Tomato LeTHIC is an Fe-Requiring HMP-P Synthase Involved in Thiamine Synthesis and Regulated by Multiple Factors

Weina Zhao1,2,3, Xudong Cheng1,2,3, Zongan Huang1,2, Huajie Fan1,2, Huilan Wu1 and Hong-Qing Ling1,*

1The State Key Laboratory of Plant Cell and Chromosome Engineering, National Centre for Plant Gene Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, No 1 West Beichen Road, Chaoyang District, Beijing 100101, PR China
2Graduate University of the Chinese Academy of Sciences, Yuquan Road, Shijingshan District, Beijing 100049, PR China
3These authors contributed equally to this work
*Corresponding author: E-mail, hqling@genetics.ac.cn; Fax, +86-10-64860377
(Received April 20, 2010; Accepted April 7, 2011)

Thiamine is a key primary metabolite which is necessary for the viability of all organisms. It is a dietary requirement for mammals because only prokaryotes, fungi and plants are thiamine prototrophs. In contrast to the well documented biosynthetic mechanism in bacteria, much remains to be deciphered in plants. In this work, a tomato thiamine-auxotrophic (thiamineless, tl) mutant was characterized. The tl mutant occurs due to inactivation of LeTHIC transcription as a result of insertion of a large unknown DNA fragment in its 5’-untranslated region. Expression of wild-type LeTHIC in tl plants was able to complement the mutant to wild type. LeTHIC possessed the same function as EcTHIC [an Escherichia coli 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate (HMP-P) synthase involved in synthesis of the pyrimidine moiety of thiamine] because expression of LeTHIC rescued THIC-deficient strains of E. coli under culture conditions without thiamine supplementation, suggesting that plants employ a bacteria-like route of pyrimidine moiety synthesis. LeTHIC is an Fe–S cluster protein localized in chloroplasts, and Fe is required for maintenance of its enzyme activity because Fe deficiency resulted in a significant reduction of thiamine content in tomato leaves. Further, we also showed that the expression of LeTHIC is tightly regulated at the transcriptional and post-transcriptional level by multiple factors, such as light, Fe status and thiamine pyrophosphate (TPP)-riboswitch. The results clearly demonstrated that a feedback regulation mechanism is involved in synthesis of the pyrimidine moiety for controlling thiamine synthesis in tomato. Our results provide a new insight into understanding the molecular mechanism of thiamine biosynthesis in plants.

Keywords: Fe-dependent HMP-P synthase • Fe–S cluster protein • Thiamine • Thiamineless • Tomato • TPP-riboswitch.

Abbreviations: AdoMet, S-adenosyl methionine; AIR, 5’-aminomimidazole ribonucleotide; BAC, bacterial artificial chromosome; CaMV, cauliflower mosaic virus; GFP, green fluorescent protein; HMP-P, 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate; IPTG, isopropyl-β-D-thiogalactopyranoside; ORF, open reading frame; RACE, rapid amplification of cDNA ends; RT–PCR, reverse transcription–PCR; THI, thiamine; TMP, thiamine monophosphate; TPP, thiamine pyrophosphate; TTP, thiamine triphosphate; UTR, untranslated region

Sequence data from this article have been deposited in the EMBL/GenBank Data Libraries under the accession number EU379346.

Introduction

Thiamine (vitamin B1) is a key primary metabolite for viability of all organisms. Its active form thiamine pyrophosphate (TPP) is a crucial co-enzyme of pyruvate dehydrogenase, α-ketoglutarate dehydrogenase and transketolase. Thiamine triphosphate (TTP) is an important regulatory factor of vertebrate neuronal activity. In the natural world, only prokaryotes, fungi and plants are thiamine prototrophs; mammals lost the capacity for de novo thiamine synthesis, and it must be obtained from dietary sources such as cereal grains, leafy green vegetables and fruits. Deficiency of thiamine disturbs the central nervous and circulatory systems and causes beriberi disease (Buridick 1998, Nosaka 2006). In plants, Tunc-Ozdemir et al. (2009) observed that thiamine is able to enhance tolerance to oxidative stress during different abiotic stress conditions. More recently, it was reported that thiamine was required for shoot meristem maintenance of maize (Woodward et al. 2010). However, thiamine biosynthesis in plants is not well elucidated.

Current understanding of thiamine biosynthesis relies primarily on genetic and biochemical studies in microorganisms, such as Escherichia coli, Bacillus subtilis and Saccharomyces cerevisiae. The de novo synthesis of thiamine is a complex, highly regulated pathway. The thiazole and pyrimidine moieties of thiamine are synthesized separately and then coupled to...
form thiamine monophosphate (TMP), which is further phosphorylated to TPP and TTP (Begley et al. 1999, Allen et al. 2002, Morett et al. 2003). In bacteria, 12 gene products distributed across 11 enzymatic steps are required for thiamine synthesis (Rodionov et al. 2002). Biosynthesis of the thiazole moiety requires five gene products (ThiF, ThiS, ThiG, ThiH and ThiI) (Rodionov et al. 2002). Biosynthesis of the thiazole moiety across 11 enzymatic steps are required for thiamine synthesis (Begley et al. 1999, Allen et al. 2002, Wachter et al. 2007). Ajjawi et al. (2007a) reported that thiamine pyrophosphate (TPP) is required for pyrimidine synthesis and catalyzes the sequential phosphorylation of HMP-P (4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate) (Chatterjee et al. 2008). In S. cerevisiae, 15 THI genes are involved in thiamine biosynthesis (Nosaka 2006). Thiazole biosynthesis requires a single enzyme THI4 and uses 1-deoxy-xylulose 5-phosphate, tyrosine and cysteine as the primary substrates (Park et al. 2003). For synthesis of the pyrimidine unit, researchers recently clarified that THIC was responsible for the reaction of AIR (5,6-diaminopurine 5′-phosphate) with AdoMet (S-adenosyl methionine) to produce HMP-P (4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate) (Chatterjee et al. 2008). In S. cerevisiae, 15 THI genes are involved in thiamine biosynthesis (Nosaka 2006). Thiazole biosynthesis requires a single enzyme THI4 and uses NAD, glycine and an unknown sulfur source (Chatterjee et al. 2007). HMP-P is formed via condensing pyridoxine with histidine with catalysis by the THIS enzyme (Maundrell 1990). In plants, some mutants defective in thiamine biosynthesis were found and characterized. Five thiamine-autotroph mutants (py, tz, th1, th2 and th3) were reported in Arabidopsis thaliana (Li and Rede 1969, Koornneef and Hanhart 1981). py and tz mutants disrupt the synthesis of the pyrimidine and thiazole moiety, respectively, while in th1, th2 and th3, some synthetic processes after pyrimidine and thiazole synthesis were blocked because the three mutants responded to spraying with thiamine, but not to spraying with pyrimidine and/or thiazole (Boynton 1966, Li and Rede 1969, Mitsuda et al. 1970a, Mitsuda et al. 1970b, Kumar and Sharma 1986, Proebsting et al. 1990, Papini-Terzi et al. 2003). The tz locus corresponds to the Thi1 gene which is homologous to S. cerevisiae THI4, and encodes an enzyme involved in the synthesis of the thiazole moiety (Praekelt et al. 1994, Godoi et al. 2006). The th1 locus resulted from mutation of the HMPