (WT), *thiM* mutant, and *thiM tz-1* double mutant plants. (C) RT-PCR analysis of the full length ThiM transcript in representative wild type and *thiM* mutant plants. M, molecular size markers (bp).

The single and double mutants, plus a wild type control, were then cultured on MS medium plus or minus thiamin or THZ. In the absence of thiamin or THZ, the THI4

and the ThiM/THI4 mutants did not grow, as was expected, and the ThiM single mutant grew normally (Figure 8). That this single mutant shows no obvious phenotype indicates that salvage of endogenous THZ is not required for normal growth of cultured *Arabidopsis* plantlets, as is the case for bacterial ThiM mutants grown in favorable conditions (Figure 3) (Mizote and Nakayama, 1989 and Petersen and Downs, 1997).

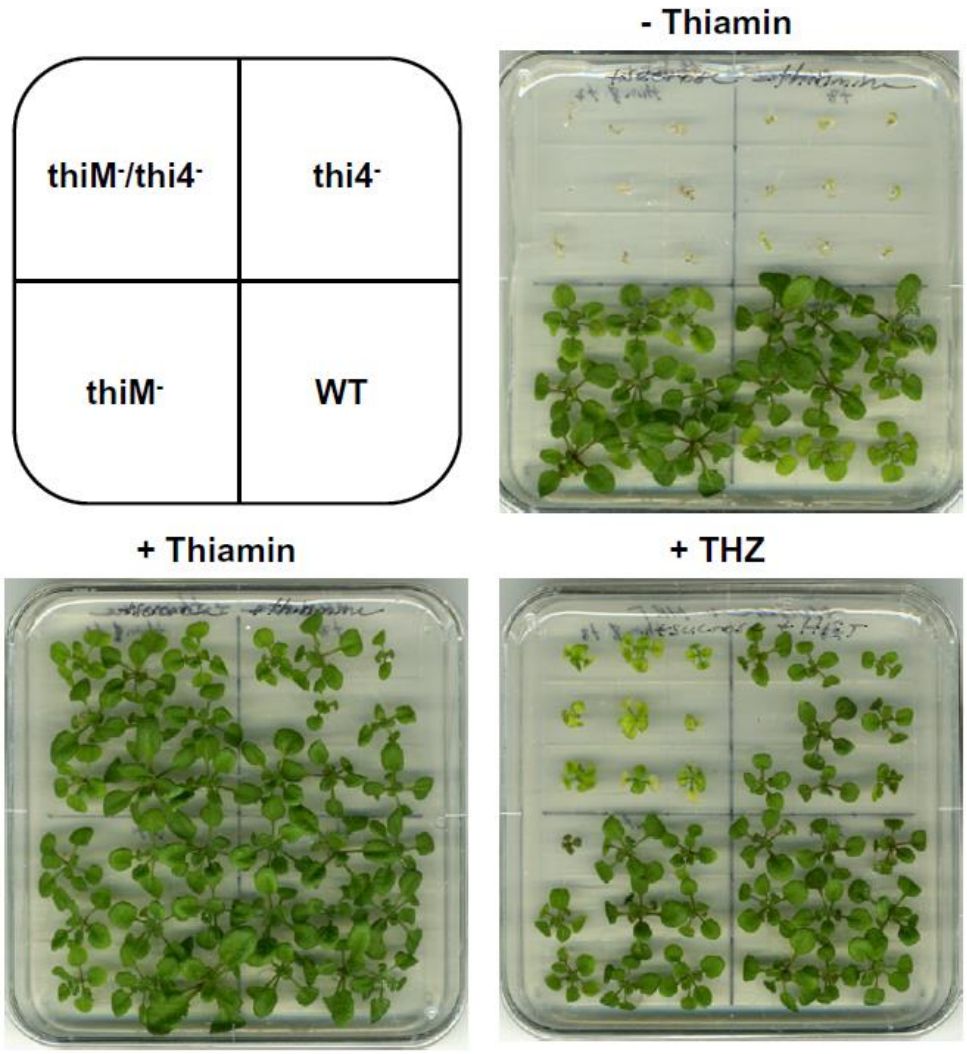


Figure 8: Effect of ablating *Arabidopsis* ThiM on salvage of supplied THZ. Wild type *Arabidopsis*, the THI4 mutant, the ThiM mutant, and the double mutant were cultured for 15 days on MS medium alone or supplemented with 0.1 mM thiamin or 0.4 mM THZ. The experiment was repeated with 0.8 mM and 1 mM THZ, with similar results; lower THZ concentrations (40 μM or 4 μM) allowed almost no growth of the double mutant (Figure 9).

The single THI4 mutant was rescued by thiamin or THZ, whereas the ThiM/THI4 double mutant was rescued completely by thiamin but only partially by THZ (Figure 8). Even at very high THZ concentrations (0.4 to 1 mM), the double mutant remained stunted and chlorotic; lower THZ levels (4 µM or 40 µM) allowed almost no growth (Figure 9).

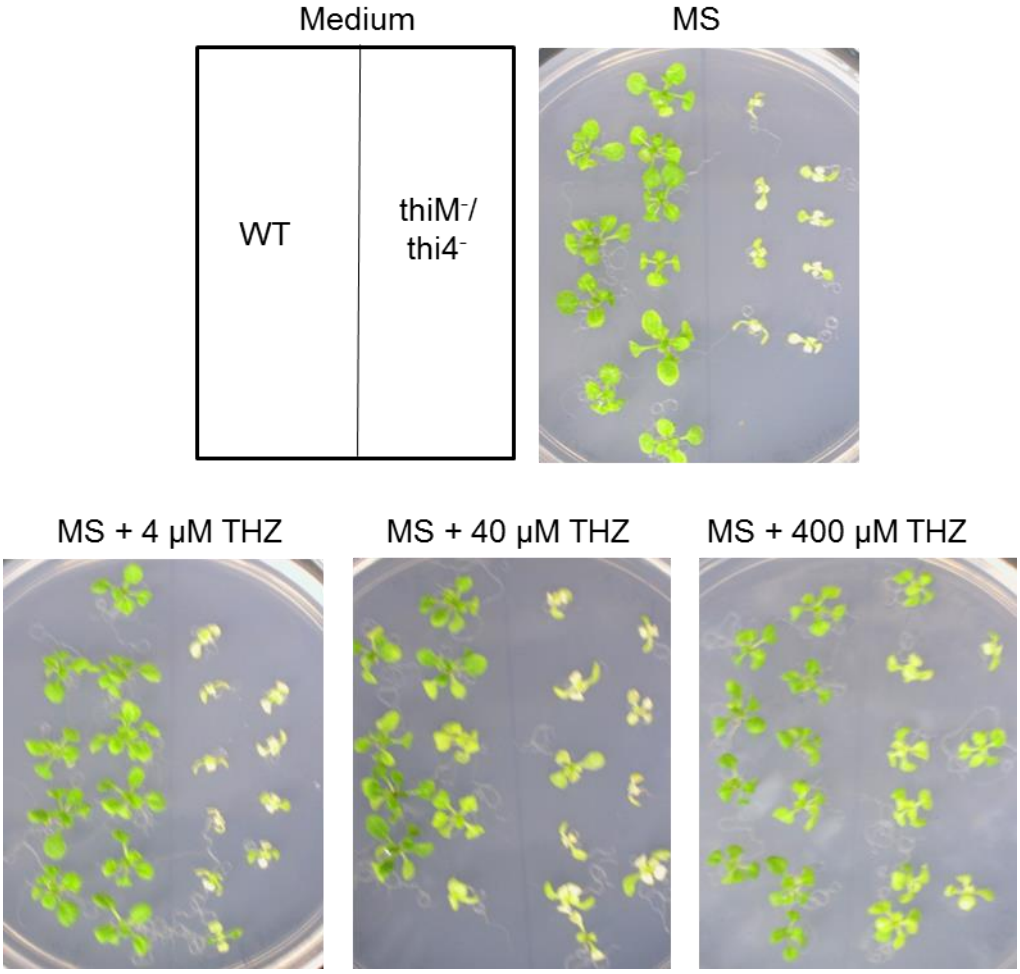


Figure 9: Effect of increasing concentrations of supplied THZ on the growth of the *Arabidopsis* thiM-/thi4- double mutant. Wild type (WT) *Arabidopsis* and double mutant plantlets were cultured for 15 days on MS medium alone or supplemented with 4, 40, or 400 µM THZ.

These results establish that ThiM has an important role in the salvage of THZ, at least when it is exogenously supplied. That the double mutant is not completely inviable when supplied with a high level of THZ indicates that it retains some vestigial ability to phosphorylate THZ. In this connection, it may be noted that another thiamin salvage kinase activity, that for 4-amino-5-hydroxymethyl-2-methylpyrimidine, resides both in a dedicated enzyme and-as a minor side-activity- in an enzyme from a different B vitamin pathway, pyridoxal kinase (Park et al., 2004). More generally, because the hydroxyethyl group that THZ kinase phosphorylates (Figure 1) is a structural motif common to various phosphorylated metabolites (e.g. phosphoethanolamine, 5-phosphomevalonate), THZ might well be a poor alternative substrate for kinases that form these metabolites. It may also be noted that the *Arabidopsis* single ThiM mutant does not require thiamin, whereas the corresponding mutant in *Chamydomonas reinhardtii* (*thi-10*) does (Ferris, 1995). While this contrast could reflect differences in the thiamin biosynthesis pathway between *C. reinhardtii* and higher plants it could also arise simply from greater reliance on salvage in the thiamin economy of *C. reinhardtii*. Consistent with the latter possibility, the *thi-10* mutant dies back less readily than other auxotrophs when thiamin is withheld (Ferris, 1995).

Arabidopsis ThiM is not transcriptionally regulated by thiamin or THZ

As transcript levels of other thiamin synthesis or salvage genes respond to thiamin in plants (Goyer, 2010) and bacteria (Petersen and Downs, 1997 and Müller et al., 2009a) we tested whether levels of the *Arabidopsis* ThiM transcript were affected by culturing plantlets in media containing thiamin or THZ at various concentrations. No significant effects were seen for either compound at any concentration tested (Figure 10). The lack

of effects cannot be ascribed to non-absorption, as exogenous thiamin and THZ are both able to rescue thiamin auxotrophs (Figure 8) (Li and Rédei, 1969).

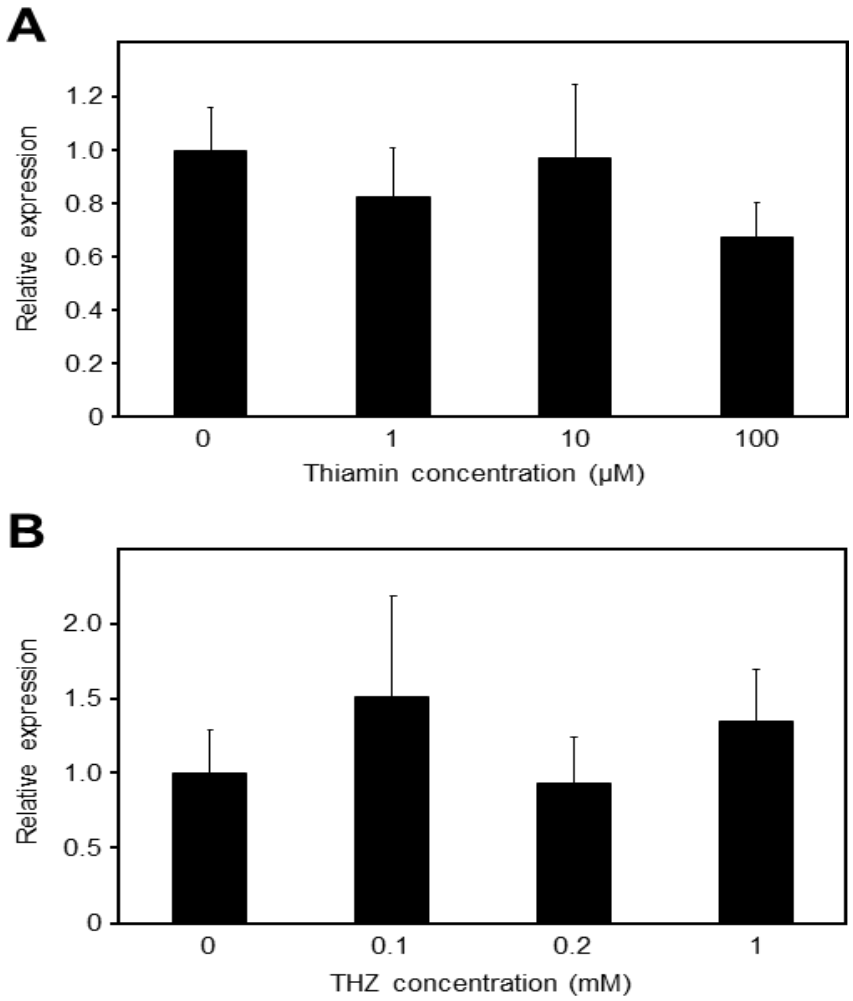


Figure 10: Transcriptional profiling of ThiM (At3g24030) expression in *Arabidopsis* leaves. Wild type plantlets were cultured on MS medium containing the indicated concentrations of thiamin (A) or THZ (B). ThiM transcript levels were determined by qRT-PCR and are expressed relative to those in untreated plants. Data are means of three biological replicates ±SE; treatment means were not significantly different from those of controls ($P = 0.05$).

CONCLUSIONS

This work demonstrates that plants have strong ThiM homologs and that those of *Arabidopsis* and maize function as thiazole kinases in vitro and in vivo. More generally,

the conservation of plant ThiM sequences enables this functional assignment to be confidently propagated by homology to other plants. Finally, since the plant thiamin synthesis pathway is localized in plastids (Goyer, 2010 and Gerdes et al., 2012) and ThiM proteins seem most likely to be cytosolic (based on lack of bioinformatic or experimental evidence to the contrary), it would appear that plastids have a THZ phosphate transporter that remains to be identified.

Table S1 Oligonucleotide primers used in this study.

Primer name	Sequence 5' to 3'	Application
EcThiG In R	GGCTTCTTCCGCTGTTTTCGC	Genotyping *Eschericha coli* mutants
EcThiG In F	CCGGTATGCGGAGGAAGCAAT	
EcThiG Out R *	GATGTTCAAAGCCCATCTCC	
EcThiG Out F *	GTCGGATTTGGCGGTCAG	
EcThiH In R	GCCAGCGGTTGCACTTCCAT	
EcThiH In F	AACTGGACTGGGACGACATCCG	
EcThiH Out R ‡	CGCCATTACCGGACTTTG	
EcThiH Out F ‡	CTGAAAGCCGCCGAAAC	
EcThiM In R	CGCCGATAATGCACAGCCAGTT	
EcThiM In F	AATCTGCGCACGCGTTACACC	
EcThiM Out R	CATTAGCACTGCGCCACA	
EcThiM Out F	CATATTCTTGTTAACCAGGTGTTGC	
EcThiM to pBAD24 F NcoI	CAGGCTCCATGGAAGTCGACCTGCTG	Creating complementation vectors
EcThiM to pBAD24 R HindIII	AAGTGGAAGCTTTCATGCCTGCACCTCCT	
AtThiM to pBAD24 F NcoI	CAGGCTCCATGGAATCAAAATCAGAACAAAA	
AtThiM to pBAD24 R HindIII	AAGTGGAAGCTTTCAACCCAACCTGGTCACAT	
ZmThiM to pBAD24 F EcoRI	CAGGCTGGAATTCACCATGGACGTTGGCGCG	
ZmThiM to pBAD24 R HindIII	AAGTGGAAGCTTTCATGGTCGCAAGGAAATC	
AtThiM to pET28b F NcoI	CACCATGGAATCAAAATCAGAACAAAACG	Creating expression vectors
AtThiM to pET28b R XhoI	CTCGAGACCCAACCTGGTCACATTCA	
ZmThiM to pET28b F BspHI	CAGGCTTCATGAACGTTGGCGCGAAGG	
ZmThiM to pET28b R SalI	AAGTGGGTCGACTGGTCGCAAGGAAATCCTAA	
ThiM1 (SALK_123358)	TTGTCCACGTGAAGCAGAGGA	Genotyping *Arabidopsis* mutants
ThiM2 (SALK_123358)	GAAGTTGTGCCTGGTGGTTGG	
LBb1	GCGTGGACCGCTTGCTGCAACT	
RT ThiM F	ATGGAATCAAAATCAGAACAAAACGAGTGG	ThiM transcript detection in *Arabidopsis* by RT
RT ThiM R	TCAACCCAACCTGGTCACATTCACA	
ThiM F	GGGACGAAGATGATGCAA	qRT-PCR of *Arabidopsis* samples
ThiM R	TAGCTTCCAGTACCCGTGAT	
ThiM probe	GGTTTGATTGTAGCGT	
eEF-1α F	GAGCCCAAGTTTTTGAAGA	
eEF-1α R	CTAACAGCGAAACGTCCCA	
eEF-1α probe	CCCCAACCAAGCCCAT	

Underlined, restriction site; F, forward; R, reverse; * and ‡, were respectively used for generating amplicons and for recombineering of EcThiG and EcThiH in the Δ*thiM* mutant background.

ACKNOWLEDGMENT

This work was supported in part by US National Science Foundation grant numbers IOS-1025398 (to A.D.H. and V. de C.) and MCB-0236210 (to D.K.S.), and by an endowment from the C.V. Griffin Sr. Foundation (to A.D.H). We thank R.J. Schmidt for the tassel primordium cDNA library and D.R. McCarty for help in analysis of maize transcript data.

REFERENCES

Atkins, C.A., Canvin, D.T., 1971. Photosynthesis and CO2 evolution by leaf discs-gas exchange, extraction, and ion-exchange fractionation of C14-labeled photosynthetic products. Can. J. Bot. 49, 1225–1234.

Aziz, R.K., Devoid, S., Disz, T., Edwards, R.A., Henry, C.S., Olsen, G.J., Olson, R., Overbeek, R., Parrello, B., Pusch, G.D., Stevens, R.L., Vonstein, V., Xia, F., 2012.SEED servers: high-performance access to the SEED genomes, annotations, andmetabolic models. PLoS One 7, e48053.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., Mori, H., 2006. Construction of *Escherichia coli* K-12 inframe, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. 2 (2006), 0008.

Campobasso, N., Mathews, I.I., Begley, T.P., Ealick, S.E., 2000. Crystal structure of 4-methyl-5-beta-hydroxyethylthiazole kinase from *Bacillus subtilis* at 1.5 Å resolution. Biochemistry 39, 7868–7877.

Chatterjee, A., Abeydeera, N.D., Bale, S., Pai, P.J., Dorrestein, P.C., Russell, D.H., Ealick, S.E., Begley, T.P., 2011. *Saccharomyces cerevisiae* THI4p is a suicide thiaminethiazole synthase. Nature 478, 542–546.

Cherepanov, P.P., Wackernagel, W., 1995. Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibioticresistance determinant. Gene 158, 9–14.

Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. USA 97, 6640–6645.

Ferris, P.J., 1995. Localization of the nic-7, ac-29 and thi-10 genes within the matingtype locus of Chlamydomonas reinhardtii. Genetics 141, 543–549.

Fitzpatrick, T.B., Basset, G.J., Borel, P., Carrari, F., DellaPenna, D., Fraser, P.D., Hellmann, H., Osorio, S., Rothan, C., Valpuesta, V., Caris-Veyrat, C., Fernie, A.R., 2012. Vitamin deficiencies in humans: can plant science help? Plant Cell 24, 395–414.

Gerdes, S., Lerma-Ortiz, C., Frelin, O., Seaver, S.M., Henry, C.S., de Crécy-Lagard, V., Hanson, A.D., 2012. Plant B vitamin pathways and their compartmentation: a guide for the perplexed. J. Exp. Bot. 63, 5379–5395.

Goyer, A., 2010. Thiamine in plants: aspects of its metabolism and functions. Phytochemistry 71, 1615–1624.

Guzman, L.M., Belin, D., Carson, M.J., Beckwith, J., 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose pBAD promoter. J. Bacteriol. 177, 4121–4130.

Jurgenson, C.T., Begley, T.P., Ealick, S.E., 2009. The structural and biochemical foundations of thiamin biosynthesis. Annu. Rev. Biochem. 78, 569–603.

Karunakaran, R., Ebert, K., Harvey, S., Leonard, M.E., Ramachandran, V., Poole, P.S., 2006. Thiamine is synthesized by a salvage pathway in Rhizobium leguminosarum bv. viciae strain 3841. J. Bacteriol. 188, 6661–6668.

Kawasaki, Y., 1993. Copurification of hydroxyethylthiazole kinase and thiaminephosphate pyrophosphorylase of *Saccharomyces cerevisiae*: characterization of hydroxyethylthiazole kinase as a bifunctional enzyme in the thiamine biosynthetic pathway. J. Bacteriol. 175, 5153–5158.

Leonardi, R., Fairhurst, S.A., Kriek, M., Lowe, D.J., Roach, P.L., 2003. Thiamine biosynthesis in Escherichia coli: isolation and initial characterisation of the ThiGH complex. FEBS Lett. 539, 95–99.

Li, S.L., Rédei, G.P., 1969. Thiamine mutants of the crucifer, Arabidopsis. Biochem. Genet. 3, 163–170.

McCourt, J.A., Nixon, P.F., Duggleby, R.G., 2006. Thiamin nutrition and catalysis-induced instability of thiamin diphosphate. Br. J. Nutr. 96, 636–638.

Mizote, T., Nakayama, H., 1989. The thiM locus and its relation to phosphorylation of hydroxyethylthiazole in Escherichia coli. J. Bacteriol. 171, 3228–3232.

Müller, I.B., Bergmann, B., Groves, M.R., Couto, I., Amaral, L., Begley, T.P., Walter, R.D., Wrenger, C., 2009a. The vitamin B1 metabolism of Staphylococcus aureus is controlled at enzymatic and transcriptional levels. PLoS One 4, e7656.

Müller, M., Gocke, D., Pohl, M., 2009b. Thiamin diphosphate in biological chemistry: exploitation of diverse thiamin diphosphate-dependent enzymes for asymmetric chemoenzymatic synthesis. FEBS J. 276, 2894–2904.

Murashige, T., Skoog, F., 1962. A revised medium for rapid growth and bio-assayswith tobacco tissue cultures. Physiol. Plant. 15, 473–497.

Nosaka, K., Nishimura, H., Kawasaki, Y., Tsujihara, T., Iwashima, A., 1994. Isolation and characterization of the THI6 gene encoding a bifunctional thiamin-phosphate pyrophosphorylase/hydroxyethylthiazole kinase from *Saccharomyces cerevisiae*. J. Biol. Chem. 269, 30510–30516.

Park, J.H., Burns, K., Kinsland, C., Begley, T.P., 2004. Characterization of two kinases involved in thiamine pyrophosphate and pyridoxal phosphate biosynthesis in Bacillus subtilis: 4-amino-5-hydroxymethyl-2methylpyrimidine kinase and pyridoxal kinase. J. Bacteriol. 186, 1571–1573.

Paul, D., Chatterjee, A., Begley, T.P., Ealick, S.E., 2010. Domain organization in Candida glabrata THI6, a bifunctional enzyme required for thiamin biosynthesis in eukaryotes. Biochemistry 49, 9922–9934.

Petersen, L.A., Downs, D.M., 1997. Identification and characterization of an operon in Salmonella typhimurium involved in thiamine biosynthesis. J. Bacteriol. 179, 4894–4900.

Reddick, J.J., Kinsland, C., Nicewonger, R., Christian, T., Downs, D.M., Winkler, M.E., Begley, T.P., 1998. Overexpression, purification and characterization of two pyrimidine kinases involved in the biosynthesis of thiamin: 4-amino-5-hydroxymethyl-2-methylpyrimidine kinase and 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate kinase. Tetrahedron 54, 15983–15991.

Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, New York.

Sun, Q., Zybailov, B., Majeran, W., Friso, G., Olinares, P.D., van Wijk, K.J., 2009. PPDB, the Plant Proteomics Database at Cornell. Nucleic Acids Res. 37, D969–D974.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.

Tanz, S.K., Castleden, I., Hooper, C.M., Vacher, M., Small, I., Millar, H.A., 2013. SUBA3: a database for integrating experimentation and prediction to define the SUBcellular location of proteins in Arabidopsis. Nucleic Acids Res. 41, D1185–D1191.

Tunc-Ozdemir, M., Miller, G., Song, L., Kim, J., Sodek, A., Koussevitzky, S., Misra, A.N., Mittler, R., Shintani, D., 2009. Thiamin confers enhanced tolerance to oxidative stress in Arabidopsis. Plant Physiol. 151, 421–432.

Wrenger, C., Eschbach, M.L., Müller, I.B., Laun, N.P., Begley, T.P., Walter, R.D., 2006. Vitamin B1 de novo synthesis in the human malaria parasite Plasmodium falciparum depends on external provision of 4-amino-5-hydroxymethyl-2-methylpyrimidine. Biol. Chem. 387, 41–51.

Zhang, Y., Taylor, S.V., Chiu, H.J., Begley, T.P., 1997. Characterization of the *Bacillus subtilis* thiC operon involved in thiamine biosynthesis. J. Bacteriol. 179, 3030–3035.

CHAPTER 3

Salvage of the thiamin pyrimidine moiety by plant TenA proteins lacking an active-site cysteine

Remi Zallot, Mohammad Yazdani, Aymeric Goyer, Michael J. Ziemak, Jiahn-Chou Guan, Donald R. Mccarty, Val´erie de Cr´ecy-Lagard, Svetlana Gerdes, Timothy J. Garrett, Jordi Benach, John F. Hunt, David K. Shintani and Andrew D. Hanson

This chapter was published in Biochemical Journal, (2014) 463:145–155.

Author contribution: Andrew Hanson, Rémi Zallot, David Shintani and Valérie de Crécy-Lagard designed the research. Andrew Hanson and Svetlana Gerdes made comparative genomic analyses. Rémi Zallot, Aymeric Goyer and Michael Ziemak carried out cloning, and Rémi Zallot characterized enzyme activities. Jiahn-Chou Guan and Donald McCarty analysed gene expression. Mohammad Yazdani, David Shintani and Aymeric Goyer carried out *Arabidopsis* mutant work. Timothy Garrett performed LC-MS analysis. Jordi Benach and John Hunt performed the structural analyses. All authors interpreted the experimental data. Andrew Hanson and Rémi Zallot wrote the paper.

ABSTRACT

The TenA protein family occurs in prokaryotes, plants and fungi; it has two subfamilies, one (TenA_C) having an active-site cysteine, the other (TenA_E) not. TenA_C proteins participate in thiamin salvage by hydrolysing the thiamin breakdown product amino-HMP (4-amino-5-aminomethyl-2-methylpyrimidine) to HMP (4-amino-5-hydroxymethyl-2-methylpyrimidine); the function of TenA_E proteins is unknown. Comparative analysis of prokaryote and plant genomes predicted that (i) TenA_E has a salvage role similar to, but not identical with, that of TenA_C and (ii) that TenA_E and TenA_C also have non-salvage roles since they occur in organisms that cannot make thiamin. Recombinant *Arabidopsis* and maize TenA_E proteins (At3g16990, GRMZM2G080501) hydrolysed amino-HMP to HMP and, far more actively, hydrolysed the *N*-formyl derivative of amino-HMP to amino-HMP. Ablating the *At3g16990* gene in

a line with a null mutation in the HMP biosynthesis gene *ThiC* prevented its rescue by amino-HMP. Ablating *At3g16990* in the wild-type increased sensitivity to paraquat-induced oxidative stress; HMP overcame this increased sensitivity. Furthermore, the expression of *TenA_E* and *ThiC* genes in *Arabidopsis* and maize was inversely correlated. These results indicate that TenA_E proteins mediate amidohydrolase and aminohydrolase steps in the salvage of thiamin breakdown products. As such products can be toxic, TenA_E proteins may also pre-empt toxicity.

Abbreviations: amino-HMP, 4-amino-5-aminomethyl-2-methylpyrimidine; formylamino-HMP, N-formyl-4-amino-5-aminomethyl-2-methylpyrimidine; HMP, 4-amino-5-hydroxymethyl-2-methylpyrimidine; qRT-PCR, quantitative reverse transcription–PCR; THZ, 4-methyl-5-(2-hydroxyethyl) thiazole; wat, crystallographically ordered water molecule

INTRODUCTION

Thiamin, in its diphosphate form, is an essential cofactor for central metabolic enzymes such as transketolase and the pyruvate and α-ketoglutarate dehydrogenase complexes [1]. Thiamin consists of HMP (4-amino-5-hydroxymethyl-2-methylpyrimidine) and THZ [4-methyl-5-(2-hydroxyethyl) thiazole] moieties (Figure 1A). Plants, fungi and most prokaryotes synthesize both of these moieties *de novo* and couple them to give thiamin [2]. These organisms can also salvage HMP and THZ derived from thiamin breakdown for re-use in thiamin synthesis [2,3], whereas certain prokaryotes rely totally on salvage of exogenous HMP and/or THZ because they do not make these compounds [4]. Even in organisms that do make HMP and THZ, salvage can

be crucial to the thiamin economy because thiamin is chemically and metabolically labile [5,6] and HMP and THZ are costly to make [2,7].

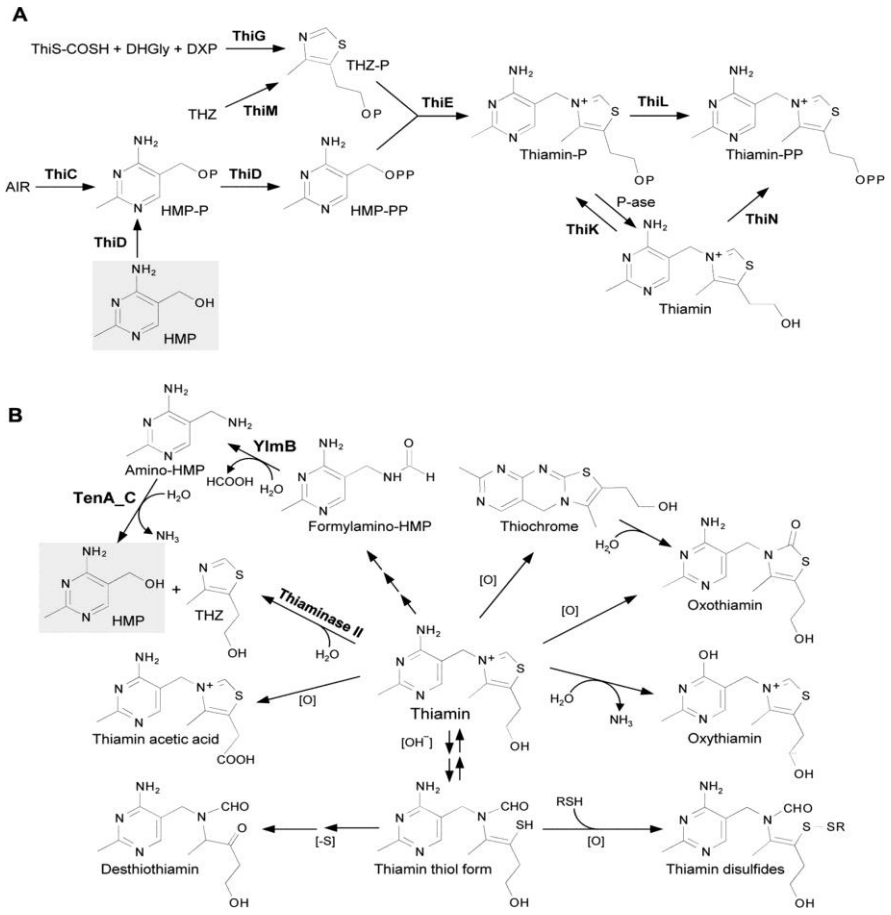


Figure 1: Thiamin synthesis, degradation and salvage routes. (**A**) Bacterial thiamin synthesis and salvage pathways. The plant pathways are the same except that the THZ moiety is synthesized by a single enzyme (THI4) from NAD and glycine, and ThiL and ThiK are absent. AIR, 5-aminoimidazole ribotide; DHGly, dehydroglycine; DXP, deoxy-D-xylulose 5-phosphate; -P, phosphate; -PP, diphosphate; ThiS-COSH, ThiS thiocarboxylate. (**B**) Selected chemical and enzymatic thiamin degradation reactions. The sequential actions of YlmB (formylamino-HMP amidohydrolase) and TenA_C (amino-HMP aminohydrolase) convert the breakdown product formylamino-HMP into HMP, which can re-enter the synthesis pathway as shown in (**A**) (HMP is highlighted in **A** and **B** to show this connection). Other reactions (not shown) include oxidation of the alcohol groups of oxothiamin, THZ and HMP to the respective acids, cleavage of oxothiamin to HMP and the oxo derivative of THZ, and cleavage of thiamin thiol form to amino-HMP and a thioketone. Various other uncharacterized breakdown products have been reported, including some that retain both rings or large parts thereof [11,52].

Thiamin biosynthesis is now largely understood in bacteria, plants and yeast [1–3] (Figure 1A). Thiamin salvage is less understood, in part because thiamin breakdown products are numerous, incompletely characterized and vary with environmental conditions [5,8–12] (Figure 1B). What is clear so far is that: (i) thiamin can be split into its HMP and THZ moieties chemically, photochemically or enzymatically; (ii) these moieties, separately or as parts of the thiamin molecule, undergo various degradation reactions; (iii) the HMP moiety is generally more stable than the THZ moiety; and (iv) intact HMP and THZ can be phosphorylated and then re-enter the synthesis pathway [2,4,6,13,14] (Figure 1A). In bacteria, yeast and plants, HMP is converted into its mono- and di-phosphates by HMP (phosphate) kinase (ThiD), and THZ is phosphorylated by THZ kinase (ThiM) [2,15,16].

Among the things that are unclear about breakdown and salvage is how HMP is formed. One route involves hydrolysis of intact thiamin by the thiaminase II activity of microbial TenA family proteins [17–19]. However, previous work in bacteria and yeast indicates that a more important source of HMP *in vivo* is hydrolysis of amino-HMP (4-amino-5-aminomethyl-2-methylpyrimidine) by the amino-HMP hydrolase activity of TenA proteins [8,19]. Amino-HMP itself can come from deconstructing thiamin, after damage to its THZ ring, via a route whose last intermediate is the *N*-formyl derivative of amino-HMP (referred to as formylamino-HMP) and whose last enzyme is formylamino-HMP amidohydrolase (YlmB) [8] (Figure 1B). Breakdown products that retain more of the THZ ring than formylamino-HMP (some of which are shown in Figure 1B) could also be TenA substrates, but this has yet to be proven [8,20].

Also unclear is which TenA family proteins have amino-HMP aminohydrolase and thiaminase II activities. TenA proteins fall into two widely distributed subfamilies: one has an active-site cysteine residue, the other does not, but often has two conserved glutamate residues [21]. We call these subfamilies TenA_C and TenA_E respectively, based on the one-letter codes for cysteine and glutamate. Crystal structures are available for representatives from both families [17,21–24]. TenA_C proteins from *Bacillus subtilis* [8,17], yeast [19] and *Helicobacter pylori* [20] have been assayed for enzyme activities; all three have amino-HMP aminohydrolase activity, whereas the first two also have low thiaminase II activity. In contrast, no TenA_E proteins seem to have been tested for activity in published literature. An activity related to thiamin salvage is, however, implied by the HMP [23,25] or HMP-phosphate [24] ligands that co-purified with the *Pyrococcus furiosus* and *Arabidopsis* TenA_E proteins and were visualized in their crystal structures.

If the TenA_E protein encoded in the *Arabidopsis* genome (At3g16990) [21,23] has such an activity, it could explain why *Arabidopsis* mutants blocked in HMP synthesis can be rescued by amino-HMP [26,27], because such rescue requires amino-HMP aminohydrolase activity. *Arabidopsis* encodes a TenA_C domain in another protein (At5g32470) [3] that alternatively could account for salvage of amino-HMP. However, this protein has a haloacid dehalogenase family domain fused to its TenA_C domain, unlike other members of the TenA family, and this unusual structure creates uncertainty as to the enzymatic function of this fusion protein.

In the present study, we present a comparative genomic analysis of the TenA family in prokaryotes and plants that predicts a thiamin salvage function for the TenA_E

subfamily. We then validate this prediction by demonstrating that: (i) *Arabidopsis* TenA_E and its maize orthologue have amino-HMP aminohydrolase activity; (ii) both proteins also have high formylamino-HMP amidohydrolase activity; and (iii) the *Arabidopsis* and maize *TenA_E* genes are strongly expressed when the HMP biosynthesis gene *ThiC* is not. The comparative genomic analysis also suggests that, besides functioning in thiamin salvage, TenA_E and TenA_C proteins can pre-empt metabolic damage from thiamin breakdown products.

EXPERIMENTAL

Bioinformatics and gene expression analysis

Protein sequences were taken from GenBank, MaizeSeqence.org and the SEED database [28]. Comparative analyses of prokaryotic and plant genomes were made using SEED tools [28]; the full results of this analysis are encoded in the SEED subsystem named 'TenA' (available at http://pubseed.theseed.org//SubsysEditor.cgi?page=ShowSpreadsheet&subsystem=TenA) . Sequences were aligned using Muscle [29]. Phylogenetic trees were constructed by the neighbour-joining method using MEGA5 [30]. *Arabidopsis* microarray gene expression data were taken from CSB.DB [31] and maize RNA-seq data were taken from qTeller (http://qteller.com/) [32]. Gene expression data for B73 endosperm tissue harvested at various times were obtained by qRT-PCR (quantitative reverse transcription–PCR) as described [33], and were expressed relative to the value at 14 days after pollination (= 1.0).

Chemicals

Thiamin hydrochloride, thiamin monophosphate chloride dihydrate, thiamin pyrophosphate, oxythiamin chloride hydrochloride and thiochrome were obtained from Sigma–Aldrich. HMP, amino-HMP and oxothiamin {3-[(4-amino-2-methyl-5-pyrimidinyl)methyl]-5-(2-hydroxyethyl)-4-methyl-2(3*H*)-thiazolone} were from Toronto Research Chemicals. Thiamin disulfide hydrate and 1*H*-benzotriazole-1-carboxal-dehyde were from TCI. Oxy-HMP [5-(hydroxymethyl)-2-methylpyrimidin-4(1*H*)-one] was from CGeneTech. Formylamino-HMP was prepared from amino-HMP as described [8]. The preparation was 92.7% pure and contained 0.4% amino-HMP, as judged from HPLC analysis. Desthiothiamin was prepared as described previously [34]. Briefly, 10 ml of a 1.5 M solution of thiamin (15 mmol) in 3 M sodium hydroxide was mixed with 7.5 ml of a 1.55 M solution of glycine (12 mmol) in 1.55 M sodium hydroxide, and stirred at room temperature for 3 days. White crystals were obtained, and washed with ice-cold ethanol. The preparation was 95.8% pure, and contained ≤0.02% thiamin, as judged from HPLC analysis.

Constructs for expression in *Escherichia coli*

The At3g16990 (AtTenA_E) and GRMZM2G080501 (ZmTenA_E) coding sequences were amplified by PCR from leaf cDNA libraries using Phusion DNA polymerase (New England BioLabs). For At3g16990, the amplicons were treated with Taq DNA polymerase to add A-overhangs before cloning into pYES2.1 TOPO (Invitrogen) and this construct served as a PCR template for subsequent cloning into pET28b (Novagen). BSU11650 (BsTenA_C) was amplified by colony PCR using Phusion DNA polymerase. All three amplified coding sequences were cloned into

pET28b between the NcoI and NotI sites, which adds a C-terminal His_6 tag. All constructs were sequence-verified. The primers used are given in Supplementary Table S1 (at http://www.biochemj.org/bj/463/bj4630145add.htm).

Production and purification of recombinant proteins

The AtTenA_E and BsTenA_C pET28b constructs were introduced into the *E. coli* strain BL21-CodonPlus (DE3)-RIPLand the ZmTenA_E pET28b construct was introduced into Rosetta-gami™ 2. Cultures (250 ml for AtTenA_E and BsTenA_C, and 2 litres for ZmTenA_E) were grown at 37°C in LB medium containing 50 $\mu g \cdot ml^{-1}$ kanamycin. When the OD_{600} reached 0.6, IPTG (final concentration of 1 mM) was added; incubation was continued for 3 h at 37°C for BsTenA_C and AtTenA_E. For ZmTenA_E, induction was preceded by a 30-min ethanol shock (final concentration of 5%, v/v), and incubation was continued overnight at 22°C. The subsequent steps were at 4°C. Cells were harvested by centrifugation, resuspended in 50 mM NaH_2PO_4, 300 mM NaCl and 10 mM imidazole, pH 8.0, and then sonicated. To the cleared supernatant, 0.5 ml of Ni^{2+}-nitrilotriacetic acid agarose 50% slurry (Qiagen) was added, followed by rotary shaking for 1 h at 4°C. The mixture was then poured into a column and allowed to drain by gravity. After washing with 16 ml of 50 mM NaH_2PO_4, 300 mM NaCl and 20 mM imidazole, pH 8.0, proteins were eluted with 2 ml of this buffer containing 250 mM imidazole, desalted on PD-10 columns (GE Healthcare) equilibrated in 50 mM Tris/HCl (pH 7.5), 1 mM DTT and 500 mM glycinebetaine, and concentrated with an Amicon Ultra 0.5 ml 10K unit (Millipore). Purified proteins were frozen in liquid N_2 and stored at −80°C. Protein was estimated by dye binding [35] with BSA as standard. As

pilot tests showed that very little ZmTenA_E could be obtained in soluble form, AtTenA_E was used for characterization work; BsTenA_C served as a benchmark.

Enzyme assays

Assays were routinely made in triplicate in 100 µl reaction mixtures containing 45 mM Tris/HCl, pH 7.5, 0.9 mM DTT, 450 mM glycinebetaine, and the specified concentrations of substrates. Assays were run at 30°C for 30–300 min and stopped on ice, and then deproteinized at 4°C using Amicon Ultra 0.5 ml 10K units. Samples (typically 40 µl) of the flow-through were analysed by HPLC with UV detection (Waters 2695 Separation module and Waters 2998 PDA detector). HPLC analysis used a C_{18} column (ACE Excel SuperC_{18}, 5 µm, 250×4.6 mm) with a column guard, equilibrated with 100 mM potassium phosphate, pH 6.6. The elution gradient (1 ml·min^{-1}) was as follows, steps being linear transitions except where noted: 0–4 min −20% potassium phosphate (100 mM, pH 6.6), 80% water; 4–10 min −20% potassium phosphate, 7% methanol, 73% water; 10–15 min, held at 20% potassium phosphate, 7% methanol, 73% water; 15–20 min −10% potassium phosphate, 60% water, 30% methanol; 20–30 min −20% potassium phosphate, 7% methanol, 73% water; and 30–35 min −20% potassium phosphate, 80% water. Detection was by absorbance at 235 nm. Values of k_{cat} and K_m were estimated by non-linear fitting using GraphPad Prism version 6.00.

LC-MS

The amino-HMP and HMP reaction products of AtTenA_E action on formylamino-HMP were analysed using a Thermo LTQ Velos mass spectrometer with an Accela 600 UPLC apparatus and Accela open autosampler (Thermo Fisher). The spectrometer was operated in positive heated-electrospray mode under the following

conditions: 3500 V for the spray needle, 325°C for the source temperature, flow rates of 40 arbitrary units sheath gas, 10 arbitrary units auxiliary gas and 325°C capillary temperature. Gradient elution was employed on the above C_{18} column equilibrated with 100 mM ammonium acetate, pH 6.6, using a flow rate of 1 $ml \cdot min^{-1}$. The flow was split to 0.3 $ml \cdot min^{-1}$ after separation to avoid salt build-up on the ion source. The elution gradient was as follows: 0–4 min −20% ammonium acetate buffer (100 mM, pH 6.6), and 80% water; 4–14 min −10% ammonium acetate, 30% water and 60% methanol; 14–18 min −10% ammonium acetate, 30% water and 60% methanol; 18–25 min −20% ammonium acetate and 80% water; 25–28 min −20% ammonium acetate and 80% water. The injections were of 5 μl and 1 μl for enzyme assays and standards diluted in enzyme assay buffer respectively. Full scan spectra were collected from *m*/*z* 100 to 600.

Experiments with *Arabidopsis* mutants

Seeds of the wild-type Columbia, the SALK-062985 homozygous *At3g16990* knockout line, and the *py-1* (*ThiC*) mutant (stock number CS3491) were obtained from the *Arabidopsis* Biological Resource Center, OH, U.S.A. The *TenA_E* knockout was verified by PCR, using gene-specific primers located 5′ or 3′ of the T-DNA and a T-DNA-specific primer. Amplicons were sequenced to confirm the insertion site. Semi-qRT-PCR was performed as follows: total RNA was extracted from triplicate samples of knockout and wild-type leaves using RNeasy Kits (Qiagen) and treated with DNase (DNA-free™ Kit, Ambion). Total RNA was used for first-strand cDNA synthesis using M-MuLV (New England BioLabs), followed by PCR with primers designed to amplify a fragment of the TenA_E transcript or a fragment of the actin transcript. Amplicons were analysed by agarose gel electrophoresis. The primers used are given in Supplementary

Table S1. Seeds of the *py-1* mutant line were grown on Murashige and Skoog medium [36] plates supplemented with 2% sucrose, 100 $mg \cdot l^{-1}$ inositol, 0.5 $mg \cdot l^{-1}$ nicotinic acid and 0.5 $mg \cdot l^{-1}$ pyridoxine, but not thiamin, and mutant plants were selected based on their phenotype [27] and crossed with the *TenA_E* knockout line. The resulting F1 plants were selfed to generate the double homozygous mutant. Seed of the wild-type, single and double mutants were cultured on Murashige and Skoog medium as above with or without 100 μM thiamin, HMP or amino-HMP. The plated seeds were vernalized at 4°C for 5 days, germinated at 21°C in continuous light (100 $\mu mol \cdot m^{-2} \cdot s^{-1}$), and imaged at 2 weeks. For paraquat (methyl viologen) stress experiments, wild-type and *TenA_E* knockout seeds were grown on Murashige and Skoog medium as above containing 0.1 μM paraquat or 0.1 μM paraquat supplemented with the specified concentrations of HMP or thiamin. Seeds were vernalized and grown as above except that plates were arranged vertically. The total root length was measured at 9 days. Stress experiments were performed with five to seven technical replications, each containing six seedlings per line, and were repeated at least three times.

Molecular modelling

The crystal structure of *Arabidopsis* TenA_E complexed with HMP at 2.1 Å (1 Å=0.1 nm) resolution (PDB ID 2F2G) [23] was used as a template for modelling, and the structure of HMP already present in the active site of *Arabidopsis* TenA_E was used as a starting point for modelling the conformations of formylamino-HMP and amino-HMP. The program COOT [37] was used to manually dock the formylamino-HMP or amino-HMP molecules in the active site of *Arabidopsis* TenA_E based on alignment of the aromatic moiety of HMP as a reference. The most favourable conformations visualized

for formylamino-HMP and amino-HMP in *Arabidopsis* TenA_E are shown. PyMOL Molecular Graphics System version 1.6 (Schrödinger) was used to generate all molecular graphics images shown in the present paper. The same approach was used to manually dock a molecule of formylamino-HMP in the active site of *B. subtilis* TenA_C; in this case, the crystal structure of the complex of *B. subtilis* TenA_C with HMP (PDB ID 1YAK) [17] was used as a template for modelling.

qRT-PCR analysis of maize gene expression

For analysis of maize *TenA_E* (GRMZM2G080501) and *ThiC* (GRMZM2G027663) expression, kernels from field-grown W22 inbred plants (Citra, Florida, spring 2013) at the stage indicated were harvested and their endosperms were immediately frozen in liquid N_2 and held at −80°C until use. After grinding in liquid N_2, total RNA was extracted with RLT buffer (2 ml per 20 mg) and purified using plant RNeasy Kits. RNAs were quantified using a NanoDrop 1000 (Thermo Fisher Scientific), and 5.5 μg of RNA from each sample was treated with RQ1 RNase-free DNase (Promega). Negative RT controls confirmed that there was no carryover of genomic DNA. For quantitative PCR, a Power SYBR green RNA-to-C_T 1-Step Kit (Applied Biosystems) was used with an iCycler iQ real-time PCR detection system (Bio-Rad Laboratories). The 18S rRNA was used as an internal reference for relative quantitative expression analysis [38]. The primers used are given in Supplementary Table S1.

RESULTS

Sequence and phylogenetic analysis of TenA proteins

We first verified that simple presence/absence of a conserved active-site cysteine residue (Cys^{135} in *B.subtilis* TenA) splits a full range of TenA proteins into two

biologically meaningful subfamilies because the correlation with the presence of cysteine had previously been proposed based on an analysis of just a small set of TenA sequences [21]. A set of 39 diverse TenA proteins from bacteria, archaea, plants and yeast was identified using Blastp and Conserved Domain tools at NCBI. These sequences were separated into plus-cysteine (TenA_C) and minus-cysteine (TenA_E) classes, aligned and subjected to phylogenetic analysis (Supplementary Figures S1 and S2 at http://www.biochemj.org/bj/463/bj4630145add.htm). The plus- and minus-cysteine sequences fell neatly into two different clades, even when they came from the same organism. This result supports the active-site cysteine residue as a convenient marker that distinguishes two distinct subfamilies. It also suggests that these subfamilies are anciently diverged and are likely to be functionally distinct. A further point worth noting is that the second *Arabidopsis* TenA protein (At5g32470) and its maize orthologues [3] are of the TenA_C class, so that plants are among the organisms that encode both TenA_E and TenA_C proteins. Plant TenA_E and TenA_C proteins are phylogenetically closest to the respective cyanobacterial proteins.

Comparative genomics connects both TenA_E and TenA_C with thiamin metabolism

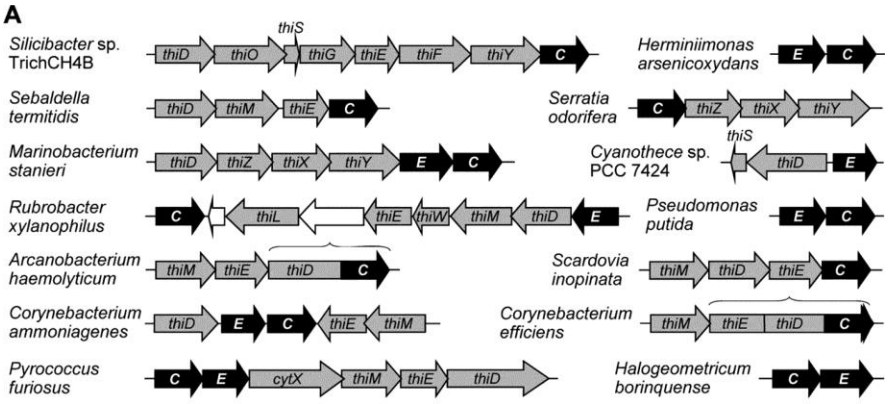
We next surveyed the distribution of TenA family genes among >12000 prokaryote genomes, using the SEED database and its tools [28]. This survey showed that TenA proteins are common in Crenarchaeota and Euryarchaeota, Bacteroidetes, Cyanobacteria, Firmicutes and Proteobacteria, and occur at least occasionally in all other major taxa except Aquificae, Acidobacteria and Thermotogae. A subset of 345 high-quality genomes from taxonomically and ecologically diverse organisms was then chosen

for further analysis. Results are summarized below, and are available in full in the SEED database (http://pubseed.theseed.org//SubsysEditor.cgi?page=ShowSpreadsheet&subsystem=TenA).

Among the 345 selected genomes, 40% had at least one *TenA* gene: 27% had only *TenA_C*, 3% had only *TenA_E* and 10% had both (Supplementary Table S2 at http://www.biochemj.org/bj/463/bj4630145add.htm). *TenA_E* is thus less common than *TenA_C*, and occurs much more often with *TenA_C* than alone. This pattern reinforces the inference from sequence and phylogeny that the two subfamilies do not have identical functions. Consistent with functions in HMP salvage, *TenA_C* and *TenA_E* are found more than twice as often in organisms whose thiamin synthesis pathway depends on salvaging HMP from exogenous sources than in those that can make HMP, i.e. in organisms that respectively lack or have *ThiC*, and have otherwise complete suites of thiamin synthesis genes.

Further evidence that *TenA_E* and *TenA_C* have non-identical functions in thiamin salvage comes from gene-clustering and gene-fusion analyses. Genes of each subfamily commonly cluster on the chromosome with genes for thiamin synthesis and salvage enzymes or transporters, often in predicted operons. Furthermore, *TenA_E* and *TenA_C* quite commonly cluster with each other, again in putative operons (Figure 2A). The co-regulation of *TenA_E* and *TenA_C* implied by these operonic structures connotes non-redundant functions even more strongly than co-occurrence in the same genomes. Genes encoding TenA_C are fused to *thiD*, or *thiD* and *thiE*, in certain Actinobacteria (Figure 1A), fungi [19], protozoa (e.g. *Perkinsus marinus*) and green algae (e.g. *Chlorella*

variabilis) (results not shown). Genes coding for TenA_E seem not to be fused to other genes in any sequenced genome.



C

Organism	TenA_C	TenA_E	ThiC	ThiD	ThiE	ThiG	ThiM	ThiL	ThiN	ThiK
Rubrobacter xylanophilus	+	+		+	+ +	+	+	+		
Cyanothece sp. PCC 7424	+	+	+	+	+	+		+		
Rhizobium leguminosarum	+	+		+	+ +		+		+	
Thermus thermophilus		+	+	+	+	+		+	+	
Serratia odorifera 4Rx13	+		+	+	+	+	+	+		+
Serratia sp. M24T3		+	+	+	+	+	+	+		+
Bacillus subtilis	+		+	+ +	+ +	+	+	+	+	
Parachlamydia acanthamoebae	+							+	+	
Truepera radiovictrix		+							+	
Oceanithermus profundus		+							+	
Wolbachia sp.	+ +								+	

Figure 2: Comparative genomic analysis of the TenA family in prokaryotes**.** (**A**) Chromosomal clustering of genes encoding TenA_C (*C*) and TenA_E (*E*) with thiamin synthesis and salvage genes, and genes encoding components of formylamino-HMP transporter ThiXYZ [8] or predicted thiamin salvage-related transporters ThiW and CytX [4]. Genes are shown as arrows pointing in the direction of transcription; overlaps denote translational coupling. Black arrows are TenA protein genes, grey arrows are thiamin-related genes and white arrows are other genes. Horizontal braces mark gene fusions. The 14 species used as examples represent seven major taxa: Actinobacteria, Cyanobacteria,

Euryarcheota, Fusobacteria, and Alphaproteobacteria, Betaproteobacteria and Gammaproteobacteria. (**B**) Categorization of prokaryotes whose genomes encode TenA_E (without or with TenA_C) according to their physiology and ecology. The 45 organisms (see Supplementary Table S2 at http://www.biochemj.org/bj/463/bj4630145add.htm) were sorted into five classes: (i) thermophiles; (ii) resistant to UV or ionizing radiation, or colonizing light-exposed surfaces; (iii) photosynthetic; (iv) plant symbionts or plant biomass degraders; and (v) other lifestyles. The chart shows the percentage of each class. (**C**) The distribution of genes coding for TenA_C, TenA_E and thiamin synthesis and salvage enzymes. The top seven rows show typical patterns for organisms that can make thiamin, either entirely *de novo* or from HMP and THZ precursors. The bottom four rows show patterns for organisms that are completely unable to make thiamin but have genes for TenA_C or TenA_E. Plus signs in grey boxes indicate presence of a gene or genes, open boxes indicate absence of a gene.

To get clues to the functional differences between *TenA_E* and *TenA_C*, we compared the lifestyles of the prokaryotes in which they occur. For *TenA_C*, no lifestyle bias was evident, but *TenA_E* was concentrated in thermophiles, phototrophs, plant-associated bacteria and organisms resistant to UV or ionizing radiation (Figure 2B). Given the established lability of thiamin to heat, light and radiation, and that these stresses lead to products that can differ from each other and from those formed metabolically or at high pH [5,10–14], the lifestyle evidence raises the possibility that TenA_E proteins prefer as substrates certain thermo-, photo- or radio-decomposition products formed from thiamin either intra- or extra-cellularly.

Comparative genomics points to non-salvage roles for TenA proteins

TenA genes usually occur in organisms that can make thiamin from HMP and THZ moieties, based on their having genes encoding ThiD, ThiE, ThiG or ThiM, and ThiL or ThiN (Figure 2C). However, *TenA_E* or *TenA_C* genes also occur in a small disparate group of bacteria that have no biosynthetic genes except those for ThiN, or

ThiN plus ThiL, and therefore have no thiamin biosynthesis capacity and require preformed thiamin (or its monophosphate) (Figure 2C). This group includes the radiation-resistant extremophile *Truepera radiovictrix* and the hydrothermal vent thermophile *Oceanithermus profundus*, which have *TenA_E*, as well as strains of the obligate intracellular bacteria *Wolbachia* and *Parachlamydia acanthamoebae*, which have *TenA_C*. *Wolbachia* genomes are highly reduced in size [39] and *P. acanthamoebae* genomes are somewhat reduced [40]. The presence of *TenA_E* or *TenA_C* in diverse organisms that cannot re-use thiamin breakdown products for thiamin biosynthesis implies that these enzymes have some function other than thiamin salvage. This implication is particularly strong for the obligate intracellular bacteria, in which genome reduction has jettisoned many other enzyme encoding genes. One possible non-salvage function might be simply to catabolize thiamin breakdown products (as discussed below).

Enzymatic activities of recombinant plant TenA_E proteins

Among the organisms with *TenA_E*, only plants such as *Arabidopsis* and maize are genetically tractable and have experimentally characterized thiamin synthesis pathways for which mutants are available [3]. We therefore targeted *Arabidopsis* TenA_E (At3g16990) and its maize orthologue (GRMZM2G080501) for experimental characterization. These proteins are predicted to be cytosolic and have not been detected in organelles [41]. Both proteins were expressed in *E. coli* in His_6-tagged form and isolated by Ni^{2+}-affinity chromatography (Supplementary Figure S3 at http://www.biochemj.org/bj/463/bj4630145add.htm), along with the well-studied *B. subtilis* TenA_C protein as a benchmark. All three proteins were then tested for activity

against most of the thiamin degradation products shown in Figure 1(B), as well as against thiamin and its phosphates (Table 1).

Table 1 Substrate ranges of *Arabidopsis* TenA_E proteins and *B. subtilis* TenA_C

Substrate concentrations were 1 mM. Activities were determined at 30°C in 45 mM Tris/HCl buffer, pH 7.5, containing 0.9 mM DTT and 0.45 M glycinebetaine. Results are means ± S.E.M. of six replicates.

	Specific activity ($nmol \cdot min^{-1} \cdot mg^{-1}$)	
Substrate	*Arabidopsis* TenA_E	*B. subtilis* TenA_C
Amino-HMP	0.34 ± 0.05	292 ± 34
Formylamino-HMP	6.52 ± 0.31	<0.05
Thiamin	<0.05	162 ± 18
Thiamin monophosphate	<0.05	<0.05
Thiamin diphosphate	<0.05	<0.05
Oxothiamin	<0.05	91 ± 20
Oxythiamin	<0.05	205 ± 30
Thiamin disulfide	<0.05	<0.05
Desthiothiamin	<0.05	<0.05
Thiochrome	<0.05	<0.05

As expected for TenA_C proteins [2,19], *B. subtilis* TenA_C hydrolysed amino-HMP, thiamin, oxythiamin and oxothiamin, but not formylamino-HMP or thiamin phosphates, thus validating our procedures. In contrast, although both plant TenA_E proteins had some activity against amino-HMP, they had 20-fold more activity against formylamino-HMP, and none against the other compounds tested (Figures 3A and 3B and Table 1). The initial product of plant TenA_E action on formylamino-HMP was amino-HMP, which was then slowly hydrolysed to HMP (Figure 3B). The identities of the successive reaction products were confirmed by MS (Figure 3C). These results demonstrate that plant TenA_E proteins are bifunctional enzymes with formylamino-HMP amidohydrolase and amino-HMP aminohydrolase activities, and that they lack detectable thiaminase II activity. Kinetic characterization of the formylamino-HMP

amidohydrolase and amino-HMP aminohydrolase activities of the *Arabidopsis* enzyme showed that formylamino-HMP was very strongly preferred over amino-HMP, as demonstrated by a 19-fold higher k_{cat} value and a 26-fold lower K_m value (Table 2). However, even for the preferred substrate formylamino-HMP, the k_{cat} value was quite low for a metabolic enzyme ($2.81\times10^{-3}s^{-1}$).

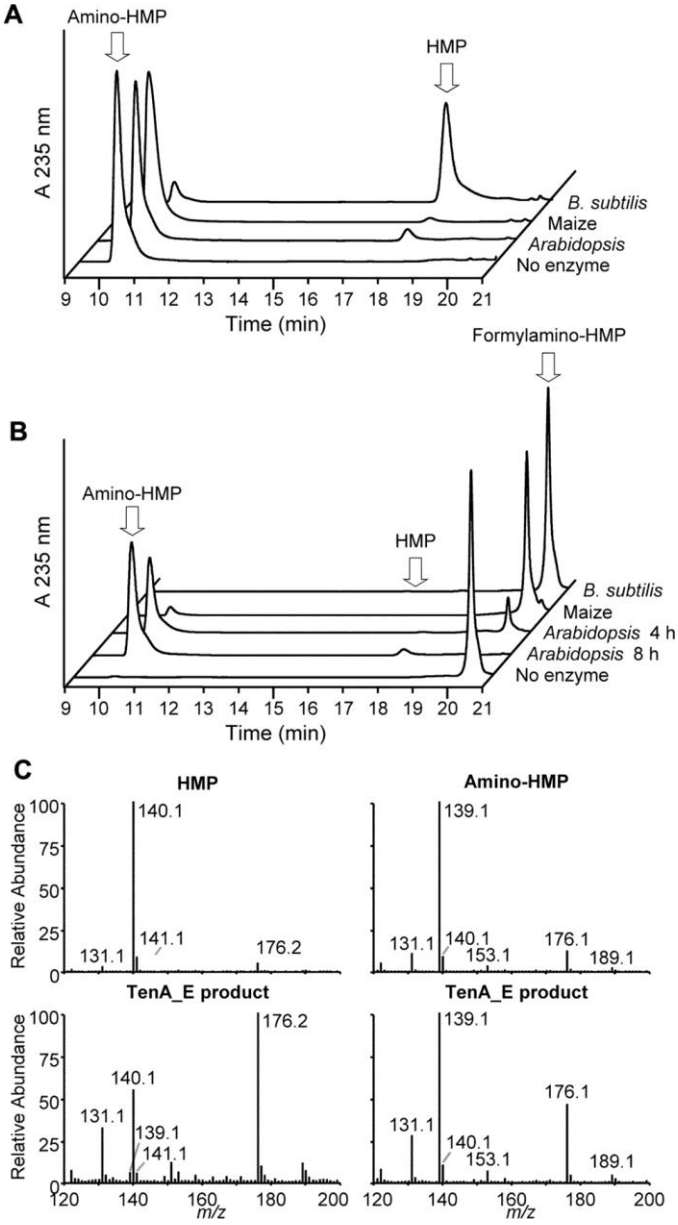


Figure 3: Evidence that plant TenA_E proteins have amino-HMP aminohydrolase and formylamino-HMP amidohydrolase activities. (**A**) HPLC analyses of reaction mixtures (40 μl) in which amino-HMP (100 nmol) was incubated for 4 h at 30°C with *Arabidopsis* or maize TenA_E (50 μg), or *B. subtilis* TenA_C (0.5 μg) as a benchmark. A control incubated without enzyme is included. (**B**) HPLC analyses of reaction mixtures (40 μl) in which formylamino-HMP (100 nmol) was incubated for 4 h at 30°C with *Arabidopsis* or maize TenA_E (50 μg), or *B. subtilis* TenA_C (0.5 μg). The *Arabidopsis* reaction mixture was also incubated for 8 h. A control incubated without enzyme is included. Note that the formylamino-HMP preparation contained a trace (0.4%) of amino-HMP. (**C**) Electrospray mass spectra of authentic HMP (molecular ion $[M^+H^+]=m/z$ 140.1) and amino-HMP (molecular ion $[M^+H^+]=m/z$ 139.1) (upper panels), and the two products formed by *Arabidopsis* TenA_E from formylamino-HMP (lower panels). The signals at m/z 176 are solvent-related.

Table 2 Kinetic constants of the *Arabidopsis* TenA_E protein

Results are means ± S.E.M. for three replicates.

Substrate	k_{cat} (s^{-1}) ($\times 10^{-3}$)	K_m (μM)	k_{cat}/K_m ($s^{-1} \cdot M^{-1}$)
Formylamino-HMP	2.81 ± 0.23	70 ± 0.06	40.9
Amino-HMP	0.15 ± 0.007	1816 ± 200	0.08

Arabidopsis TenA_E acts as an amino-HMP aminohydrolase in planta

To test the *in vivo* function of TenA_E in thiamin salvage, we first identified a homozygous *TenA_E* knockout line (062985.53.80.x) in the Salk *Arabidopsis* T-DNA mutant collection. Plants of this mutant line were confirmed to have a T-DNA insertion in exon two and to lack detectable TenA_E mRNA, i.e. to be knockouts (Supplementary Figure S4 at http://www.biochemj.org/bj/463/bj4630145add.htm). The *TenA_E* mutant was then crossed with a *ThiC* (*py*) mutant (which cannot synthesize the thiamin HMP moiety) [27] to give the double mutant.

The growth of the single and double mutants and of the wild-type (Col-0) was then compared on Murashige and Skoog medium alone, or supplemented with thiamin, HMP or amino-HMP (Figure 4A). The growth of the *TenA_E* mutant was indistinguishable from that of the wild-type on all four media. The HMP-supplemented medium gave slightly poorer growth than the other three, possibly because HMP is phosphorylated by ThiD and the resulting phospho-HMP, an analogue of the cofactor pyridoxal 5′-phosphate, can be toxic at high levels [42]. As expected [27], the *ThiC* single and *TenA_E ThiC* double mutants showed little or no growth on Murashige and Skoog medium alone, and fairly good growth on Murashige and Skoog medium containing HMP. Most importantly, the *ThiC* mutant grew well on medium containing amino-HMP, whereas the double mutant hardly grew at all on this medium. This result demonstrates that TenA_E is required for the hydrolysis of amino-HMP to HMP *in vivo*. This requirement prevented use of the *TenA_E* mutant to test whether TenA_E is also needed for formylamino-HMP hydrolysis because this step is upstream of amino-HMP hydrolysis.

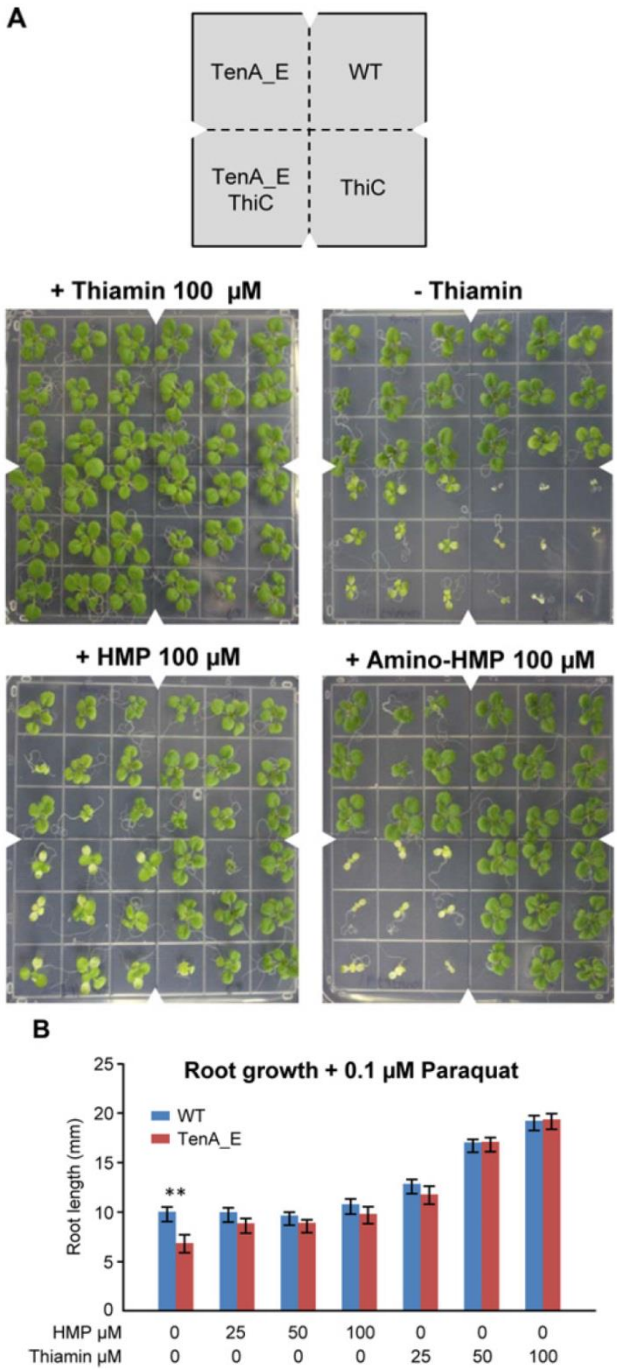


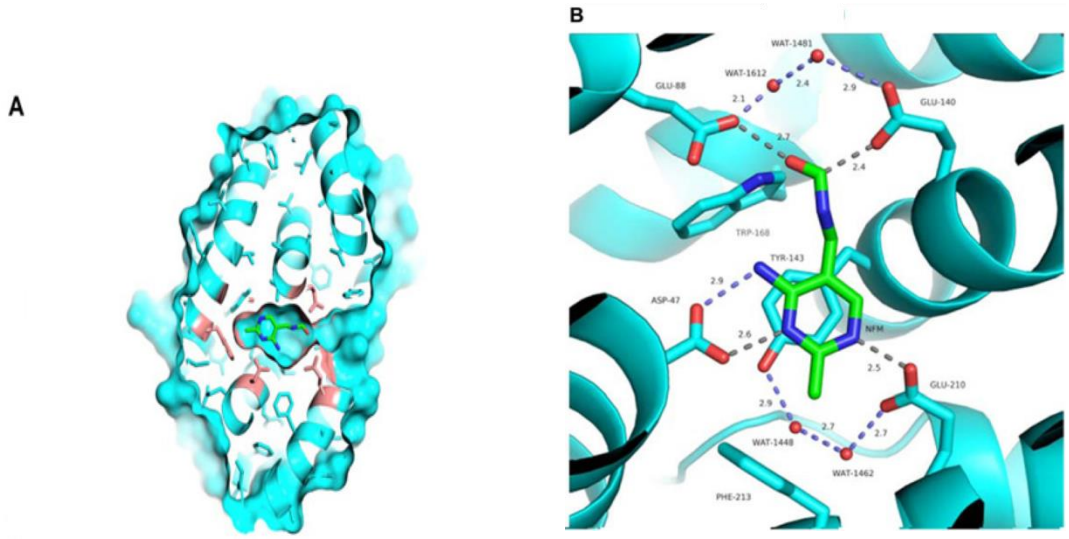
Figure 4: Genetic evidence that *Arabidopsis* TenA_E can convert amino-HMP into HMP in vivo. (A) WT (wild-type), TenA_E single mutant, ThiC single mutant and TenA_E ThiC double mutant plants were cultured for 2 weeks on Murashige and Skoog medium alone or containing 100 µM thiamin, HMP or amino-HMP. Nine plants of each genotype occupy a quadrant on each plate. That the double mutant showed somewhat more growth in the absence of thiamin than the ThiC single mutant may be attributable to differences in thiamin content of the seeds, i.e. the amount of thiamin received from the maternal

parent plants, which were irrigated with a 0.01% thiamin solution. (**B**) WT and TenA_E single mutant seeds were germinated on Murashige and Skoog medium containing 0.1 μM paraquat. Seedlings were grown vertically, and root length was measured after 9 days. Results are means±S.E.M. for 12–21 replicate plants. Root lengths of WT and TenA_E mutant plants grown on Murashige and Skoog medium without paraquat were 42.5±8.7 and 44.0±6.7 mm respectively. ***P*<0.002 was considered significant for the difference between WT and TenA_E mutant.

Because paraquat-induced oxidative stress and thiamin status are intertwined [43] and thiamin can be oxidatively degraded (Figure 1B), we compared the effect of paraquat on wild-type and *TenA_E* mutant plants by measuring root growth, which paraquat strongly inhibits [43]. The *TenA_E* mutant was found to be significantly more sensitive to paraquat than the wild-type (Figure 4B). Supplementation of the medium with 25 μM HMP eliminated this difference, as did supplementation with 25 μM thiamin (Figure 4B). However, although increasing the HMP concentration to 50 or 100 μM had no further effect on the mutant or wild-type, increasing the thiamin concentration caused a proportional increase in growth, as noted previously for wild-type roots [43]. Collectively, these data are consistent with the salvage of the thiamin HMP moiety by TenA_E. Thus oxidative stress is likely to promote thiamin degradation, so that salvage of the degradation products becomes a larger factor in maintaining thiamin pool levels, and TenA_E-mediated reclamation of HMP becomes more crucial. The observation that thiamin increases growth of both wild-type and mutant, whereas HMP does not, suggests that THZ becomes more limiting than HMP during oxidative stress, which fits with the greater stability of the latter compound [5].

Modelling of formylamino-HMP and amino-HMP in the active site of *Arabidopsis* TenA_E and *B. subtilis* TenA_C

Because a crystal structure is available for *Arabidopsis* TenA_E with bound HMP [23,25], it was straightforward to manually model formylamino-HMP (Figure 5A) and amino-HMP (Figure 5B) in the active site. As for other TenA proteins [17,20,24], the active site of *Arabidopsis* TenA_E is largely sequestered from the bulk solvent and only a very narrow passage into the active site is observed when calculating the solvent accessible surface of the protein (Figure 5A). This ground-state structure implies that the enzyme is likely to have to undergo a conformational fluctuation of substantial magnitude in order to bind the substrate and release the product. Such a requirement could explain the relatively low turnover rates (k_{cat} values) exhibited by the enzyme (Table 2).



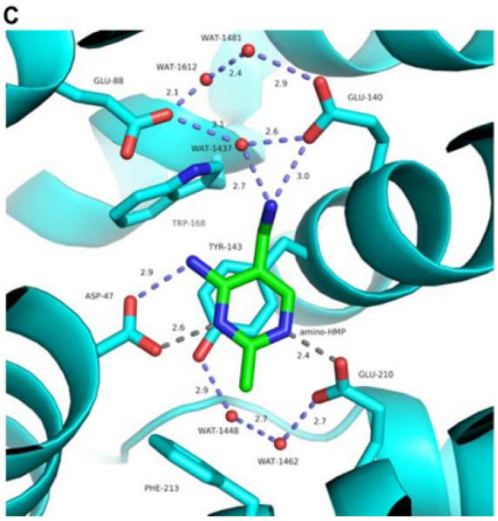


Figure 5: Formylamino-HMP or amino-HMP modelled in the active site of *Arabidopsis* TenA_E. The indicated HMP derivatives were manually docked into the experimentally determined crystal structure of *Arabidopsis* TenA_E containing bound HMP that co-purified with the enzyme (PDB code 2F2G). (**A**) The solvent-accessible surface of AtTenA_E with formylamino-HMP docked in the active site. The protein (cyan) is represented in both cartoon and stick formats. Note the narrow channel into the active site on the right. (**B**) The environment of formylamino-HMP docked in the active site of *Arabidopsis* TenA_E. Residues within 4 Å of the formylamino-HMP molecule are shown in stick representation. The blue broken lines show possible hydrogen bond interactions and the grey broken lines show other interatomic distances (in Å). Carbon atoms are green, nitrogen atoms are blue and oxygen atoms are red. Crystallographically ordered water (WAT) molecules are depicted as red spheres. (**C**) The environment of amino-HMP docked in the active site of *Arabidopsis* TenA_E, displayed in an equivalent manner to that shown in (**A**).

The structural models preserve the previously observed π–π interactions between Tyr^{143}, Phe^{50} and the pyrimidine ring, and the hydrogen bonds between Asp^{47} and the ring N4 and amino groups of the pyrimidine [23,25]. In the case of formylamino-HMP (Figure 5B), our model shows that the carboxylate groups of residues Glu^{88} and Glu^{140} are located within hydrogen-bonding distances of the carbonyl moiety of the formyl group of this ligand. This stereochemistry would enable Glu^{88} and Glu^{140} to participate in

the amidohydrolase reaction mechanism and potentially act as general base catalysts. Although neither of these glutamate residues is conserved in all TenA_E orthologues, one or the other is always present (Supplementary Figure S1). Furthermore, a pair of crystallographically ordered water molecules (wat-1612 and wat-1481) proximal to residues Glu^{88} and Glu^{140} in the active site could also participate in the amidohydrolase reaction mechanism.

These water molecules are preserved in their locations in our manual model with bound amino-HMP (Figure 5C), which also contains an additional crystallographically ordered water molecule (wat-1437) that had to be removed to model the formylamino-HMP molecule in the active site. This third water molecule could play a functional role in the amidohydrolase reaction, because it makes strong interactions with Glu^{88}/Glu^{140} and also the amino group of the aminomethyl moiety of amino-HMP.

Because *B. subtilis* TenA_C cannot hydrolyse formylamino-HMP, but can hydrolyse the larger molecule thiamin (Table 1), it was of interest to model formylamino-HMP in the TenA_C active site. Manual modelling of formylamino-HMP into the HMP-bound structure of *B. subtilis* TenA_C [17] showed that it is possible to form a good hydrogen bond between the formyl group of formylamino-HMP and the hydroxyl group on the side chain of residue Tyr^{163}. Formation of this hydrogen bond results in the nitrogen atom of formylamino-HMP being positioned 4.5 Å away from thiol group of the catalytic Cys^{135} (Supplementary Figure S5 at http://www.biochemj.org/bj/463/bj4630145add.htm), which is likely to be too far away to catalyse cleavage. If this modelled structure is stable, formylamino-HMP could act as a competitive inhibitor of amino-HMP or thiamin hydrolysis by TenA_C enzymes. (No

equivalent *in silico* docking exercise could be performed for a plant TenA_C because a structure has not yet been determined for any of these enzymes.)

Complementary expression patterns of TenA_E and ThiC genes in *Arabidopsis* and maize

Because salvage of formylamino-HMP and amino-HMP via TenA_E is a potential alternative to *de novo* HMP synthesis via ThiC, and because *ThiC* expression varies greatly between tissues [44,45], we investigated whether the expression of *TenA_E* is inversely correlated with that of *ThiC*. We used CSB.DB microarray data for *Arabidopsis* [31], and a combination of qTeller RNA-seq [32] data and our own qRT-PCR data for maize. The *Arabidopsis* data showed that ThiC is expressed far more strongly in shoots than in roots, as previously reported [44], and that *TenA_E* is expressed modestly more in roots than in shoots (Figure 5A). The same pattern was evident in maize shoots and roots (Figure 6B). During *Arabidopsis* seed development, ThiC expression declined markedly, whereas *TenA_E* expression increased (Figure 6A), and the same pattern was seen in the maize endosperm (Figure 6B). These opposite trends in *TenA_E* and *ThiC* expression in various organs are consistent with the idea that TenA_E has an *in vivo* salvage function that complements, and may sometimes partially replace, the capacity for *de novo* HMP synthesis.

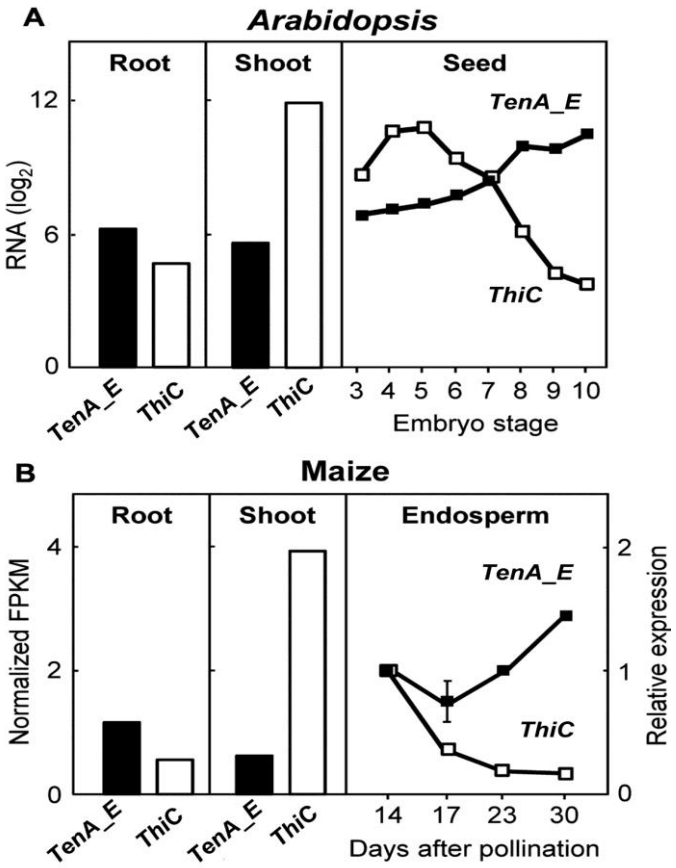


Figure 6: Complementary expression patterns of plant TenA_E and ThiC genes. (**A**) Expression of *TenA_E* and *ThiC* (At2g29630) genes in roots and shoots (above ground tissues) of 3-day-old plants, and developing seeds of *Arabidopsis*. Data were extracted from the CSB. DB microarray database and are expressed on a $\log_2$ scale. (**B**) Expression of *TenA_E* and *ThiC* (GRMZM2G027663) genes in roots and shoots of seedlings, and developing endosperm of maize. FPKM (fragments per kilobase of exon per million fragments mapped) values were from seedling root and shoot libraries in the qTeller RNA-seq database and were normalized to the mean value for each gene in 25 qTeller datasets derived from diverse maize tissues. Endosperm data were obtained by qRT-PCR and are expressed relative to the value at 14 days after pollination (= 1.0). Data points are means±S.E.M. for three replicates. Where no error bars appear they were smaller than the symbols.

DISCUSSION

To our knowledge, the present study is the first to demonstrate a biochemical function for the TenA_E subfamily of TenA proteins. The plant TenA_E proteins we tested exhibit dual formylamino-HMP amidohydrolase and amino-HMP aminohydrolase activities. Therefore, they are able to carry out two successive steps in the salvage of the thiamin breakdown product formylamino-HMP, whereas two separate enzymes, the amidohydrolase YlmB and the aminohydrolase TenA_C, mediate these steps in *Bacillus* species [8] (Figure 7). The amidohydrolase and aminohydrolase reactions catalysed by plant TenA_E proteins are chemically quite distinct, as discussed further below.

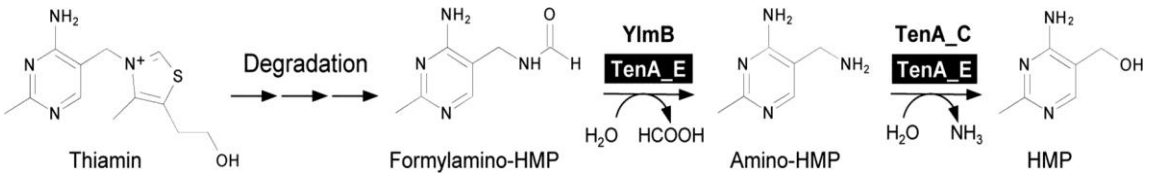


Figure 7: Salvage of the HMP moiety of thiamin in plants and *Bacillus* species. Formylamino-HMP derived from thiamin degradation is hydrolysed first to amino-HMP, and then to HMP, by different enzymes in plants and *Bacillus* species. In plants, the TenA_E protein catalyses both steps. In *Bacillus*, the YlmB protein mediates the first (amidohydrolase) step, whereas the TenA_C protein mediates the second (aminohydrolase) step [8].

Although both plant TenA_E proteins prefer formylamino-HMP to amino-HMP *in vitro*, the *Arabidopsis* mutant data show that the activity against amino-HMP is physiologically significant. Thus that the *TenA_E ThiC* double mutant can be rescued by HMP, but not amino-HMP, proves that the amino-HMP aminohydrolase activity of TenA_E is functionally important *in vivo*, at least when HMP synthesis is blocked by a ThiC mutation. That TenA_E functions similarly during normal development, waxing

important in HMP salvage when HMP synthesis wanes, is suggested by the opposing expression patterns of *TenA_E* and *ThiC*. Because *Arabidopsis* has *TenA_C* as well as *TenA_E*, the failure of amino-HMP to rescue the double mutant also indicates that TenA_C cannot replace TenA_E amino-HMP aminohydrolase activity. Whether this is because *Arabidopsis* TenA_C lacks aminohydrolase activity, lacks expression or lacks access to exogenous amino-HMP remains to be tested.

Our comparative analysis of prokaryotic genomes predicted the function of *TenA_E* so neatly that it could serve as a textbook example. Thus the analysis predicted a role for TenA_E in thiamin salvage that is similar, but not identical, to that of TenA_C, and this exactly matches the experimental findings that plant TenA_E proteins have both aminohydrolase activity (like TenA_C) and amidohydrolase activity (unlike TenA_C). This correspondence between prediction and observation makes it probable that prokaryotic TenA_E proteins are also formyl-HMP amidohydrolases. TenA_E proteins may therefore be the 'missing' amidohydrolases [8] that substitute for YlmB in the many prokaryotes whose genomes encode TenA, but not YlmB. If prokaryotic TenA_E proteins are amidohydrolases, and TenA_C proteins aminohydrolases (as in *B. subtilis* and *H. pylori* [8,20]), then the co-occurrence and chromosomal clustering of *TenA_E* and *TenA_C* genes would be explicable based on their specifying adjacent steps in thiamin salvage. Our data may also explain why the yeast TenA_E protein Pet18 can functionally replace *E. coli* ThiC when cells are cultured on minimal agar medium [46]. Agar has been shown to contain traces of HMP moieties [47], and these could be salvaged by amidohydrolase and/or aminohydrolase activities of Pet18.

Besides predicting a role for *TenA_E* in thiamin salvage, comparative genomic analysis pointed to connections between *TenA_E* and prokaryote lifestyles involving light, heat, ionizing radiation or plants. If prokaryotic TenA_E proteins have activities similar to those of their plant counterparts, these connections are readily explained because: (i) UV photolysis [5] and probably thermolysis [13] of thiamin produce amino-HMP; (ii) radiolysis produces a formylamino-HMP analogue [48,49]; and (iii) amino-HMP and formylamino-HMP from plants could be available to phytobacteria. Comparative genomics also pointed to a non-salvage role for TenA_E (and TenA_C) because they occur in thiamin auxotrophs that rely on thiamin uptake from the environment or host. A plausible rationale for such occurrences is that some thiamin breakdown products are toxic and TenA renders them harmless. Amino-HMP is known to inhibit the thiamin transporter in yeast [50]; were this true of thiamin auxotrophs, hydrolysing amino-HMP could clearly benefit them. Whatever the case, as amino-HMP and formylamino-HMP are analogues of thiamin and of pyrimidine nucleobases, they are potential metabolic inhibitors. It is therefore reasonable to infer that TenA proteins can serve a damage pre-emption function [51] by hydrolysing products that would otherwise do harm.

Although formylamino-HMP and amino-HMP were the only TenA_E substrates found in the panel of thiamin degradation products tested (Figure 1B), there is a wide range of other potential products, of which many may have a corrupt THZ moiety but an intact, and therefore salvageable, HMP moiety [5,8,11]. Such products could also be TenA_E substrates. Unfortunately, they are complex, poorly known and in some cases reactive, making them problematic to test in enzymatic assays *in vitro*.

FUNDING

This study was supported by the US National Science Foundation [grant numbers MCB-1153413 and IOS-1025398 (to A.D.H.) and MCB-0236210 (to D.K.S.)], an endowment from the C.V. Griffin Sr Foundation and a General Research Fund grant from the Oregon State University Research Office (to A.G.). This study was also supported by the Southeast Center for Integrated Metabolomics via the National Institutes of Health [grant number U24 DK097209-01A1 (to T.J.G.)] and the Northeast Structural Genomics Consortium [grant number 2U54GM75026 (to J.F.H. and J.B.)].

ACKNOWLEDGMENT

We thank Dr T.P. Begley for insightful discussion and Dr Y. Kamiyoshihara for help with translation from Japanese.

REFERENCES

1 Goyer, A. (2010) Thiamine in plants: aspects of its metabolism and functions. Phytochemistry 71, 1615–1624

2 Jurgenson, C. T., Begley, T. P. and Ealick, S. E. (2009) The structural and biochemical foundations of thiamin biosynthesis. Annu. Rev. Biochem. 78, 569–603

3 Gerdes, S., Lerma-Ortiz, C., Frelin, O., Seaver, S. M., Henry, C. S., de Cr´ecy-Lagard, V. and Hanson, A. D. (2012) Plant B vitamin pathways and their compartmentation: a guide for the perplexed. J. Exp. Bot. 63, 5379–5395

4 Rodionov, D. A., Vitreschak, A. G., Mironov, A. A. and Gelfand, M. S. (2002) Comparative genomics of thiamin biosynthesis in procaryotes: new genes and regulatory mechanisms. J. Biol. Chem. 277, 48949–48959

5 Dwivedi, B. K. and Arnold, R. G. (1973) Chemistry of thiamine degradation in food products and model systems: a review. J. Agric. Food Chem. 21, 54–60

6 McCourt, J. A., Nixon, P. F. and Duggleby, R. G. (2006) Thiamin nutrition and catalysis-induced instability of thiamin diphosphate. Br. J. Nutr. 96, 636–638

7 Chatterjee, A., Abeydeera, N. D., Bale, S., Pai, P. J., Dorrestein, P. C., Russell, D. H., Ealick, S. E. and Begley, T. P. (2011) *Saccharomyces cerevisiae* THI4p is a suicide thiamine thiazole synthase. Nature 478, 542–546

8 Jenkins, A. H., Schyns, G., Potot, S., Sun, G. and Begley, T. P. (2007) A new thiamin salvage pathway. Nat. Chem. Biol. 3, 492–497

9 Pribat, A., Blaby, I. K., Lara-N´u˜nez, A., Jeanguenin, L., Fouquet, R., Frelin, O., Gregory, 3rd, J. F., Philmus, B., Begley, T. P., de Cr´ecy-Lagard, V. and Hanson, A. D.

(2011) A 5-formyltetrahydrofolate cycloligaseparalog from all domains of life: comparative genomic and experimental evidence for a cryptic role in thiamin metabolism. Funct. Integr. Genomics 11, 467–478

10 Amos, Jr, W. H. and Neal, R. A. (1970) Isolation and identification of 3-(2-methyl-4-amino-5-pyrimidylmethyl)-4-methylthiazole-5-acetic acid (thiamine acetic acid) and 2-methyl-4-amino-5-formylaminomethylpyrimidine as metabolites of thiamine in the rat. J. Biol. Chem. 245, 5643–5648

11 Chijate, C., Albarran, G. and Negron-Mendoza, A. (1998) Radiolysis of aqueous solutions of thiamine. Radiat. Phys. Chem. 52, 401–404

12 Parkhomenko, I. M., Stepuro, I. I., Donchenko, G. V. and Stepuro, V. I. (2012) Oxidized derivatives of thiamine: formation, properties, biological role. Ukr. Biokhim. Zh. 84, 5–24

13 Hartman, G. J., Carlin, J. T., Scheide, J. D. and Ho, C. T. (1984) Volatile products formed from the thermal degradation of thiamin at high and low moisture levels. J. Agric. Food Chem. 32, 1015–1018

14 Vaid, F. H. M., Shaikh, R. H., Ansari, I. A. and Ahmad, I. (2006) Chromatographic study of the photolysis of aqueous thiamine hydrochloride and ascorbic acid solutions in the presence and absence of riboflavin. J. Chem. Soc. Pakistan 28, 461–464

15 Yazdani, M., Zallot, R., Tunc-Ozdemir, M., de Cr´ecy-Lagard, V., Shintani, D. K. and Hanson, A. D. (2013) Identification of the thiamin salvage enzyme thiazole kinase in Arabidopsis and maize. Phytochemistry 94, 68–73

16 Rapala-Kozik, M., Olczak, M., Ostrowska, K., Starosta, A. and Kozik, A. (2007) Molecular characterization of the thi3 gene involved in thiamine biosynthesis in Zea mays: cDNA sequence and enzymatic and structural properties of the recombinant bifunctional protein with 4-amino-5-hydroxymethyl-2-methylpyrimidine (phosphate) kinase and thiamine monophosphate synthase activities. Biochem. J. 408, 149–159

17 Toms, A. V, Haas, A. L., Park, J. H., Begley, T. P. and Ealick, S. E. (2005) Structural characterization of the regulatory proteins TenA and TenI from *Bacillus subtilis* and identification of TenA as a thiaminase II. Biochemistry 44, 2319–2329

18 Haas, A. L., Laun, N. P. and Begley, T. P. (2005) Thi20, a remarkable enzyme from *Saccharomyces cerevisiae* with dual thiamin biosynthetic and degradation activities. Bioorg. Chem. 33, 338–344

19 Onozuka, M., Konno, H., Kawasaki, Y., Akaji, K. and Nosaka, K. (2008) Involvement of thiaminase II encoded by the THI20 gene in thiamin salvage of *Saccharomyces cerevisiae*. FEMS Yeast Res. 8, 266–275

20 Barison, N., Cendron, L., Trento, A., Angelini, A. and Zanotti, G. (2009) Structural and mutational analysis of TenA protein (HP1287) from the Helicobacter pylori thiamin salvage pathway: evidence of a different substrate specificity. FEBS J. 276, 6227–6235

21 French, J. B., Begley, T. P. and Ealick, S. E. (2011) Structure of trifunctional THI20 from yeast. ActaCrystallogr. D. Biol. Crystallogr. 67, 784–791

22 Itou, H., Yao, M., Watanabe, N. and Tanaka, I. (2004) Structure analysis of PH1161 protein, a transcriptional activator TenA homologue from the hyperthermophilic archaeon Pyrococcus horikoshii. ActaCrystallogr. D. Biol. Crystallogr. 60, 1094–1100

23 Blommel, P. G., Smith, D. W., Bingman, C. A., Dyer, D. H., Rayment, I., Holden, H. M., Fox, B. G. and Phillips, Jr, G. N. (2004) Crystal structure of gene locus At3g16990 from *Arabidopsis thaliana*. Proteins 57, 221–222

24 Benach, J., Edstrom, W. C., Lee, I., Das, K., Cooper, B., Xiao, R., Liu, J., Rost, B., Acton, T. B., Montelione, G. T. and Hunt, J. F. (2005) The 2.35A° structure of the TenA homolog from Pyrococcus furiosus supports an enzymatic function in thiamine metabolism. Acta Crystallogr. D. Biol. Crystallogr. 61, 589–598

25 Levin, E. J., Kondrashov, D. A., Wesenberg, G. E. and Phillips, G. N. (2007) Ensemble refinement of protein crystal structures: validation and application. Structure 15, 1040–1052

26 R´edei, G. P. (1965) Genetic blocks in the thiamine synthesis of the angiosperm Arabidopsis. Am. J. Bot. 52, 834–841

27 Li, S. L. and R´edei, G. P. (1969) Thiamine mutants of the crucifer, Arabidopsis. Biochem. Genet. 3, 163–170

28 Overbeek, R., Begley, T., Butler, R. M., Choudhuri, J. V, Chuang, H. Y., Cohoon, M., de Cr´ecy-Lagard, V., Diaz, N., Disz, T., Edwards, R. et al. (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res. 33, 5691–5702

29 Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797

30 Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739

31 Steinhauser, D., Usadel, B., Luedemann, A., Thimm, O. and Kopka, J. (2004) CSB.DB: a comprehensive systems-biology database. Bioinformatics 20, 3647–3651

32 Wang, X., Elling, A. A., Li, X., Li, N., Peng, Z., He, G., Sun, H., Qi, Y., Liu, X. S. and Deng, X. W. (2009) Genome-wide and organ-specific landscapes of epigenetic modifications and their relationships to mRNA and small RNA transcriptomes in maize. Plant Cell 21, 1053–1069

33 Guan, J. C., Koch, K. E., Suzuki, M., Wu, S., Latshaw, S., Petruff, T., Goulet, C., Klee, H. J. and McCarty, D. R. (2012) Diverse roles of strigolactonesignaling in maize architecture and the uncoupling of a branching-specific subnetwork. Plant Physiol. 160, 1303–1317

34 Kurata, G.-I, Sakai, T. and Miyara, T. (1968) Studies on antagonists of thiamine: (XVIII) Reaction condition in formation of desthiothiamine from alkaline thiamine solution with amino acids. Vitamins 37, 398–402

35 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

36 Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol. Plant. 15, 473–497

37 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D. Biol. Crystallogr. 60, 2126–2132

38 Pfaffl, M. W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29, e45

39 Bordenstein, S. R., Paraskevopoulos, C., Dunning Hotopp, J. C., Sapountzis, P., Lo, N., Bandi, C., Tettelin, H., Werren, J. H. and Bourtzis, K. (2009) Parasitism and mutualism in Wolbachia: what the phylogenomic trees can and cannot say. Mol. Biol. Evol. 26, 231–241

40 Collingro, A., Tischler, P., Weinmaier, T., Penz, T., Heinz, E., Brunham, R. C., Read, T. D., Bavoil, P. M., Sachse, K., Kahane, S. et al. (2011) Unity in variety: the pan-genome of the Chlamydiae. Mol. Biol. Evol. 28, 3253–3270

41 Sun, Q., Zybailov, B., Majeran, W., Friso, G., Olinares, P. D. and van Wijk, K. J. (2009) PPDB, the Plant Proteomics Database at Cornell. Nucleic Acids Res. 37, D969–D974

42 Haughton, B. G. and King, H. K. (1958) Toxo-pyrimidine phosphate as an inhibitor of bacterial enzyme systems that require pyridoxal phosphate. Biochem. J. 70, 660–665

43 Tunc-Ozdemir, M., Miller, G., Song, L., Kim, J., Sodek, A., Koussevitzky, S., Misra, A. N., Mittler, R. and Shintani, D. (2009) Thiamin confers enhanced tolerance to oxidative stress in Arabidopsis. Plant Physiol. 151, 421–432

44 Raschke, M., B¨urkle, L., M¨uller, N., Nunes-Nesi, A., Fernie, A. R., Arigoni, D., Amrhein, N. and Fitzpatrick, T. B. (2007) Vitamin B1 biosynthesis in plants requires the essential iron sulfur cluster protein, THIC. Proc. Natl. Acad. Sci. U.S.A. 104, 19637–19642

45 Bocobza, S. E., Malitsky, S., Ara´ujo, W. L., Nunes-Nesi, A., Meir, S., Shapira, M., Fernie, A. R. and Aharoni, A. (2013) Orchestration of thiamin biosynthesis and central metabolism by combined action of the thiamin pyrophosphate riboswitch and the circadian clock in Arabidopsis. Plant Cell 25, 288–307

46 Morett, E., Korbel, J. O., Rajan, E., Saab-Rincon, G., Olvera, L., Olvera, M., Schmidt, S., Snel, B. and Bork, P. (2003) Systematic discovery of analogous enzymes in thiamin biosynthesis. Nat. Biotechnol. 21, 790–795

47 Karunakaran, R., Ebert, K., Harvey, S., Leonard, M. E., Ramachandran, V. and Poole, P. S. (2006) Thiamine is synthesized by a salvage pathway in *Rhizobium leguminosarumbv*. viciae strain 3841. J. Bacteriol. 188, 6661–6668

48 Xuetong, F. (2012) Radiation chemistry of food components. In Food Irradiation Research and Technology (Fan, X. and Sommers, C. H., eds), pp. 75–98, Wiley-Blackwell, NJ

49 Kawasaki, C. (1964) Modified thiamine compounds. In Vitamins and Hormones, Vol. 21 (Harris, R. S, Wool, I. G. and Loraine, J. A., eds), pp. 69–111, Academic Press, NY

50 Iwashima, A., Kawasaki, Y. and Kimura, Y. (1990) Transport of 2-methyl-4-amino-5-hydroxymethylpyrimidine in *Saccharomyces cerevisiae*. Biochim. Biophys. Acta 1022, 211–214

51 Linster, C. L., Van Schaftingen, E. and Hanson, A. D. (2013) Metabolite damage and its repair or pre-emption. Nat. Chem. Biol. 9, 72–80

52 Balaghi, M. and Pearson, W. N. (1967) Comparative studies of the metabolism of 14C-pyrimidine-labeled thiamine, 14C-thiazole-labeled thiamine and 35S-labeled thiamine in the rat. J. Nutr. 91, 9–19

SUPPLEMENTARY DATA

Table S1 Oligonucleotide primers used in the present study

F, forward; R, reverse.

Primer name	Sequence 5′→3′	Application
AtTenA_E to pYES2.1 TOPO F	ATGGAGAAGAGAGGGGTGATC	At3g16990 cloning intopYES2.1TOPO
AtTenA_E to pYES2.1 TOPO R	CTACTGTCCTCCACGGCTCAT	At3g16990 cloning intopYES2.1 TOPO
AtTenA_E to pET28B F NcoI	ATGCATCCATGGAGAAGAGAGGGGTGAT	At3g16990 cloning intopYES2.1
AtTenA_E to pET28B R NotI	ATGCATGCGGCCGCCTGTCCTCCACGGCTCAT	At3g16990 cloning intopYES2.1
ZmTenA_E to pET28B F NcoI	ATGCATCCATGGACGGCGGCG	GRMZM2G080501 cloninginto pET28B
ZmTenA_E to pET28B R NotI	ATGCATGCGGCCGCGGACCGAGAAGAGCTCATATC	GRMZM2G080501 cloninginto pET28B
BsTenA_E to pET28B F NcoI	ATGCATCCATGGCGTTTTCAGAAGAATGCCGC	BSU11650 cloning intopET28B
BsTenA_E to pET28B R NotI	ATGCATGCGGCCGCTCCGTTATGGCGTGAAG	BSU11650 cloning intopET28B
AtTenA_E genotyping F	TGGAACTTGTTTCTCTTTGGC	Genotyping *Arabidopsis*SALK_062985
AtTenA_E genotyping R	TTTAAAGCCCAATGAACGTG	Genotyping *Arabidopsis* SALK_062985
LB1.3 SALK genotyping	ATTTTGCCGATTTCGGAAC	Genotyping *Arabidopsis* SALK_062985
RT PCR AtTenA_E F	ATCTATACGGCGGCTACACG	RT-PCR to verify KOSALK_062985 (At3g16990)
RT PCR AtTenA_E R	TGCTATACCACCAAGCACCA	RT-PCR to verify KOSALK_062985 (At3g16990)
RT PCR Actin F	CTTGCACCAAGCAGCATGAA	RT-PCR actin positive
RT PCR Actin R	CCGATCCAGACACTGTACTTCCTT	RT-PCRactin positive control (AT3G18780)
qRT PCR maize ZmTenA_E F	CGTCTATCAGCGACGAGATCTC	qRT-PCR maize TenA_E(GRMZM2G080501)
qRT PCR maize ZmTenA_E R	ACTCCAACTGTTTCTCCATG	qRT-PCR maize TenA_E(GRMZM2G080501)
qRT PCR maize ZmThiC F	GATGCAAATGATAGTGCACAG	qRT-PCR maize ThiC(GRMZM2G027663)
qRT PCR maize ZmThiC R	ACTCCAACTGTTTCTCCATG	qRT-PCR maize ThiC(GRMZM2G027663)
qRT PCR maize Zm18S rRNA F	ATTCTATGGGTGGTGGTGCAT	qRT-PCR maize 18S rRNA
qRT PCR maize Zm18S rRNA R	TCAAACTTCGCGGCCTAAA	qRT-PCR maize 18S rRNA

Table S2 Occurrence of TenA genes in 345 genomes representing major prokaryote taxa

Full results are available online in the TenA subsystem of the SEED database (at http://pubseed.theseed.org//SubsysEditor.cgi?page=ShowSpreadsheet&subsystem=TenA)

Taxa	Total genomes	TenA absent	TenA_C only	TenA_C + TenA_E	TenA_E only
Crenarcheoata (Archaea)	6	2	1	3	0
Euryarcheota (Archaea)	4	2	0	2	0
Actinobacteria	23	8	8	7	0
Aquificae	5	5	0	0	0
Bacteriodetes	9	4	5	0	0
Chlorobi	5	4	1	0	0
Chlamydiae/Verrucomicrobia	13	11	2	0	0
Chloroflexi	4	2	2	0	0
Cyanobacteria	39	22	11	6	0
Deinococcus-Thermus	8	3	0	0	5
Acidobacteria	3	3	0	0	0
Firmicutes	10	3	7	0	0
Fusobacteria	10	9	1	0	0
Planctomycetes	12	11	1	0	0
Alphaproteobacteria	40	20	12	6	2
Betaproteobacteria	30	20	6	4	0
Gammaproteobacteria	59	37	12	6	4
Deltaproteobacteria	10	8	2	0	0
Epsilonproteobacteria	37	19	18	0	0
Spirochaetes	14	9	5	0	0
Thermotogae	4	4	0	0	0
Totals	345	206	94	34	11
Percentages	100	59.7	27.2	9.9	3.2

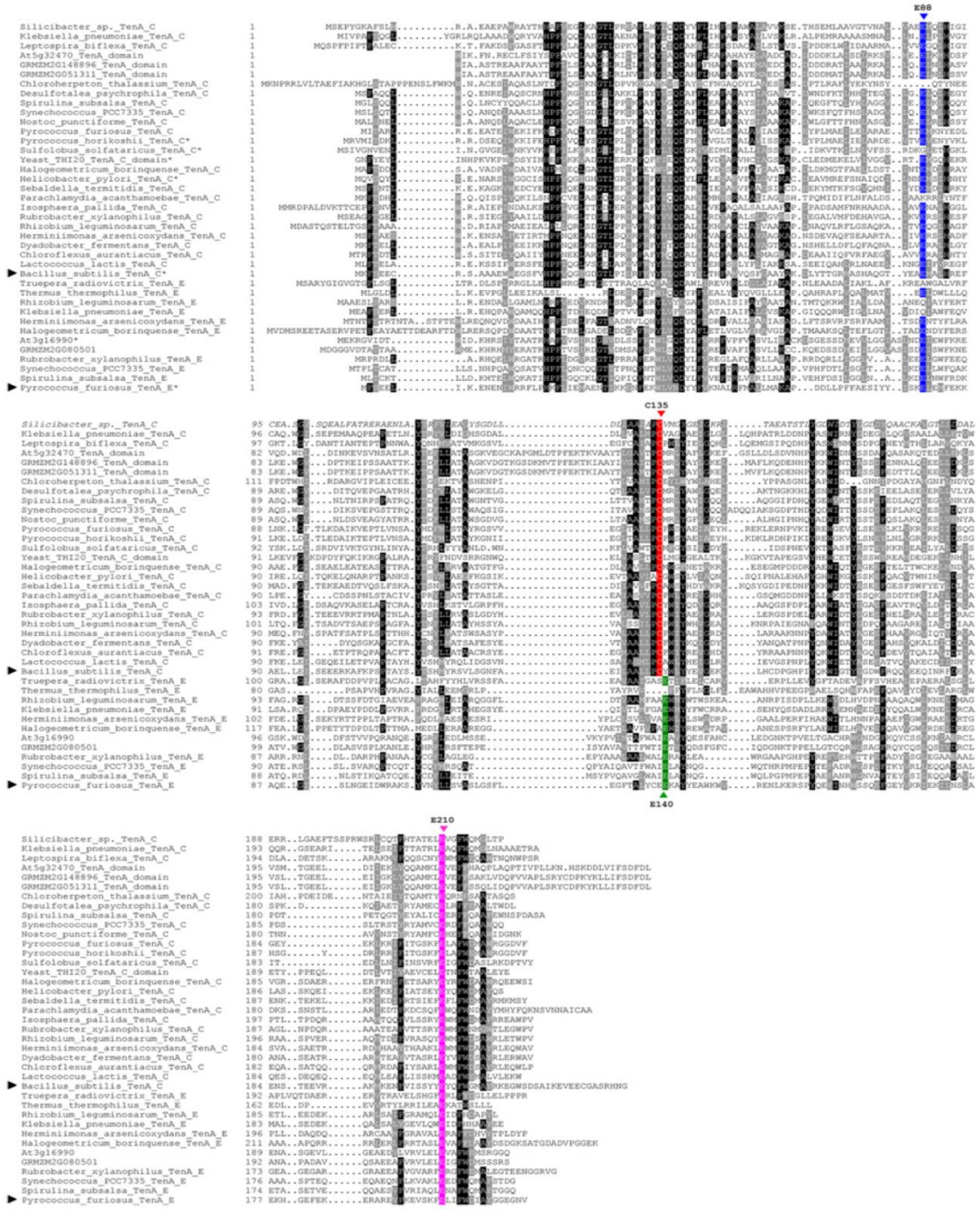


Figure S1: Alignment analysis of representative TenA family proteins. The sequences analysed come from genomes representing the major taxa in which TenA proteins occur. Ten of these genomes (eight prokaryotic and two plant genomes) encode both TenA_C and TenA_E proteins. Proteins for which crystal structures are available are indicated with an asterisk. Black arrowheads mark the foundational *Bacillus subtilis* TenA_C and Pyrococcusfuriosus TenA_E sequences [21]. Red, blue and magenta colours indicate the diagnostic active site cysteine residue (Cys135 in *B. subtilis* TenA_C) and two glutamate

residues, of which one (Glu210 in *Arabidopsis* TenA_E) is conserved in all TenA_E and TenA_C proteins and the other (Glu88 in *Arabidopsis* TenA_E) is not.

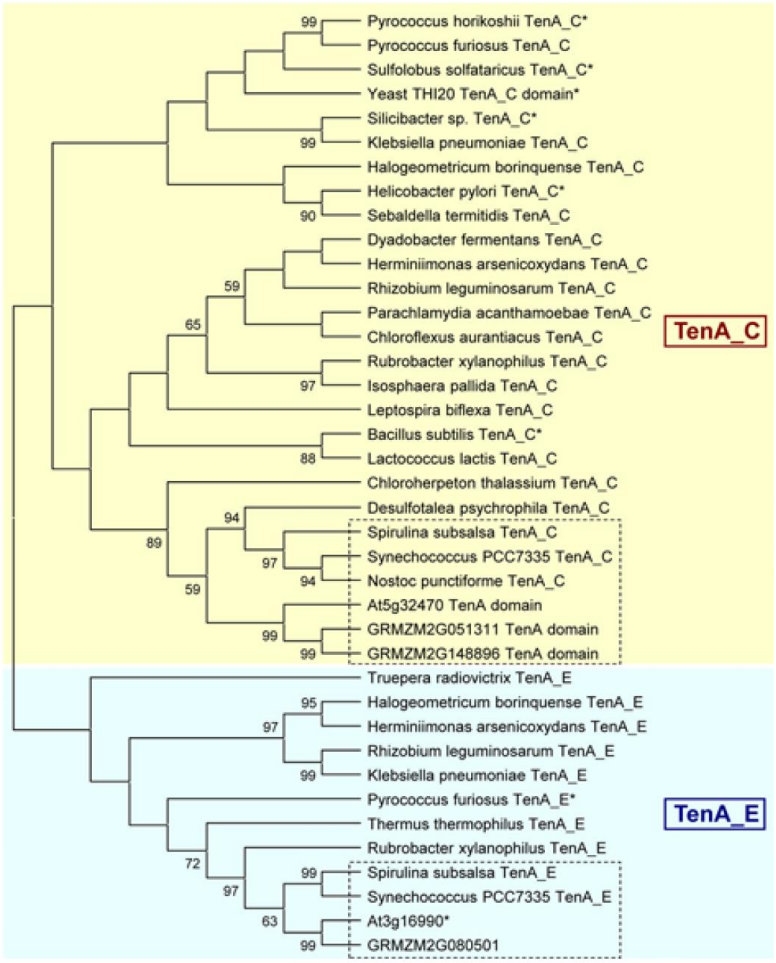


Figure S2: Phylogenetic analysis of representative TenA family proteins. Neighbour-joining phylogenetic tree. The sequences analysed come from genomes representing the major taxa in which TenA proteins occur. Ten of these genomes (eight prokaryotic and two plant genomes) encode both TenA_C and TenA_E proteins. Proteins for which crystal structures are available are indicated with an asterisk. Bootstrap values (%) for 1000 replicates are shown next to nodes; values <50% are omitted. Only tree topology is shown; branch lengths are not proportional to estimated numbers of amino acid substitutions. Plant and cyanobacterial TenA_C and TenA_E sequences are boxed.

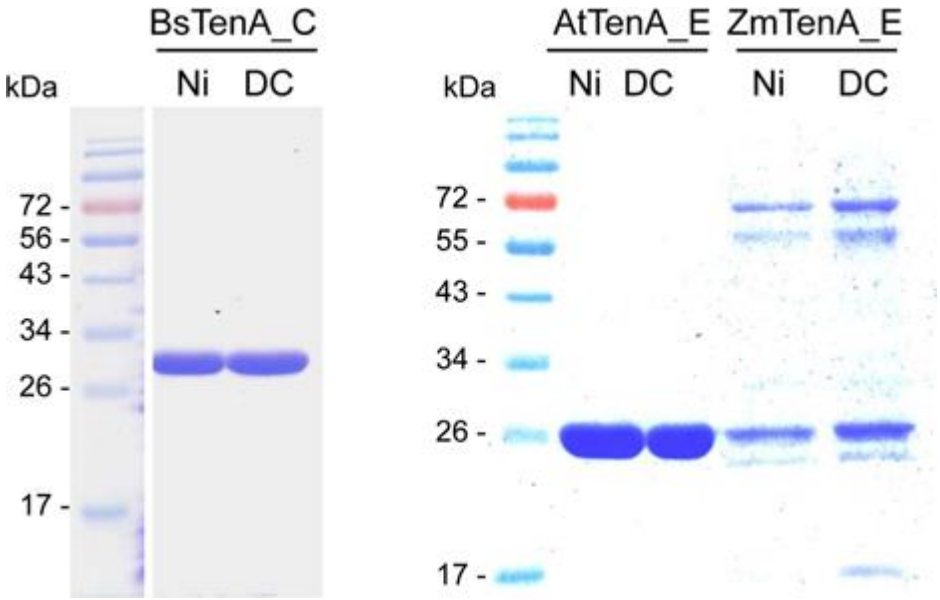


Figure S3: Isolation of recombinant *B. subtilis* TenA_C and plant TenA_E proteins. Purification to near-homogeneity of recombinant proteins used in the present study. Fiven micrograms of Ni2+ -affinity purified (Ni) and desalted and concentrated (DC) *B. subtilis* TenA_C (Bs), *Arabidopsis* TenA_E (At) and maize TenA_E (Zm) were analysed by SDS/PAGE (12% gel) with Coomassie Blue staining. The *B. subtilis* and *Arabidopsis* protein preparations were estimated to be_90% pure. Molecular mass markers are shown on the left of each gel.

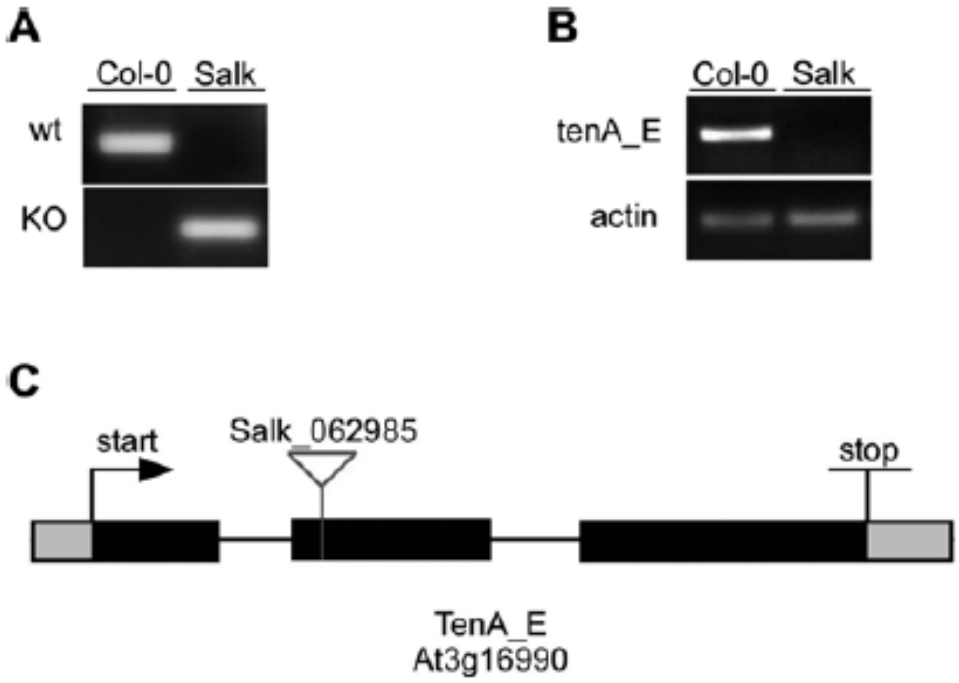


Figure S4: Characterization of an *Arabidopsis* T-DNA insertional knockout mutant of TenA_E. (**A**) Genotyping the TenA_E (At3g16990) homozygous T-DNA knockout line (Salk 062985.53.80.x). Genomic DNA from the knockout line and the corresponding wild-type (wt) (Col-0) was used as a template for PCR with wild-type allele primers or T-DNA insertion allele primers. Amplicons were analysed by agarose-gel electrophoresis. (**B**) Testing for the presence of the TenA_E transcript. Total RNA was used for first-strand cDNA synthesis, followed by PCR with primers designed to amplify a fragment of the TenA_E transcript or a fragment of the actin transcript. Amplicons were analysed by agarose-gel electrophoresis. (**C**) Confirmation of the T-DNA insertion site in the TenA_E gene. PCR amplicons obtained with the T-DNA insertion allele primers in (**A**) above were cloned and sequenced. The gene model is drawn to scale with exons as black boxes, introns as black lines, and 5_- and 3_-UTRs as grey boxes. The T-DNA insertion site is indicated (Salk 062985). The positions of the start and stop codons are indicated.

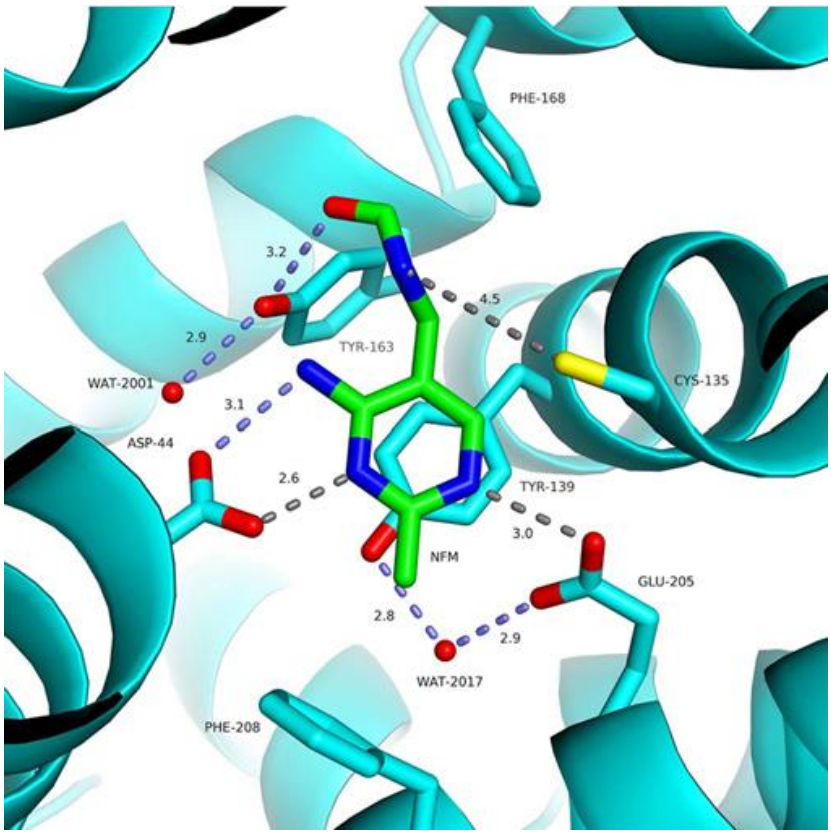


Figure S5: The environment of a formylamino-HMP molecule bound to *Bacillus subtilis* TenA_C. The indicated formylamino-HMP molecule was manually docked into the experimentally determined crystal structure of *B. subtilis* TenA_C containing bound HMP (PDB code 1YAK) [17]. Residues within 4A° of the formylamino-HMP molecule are shown in stick representation. Blue broken lines show possible hydrogen bond interactions; grey broken lines are interatomic distances. Distances are in A°. Carbon atoms are green, nitrogen atoms are blue and oxygen atoms are red. Water (WAT) molecules are depicted as red spheres. Note the hydrogen bond involving the formyl group of formylamino-HMP and Tyr163 that could position the nitrogen atom of formylamino-HMP too far (4.5A°) from the catalytic residue Cys135 to allow the cleavage reaction to take place.

CHAPTER 4

Genetic engineering of *Arabidopsis* plant to increase thiamin and stress tolerance

ABSTRACT

Thiamin and thiamin pyrophosphate (TPP) are essential components for the function of numerous enzymes involved in the metabolism of carbohydrates and amino acids in living organisms. In addition to its role as a cofactor, thiamin plays a key role in resistance against biotic and abiotic stresses in plants. Most of the studies used exogenous thiamin to enhance stress tolerance in plants. However, we could achieve this objective through the genetic engineering of the *Arabidopsis* (*Arabidopsis thaliana*) by overexpressing of the thiamin biosynthetic genes *Thi4*, *ThiC*, and *ThiE* using strong seed-specific promoters. Elevated thiamin content in transgenic plants was accompanied by enhanced expression of transcripts encoding thiamin cofactor-dependent enzymes. Furthermore, seed germination and root growth in thiamin over-producing lines were more tolerant to oxidative stress caused by salt and paraquat treatments. The transgenic seeds could also accumulate more oil and carbohydrate but less protein than the control plants. The same results were also observed in TPP over-producing plants generated by the seed-specific overexpression of *TPK1* gene in this study. Taken together, our findings suggest that thiamin and TPP over-production in transgenic lines could confer a boosted abiotic stress tolerance as well as could alter the seed carbon partitioning.

INTRODUCTION

Thiamin (Vitamin B_1) in the form of thiamin pyrophosphate (TPP) is an essential component for the function of numerous enzymes such as transketolase (TK), pyruvate dehydrogenase (PDH), and α-ketoglutarate dehydrogenase (α-KGDH) which are involved in central metabolism in all organisms (Jordan, 2003; Nosaka, 2006; Zallot et al., 2014). Human and animals can synthesize TPP from thiamin but they are not able to synthesize thiamin *de novo*. Hence, they must take it up from their diet to maintain a normal metabolism (Roje, 2007; Kowalska and Kozik, 2008). Severe vitamin B_1 deficiency causes the lethal disease beriberi in humans (Lonsdale, 2006; Roje, 2007). Most of the key enzymes involved in thiamin *de novo* biosynthesis were identified in plants and bacteria (Jurgenson et al., 2009; Goyer, 2010). In plants, thiamin is synthesized by the condensation of 4-amino-5-hydroxymethyl-2-methylpyrimidine pyrophosphate (HMP-PP) and 4-methyl-5-(β-hydroxyethyl)thiazole phosphate (HET-P) which is mediated by a bifunctional enzyme known as thiamin-phosphate pyrophosphorylase or ThiE (Roje, 2007; Goyer, 2010; Yazdani et al., 2013). These precursors of TPP are biosynthesized through two independent pathways mediated by thiazolephosphate synthase and phosphomethylpyrimidine synthase enzymes which are known as Thi4 and ThiC, respectively. Likewise, at the end of the pathway, free thiamin is pyrophosphorylated by thiamin pyrophosphokinase (TPK) enzyme (Ajjawi et al, 2007) to produce its active form TPP.

TPP has been demonstrated to play a vital role other than cofactor function in resistance against biotic and abiotic stress conditions in different plant species (Sayed and Gadallah, 2002; Ahn et al., 2005; Tunc-Ozdemir et al., 2009). Abiotic stress, which is the

major cause of crop loss worldwide, can reduce average yield for most staple crops by more than 50% (Boyer, 1982; Bray et al., 2000). Recent studies showed that total thiamin content is elevated in plant seedlings subjected to osmotic, salt, and oxidative stress conditions (Rapala-Kozik et al., 2008; Tunc-Ozdemir et al., 2009).

Plants subjected to high salinity can suffer from water stress, ion toxicity, nutritional imbalance, oxidative stress, and/or a combination of these adverse factors (Ashraf, 1994; Hernandez and Almansa, 2002). When exposed to high salinity, vital activities in plants such as photosynthesis, protein, and lipid metolisms can be affected (Parida and Das 2005) by the production of reactive oxygen species (ROS) (Imlay, 2003). ROS production have shown to have detrimental effects on lipids, proteins and nucleic acids (Mc Kersie and Leshem, 1994). Salt stress can produce superoxide radicals, which can be removed by antioxidant enzymes such as superoxide dismutase and ascorbate peroxidase (Hernandez and Almansa, 2002). It is also well-documented that high salinity can cause osmotic and ionic stresses wich trigger oxidative stress and plants combat with these unfavorable conditions by the induction of their antioxidant systems (Parida and Das, 2005).

Paraquat can also induce oxidative stress and reduce plant productivity (Bowler et al., 1992; Koca et al., 2007). In chloroplasts, the herbicide paraquat (1,1'-dimethyl-4,4'-bipyridylium) plays its role in the light reaction of photosynthesis by the production of superoxide radicals (Dodge, 1994). These noxious molecules are then scavenged by a chain reaction mediated by antioxidant enzymes such as superoxide dismutase (Asada and Kiso, 1973), ascorbate peroxidase, and glutathione reductase (Foyer and Mullineaux, 1994).

In addition to the accumulation of thiamin in plants subjected to various abiotic stresses, its exogenous application conferred some degree of resistance to salt and oxidative stresses. Sayed and Gadallah (2002) reported that in sunflower plants, either sprayed on shoots or applied to roots, thiamin alleviates the adverse effects of salinity on shoot and root growth.

Transcriptome and proteome analysis were also performed in recent years to gain more insight about the relationship between thiamin biosynthesis and stress tolerance in plants. Transcriptomic studies have been reported on the accumulation of some thiamin biosynthesis enzymes transcripts under heat and drought stress conditions in plants (Rizhsky et al., 2004). Moreover, leaf proteome analysis of *Populus euphratica* plant subjected to heat stress condition showed changes in the abundance of thiamin biosynthesis enzymes (Ferreira et al., 2006). Proteomic analysis on rice seedlings has also revealed that thiamin has an important role in this plant in response to cold stress (Cui et al., 2005).

Although the role of thiamin in the elevation of plant responses to biotic and abiotic stress conditions has been confirmed, there are fewer studies regarding the evaluation of the response of thiamin biosynthetic genes in stress conditions (Pourcel et al., 2013). It has been shown that combinational application of hydroxyethylthiazole (HET) and hydroxymethylpyrimidine (HMP) in the medium of wild type and *ThiC* mutants could augment the level of thiamin and its phosphate esters in *Arabidopsis* plant, while the solo application of these precursors did not show significant increase. These findings propose that thiamin biosynthesis is limited by the availability of both thiazole and pyrimidine (Pourcel et al., 2013). In contrast, the level of thiamin monophosphate

(TMP) and TTP did not show dramatic difference in the wild type plants by dual application of HET and HMP suggesting that ThiE and phosphatase are not limiting in thiamin biosynthesis pathway (Pourcel et al., 2013).

Taken together, these evidences confirm that thiamin plays an important role to sustain normal functioning of plants subjected to various environmental stresses. Therefore, the aim of this research is to generate transgenic plants with improved thiamin content in the seeds in order to: 1- boost seed nutritional value, 2- increase abiotic stress tolerance and 3- determine the possible impact of excess amount of thiamin on some metabolic pathways in plants. To achieve these goals, we took advantage of molecular biology and biochemistry approaches for overexpression of *Thi4*, *ThiC*, *ThiE*, and *TPK1* genes under the control of seed specific promoters and analyzing the transgenic *Arabidopsis* seeds. Ultimately, we will have the transgenic plants with improved nutritional content and also increased abiotic stress tolerance which are important for food crops.

EXPERIMENTAL

Plant materials and growth conditions

Seeds of wild type *Arabidopsis* plants (ecotype Columbia) were provided from *Arabidopsis* Biological Resource Center (ABRC). Wild type seeds were then being grown on 1x MS (Murashige and Skoog, 1962) agar plates at 21°C under 100 µmol m^{-2} s^{-1} constant light for 10 days. They were then transplanted to the soil for growth under 130 µmol m^{-2} s^{-1} light with 16 h light/8 h dark period in a growth chamber. Twenty five-

day old plants were used for floral dip transformation by *Agrobacterium tumefaciens* (Clough and Bent., 1998).

Constructs for overexpression of thiamin biosynthesis genes

To overexpress the cDNAs of the *Thi4* (At5G54770) and *ThiC* (At2G29630) genes in wild type *Arabidopsis* seed, two gene expression cassettes were made (Figure 1) in which the *Thi4* and *ThiC* open reading frames (ORF) were cloned into the NotI site of pMS4 and pKMS2 under the control of Brassica napin (Ellerström et al., 1996) and oleosin (Keddie et al., 1994) promoters, respectively. The pMS4 and pKMS2 vectors were then excised in their AscI sites to remove the fragments of napin promoter::*Thi4*::glycinin 3′UTR and oleosin promoter::*ThiC*::oleosin terminater, respectively. These two expression cassettes were then sub-cloned into the AscI and MluI sites of the RS3GseedDSredMSC binary vector, respectively. The *ThiE* (At1G22940) ORF was cloned directly into the XmaI site of RS3GseedDSredMSC vector under the control of soybean glycinin promoter (Iida et al., 1995). The binary vector harboring these 3 genes was transformed into *Agrobacterium tumefaciens* strain GV3101. Transgenic *Arabidopsis* seeds (T_1) were produced using *Agrobacterium*-mediated transformation by the floral dip method (Clough and Bent., 1998). The T_1 seeds were then selfed to produce the T_2 seeds. To identify the homozygous lines, T_2 plants were grown and segregation analysis was performed on the siliques containing the T_3 seeds against the DSRed gene using fluorescent microscope. More than twenty homozygous lines were obtained and eight independent homozygous lines were used for this study. In parallel, homozygous plants harboring the empty vector control were produced to be used for further analysis.

For the overexpression of *TPK1* (At1g02880) gene, *TPK1* ORF was cloned into the XmaI site of the RS3GseedDSredMSC binary vector to have soybean glycinin promoter::*TPK1*::glycinin terminator. PCR was performed to amplify the *TPK1* gene with the glycinin promoter and terminator. This fragment was next sub-cloned into the NotI site of the pART27 binary vector (Figure 12). The homozygous transgenic and empty vector control lines were produced as described for 3-gene overexpressor plants, except we used kanamycin to select the transgenic lines.

HPLC analysis

Dry seeds (~ 5 mg) from wild type, empty vector control, and transgenic lines were ground in 300 mL of 2% (w/v) trichloroacetic acid (TCA). Samples were incubated at 95 °C for 30 min and kept on ice for 2 min, and then centrifuged at 14,000g for 5 min. The supernatant was centrifuged using Nanosep Centrifugal Filter Device (0.2 mm) columns (Pall Life Sciences) for 3 minutes. Free thiamin and TPP in 2% TCA were then converted into thiochrome using cyanogen bromide (Kim et al., 1998). Thiochrome peaks were identified by HP 1046A fluorescence detector (Agilent Technologies, Palo Alto, CA) at 370 nm excitation and 430 nm emission using HP 1100 HPLC (Agilent Technologies, Palo Alto, CA) equipped with a Capcell Pak NH_2 column (150 mm × 4.6 mm i.d., Shiseido, Tokyo). A 4:6 (v/v) solution of 90 mM potassium phosphate buffer, pH 8.4 and acetonitrile was used as the mobile phase (Ajjawi et al. 2007a).

RNA isolation and Real-Time Quantitative RT-PCR

Total RNA was isolated from 9-day old developing *Arabidopsis* siliques (Van Erp et al., 2014) using the Spectrum^{TM} Plant Total RNA Kit (Sigma-Aldrich). The isolated

total RNA was then treated with DNase I (Amplification Grade, Sigma-Aldrich) according to manufacturer's protocol. First strand cDNA was synthesized utilizing 2 µg of total RNA and iScript^{TM} cDNA Synthesis Kit (Bio-Rad). A fraction (0.75 µg) of the cDNA was used as the template in 10 µl reaction mixture per well in real-time PCR. TaqMan probes were purchased for *Thi4*, *ThiC*, *ThiE*, *TPK1*, *TK* (AT245290), *TK* (AT3G60750), *PDH* (AT1G24180), *PDH* (AT2G34590), *α-KGDH* (AT3G55410), *α-KGDH* (AT5G65750), *DXPS* (AT5G11380) and *DXPS* (AT4G15560) (Assay ID. No. At02243600_g1, At02252013_g1, At02305767_g1, At02333096_gH, At02263110_g1, At02197798_g1, At02306409_g1, At02216989_g1, At02189484_g1, At02283104_g1, At02273275_g1 and At02221416_g1, respectively) from Applied Biosystems (Life Technologies, Carlsbad, CA). The *Arabidopsis eEF-1 α* (Assay ID. No. At02337969_g1) gene tagged with VIC dye (Life Technologies, Carlsbad, CA) was added in each well as an internal control. PCR thermal cycling conditions used for amplification was 95 °C for 2 min followed by 39 cycles of 95 °C for 15 sec., 60 °C for 1 min. Gene expression levels were quantified by CFX96 Real-Time System (Bio-Rad). Data was analyzed using $2^{-\Delta\Delta CT}$ method (Tunc-Ozdemir et al., 2009).

Stress assays

In all stress experiments the seeds were surface sterilized and kept at 4 °C for 2 days in complete darkness and then germinated at 22 °C under 100 µmol m^{-2} S^{-1} constant light in growth chamber. Germination assays were performed with at least five technical replications, each containing 50 seeds per line, and repeated at least three times. For seed germination assay, the seeds were grown on 1× MS agar plates without thiamin and containing the different concentrations of NaCl and paraquat (methyl viologen) for 5 days

and the number of germinated seeds were counted each day. As a morphological marker, visible penetration of the radicles into the seed coat was considered for seed germination (Kim et al., 2008). These plates were also kept in the growth chamber for 4 more days to evaluate the number of viable seedlings for each genotype.

For root growth assay, the seedlings (5 seeds per line with at least 10 technical replications) were grown vertically on 1× MS agar plates without thiamin and supplemented with various concentrations of NaCl and paraquat and the increase in root length was measured each day for 9 days. The RGR was then calculated as described by Lupini et al. (2010).

Lipid analysis

Seed oil analysis was performed according to Li et al. (2006). To measure seed oil content, 2 mg of seed was measured out and aliquoted to glass tubes. A 300 µl of the mixture of toluene and triheptadecanoin (total of 96 µg triheptadecanoin per sample) was added to each sample, then 25 µl of butylated hydroxytoluene solution (0.2% butylated hydroxytoluene in MeOH), and finally 1 ml of sulfuric acid (5% H_2SO_4 in MeOH). The samples were incubated at 95 °C for 1.5 h and then allowed to cool to room temperature for 30 minutes. A 1.5 ml of 0.9% NaCl (W/V) was added to each sample. Finally, a 2 ml hexane extraction was performed on each sample twice by vortexing to mix and then centrifugation for phase separation. The top supernatant layer of each extraction was removed and pooled into a fresh glass tube and then dried down using a stream of N_2 (g). Samples were then re-suspended in 400 µl of the hexane and transferred to a GC vial. Oxidative air was purged from each GC vial using a gentle stream of N2 (g).

Samples were run through a DB-23 column (Agilent Technologies, Santa Clara, CA) using 2 µl injections in a 25:1 split ratio. Injector and FID temps set to 260 °C. Prior to starting the run, the column was conditioned at 240 °C for 2 hours. For each sample, the oven temperature started at 50 °C for each run and was ramped up to 150 °C at 25°C/minute and then held at 150 °C for 3 minutes. The oven was then ramped to 240 °C at 10 °C/minute, and then the temperature was held at 240 °C for 5 minutes.

For oil data analysis, the raw data for each chromatogram was transferred to an excel format. The data was processed by labeling each Major FAME peak using the external standards as a guide. Analysis was performed by normalizing the area under each peak to the area of the C17 internal standard peak to produce a normalized peak area. The normalized peak area was multiplied by the mass of internal standard added to each sample (0.096 mg) for an amount of lipid per peak, which was then divided by the total sample weight (2 mg) and multiplied by 100% to produce an oil content relative to dry weight. The peaks for each sample were summed to produce a total lipid content for that sample. The data for each sample was analyzed for an average value, standard error, and a 95% confidence interval.

Protein and carbohydrate assay

Seed protein content was measured as described by Focks and Benning (1998). Briefly, 20 to 50 seeds were weighed out and homogenized in 250 µl of acetone in a 1.5-ml tube. Following centrifugation at 16,000*g*, the supernatant was discarded and the vacuum-dried pellet was resuspended in 250 µl of extraction buffer containing 50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 1 mM EDTA, and 1% (w/v) SDS. Following a 2 h incubation at 25 °C, the homogenate was centrifuged at 16,000*g* for 5 min and 5 µl of the

supernatant was used for protein measurement in a microplate format utilizing the Lowry DC protein assay kit and γ-globulin for generating standard curve (Bio-Rad).

For quantification of total carbohydrate, 50 seeds were homogenized in 500 µL of 80% (v/v) ethanol in a 1.5-ml tube and incubated at 70 °C for 90 min (Focks and Benning 1998). The total sugar was measured using the phenol-sulfuric acid method in microplate format and glucose was used to generate standard curve (Masuko et al., 2005). Soluble sugars were analyzed according to Focks and Benning (1998).

Hypocotyl length assay

Hypocotyl length assay was performed as described in Gantlet website (www.gantlet.org). In brief, surface sterilized seeds were plated on agar plates containing 0.5 × MS salt without thiamin and sucrose. The seeds of each genotype were plated in different positions to minimize the plate position effects. The edge of the plates was then wrapped with 3M surgical tape. Following wrapping the plates in aluminum foil, they were kept in 4 °C for 2 days in complete darkness to break the seed dormancy. The aluminum foil was then removed and the plates were placed in 22 °C under 100 µmol m^{-2} S^{-1} constant light for 24 h. After 24 h, the plates re-wrapped in aluminum foil individually and placed vertically in 22 °C. The hypocotyl length was measured for 8 days.

Statistical analysis

Data analysis was performed using Student's *t* test (Suzuki et al., 2008) to show the significance of differences between data sets.

Table 1: Oligonucleotide primers used in this study.

PCR primer name	Sequence 5' to 3'	Application
AtThi4-Forward-NotI	ATAGCGGCCGCATGGCTGCCATAGCTTCTAC	To clone into pMS4 vector
AtThi4-Reverse-NotI	ATAGCGGCCGCTTAAGCATCTACGGTTTCAGC	To clone into pMS4 vector
AtThiC-Forward-NotI	ATAGCGGCCGCATGGCTGCTTCAGTACACTG	To clone into pKMS2 vector
AtThiC-Reverse-NotI	ATAGCGGCCGCTTATTTCTGAGCAGCTTTGAC	To clone into pKMS2 vector
AtThiE-Forward-XmaI	ATACCCGGGATGAATAGCTTAGGAGGAATT	To clone into RS3GseedDSredMSC vector
AtThiE-Reverse-XmaI	ATACCCGGGTCAAATTCCCCTTTTGCTCTC	To clone into RS3GseedDSredMSC vector
AtTPK1-Forward-XmaI	ATACCCGGG ATGTCAGCCATGGATGTTATGATT	To clone into RS3GseedDSredMSC vector
AtTPK1-Reverse-XmaI	ATACCCGGG CTAAGGTGTATGGTCTTGTATGGA	To clone into RS3GseedDSredMSC vector
Glycinin promoter::AtTPK1::Glycinin terminator-Forward-PspOMI*	ATAGGGCCCGATCCGTACGTAAGTACGTAC	To clone into pART27 vector
Glycinin promoter::AtTPK1::Glycinin terminator-Reverse-PspOMI*	ATAGGGCCCTAAGTCATGAAGAACCTGATA	To clone into pART27 vector

Underlined: restriction site; *: PspOMI and NotI have compatible ends.

RESULTS

Overexpression of thiamin biosynthesis genes in the seed of wild type *Arabidopsis*

In order to overexpress the *Thi4*, *ThiC*, and *ThiE* cDNAs in wild type *Arabidopsis* seed, two gene expression cassettes were made (Figure 1) in which the *Thi4*, *ThiC* genes were under the control of Brassica napin (Ellerström et al., 1996) and Oleosin promoters (Keddie et al., 1994) promoters, respectively, as described in the "Materials and Methods" section. The *ThiE* gene was cloned directly into the binary vector under the control of soybean glycinin promoter (Iida et al., 1995). These three promoters were chosen because all three promoters are known to be expressed in both the endosperm and embryo tissues of seeds from various dicot and monocot species (Ellerström et al., 1996; Iida et al., 1995; Keddie et al., 1994).Transgenic *Arabidopsis* seeds (T_1) were produced using *Agrobacterium*-mediated transformation. The T_1 seeds were grown to have the T_2 seeds. To identify the homozygous lines, T_2 plants were grown and segregation analysis was performed on the T_3 generation. More than twenty homozygous lines were obtained and eight independent homozygous lines were used for this study. In parallel,

homozygous plants harboring the empty vector control were produced to be used for further analysis.

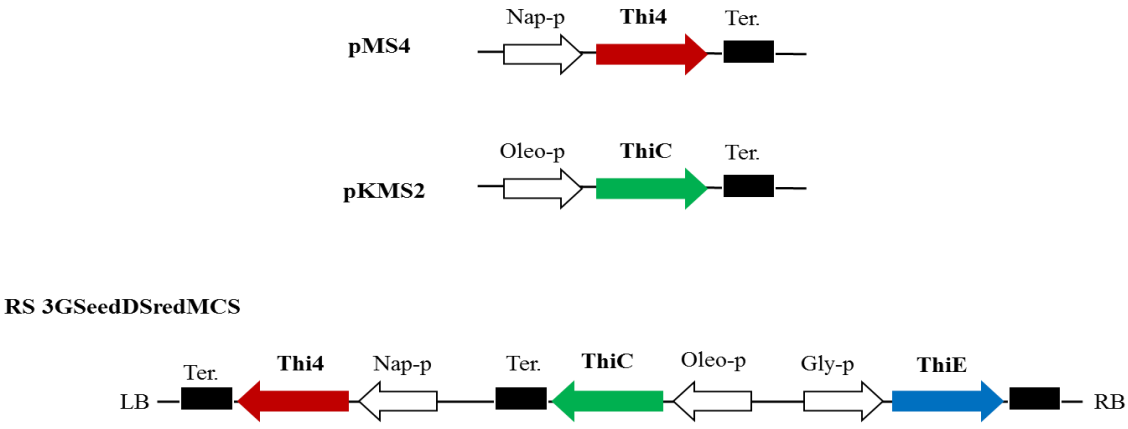


Figure 1: Simplified diagram depicting Thi4 and ThiC gene cassettes in pMS4 and pKMS2 vectors. RS 3GSeedDSredMCS binary vector was used to clone *ThiE* gene. *Thi4* and *ThiC* gene cassettes were also sub-cloned into RS 3GSeedDSredMCS binary vector. Nap-p, Oleo-p, and Gly-p are napin, oleosin, and glycinin promoters, respectively. Ter.: terminator, LB: left border, RB: right border.

Transgenic *Arabidopsis* plants showed greater thiamin content in the seeds

Thiamin content analysis of control and transgenic plants by HPLC method revealed that overexpression of *Thi4*, *ThiC*, and *ThiE* genes under the control of seed-specific promoters could significantly boost the thiamin level in transgenic *Arabidopsis* seeds. All transgenic lines have greater abundance of this metabolite in comparison to the wild type and empty vector control plants (Figure 2). The thiamin level in the seed of transgenic lines ranged from 81.2 (line 6-5) to 130.8 ng per mg seed weight (line 13-15) which is approximately 4.2 to 7-fold greater of those in control plants. Interestingly, thiamin in the form of cofactor was not detectable in either transgenic or control seeds.

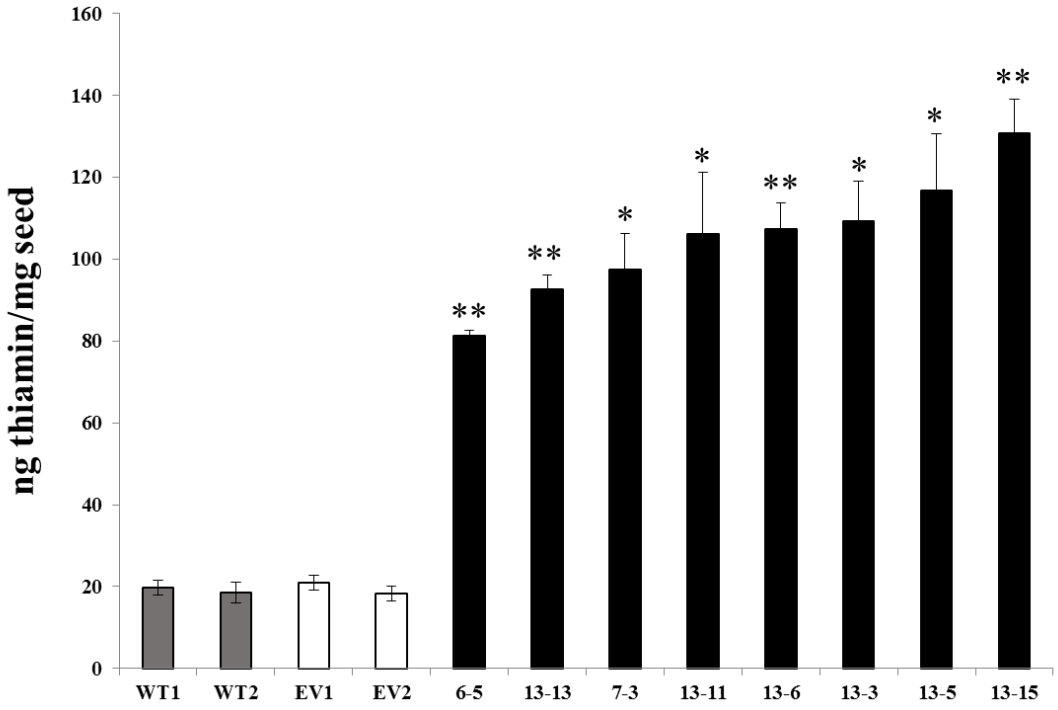


Figure 2: HPLC analysis shows increased seed thiamin content of homozygous transgenic *Arabidopsis* plants (3-gene overexpressors). Values represent mean±SE. Student *t* tests were carried out to compare the seed thiamin content in transgenic and WT plants. ** indicates significant difference at $P<0.001$ and * indicates significant difference at $P<0.05$. WT: wild type, EV: empty vector control.

Expression pattern of *Thi4*, *ThiC*, and *ThiE* genes in transgenic plants

We performed quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis to determine the expression pattern of thiamin biosynthesis genes in 9-day-old developing siliques. The results showed that all three transcripts were increased significantly in transgenic plants. The highest expression (>7-fold) was observed for the *ThiC* and *ThiE* genes while for the *Thi4* the transcript abundance level was approximately 4-fold more in comparison to the control plants (Figure 3).

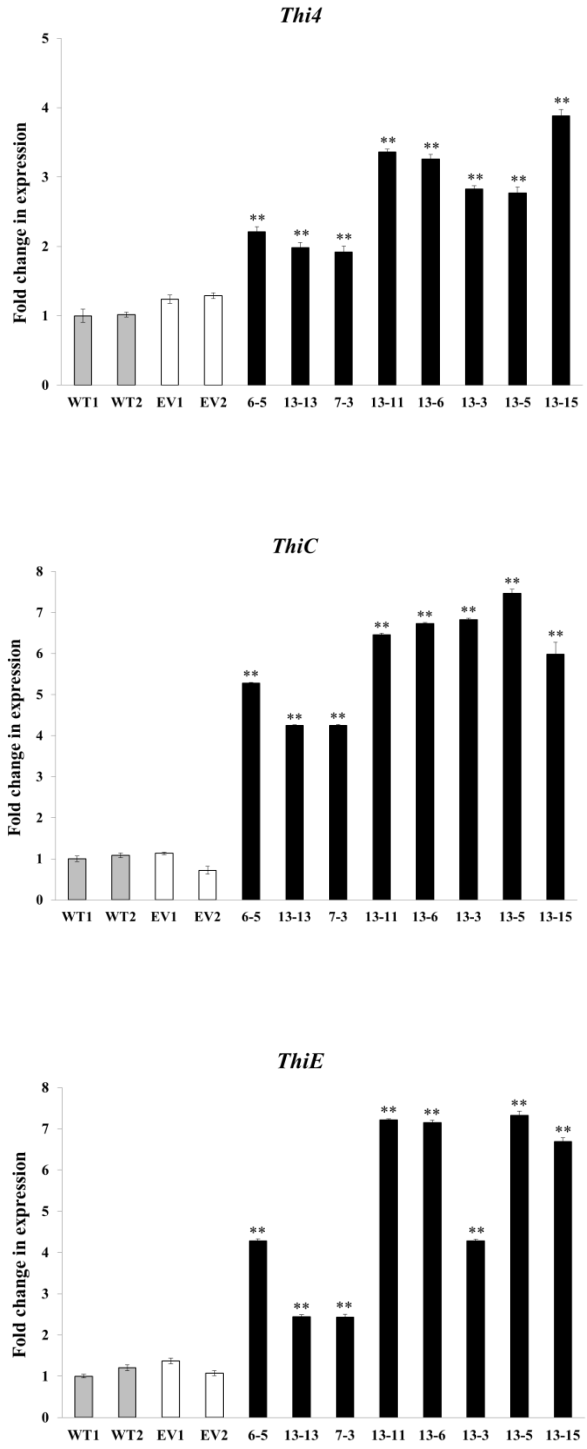


Figure 3: Transcript abundance of thiamin biosynthetic genes in 3-gene overexpressing *Arabidopsis* plants. Values are mean±SE. Student *t* tests were carried out to compare the transcript levels in transgenic and WT plants. ** indicates significant difference at $P<0.001$. WT: wild type, EV: empty vector control.

Transgenic *Arabidopsis* plants have enhanced oxidative stress tolerance

To test the hypothesis that transgenic plants are more tolerant to oxidative stress, root growth assay was carried out using the seeds of transgenic and control plants grown on agar medium supplemented with various concentration of paraquat. Root growth assay was chosen because it is fast, accurate, and easy to perform (Ciftci-Yilmaz et al., 2007; Verslues et al., 2006). Paraquat was also used because it can induce the production of superoxide radical which has detrimental effects on electron transfer mechanism in the chloroplast and mitochondrion (Bowler et al., 1983; Tunc-Ozdemir et al., 2009).

Our results showed that root growth rate (RGR) in transgenic seedlings was significantly increased relative to the control plants. The RGR for the transgenic plants grown on agar medium supplemented with 0.05 and 0.1 µM paraquat showed approximately 26% increase compared to control plants, while this value for the transgenic plants grown on agar medium supplemented with 0.25 µM paraquat was approximately 58% (Figure 4).

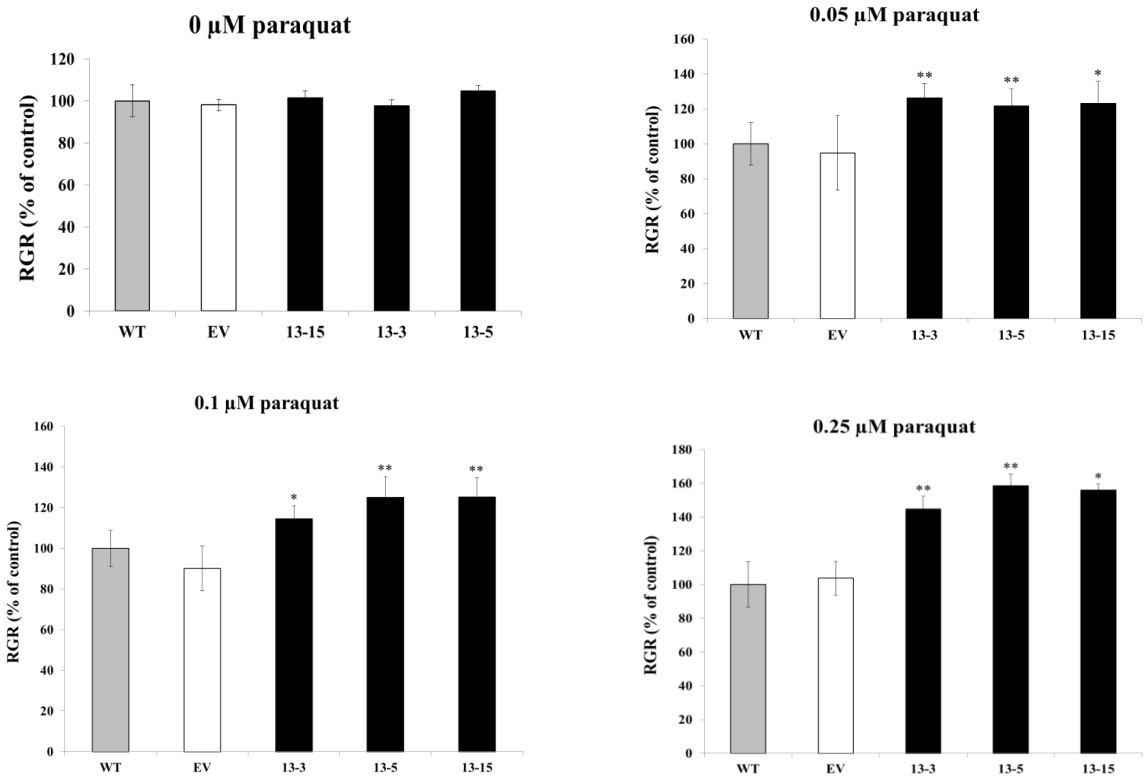


Figure 4: Effect of paraquat on root growth rate (RGR) of 9-day old *Arabidopsis* seedlings overexpressing 3 genes in the seed. 100% of the WT in 0.05, 0.1, and 0.25 µM paraquat equals 18.9, 14.2, and 6.6 mm, respectively. Values are mean±SE. Student *t* tests were carried out to compare the RGR in transgenic and WT plants. ** indicates significant difference at $P<0.001$ and * indicates significant difference at $P<0.05$. WT: wild type, EV: empty vector control.

Transgenic seeds accumulating high-thiamin levels show higher germination rate and seedling viability in salt and oxidative stress conditions

To investigate how thiamin over-producing lines react to the abiotic stress conditions, the seeds of transgenic plants were grown on agar plates supplemented with various concentrations of NaCl and paraquat, separately. The results showed that the germination of transgenic seeds was more tolerant to salt treatment than wild type and empty vector control seeds in the presence of 50 and 100 mM NaCl. At 50 mM NaCl,

more than 97% of the seeds in line 13-15 were germinated at day 2 compared with approximately 55 and 58% seed germination for the wild type and empty vector controls, rpectively (Figure 5). These values for the seeds treated with 100 mM NaCl were approximately 15, 16, and 46% for the wild type, empty vector control, and transgenic line 13-15, respectively (Figure 5). At day 3, the germination rate of transgenic seeds was significantly higher than that of the controls and it could reach to about 100% in transgenic lines grown on 50 mM NaCl and 88-96% in high-thiamin lines grown on 100 mM NaCl relative to the 88% in 50 mM NaCl and 65% in 100 mM NaCl in control plants (Figure 5).

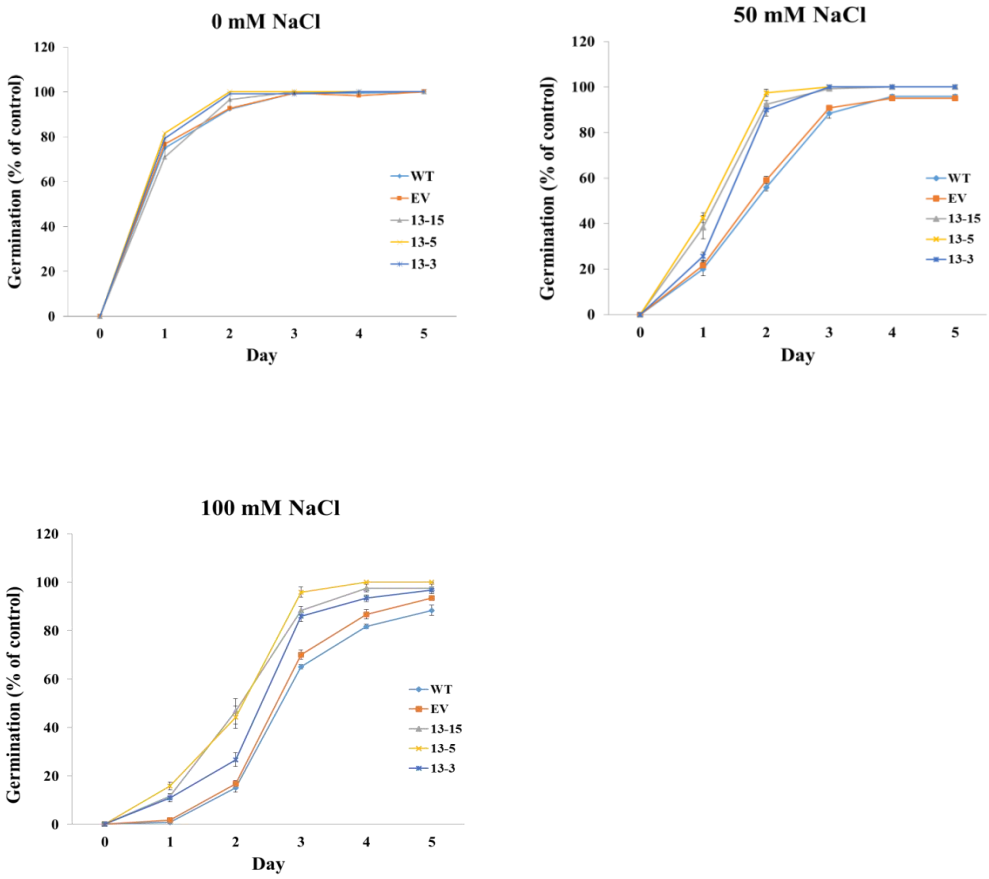


Figure 5: Effects of salt on the germination of 3-gene overexpressing *Arabidopsis* seeds. Radicle emergence was used as a marker for seed germination. Germination assays were

carried out using four replicates each consisting of 30-45 seeds. Student *t* tests were performed to compare the germination in transgenic and WT plants. Bars denote mean±SE. WT: wild type, EV: empty vector control.

Regarding paraquat treatment, germination of control seeds was inhibited significantly at day 1. The results showed that about 46% of control seeds were germinated compared to more than 80% germination in transgenic lines grown on agar medium supplemented with 0.1 µM paraquat (Figure 6). The germination was continued at day 2 with significant difference between the controls and transgenic seeds (approximately 78 % vs. 99% in line 13-5) then it reached to about 100% germination in all genotypes at day 3. The seeds grown on 0.25 µM paraquat were germinated approximately 53% for control ones and 81-94% for high-thiamin lines (Figure 6). Although seeds of all genotypes continued to germinate at day 2 and 3, the germination rate of high-thiamin lines were significantly higher than the controls (Figure 6). Nonetheless, in contrast to 0.1 µM paraquat treatment, both wild type and empty vector control seeds did not show 100% germination for 0.25 µM paraquat treatment at day 3, the germination was approximately 80 and 100% at day 3 for controls and transgenic lines, respectively (Figure 6).

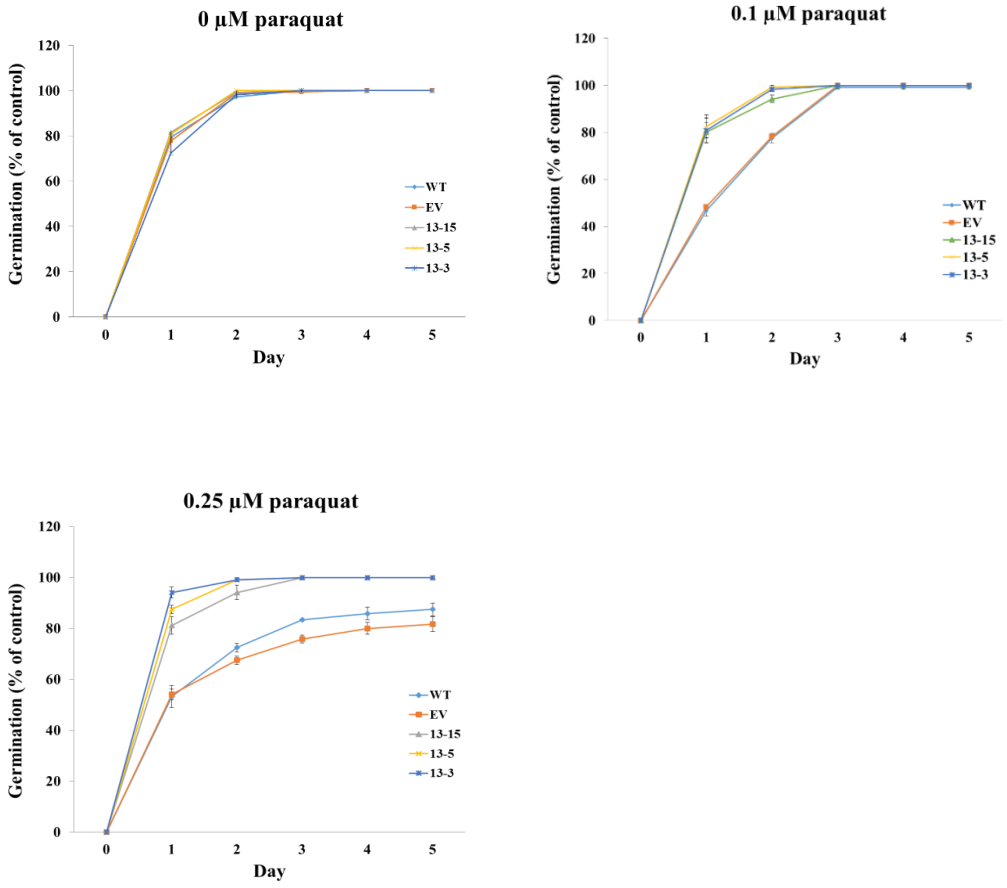


Figure 6: Effects of paraquat on the germination of 3-gene overexpressing *Arabidopsis* seeds. Radicle emergence was used as a marker for seed germination. For germination assays, four replicates each consisting of 30-45 seeds, were averaged and statistically treated using the Student's *t* test. Bars denote mean±SE. WT: wild type, EV: empty vector control.

To assess whether the seeds are able to generate seedling after day 5 on agar plates supplemented with 100 mM NaCl and 0.25 µM paraquat, the corresponding plates were kept in growth chamber for 9 days and the number of seedlings were counted. The results revealed that although about 88% of control seeds could germinate by day 5 in both 100 mM NaCl and 0.25 µM paraquat treatments, only 18% of them were able to

form seedlings under 100 mM NaCl condition (Figure 7). In contrast, approximately 42-47% of transgenic seeds could form seedling. These values for the seeds grown on agar plate containing 0.25 µM paraquat were 65 and 99% for control and high-thiamin lines, respectively (Figure 7).

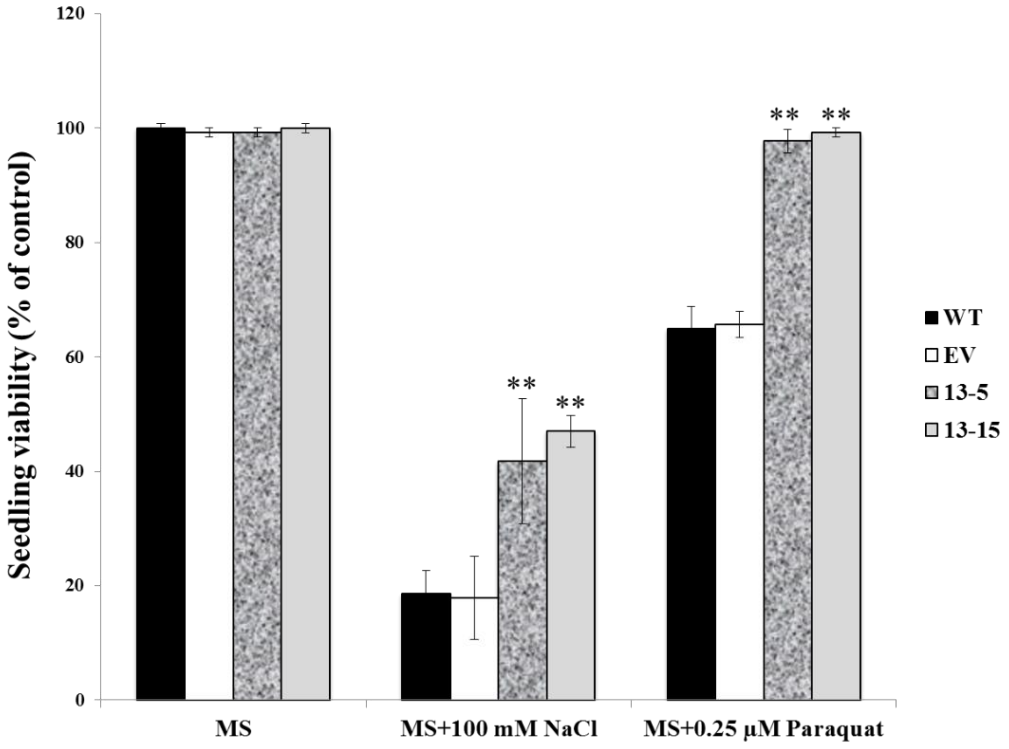


Figure 7: Effects of salt and paraquat on seedling formation of 3-gene overexpressing *Arabidopsis* seeds. Seeds were grown on MS plates containing 100 mM NaCl and 0.25 µM paraquat, separately. The assays were performed in three replicates each consisting of 45 seeds. 100% of the WT in MS+100 mM NaCl and MS+0.25 µM Paraquat were 8.3 and 29 seeds, respectively. Student *t* tests were performed to compare the seedling viability in transgenic and WT plants. Bars denote mean±SE. ** indicates significant difference at $P<0.001$. WT: wild type, EV: empty vector control.

High-thiamin transgenic lines showed altered seed phenotype in comparison to control seeds

To the best of our knowledge, this is the first report regarding the effect of thiamin on fatty acid biosynthesis pathway in plants. Seeds of *Arabidopsis* plants produce triacylglycerols (TAG) which are the esterified form of the fatty acids (Li et al, 2006). To evaluate the seed oil content in our study, we used direct methylation method of intact seeds described by Li et al., (2006) followed by gas chromatography (GC) analysis of fatty acid methyl esters (FAMSs). Surprisingly, we observed that transgenic seeds could accumulate more oil in the form of FAME than those of the control seeds. The level of oil content ranged from 42.1% in line 6-5 to 46.1% in line 13-15 per seed dry weight which are approximately 6-16.5% over the wild type seeds (Table 2). To assess whether over-production of thiamin could alter the metabolic pathways in high-thiamin lines, seed total protein and carbohydrate content were also measured. As shown in Table 2, the striking difference in the amount of total protein was observed between high-thiamin lines and controls. For instance, the total protein content of the high-thiamin line 13-15 was approximately 2.6-fold less than the wild type seeds which represent 21.9 % decrease per seed dry weight compare to its wild type counterpart. In contrast, in transgenic lines the total carbohydrate level was elevated. The level of total sugar in line 13-15 was almost 1.6-fold greater than the wild type which is 4.4% increase in total sugar per seed dry weight (Table 2). In addition, soluble sugars and seed weight analyses showed that thiamin over-producing lines significantly contained more soluble sugars (Figure 8) and had greater weight than the control plants (Figure 9).

Table 2: Seed compositions of transgenic plants overexpressing 3 genes (% per seed weight). Values are mean ± SE.

Genotypes	Lipid	*P*-value	Protein	*P*-value	Carbohydrate	*P*-value
WT1	39.6 ± 0.84		35.2 ± 2.01		7.2 ± 0.46	
WT2	39.9 ± 0.27		35.8 ± 2.18		8.2 ± 0.99	
EV1	40.1 ± 1.22	0.92	34.7 ± 0.66	0.34	8.6 ± 0.52	0.22
EV2	40.0 ± 0.56	0.88	36.3 ± 1.40	0.98	7.3 ± 0.52	0.65
6-5	**42.1 ± 0.61**	0.04	**14.8 ± 0.69**	0.0001	**9.8 ± 0.42**	0.001
13-13	**43.7 ± 0.56**	0.004	**8.7 ± 0.96**	0.0004	**9.9 ± 0.74**	0.03
7-3	**43.4 ± 0.03**	0.0001	**14.4 ± 0.53**	0.0008	**10.5 ± 0.43**	0.0008
13-11	**43.1 ± 1.00**	0.002	**13.7 ± 0.81**	0.0003	**10.7 ± 0.22**	0.0002
13-6	**43.9 ± 0.78**	0.01	**18.1 ± 0.50**	0.0004	**11.3 ± 0.76**	0.01
13-3	**44.8 ± 0.60**	0.002	**12.3 ± 1.95**	0.001	**11.7 ± 0.35**	0.0005
13-5	**43.9 ± 0.70**	0.01	**13.0 ± 0.64**	0.0006	**11.4 ± 0.11**	0.0002
13-15	**46.1 ± 0.71**	0.003	**13.3 ± 0.10**	0.0002	**11.6 ± 0.93**	0.02

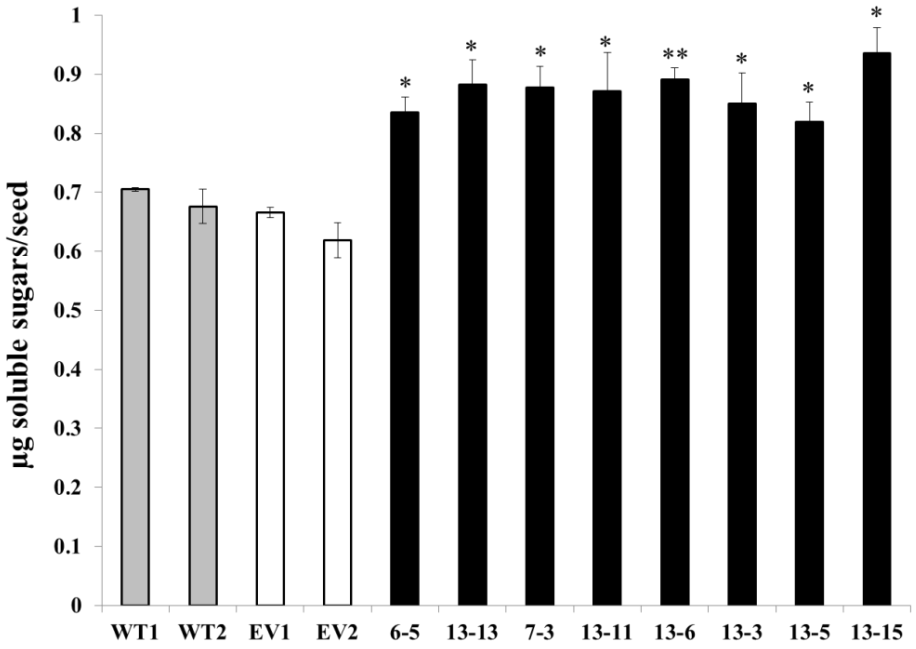


Figure 8: Seed soluble sugars content of transgenic *Arabidopsis* plants overexpressing 3 thiamin biosynthetic genes. Values are mean±SE. Student *t* tests were carried out to compare the soluble sugars content in transgenic and WT plants. ** indicates significant difference at $P<0.001$ and * indicates significant difference at $P<0.05$. WT: wild type, EV: empty vector control.

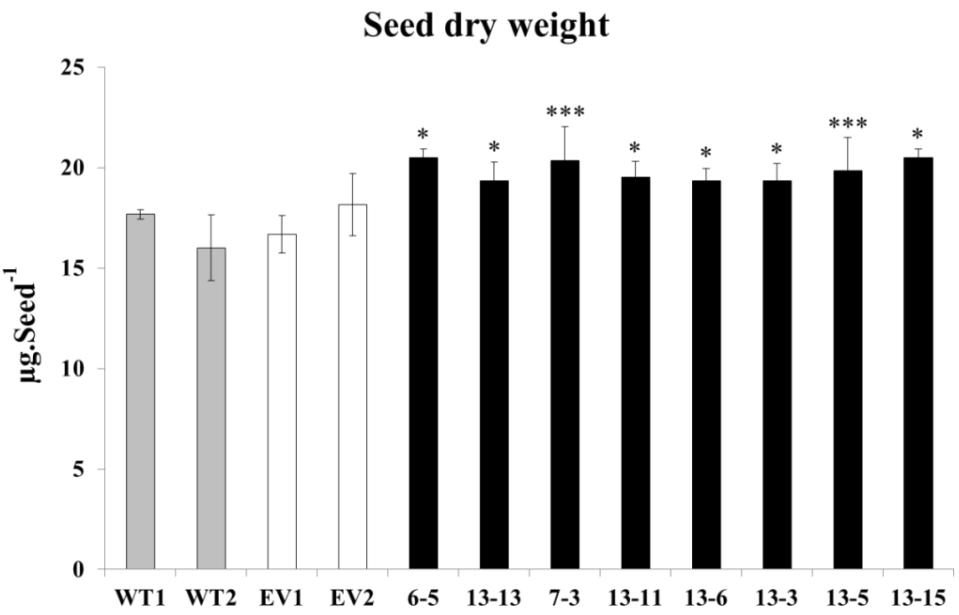
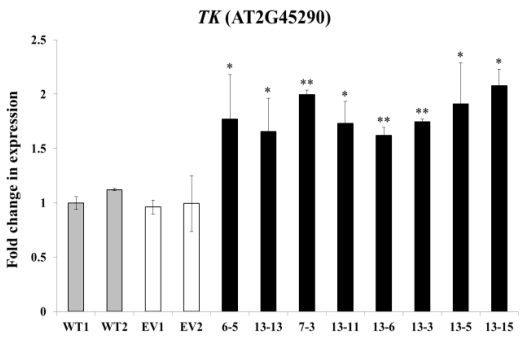


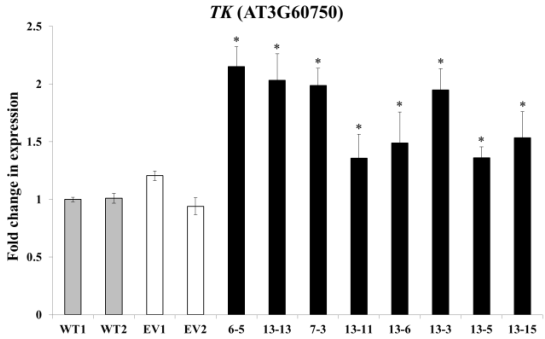
Figure 9: Seed weight analysis of thiamin-overproducing lines overexpressing 3 genes. The analysis was performed using 100 seeds for each line in triplicates. Values are mean±SE. Student *t* tests were carried out to compare the seed weight in transgenic and WT plants. * indicates significant difference at $P<0.05$ and *** indicates significant difference at $P<0.05<0.1$. WT: wild type, EV: empty vector control.

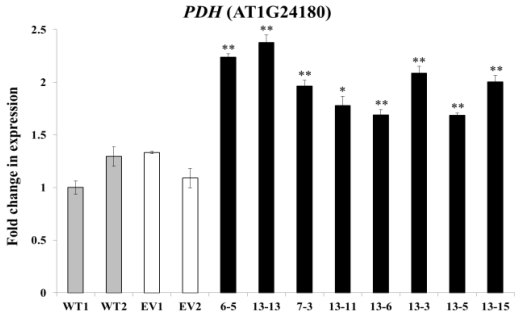
Thiamin cofactor-dependent enzymes showed enhanced gene expression pattern in 3-gene overexpressor plants

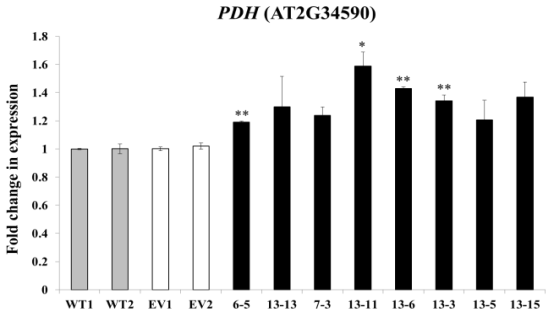
We next addressed whether over-production of thiamin in transgenic lines is able to up-regulate the gene expression of some essential thiamin cofactor-requiring enzymes, which are involved in plant core metabolism such as TCA cycle, pentose phosphate pathway, and isoprenoid pathway. To test this hypothesis, qRT-PCR analysis was carried out for transketolase (TK; AT2G45290 and AT3G60750), pyrovate dehydrogenase (PDH; AT1G24180 and AT2G34590), α-ketoglutarate dehydrogenase (α-KGDH; AT3G55410 and AT5G65750), and deoxy-xylulose phosphate synthase (DXPS; AT5G11380 and AT4G15560) using 9-day-old developing siliques. Our results indicated

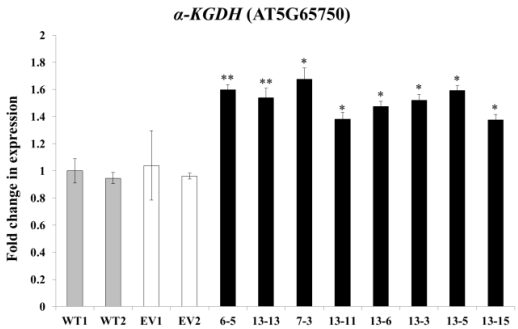
that transcript abundance level in high-thiamin lines for all aforementioned enzymes were approximately 1.5 to 2.5-folds greater than the corresponding controls (Figure 10).

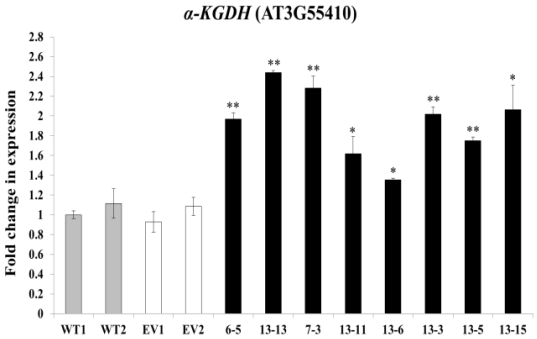












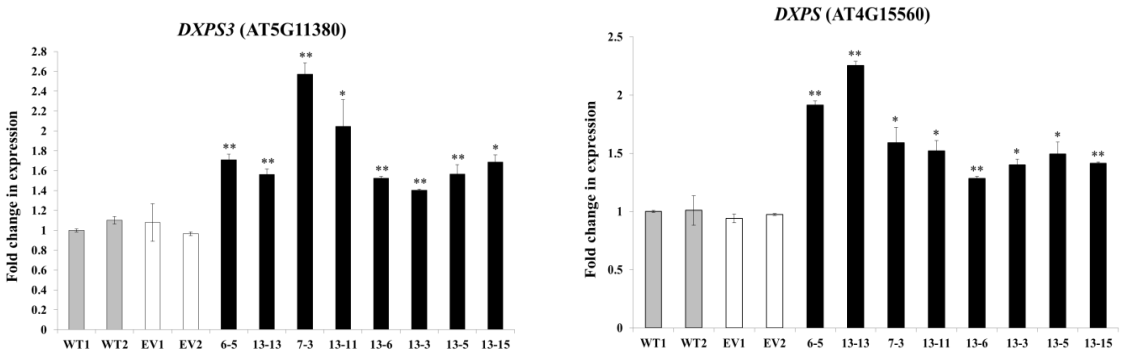


Figure 10: Transcript levels of thiamin cofactor-dependent enzymes. *TK*: transketolase, *PDH*: pyruvate dehydrogenase, *α-KGDH*: α-ketoglutarate dehydrogenase, *DXPS*: deoxy-xylulose phosphate synthase. Bars represent mean±SE. ** indicates significant difference at $P<0.001$ and * indicates significant difference at $P<0.05$. WT: wild type, EV: empty vector control.

Hypocotyl elongation assay

The requirement of the plants to thiamin for their growth and development raised the question whether the high-thiamin phenotypes in transgenic seeds can help their germination and growth on the medium which is lack of thiamin and sucrose. To address this question, we performed hypocotyl elongation assay according to Gantlet website (www.gantlet.org). The results obtained from this assay revealed that there was not a significant difference in hypocotyle length between transgenic and wild type seeds grown on MS plates supplemented with thiamin and sucrose (Figure 11A), but the hypocotyl elongation in high-thiamin lines were affected by their higher thiamin pools when grown on MS plates without thiamin and sucrose as they could form longer hypocotyls compared to the wild type controls (Figure 11B and C). As shown in Figure 11B, although the hypocotyl elongation was continued from day 2 through day 6 for both transgenic and control seeds, the transgenic lines could form longer hypocotyls than the

controls. In line 13-15 for instance, at day 2 hypocotyl length was approximately 62% longer than the controls. This value for day 4, 6, and 8 was 43, 43, and 30%, respectively. After day 6, hypocotyl elongation reached to plateau for all genotypes (Figure 11B and C).

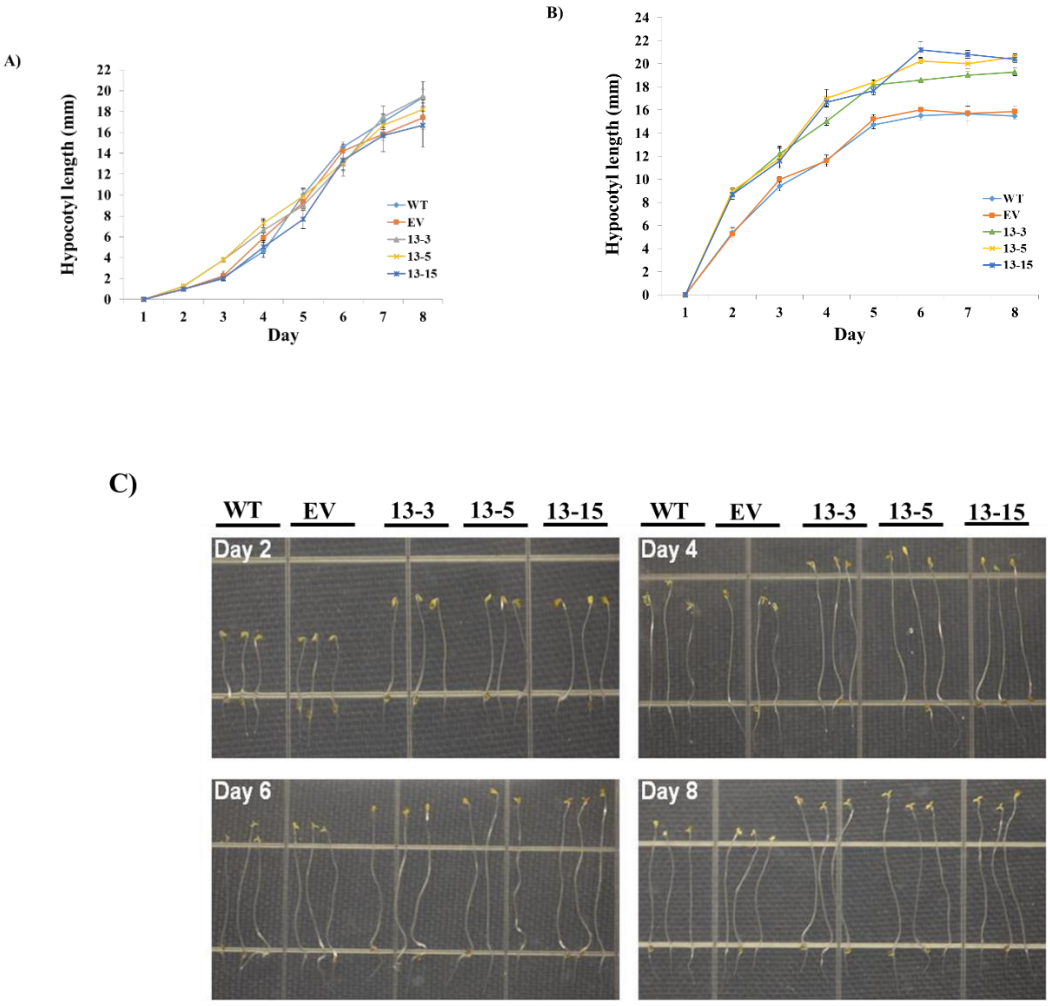


Figure 11: Hypocotyl elongation in *Arabidopsis* seeds. Seeds were plated on agar plates containing 0.5× MS salt with thiamin and sucrose (A) or without thiamin and sucrose (B and C) and were kept in 4 °C for 2 days in complete darkness. Plates were then placed in 22 °C under 100 µmol m^{-2} S^{-1} constant light for 24 h. After 24 h, the plates wrapped in aluminum foil individually and placed vertically in 22 °C. The hypocotyl length was measured for 8 days.

Overexpression of thiamin biosynthesis gene *TPK1* in the seed of wild type *Arabidopsis*

In order to confirm our previous results regarding seed phenotypes obtained from 3-gene overexpressor plants, we tried to overexpress the *TPK1* gene which is the last gene in thiamin biosynthesis pathway (Figure 1 in chapter1) and catalyzes the conversion of free thiamin to its cofactor form, TPP. To achieve this objective, a gene cassette was made in which *TPK1* cDNA was under the control of the glycinin promoter (Figure 12).

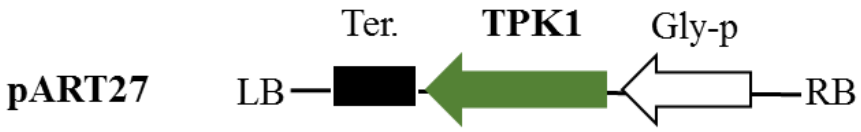


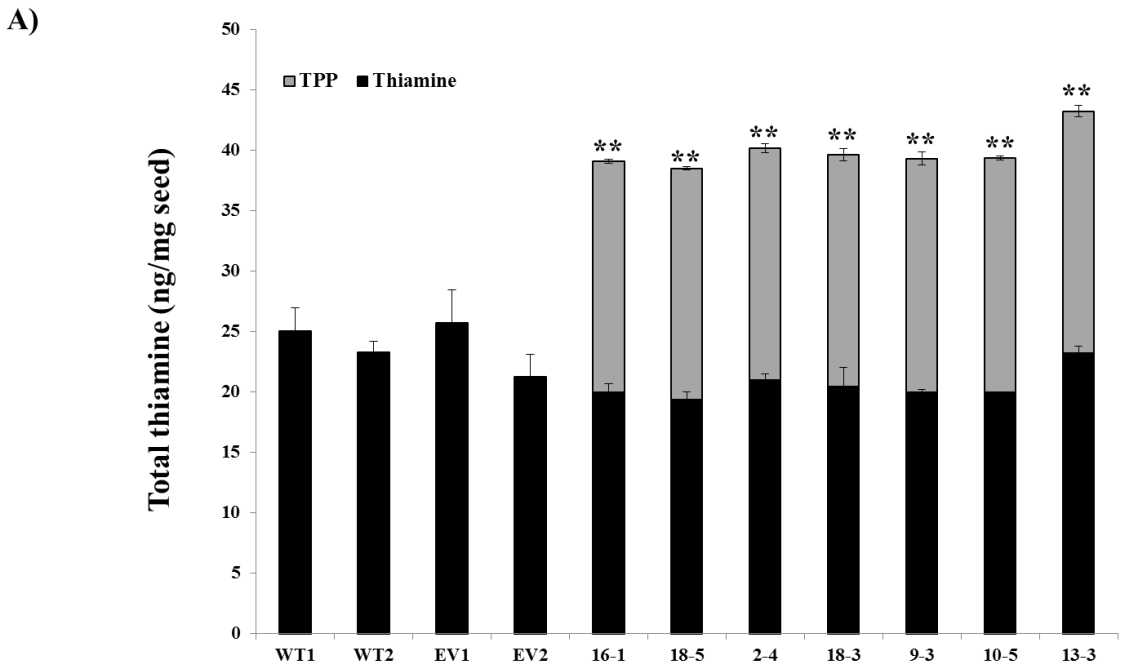
Figure 12: Simplified diagram representing *TPK1* gene cassette in pART27 binary vector under the control of glycinin promoter (Gly-p) and terminator (Ter.). LB: left border, RB: right border.

We chose *TPK1* gene because it has been shown that this gene has stronger expression in various parts of the *Arabidopsis* plant than *TPK2* (Ajjawi et al., 2007). The homozygous transgenic and empty vector lines were also produced as described for 3-genes overexpressor plants for further analysis.

***TPK1* transgenic plants accumulate both free thiamin and TPP in their seeds**

HPLC analysis was performed to determine the thiamin content of the *TPK1* overexpressors and control plants. The HPLC results showed that unlike *TPK1* transgenic lines, TPP was not detectable in control plants (Figure 13A). The TPP content in transgenic seeds was close to each other ranged from 2.2 in line 16-1 to 2.3 ng per mg

seed weight in line 13-3. Surprisingly, the total thiamin content in *TPK1* transgenic lines was about 2-folds greater than the controls (Figure 13A). Additionally, qRT-PCR analysis showed that transcript abundance in *TPK1* transgenic lines is approximately 4 to 7-folds greater than the controls (Figure 13B).



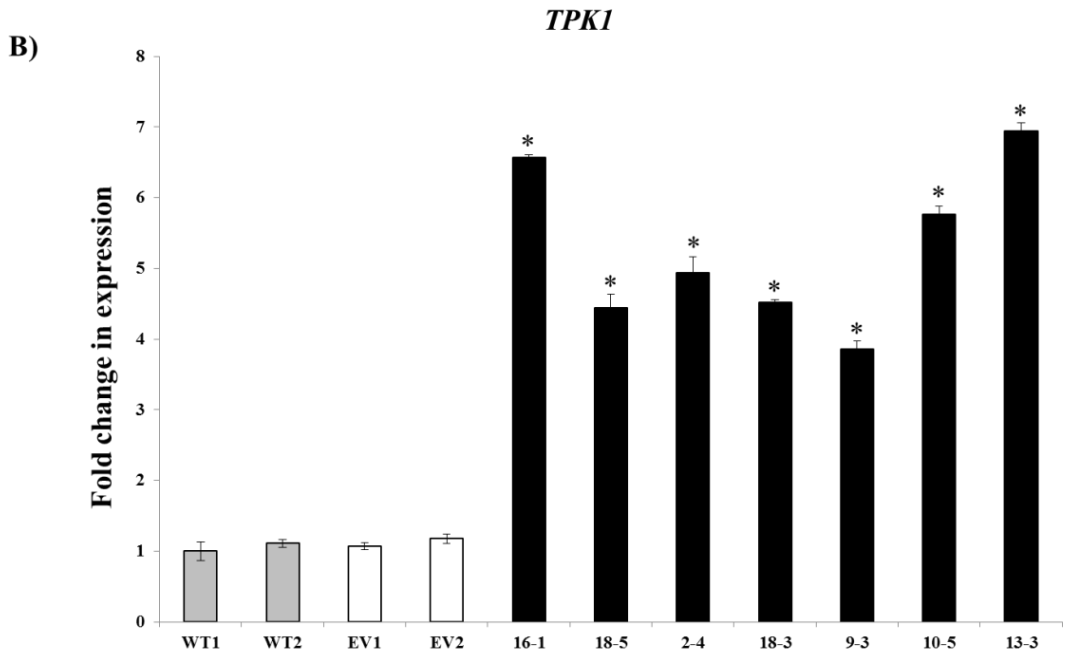


Figure 13: A) HPLC analysis of seed thiamin content of homozygous transgenic *Arabidopsis* plants (*TPK1* overexpressors). B) Transcript abundance in *TPK1* overexpressing *Arabidopsis* plants. Values are mean±SE. Student *t* tests were carried out to compare the seed thiamin content and transcript levels in transgenic and control plants. ** represent significant difference at $P<0.001$ and * indicates significant difference at $P<0.05$. WT: wild type, EV: empty vector control.

Overexpression of *TPK1* gene up-regulated the expression of the upstream genes involved in thiamin biosynthesis pathway

As described in previous section, the sum of free thiamin and TPP content in *TPK1* transgenic lines was about 2-folds greater than the controls. These results raised the question if the overexpression of *TPK1* gene is able to alter the expression of the other genes involved in thiamin biosynthesis pathway. To find the answer, we carried out qRT-PCR to determine the *Thi4*, *ThiC*, and *ThiE* genes expression level in transgenic and control plants. The results revealed that the expression pattern of *Thi4* and *ThiC* genes could be affected by the overexpression of *TPK1* gene and they were up-regulated

between 1.5 to 2.5-folds more than that of the controls. On the contrary, no significant changes were observed in *ThiE* gene expression (Figure 14).

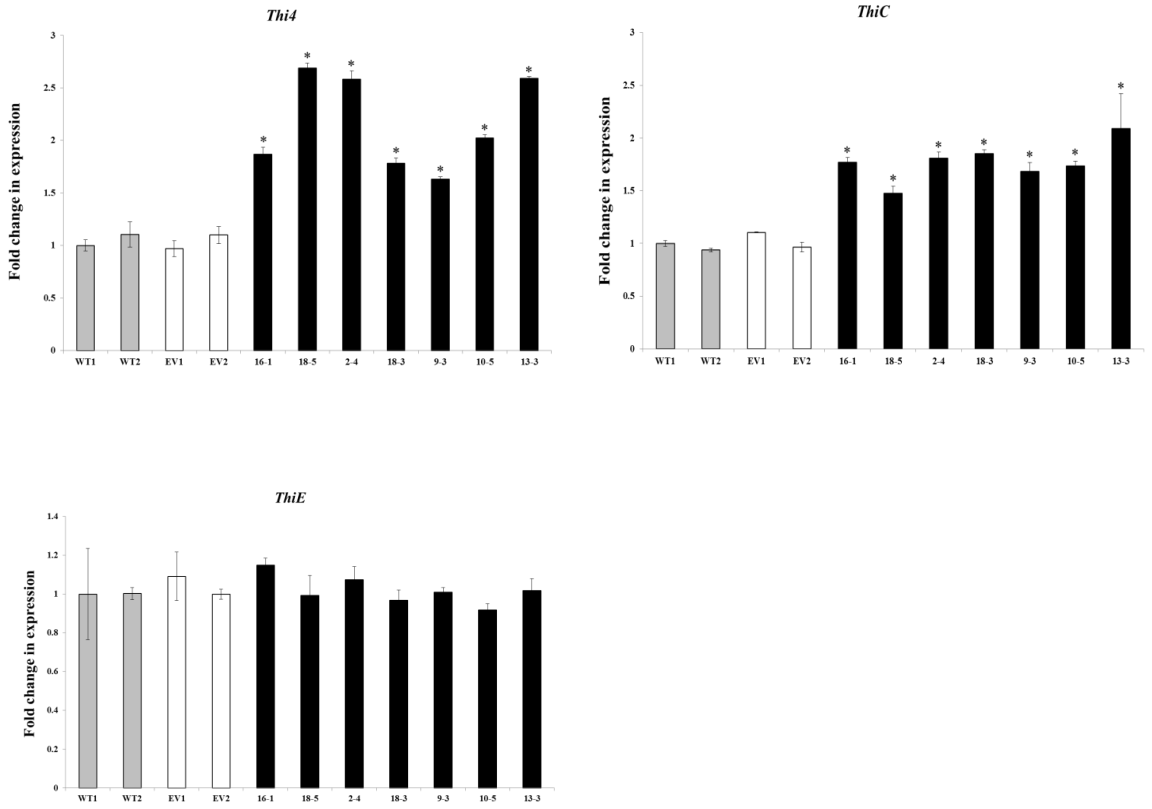


Figure 14: Transcript abundance of thiamin biosynthetic genes in *TPK1* overexpressing *Arabidopsis* plants. Values are mean±SE. Student *t* tests were carried out to compare the transcript levels in transgenic and control plants. * indicates significant difference at $P<0.05$. WT: wild type, EV: empty vector control.

TPK1 transgenic plants show altered seed oil, protein, and sugar content in comparision to control seeds

As we described previously regarding the seed phenotypes of 3-gene overexpressor plants, they showed greater amount of oil and carbohydrate but lesser

amount of protein in the seeds. To confirm these results and see whether thiamin or TPP are responsible for these results, we performed oil, carbohydrate, and protein analysis of *TPK1* transgenic seeds. Interestingly, oil and carbohydrate level were boosted in *TPK1* transgenic seeds. The GC results showed that although it was in lesser amount than 3-gene overexpressors, the oil level of *TPK1* transgenic seeds were elevated approximately between 2.1% in line 18-5 to 9% in line 13-3 more than control seeds. The carbohydrate level was also increased between 20.2 to 58% more than the controls in lines 18-3 and 2-4, respectively. Consistent with the results for protein level in 3-gene overexpressor plants, *TPK1* transgenic plants stored 1.3-folds (line 18-5) to 1.6-folds (line 2-4) lesser amount of protein than the controls in their seeds (Table 3).

Table 3: Seed compositions of transgenic plants overexpressing *TPK1* gene (% per seed weight). Values are mean ± SE.

Genotypes	Lipid	*P*-value	Protein	*P*-value	Carbohydrate	*P*-value
WT1	41.4 ± 0.28		41.2 ± 3.61		12.5 ± 1.45	
WT2	39.7 ± 1.59		42.5 ± 4.67		11.3 ± 2.95	
EV1	41.0 ± 0.31	0.29	46.0 ± 7.09	0.32	12.3 ± 0.40	0.41
EV2	40.9 ± 0.40	0.35	44.9 ± 5.22	0.33	12.8 ± 0.27	0.27
16-1	**42.8 ± 0.75**	0.05	**28.4 ± 4.84**	0.004	**16.8 ± 0.10**	0.01
18-5	**41.4 ± 0.65**	0.19	**33.4 ± 2.13**	0.01	**17.9 ± 0.10**	0.002
2-4	**43.1 ± 0.58**	0.03	**25.8 ± 3.49**	0.0001	**18.8 ± 0.20**	0.04
18-3	**42.2 ± 0.30**	0.05	**25.1 ± 1.67**	0.0009	**14.3 ± 0.21**	0.05
9-3	**43.4 ± 0.92**	0.03	**27.8 ± 2.40**	0.002	**15.9 ± 0.52**	0.007
10-5	**42.8 ± 0.14**	0.01	**31.1 ± 2.96**	0.01	**16.4 ± 0.61**	0.004
13-3	**44.4 ± 1.10**	0.02	**31.3 ± 3.10**	0.02	**16.1 ± 0.55**	0.006

DISCUSION

Thiamin in the form of its cofactor, TPP, plays a key role in the core metabolism of living organisms through the participation in metabolic pathways such as: Calvin

cycle, acetyl-CoA formation, TCA cycle, pentose phosphate pathway (PPP), and the branched-chain amino acid biosynthesis pathway (Friedrich, 1987). Although the thiamin biosynthetic pathway and its regulation mechanism have been recently well-characterized (Begley et al., 1999; Nosaks, 2006; Wachter et al., 2007; Ajjawi et al., 2007; Yazdani et al., 2013; Zallot et al., 2014), there have been no reports regarding the generation of the transgenic plants with the elevated level of thiamin in the seed through genetic engineering. In this study, we showed that using strong seed-specific promoters, the free thiamin level in the plant seeds could be boosted significantly by the overexpression of the *Thi4* and *ThiC* genes which are responsible to produce thiazole and pyrimidine moieties of the thiamin and also *ThiE* gene which can combine these two products into the TMP (Figure 2). These results suggest that overexpressing these three genes is sufficient to elevate the thiamin level in plant seeds. On the other hand, overexpression of the *TPK1* gene led to accumulation of TPP in transgenic seeds (Figure 13A). Indeed, *TPK1* gene overexpression could up-regulate the expression of *Thi4* and *ThiC* genes (Figure 14) to increase free thiamin level required for producing more TPP in *TPK1* gene overexpressors. These findings propose that *Thi4* and *ThiC* genes are limiting the thiamin biosynthesis pathway (Pourcel et al., 2013). Like the wild type plants, thiamin over-producing plants also showed the normal growth. This suggests that thiamin in higher amounts does not have adverse effects on plant growth and development. This result is consistent with the results obtained from the feeding studies reported by Pourcel et al, (2013).

Our data shows that both control plants and transgenic lines (3-gene overexpressors) prefer to store the thiamin in the form of free thiamin (Figure 2). These

results are in agreement with the previously reported data indicating the absence of phosphorylated forms of thiamin in mature seeds of some monocots and dicots such as rice, corn, soybean, and pea (Yusa, 1961; Molin et al, 1980; Golda et al, 2004).

Oxidative stress is the major damage for plants under environmental stresses because of the imbalance between reactive oxygen species (ROS) production and the activity of antioxidant systems (Scandalios 1993). When plants are exposed to abiotic stress conditions, they switch on some genes that increase the level of certain metabolites which contest these stresses. Thiamin has been proven to be a metabolite that enhances tolerance to abiotic stresses such as heat, cold, salinity, and paraquat (PQ) in plants (Tunc-Ozdemir et al. 2009). Although several studies in animal cells (Lukienko et al. 2000; Ba, 2008) and yeast (Wolak et al. 2014) have given credit to thiamin as a potent antioxidant, there is not enough evidence to support this idea in plants. Most of the data regarding the role of thiamin in biotic and abiotic stress conditions is relied on feeding studies by using exogenous thiamin in the plants medium (Sayed and Gadallah, 2002; Conrath et al., 2002; Ahn et al., 2005; Tunc-Ozdemir et al., 2009). This suggests an indirect role for thiamin in scavenging ROS in plants by provision of NADH and NADPH to combat oxidative stress (Tunc-Ozdemir et al., 2009; Asensi-Fabado and Munne-Bosch, 2010). In our study the root growth assay showed that genetically engineered *Arabiopsis* seeds (3-gene overexpressors) had significantly increased RGR than that of the control plants under the PQ treatment (Figure 4). PQ is a well-known wildly used herbicide to control weeds in agriculture (Babbs et al., 1989). As a powerful ROS producer PQ can cause loss of photosynthetic activity via inhibition of ferredoxin reduction, generating the NADPH, regeneration of antioxidant enzymes (Lascano et al.,

2012.), and loss of cell membrane integrity as well (Kunert and Dodge 1989). The increased RGR in transgenic plants might be due to the elevation of free thiamin pool which can be subsequently converted to TPP cofactor (Tunc-Ozdemir et al., 2009). It has been shown that plants under stress conditions have a high demand for TPP, which activates enzymes such as TK, PDH, and α-KGDH that are involved in plants core metabolism including PPP and TCA cycle to produce more NADH and NADPH required for scavenging the PQ-generated ROS (Tunc-Ozdemir et al., 2009; Rapala-Kozik, 2012). In our study it is noteworthy that in thiamin over-producing lines, the *TK*, *PDH*, and *α-KGDH* gene expression were also up-regulated (Figure 10). It is consistent with the fact that TPP-dependent enzymes have a high demand for this cofactor to produce reducing molecules in stress conditions (Rapala-Kozik et al., 2008). It has been shown that when plants are exposed to abiotic stress conditions these reducing molecules are required for the activity of several NAD(P)H-dependent enzymes such as glutathione reductase and mono-dehydro-ascorbate reductases as well as the recycling of the oxidized form of vitamin E, which play a crucial role to combat oxidative stresses (Arora et al. 2002; Miller et al. 2010). Furthermore, Tunc-Ozdemir et al. (2009) reported that the application of thiamin alleviated PQ sensitivity in both wild type and *ascorbate peroxidase1* (*apx1*) mutant which is highly sensitive to oxidative damage. These results suggest that thiamin can complement the function of APX1 enzyme and the thiamin may have a direct role in reducing oxidative stresses (Tunc-Ozdemir et al., 2009).

Interestingly, our transgenic seeds showed the greater total carbohydrates and soluble sugar content (Table 1and Figure 8, respectively) compare to the control seeds possibly due to the higher activity of TK in Calvin cycle. Hare et al. (1998) reported that

accumulation of sugar is a common response of organisms to abiotic stress and can function as osmoprotectant and biomolecule stabilizer. Aditionally, Kerepesi and Galiba (2000) indicated that wheat plants accumulate more water-soluble carbohydrates during osmotic and salt stresses. Although the role of some soluble sugars including glucose and fructose in stress tolerance is controversial (Kerepesi and Galiba, 2000), it has been shown that other soluble carbohydrates such as sucrose and fructans have potential role in adaptation to drought and salt stresses (McKersie and Leshem, 1994). Sucrose can prevent structural changes in soluble proteins in abiotic stress conditions (Kerepesi and Galiba, 2000). The exact mechanism of fructans in conferring stress tolerance is not clear. It may protect the cell membrane or other cellular component from the detrimental impacts of abiotic stresses (Pilon-Smits et al., 1995).

It has also been shown that ROS scavenging enzymes in animal cells and yeast have high demand for NADPH generated by the PPP under oxidative stress conditions (Palmer, 1999; Larochelle et al., 2006). These results indicate that TK, which is a TPP-dependent enzyme in PPP, has an important role in producing NADPH required by antioxidant enzymes. On the other hand, our results showed up-regulation of the 1-deoxy-D-xylulose-5-phosphate syntase (DXPS) in thiamin over-producing lines (Figure 10). DXPS can also serve as a defense agent in high-thiamin lines against ROS in stress conditions. It is well documented that carotenoids, which are a vital class of antioxidants are produced in chloroplasts by a TPP-dependent enzyme DXPS, play an important role in detoxifying the ROS within thylakoid membranes (Asensi-Fabado and Munne-Bosch, 2010).

Our results showed that seed germination was delayed by both salt and PQ stresses (Figure 5 and 6, respectively). Transgenic lines also showed increased seedling viability under salt and PQ treatments (Figure 7). Inhibition of seed germination could be attributed to osmotic stress and ion toxicity caused by NaCl (Huang and Redmann, 1995). Likewise, Agarwal and Pandey, (2004) reported that salinity has detrimental effects on seed germination and seedling root and shoot length. Water stress caused by salt stress has been reported to reduce stromal volume in chloroplasts as well as to generate ROS, which both play a critical role in photosynthesis inhibition (Price and Hendry, 1991). Oxidative damage generated by ROS can be alleviated by antioxidant systems that can elevate plant tolerance to salt stress. Yoshimura et al. (2004) reported that overexpression of *Chlamydomonas* glutathione peroxidase in the cytosol and chloroplast of the tobacco plants could boost oxidative stress tolerance brought by salt and PQ stresses. This would suggest that excess amounts of thiamin in our transgenic lines could help plants to alleviate the oxidative stress caused by salt and PQ.

We used *Arabidopsis* in this study because it is a model plant for biology research (Meyerowitz and Somerville, 1994; Li et al., 2006) and it is also a close relative of rapeseed (*Brassica napus*) which is an important oilseed crop (Li et al., 2006). We showed that free thiamin and TPP-over-producing lines had altered carbon partitioning in the seeds. Seed phenotype analysis revealed that these seeds accumulate more oil and carbohydrate, but less protein than the control plants (Table 2 and 3). To date, the attempts in order to increase seed oil content in plants were mostly focused on changing the expression level of the genes involved in oil biosynthetic pathway (Zou et al., 1997; Jako et al., 2001; Vigeolas et al., 2007; Kim et al., 2014; Van Erp et al., 2014). It has

been shown that the fatty acid biosynthesis pathway is highly dependent upon NADH and NADPH supplies (Geer et al., 1979; Slabas and Fawcett, 1992). However, the sources of reducing molecules have not been identified yet (Schwender et al., 2003). Rawsthorne (2002) reported that NADPH generated in the chloroplasts might be a potential source for oil biosynthesis. Additionally, Kang and Rawsthorne (1996) reported that NADPH generated by oxidative PPP possibly provides reducing power for fatty acid biosynthesis. On the other hand, using a quantitative metabolic flux model, Schwender et al. (2003) suggested that glycolysis and the oxidative PPP can provide the reductant required for fatty acid biosynthesis. In our study the significant increase observed in *Arabidopsis* seed oil content in thiamin and TPP over-producing lines might be due to the presence of the high amounts of reductants produced by thiamin cofactor-dependent enzymes such as TK, PDH, and α-KGDH (Jordan, 2003; Nosaka, 2006; Zallot et al., 2014), as their transcript level was also significantly increased in high-thiamin lines (Figure 10). Additionally, the oil biosynthetic pathway is dependent on acetyl-CoA supply which is required for growing the acyl chain in the pathway (Slabas and Fawcett, 1992). The significant increase of PDH gene expression indicated in this study might be a possible way to elevate the acetyl-CoA level needed for producing more oil in transgenic lines.

It has been shown that overexpression of *DIACYLGLYCEROL ACYL TRANSFERASE1* (*DGAT1*) and *WRINKLED1* (*WRI1*) along with the RNAi suppression of the lipase *SUGAR-DEPENDENT1* (*SDP1*) could increase the oil content of the *Arabidopsis* seeds; however, a significant decrease in seed protein and sugar content was observed (Van Erp et al., 2014). Additionally, using mutational analysis Focks and Benning (1998) reported that *WRI1* has an indirect effect on fatty acid biosynthesis. They

suggested that *wri1* mutants have aborted the conversion of glucose into fatty acid precursors and sucrose and hexoses concentrations were increased in developing *wri1* mutant seeds. These results indicate that plants use protein and sugar as the carbon source for oil biosynthesis (Focks and Benning, 1998) and also show the competition between the metabolic pathways for the substrates (Schwender and Hay, 2012; Van Erp et al., 2014). Increased seed oil content in both thiamin and TPP over-producing lines in this study might also be a consequence of increased carbohydrate levels generated by the up-regulation of *TK* gene expression due to high level of thiamin cofactor. In Calvin cycle, it has been demonstrated that TK has a significant role to carbon flux (Stitt and Schulze 1994; Haake et al. 1999; Raines et al. 2000). Using antisense technology in tobacco plants, Henkes et al. (2001) suggested a pivotal role for TK in plant metabolism and showed that its activity is a limiting factor for photosynthesis and sucrose production. They showed that 20 to 50% suppression in *TK* expression in tobacco plants could remarkably inhibit photosynthesis rate, sugar content, and growth rate as well. Elevated sugar content in our transformants might be the consequence of the increased regeneration level of ribulose-1, 5-bisphosphate which is a substrate for the activity of Rubisco enzyme in Calvin cycle by TK enzyme. This is consistent with the fact that in tobacco plants antisense TK transformants displayed reduced photosynthesis due to the inhibition of ribulose-1, 5-bisphosphate regeneration (Henkes et al. 2001). These results suggest that plastid TK is a key enzyme in the core metabolism of plants and several metabolic pathways which rely on TK are highly sensitive to slight changes in plastid *TK* expression (Henkes et al. 2001).

Because of the participation of N and S atoms in thiamin structure (Jurgenson et al. 2009), the substantial decrease in seed protein level in our high-thiamin and TPP transformants can be partly attributed to the fact that these plants use protein as N and S sources to produce pyrimidine and thiazole rings of the thiamin. In fact, these results suggest that various metabolic pathways compete together for the substrates (Schwender and Hay, 2012; Van Erp et al., 2014). Maisonneuve et al. (2010) also reported that increased oil content in transgenic plants overexpressing rapeseed lysophosphatidic acid acyltransferase (LPAAT) could be at the expense of seed storage reserve content such as storage proteins.

Oil, storage proteins, and carbohydrates in the form of starch are the major constituent in the seeds of flowering plants such as *Arabidopsis* (Baud et al. 2002; Wang et al. 2007). Seed reserves provide initial carbon and energy source for growing seedlings (Bradbeer, 1988). During morphogenesis, oil in the form of triacylglycerol (TAG) is accumulated in *Arabidopsis* embryos to be used for seed germination and seedlings establishment (Mansfield and Briarty, 1992; Eastmond et al., 2000). Analysis of hypocotyl growth in control and thiamin over-producing lines revealed that transgenic plants could produce a longer hypocotyl in comparison to the control plants (Figure 11). These results can be attributed to the fact that higher levels of thiamin provide more TPP essential for carbohydrate, oil, and protein metabolism which is required for hypocotyl growth when photosynthesis is blocked by complete darkness in transgenic lines.

Seed mass was another trait that was analyzed in this study. Indeed, seed weight of the thiamin over-producing genotypes was significantly greater than the control lines (Figure 9). Using seed-specific promoters, Van Erp et al. (2014) reported that

overexpression of the genes pertaining to the oil biosynthetic pathway such as *DGAT1* and *WRI1* could significantly increase the seed weight. Additionally, Jako et al. (2001) for the first time reported that increasing in seed weight ranged from 40% to 100% was achieved by the overexpression of *DGAT* gene in *Arabidopsis*. Moreover, overexpression of rapeseed lysophosphatidic acid acyltransferase (LPAAT), an important enzyme to synthesize phosphatidic acid required for storage lipids biosynthesis in developing seeds, led to a boost in oil content and seed mass in *Arabidopsis* without any significant difference in fatty acid composition of TAG between transgenic and control plants (Maisonneuve et al., 2010). Likewise, expression of yeast LPAAT isozymes in *Arabidopsis* and rapeseed remarkably enhanced seed oil and seed weight (Zou et al., 1997). On the other hand, Zhang et al. (2005) reported that silencing of *DGAT1* gene in tobacco reduced mature seed oil content accompanied by reduction in the average seed weight. On the contrary, an increase in protein and soluble sugars content of the mature transgenic seeds was observed. These results suggest that plants use protein and carbohydrate to produce more oil.

In conclusion, our data demonstrates a critical role for thiamin and TPP cofactor in the alleviation of oxidative stress in plant cells. It will be interesting to determine whether thiamin functions as a direct antioxidant against abiotic stresses. Additionally, our results show that thiamin and its cofactor are able to alter the carbon partitioning in plant cells which can have biotechnological application to produce crop plants with favorite traits.

REFERENCES:

Agarwal, S., Pandey, V. (2004). Antioxidant enzyme responses to NaCl stress in *Cassia angustifolia*. Biol. Plantarum. 48: 555-560.

Ahn, I.P., Kim, S., Lee, Y.H. (2005). Vitamin B1 functions as an activator of plant disease resistance. Plant Physiol. 138: 1505-1515.

Ajjawi, I., Rodriguez Milla, M.A., Cushman, J., Shintani, D.K. (2007). Thiamine pyrophosphokinase is required for thiamine cofactor activation in *Arabidopsis*. Plant Mol. Biol. 65: 151-162.

Ajjawi, I., Tsegaye, Y., Shintani, D. (2007a). Determination of the genetic, molecular, and biochemical basis of the *Arabidopsis thaliana* thiamine auxotroph *th*1. Arch. Biochem. Biophys. 459: 107-114.

Arora, A., Sairam, R.K., Srivastava, G.C. (2002). Oxidative stress and antioxidative system in plants. Curr. Sci. 82: 1227-1238.

Asada, K., Kiso, K. (1973). Initiation of aerobic oxidation of sulfite by illuminated spinach chloroplasts. Eur. J. Biochem. 33: 253-257

Asensi-Fabado, M.A, Munné-Bosch, S. (2010). Vitamins in plants: occurrence, biosynthesis and antioxidant function. Trends plant Sci. 15: 582-592.

Ashraf, M. (1994). Organic substances responsible for salt tolerance in *Eruca sativa*. Biol. Plant. 36: 255-259.

Ba, A. (2008). Metabolic and structural role of thiamine in nervous tissues. Cell Mol. Neurobiol. 28: 923-931.

Babbs, C.F., Pham, J.A., Coolbaugh, R.C. (1989). Lethal Hydroxyl Radical Production in Paraquat-Treated Plants. Plant Physiol. 90: 1267-1270.

Baud, S., Boutin, J.P., Miquel, M., Lepiniec, L., Rochat, C. (2002). An integrated overview of seed development in *Arabidopsis thaliana* ecotype WS. Plant Physiol. Biochem. 40: 151-160.

Begley, T.P., Downs, D.M, Ealick, S.E., McLafferty, F.W., Van Loon, A.P.G.M., Taylor, S., Campobasso, N., Chiu, H.J., Kinsland, C., Reddick, J.J., Xi, J. (1999). Thiamine biosynthesis in prokaryotes. Arch. Microbiol. 171: 293-300.

Bowler, C., Slooten, L., Vandenbranden, S., De Rycke, R., Botterman, J., Sybesma, C., Van Montagu, M., Inze, D. (1983). Manganese superoxide dismutase can reduce cellular damage mediated by oxygen radicals in transgenic plants. EMBO J. 10: 1723-1732.

Bowler, C., van Montagu, M., Inze, D. (1992). Superoxide dismutase and stress tolerance. Annu. Rev. Plant Physiol. Plant Mol. Biol. 43: 83-116

Boyer, J.S. (1982). Plant productivity and environment. Science, 218: 443-448.

Bradbeer, J.W. (1988). Seed Dormancy and Germination. Blackie, Glasgow.

Bray, E.A., Bailey-Serres, J., Weretilnyk, E. (2000). Responses to abiotic stresses. In: Gruissem W, Buchannan B, Jones R (eds) Biochemistry and molecular biology of plants. American Society of Plant Physiologists, Rockville, MD, pp 1158-1249.

Ciftci-Yilmaz, S., Morsy, M. R., Song, L., Coutu, A., Krizek, B. A., Lewis, M.W., Warren, D., Cushman, J., Connolly, E.L., and Mittler, R. (2007). The EAR-motif of the Cys2/His2-type zinc finger protein Zat7 plays a key role in the defense response of Arabidopsis to salinity stress. J. Biol. Chem. 282*:* 9260-9268.

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

Conrath, U., Pieterse, C.M., Mauch-Mani, B. (2002). Priming in plant-pathogen interactions. Trends Plant Sci. 7: 210-216.

Cui, S., Huang, F., Wang, J., Ma, X., Cheng, Y., Liu, J. (2005). A proteomic analysis of cold stress responses in rice seedlings. Proteomics, 5: 3162-3172.

Dodge, A. (1994). Herbicide action and effects on detoxification processes. *In* CH Foyer, PM Mullineaux, eds, Causes of Photooxidative Stress and Amelioration of Defense Systems in Plants. CRC, Boca Raton, FL, pp 219-236

Eastmond, P.J., Germain, V., Lange, P.R., Bryce, J.H., Smith, S.M., Graham, I.A. (2000). Postgerminative growth and lipid catabolism in oilseeds lacking the glyoxylate cycle. PNAS. 97: 5669-5674.

Ellerström, M., Stålberg, K., Ezcurra, I., Rask, L. (1996). Functional dissection of a napin gene promoter: identification of promoter elements required for embryo and endosperm-specific transcription. Plant Mol. Biol. 32: 1019-1027.

Ferreira, S., Hjerno, K., Larsen, M., Wingsle, G., Larsen, P., Fey, S., Roepstorff, P., Salome Pais, M. (2006). Proteome profiling of *Populus euphratica* Oliv. upon heat stress. Ann. Bot. 98: 361-377.

Focks, N., Benning, C. (1998). *wrinkled1*: A Novel, Low-Seed-Oil Mutant of Arabidopsis with a Deficiency in the Seed-Specific Regulation of Carbohydrate Metabolism. Plant Physiol. 118: 91-101.

Foyer, C., Mullineaux, P. (1994). Causes of photooxidative stress and amelioration of defense systems in plants. CRC, Boca Raton, FL

Friedrich, W. Handbuch der Vitamine. Munich: Urban and Schwartzenberg; (1987). Thiamine (Vitamin B1, aneurin); p. 240-258.

Geer, B.W., Lindel, D.L., Lindel, D.M. (1979). Relationship of the oxidative pentose shunt pathway to lipid synthesis in Drosophila melanogaster. Biochem. Genet. 17: 881-895.

Golda, A., Szyniarowski, P., Ostrowska, K., Kozik, A., Rapala-Kozik, M. (2004). Thiamine binding and metabolism in germinating seeds of selected cereals and legumes. Plant Physiol. Biochem. 42: 187-195.

Goyer, A. (2010). Thiamine in plants: Aspects of its metabolism and functions. Phytochemistry, 71: 1615-1624.

Haake, V., Geiger M, Walch-Liu, P., Engels, C., Zrenner, R., Stitt, M. (1999). Changes in aldolase activity in wild-type potato plants are important for acclimation to growth irradiance and carbon dioxide concentration, because plastid aldolase exerts control over the ambient rate of photosynthesis across a range of growth conditions. Plant J. 17: 479-489.

Hare, P.D., Cress, W.A., Van Staden, J. (1998). Dissecting the roles of osmolyte accumulation during stress. Plant Cell Environ. 21: 535-553.

Henkes, S., Sonnewald, U., Badur, R., Flachmann, R., Stitt, M. (2001). A small decrease of plastid transketolase activity in antisense tobacco transformants has dramatic effects on photosynthesis and phenylpropanoid metabolism. Plant Cell. 13: 535-551.

Hernandez, J.A., Almansa, M.S. (2002). Short-term effects of salt stress on antioxidant systems and leaf water relations of pea leaves. Physiol. Plantarum. 115: 251-257.

Huang, J., Redmann, R.E. (1995). Salt tolerance of *Hordeum* and *Brassica* species during germination and early seedling growth. Can. J. Plant Sci. 75: 815-819.

Iida, A., Nagasawa, A., Oeda, K. (1995). Positive and negative cis-regulatory regions in the soybean glycinin promoter identified by quantitative transient gene expression. Plant Cell Rep. 14: 539-544.

Imlay, J.A. (2003). Pathways of oxidative damage. Annu. Rev. Microbiol. 57: 395-418.

Jako, C., Kumar, A., Wei, Y., Zou, J., Barton, D.L., Giblin, E.M., Covello, P.S., Taylor, D.C. (2001). Seed-specific over-expression of an *Arabidopsis* cDNA encoding a

diacylglycerol acyltransferase enhances seed oil content and seed weight. Plant Physiol. 126: 861-874.

Jordan, F. (2003). Current mechanistic understanding of thiamine diphosphate-dependent enzymatic reactions. Nat. Prod. Rep. 20: 184-201.

Jurgenson, C.T., Begley, T.P., Ealick, S.E. (2009). The Structural and Biochemical Foundations of Thiamine Biosynthesis Annu. Rev. Biochem. 78:569-603.

Kang, F., Rawsthorne, S. (1996). Metabolism of glucose-6-phosphate and utilization of multiple metabolites for fatty acid synthesis by plastids from developing oil seed rape embryos. Planta, 199: 321-327.

Keddie, J.S., Tsiantis, M., Piffanelli, P., Cella, R., Hatzopoulos, P., Murphy, D.J. (1994). A seed-specific Brassica napus oleosin promoter interacts with a G-box-specific protein and may be bi-directional . Plant Mol. Biol. 24: 327-340.

Kerepesi, I., Galiba, G. (2000). Osmotic and Salt Stress-Induced Alteration in Soluble Carbohydrate Content in Wheat Seedlings. Crop Sci. 40: 482-487.

Kim, M.J., Yang, S.W, Mao, H.Z., Veena, S.P., Yin, J.L., Chua, N.H. (2014). Gene silencing of *Sugar-dependent 1* (*JcSDP1*), encoding a patatin-domain triacylglycerol lipase, enhances seed oil accumulation in *Jatropha curcas*. Biotechnol. Biofuels. 7: 36.

Kim, S.G., Lee, A.K., Yoon, H.K., Park, C.M. (2008). A membrane-bound NAC transcription factor NTL8 regulates gibberellic acid-mediated salt signaling in Arabidopsis seed germination. Plant J. 55:77-88.

Kim, Y.S., Nosaka, K., Downs, D.M., Kwak, J.M., Park, D., Chung, I.K., Nam, H.G. (1998). A Brassica cDNA clone encoding a bifunctional hydroxymethylpyrimidine kinase/thiamin-phosphate pyrophosphorylase involved in thiamine biosynthesis. Plant Mol. Biol. 37: 955-966.

Koca, M., Bor, M., Ozdemir, F., Turkan, I. (2007). The effect of salt stress on lipid peroxidation, antioxidative enzymes and proline content of sesame cultivars. Environ. Exp. Bot. 60: 344-351.

Kowalska, E., Kozik, A. (2008). The genes and enzymes involved in the biosynthesis of thiamine and thiamine diphosphate in yeasts. Cell Mol. Biol. Lett. 13: 271-282.

Kunert, K.J., Dodge, A.D. (1989) Herbicide-induced radical damage and antioxidative systems. In: Boger P, Sandmann G (eds) Target Sites of Herbicide Action. CRC Press Inc, Boca Raton, Fl, pp 45-63.

Larochelle, M., Drouin, S., Robert, F., Turcotte, B. (2006). Oxidative stress-activated zinc cluster protein Stb5 has dual activator/repressor functions required for pentose phosphate pathway regulation and NADPH production. Mol. Cell Biol. 26: 6690-6701.

Lascano, R., Munoz, N., Robert, G., Rodriguez, M., Melchiorre, M., Trippi, V., Quero, G. (2012). Paraquat: an oxidative stress inducer. INTECH Open Access Publisher.

Li, Y., Beisson, F., Pollard, M., Ohlrogge, J. (2006). Oil content of Arabidopsis seeds: The influence of seed anatomy, light and plant-to-plant variation. Phytochsemistry, 67: 904-915.

Lonsdale, D. (2006). A review of the biochemistry, metabolism and clinical benefits of thiamine and its derivatives. eCAM, 3: 49-59.

Lukienko, P.I., Mel'nichenko, N.G., Zverinskii, I.V., Zabrodskaya, S.V. (2000). Antioxidant properties of thiamine. B. Exp. Biol. Med. 130: 874-876.

Lupini, A., Sorgonà, A., Miller, A.J., Abenavoli, M.R. (2010). Short-term effects of coumarin along the maize primary root axis. Plant Signal. Behav. 5: 1395-1400.

Maisonneuve, S., Bessoule, J.J., Lessire, R., Delseny, M., Roscoe, T.J. (2010). Expression of Rapeseed Microsomal Lysophosphatidic Acid Acyltransferase Isozymes Enhances Seed Oil Content in Arabidopsis. Plant Physiol. 152: 670-684.

Mansfield, S.G., Briarty, L.G. (1992). Cotyledon cell development in *Arabidopsis thaliana* during reserve deposition. Can. J. Bot. 70: 151-164.

Masuko, T., Minami, A., Iwasaki, N., Majima, T., Nishimura, S.I., Lee, Y.C. (2005). Carbohydrate analysis by a phenol-sulforic acid method in microplate format. Anal. Biochem. 339: 69-72.

Meyerowitz, E.M., Somerville, C.R. (1994). *Arabidopsis*. CSHL Press, New York.

Mc Kersie, B.D., Leshem, Y.Y. (1994). Stress and stress coping in cultivated plants. Kluwer Academic Publishers, Dordrecht, The Netherlands, 256 pp.

Miller, G., Suzuki, N., Ciftci-Yilmaz, S., Mittler, R. (2010). Reactive oxygen species homeostasis and signaling during drought and salinity stresses. Plant Cell Environ. 33: 453-467.

Molin, W.T., Wilkerson, C.G., Fites, R.C. (1980). Thiamine Phosphorylation by thiamine pyrophosphotransferase during seed germination. Plant Physiol. 66: 313-315.

Murashige, T., Skoog, F. (1962). A revised medium for rapid growth and bio-assays with tobacco tissue cultures. Physiol. Plant. 15: 473-497.

Nosaka, K. (2006). Recent progress in understanding thiamine biosynthesis and its genetic regulation in *Saccharomyces cerevisiae*. Appl. Microbiol. Biot. 72: 30-40.

Palmer, A.M. (1999). The activity of pentose phosphate pathway is increased in response to oxidative stress in Alzheimer's disease. J. Neural Transm. 106:317-328.

Parida, A.K., Das, A.B. (2005). Salt tolerance and salinity effect on plants: a review. Ecotoxicol. Environ. Saf. 60: 324-349.

Pilon-Smits, E.A.H., Ebskamp, M.J.M., Paul, M.J., Jeuken, M.J.W., Weisbeek, P.J., Smeekens, S.C.M. (1995). Improved performance of transgenic fructan-accumulating tobacco under drought stress. Plant Physiol. 107: 125-130.

Pourcel, L., Moulin, M., Fitzpatrick, T.B. (2013). Examining strategies to facilitate vitamin B_1 biofortification of plants by genetic engineering. Front. Plant Sci. 4: 1-8.

Price, A.H., Hendry, G.A.F. (1991). Iron-catalyzed oxygen radical formation and its possible contribution to drought damage in nine native grasses and three cereals. Plant Cell Environ. 14: 477-484.

Raines, C.A., Harrison, E.P., Olcer, H., Lloyd, J.C. (2000). Investigating the role of the thiol-regulated enzyme sedoheptulose-1, 7-bisphosphatase in the control of photosynthesis. Physiol. Plant. 110: 303-308.

Rapala-Kozik, M., Kowalska, E., Ostrowska, K. (2008). Modulation of thiamine metabolism in *Zea mays* seedlings under conditions of abiotic stress. J. Exp. Bot. 59: 4133-4143.

Rapala-Kozik, M., Wolak, N., Kujda, M., Banas, A.K. (2012). The upregulation of thiamine (vitamin B1) biosynthesis in *Arabidopsis thaliana* seedlings under salt and osmotic stress conditions is mediated by abscisic acid at the early stages of this stress response. BMC Plant Biol. 12: 1-14.

Rawsthorne, S. (2002). Carbon flux and fatty acid synthesis in plants. Prog. Lipid Res. 41: 182-196.

Rizhsky, L., Liang, H., Shuman, J., Shulaev, V., Davletova, S., Mittler, R. (2004). When defense pathways collide. The response of Arabidopsis to a combination of drought and heat stress. Plant Physiol. 134: 1683-1696.

Roje, S. (2007). Vitamin B biosynthesis in plants. Phytochemistry, 68: 1904-1921.

Sayed, S.A., Gadallah, M.A.A. (2002). Effects of shoot and root application of thiamine on salt-stressed sunflower plants. Plant Growth Regul. 36: 71-80.

Scandalios, J.G. (1993). Oxygen stress and superoxide dismutase. Plant Physiol. 101: 7-12.

Schwender, J., Hay, J.O. (2012). Predictive modeling of biomass component tradeoffs in *Brassica napus* developing oilseeds based on in silico manipulation of storage metabolism. Plant Physiol. 160: 1218-1236.

Schwender, J., Ohlrogge, J.B., Shachar-Hill, Y. (2003). A Flux Model of Glycolysis and the Oxidative Pentosephosphate Pathway in Developing Brassica napus Embryos. J. Biol. Chem. 278: 29442-29453.

Slabas, A.R., Fawcett, T. (1992). The biochemistry and molecular biology of plant lipid biosynthesis. Plant Mol. Biol. 19: 161-191.

Stitt, M., Schulze, E.D. (1994). Does Rubisco control the rate of photosynthesis and plant growth? An exercise in molecular ecophysiology. Plant Cell Environ. 17: 465-487.

Suzuki, N., Bajad, S., Shuman, J., Shulaev, V., Mittler, R. (2008). The transcriptional co-activator MBF1c is a key regulator of thermotolerance in *Arabidopsis thaliana*. J. Biol. Chem. 283: 9269-9275.

Tunc-Ozdemir, M., Miller, G., Song, L., Kim, J., Sodek, A., Koussevitzky, S., Misra, A.N., Mittler, R., Shintani, D. (2009). Thiamine confers enhanced tolerance to oxidative stress in *Arabidopsis*. Plant Physiol. 151: 421-432.

Van Erp, H., Kelly, A.A., Menard, G., Eastmond, P.J. (2014). Multigene engineering of triacylglycerol metabolism boosts seed oil content in Arabidopsis. Plant Physiol. 165: 30-36.

Verslues, P.E., Agarwal, M., Katiyar-Agarwal, S., Zhu, J., Zhu, J.K. (2006). Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. Plant J. 45: 523-539.

Vigeolas, H., Waldeck, P., Zank, T., Geigenberger, P. (2007). Increasing seed oil content in oil-seed rape (*Brassica napus* L.) by over-expression of a yeast glycerol-3-phosphate dehydrogenase under the control of a seed-specific promoter. Plant Biotechnol. J. 5: 431-441.

Wachter, A., Tunc-Ozdemir, M., Grove, B.C., Green, P.J., Shintani, D.K., Breaker, R.R. (2007). Riboswitch control of gene expression in plants by splicing and alternative 3′ end processing of mRNAs. Plant Cell. 19: 3437-3450.

Wang, H., Guo, J., Lambert, K.N., Lin, Y. (2007). Developmental control of *Arabidopsis* seed oil biosynthesis. Planta, 226: 773-783.

Wolak, N., Kowalska, E., Kozik, A., Rapala-Kozik, M. (2014). Thiamine increases the resistance of baker's yeast *Saccharomyces cerevisiae* against oxidative, osmotic and thermal stress, through mechanisms partly independent of thiamine diphosphate-bound enzymes. FEMS Yeast Res. 14:
1249-1262.

Yazdani, M., Zallot, R., Tunc-Ozdemir, M., De Crécy-Lagard, V., Shintani D.K., Hanson, A.D. (2013). Identification of the thiamine salvage enzyme thiazole kinase in *Arabidopsis* and maize. Phytochemistry, 94: 68-73.

Yoshimura, K., Miyao, K., Gaber, A., Takeda, T., Kanaboshi, H., Miyasaka, H., Shigeoka, S. (2004). Enhancement of stress tolerance in transgenic tobacco plants overexpressing *Chlamydomonas* glutathione peroxidase in chloroplasts or cytosol. Plant J. 37: 21-33.

Yusa, T. (1961). Thiamine triphosphate in yeasts and some plant materials. Plant Cell Physiol. 2: 471-474.

Zallot, R., Yazdani, M., Goyer, A., Ziemak, M.J., Guan, J.C., McCarty, D.R., De Crécy-Lagard, V., Gerdes, S., Garrett, T.J., Benach, J., Hunt, J.F., Shintani, D.K., Hanson, A.D. (2014). Salvage of the thiamine pyrimidine moiety by plant TenA proteins lacking an active-site cysteine. Biochem. J. 463: 145-155.

Zhang, F.Y., Yang, M.F., Xu, Y.N. (2005). Silencing of *DGAT1* in tobacco causes a reduction in seed oil content. Plant Sci. 169: 689-694.

Zou, J., Katavic, V., Giblin, E.M., Barton, D.L., MacKenzie, S.L., Keller, W.A., Hu, X., Taylor, D.C. (1997). Modification of seed oil content and acyl composition in the Brassicaceae by expression of a yeast sn-2 acyltransferase gene. Plant Cell. 9: 909-923.

CHAPTER 5

Concluding remarks and future directions

Elucidation of thiamin salvage pathway in plants

Although the genes and enzymes involved in thiamin *de novo* biosynthetic pathways have been identified and characterized in microorganisms and plants in several studies (Jurgenson et al., 2009; Goyer, 2010; Gerdes et al, 2012), there is not enough information regarding thiamin breakdown and the fate of its degradation products i.e. thiazole and pyrimidine (Goyer, 2010). Because thiamin and its phosphate esters are chemically and enzymatically unstable (McCourt et al., 2006, Goyer, 2010; Fitzpatrick et al., 2012), plants and microorganisms have developed alternative pathways to re-use the thiamin degradation products for thiamin biosynthesis (Li and Rédei, 1969; Jurgenson et al., 2009).

Our study showed that the *Arabidopsis* and maize genomes encode homologs of ThiM with HET kinase activity. These results were confirmed via enzymatic assays using purified recombinant proteins. The results demonstrated that ThiM homologs in plants play an important role in thiamin salvage.

Because thiamin biosynthesis in plants occurs in plastids (Goyer, 2010 and Gerdes et al., 2012) and ThiM proteins seem to be cytosolic, plastids might have a transporter for thiazole phosphate that remains to be identified. Finally, it is also worth to investigate the effects of ThiM on the expression pattern of thiamin biosynthetic genes involved in *de novo* pathway.

Thiaminase II (TenA) is another enzyme with a dual role in thiamin degradation and salvage pathway. However, it has been reported that TenA function in bacteria is mostly the regeneration of the pyrimidine moiety of thiamin rather than thiamin degradation (Jenkins et al., 2007).

Our study showed that plants possess TenA_E subfamily of TenA proteins. Using amino-HMP in plant growth medium, our experiments demonstrated that *ThiC/TenA_E* double mutant *Arabidopsis* plants were not able to grow. However, *ThiC* and *TenA_E* single mutants could grow normally on medium supplemented with amino-HMP. These results show that TenA_E protein is an important factor in thiamin salvage pathway by converting of the amino-HMP to HMP. Additionally, stress experiments using paraquat-induced oxidative stress showed that in *TenA_E* single mutants root growth was significantly more sensitive to oxidative stress compare to the wild type plants. Although this sensitivity could be eliminated by the adding of HMP or thiamin to the growth medium, thiamin had stronger effect in removing the oxidative stress.

In conclusion, our data demonstrated that salvage pathways mediated by ThiM and TenA_E enzymes play an important role in both conserving energy and stress tolerance in plants. Additionally, it is worth investigating whether ThiM protein would be able to boost plant stress tolerance like what was shown for TenA_E protein in our study. It would also be interesting to investigate if plants are able to store more thiamin by the suppression of TenA_E, which has been shown to have thiamin degradation activity in bacteria in addition to its role in thiamin salvage pathway.

Genetic engineering of *Arabidopsis* seeds to increase nutritional value and abiotic stress tolerance

Plants are one of the main sources of nutrients for animals and humans. Because the world population is booming rapidly, it is critical to improve not only the quantity but also the quality of the crop plants. Genetic engineering is one of the ways to boost valuable nutrient content in plants especially staple crop plants such as rice, maize, and wheat. It can also be used to increase plant resistant against unfavorable conditions such as biotic and abiotic stresses.

In our research, by seed overexpression of thiamin biosynthetic genes (*Thi4*, *ThiC* and *ThiE*) we could achieve to up to 7-fold increase in total thiamin content in transgenic seeds. This achievement was much higher than the thiamin-overproduction using the overexpression of thiamin biosynthetic genes (*Thi4* and *ThiC*) by constitutive promoters reported by Dong et al, (2015) with up to 2.6-fold increase in seed thiamin content. The greater amount of thiamin in our study could be as a result of using strong seed-specific promoters i.e. glycinin, oleosin, and napin (Ellerström et al., 1996; Iida et al., 1995; Keddie et al., 1994) and overexpression of *ThiE* gene beside the overexpression of the two other genes, as well.

The fact that thiamin can alleviate the adverse effects of abiotic stresses in plants has been proven in several studies (Sayed and Gadallah, 2002; Rapala-Kozik et al., 2008; Tunc-Ozdemir et al., 2009). These results were obtained from the feeding studies by the adding of thiamin to the plants growth medium. To the best of our knowledge, this is the first report showing that the over-production of thiamin in plant seeds could increase the seeds germination and seedlings viability against salt and paraquat stresses. These results were inconsistent with the results reported by Dong et al, (2015) in which 2.6-fold

increase in *Arabidopsis* seed thiamin could not confer stress tolerance to the plants under salt and paraquat treatments. One possibility could be that 2.6-fold increase might be below the minimum thiamin threshold required for abiotic stress resistance. Additionally, using tissue-specific promoters can reduce energy drain on the whole plant (Pourcel et al., 2013) which has not been considered in Dong et al, (2015) study in which constitutive promoters were used to overexpress the *Thi4* and *ThiC* genes in *Arabidopsis*.

As mentioned in chapter 4, thiamin over-producing seed showed the altered seed phenotypes compared to the control plants. This is a novel finding showing that the high-thiamin lines contained more oil and carbohydrates but less total protein. These results clearly show that the carbon flux in transgenic seeds could be significantly affected by the increased thiamin levels in the seeds. This shift in carbon flux could be a result of the increased levels of reducing power in the cell by the thiamin cofactor-requiring enzymes such as transketolase, pyruvate dehydrogenase, and α-ketoglutarate dehydrogenase which are involved in the production of NAD(P)H, as the expression of the genes corresponding to these enzymes in high-thiamin lines were significantly higher than the control plants.

In addition, we hypothesized that altered seed phenotypes in transgenic lines could be an impact of high levels of thiamin in these seeds which could led to an increase in TPP levels. TPP is an important cofactor for the activity of TPP-requiring enzymes. To test this hypothesis *TPK1* transformants were analyzed for oil, carbohydrate, and protein content. Interestingly, the results were in agreement with the results obtained from 3-gene transformants. In *TPK1* transformants the total seed thiamin content was increased in comparison to the control seeds. Expression pattern analysis of thiamin biosynthetic genes revealed that overexpression of *TPK1* could upregulate the expression of *Thi4* and

ThiC genes in *TPK1* transgenic plants. These results confirm that both branches of the thiamin biosynthetic pathway are important to produce high-thiamin transgenic plants (Pourcel et al., 2013).

Taken together, this thesis focuses on plant metabolic engineering to increase both seed nutritional value by the overexpression of thiamin biosynthetic genes and plant abiotic stress tolerance. The results obtained from this study will provide valuable tools not only to improve the nutritional content of food crops such as rice, wheat, maize, and soybean but also to increase the tolerance of plants against abiotic stress conditions which are important in agriculture and industry in both developed and developing counties. It is also important to consider the seed thiamin binding proteins which play a critical role during seed maturation and germination (Watanabe et al., 2004; Pourcel et al., 2013). In addition, it is worth investigating the cross of the 3-gene overexpressors with the *TPK1* transformants to assess thiamin, TPP, and the other metabolite derivatives, and its possible effect on plant stress tolerance. It would also be interesting to investigate if thiamin plays a direct role in abiotic stress tolerance by the detoxification of ROS. This idea could be tested by the exposuring plants to abiotic stress conditions and then measuring the levels of oxidized form of thiamin, thiochrome by HPLC, as Lukienko et al. (2000) showed that in animal cells thiamin could be oxidized to thiochrome by the direct interaction with free radicals and hydroperoxides. Furthermore, it is important to investigate how high-thiamin levels can affect the signal transduction pathways in plants in both biotic and abiotic stress conditions. This idea could be evaluated by the measuring of the activity of plasma membrane NAD(P)H-oxidase, levels of produced ROS, and the

activity of Mitogen-Activated Protein Kinases (MAPKs) in wild type and transgenic plants subjected to various biotic and abiotic stresses.

Last but not least, transcriptome and metabolome profiling of thiamin over-producing seeds will help us to have better insight about the impact of high-thiamin on the expression pattern of the genes and metabolic flux in plant cells.

REFERENCES:

Dong, W., Stockwell, V.O., Goyer, A. (2015). Enhancement of thiamin content in *Arabidopssis thaliana* by metabolic engineering. Plant Cell Physiol. doi: 10.1093/pcp/pcv148.

Ellerström, M., Stålberg, K., Ezcurra, I., Rask, L. (1996). Functional dissection of a napin gene promoter: identification of promoter elements required for embryo and endosperm-specific transcription. Plant Mol. Biol. 32: 1019-1027.

Fitzpatrick, T.B., Basset, G.J., Borel, P., Carrari, F., DellaPenna, D., Fraser, P.D., Hellmann, H., Osorio, S., Rothan, C., Valpuesta, V., Caris-Veyrat, C., Fernie, A.R. (2012). Vitamin deficiencies in humans: can plant science help? Plant Cell. 24: 395-414.

Gerdes, S., Lerma-Ortiz, C., Frelin, O., Seaver, S.M., Henry, C.S., de Crécy-Lagard, V., Hanson, A.D., (2012). Plant B vitamin pathways and their compartmentation: a guide for the perplexed. J. Exp. Bot. 63: 5379-5395.

Goyer, A. (2010). Thiamine in plants: Aspects of its metabolism and functions. Phytochemistry, 71: 1615-1624.

Iida, A., Nagasawa, A., Oeda, K. (1995). Positive and negative cis-regulatory regions in the soybean glycinin promoter identified by quantitative transient gene expression. Plant Cell Rep. 14: 539-544.

Jenkins, A.H., Schyns, G., Potot, S., Sun, G., Begley, T.P. (2007). A new thiamin salvage pathway. Nat. Chem. Biol. 3: 492-497.

Jurgenson, C.T., Begley, T.P., Ealick, S.E. (2009). The structural and biochemical foundations of thiamin biosynthesis. Annu. Rev. Biochem. 78: 569-603.

Keddie, J.S., Tsiantis, M., Piffanelli, P., Cella, R., Hatzopoulos, P., Murphy, D.J. (1994). A seed-specific Brassica napus oleosin promoter interacts with a G-box-specific protein and may be bi-directional . Plant Mol. Biol. 24: 327-340.

Li, S.L., Rédei, G.P. (1969). Thiamine mutants of the crucifer, Arabidopsis. Biochem. Genet. 3:163-170.

Lukienko, P.I., Mel'nichenko, N.G., Zverinskii, I.V., Zabrodskaya, S.V. (2000). Antioxidant properties of thiamine. B. Exp. Biol. Med. 130: 874-876.

McCourt, J.A., Nixon, P.F., Duggleby, R.G., (2006). Thiamin nutrition and catalysis-induced instability of thiamin diphosphate. Br. J. Nutr. 96: 636-638.

Pourcel, L., Moulin, M., Fitzpatrick, T.B. (2013). Examining strategies to facilitate vitamin B_1 biofortification of plants by genetic engineering. Front. Plant Sci. 4: 1-8.

Rapala-Kozik, M., Kowalska, E., Ostrowska, K. (2008). Modulation of thiamine metabolism in *Zea mays* seedlings under conditions of abiotic stress. J. Exp. Bot. 59: 4133-4143.

Sayed, S.A., Gadallah, M.A.A. (2002). Effects of shoot and root application of thiamin on salt-stressed sunflower plants. Plant Growth Regul. 36: 71-80.

Tunc-Ozdemir, M., Miller, G., Song, L., Kim, J., Sodek, A., Koussevitzky, S., Misra, A.N., Mittler, R., Shintani, D. (2009). Thiamin confers enhanced tolerance to oxidative stress in *Arabidopsis*. Plant Physiol. 151: 421-432.

Watanabe, K., Nishida, N., Adachi, T., Ueda, M., Mitsunga, T., Kawamura, Y. (2004). Accumulation and degradation of thiamin-binding protein and level of thiamine in wheat seeds during seed maturation and germination. Biosci. Biotechnol. Biochem. 68: 1243-1248.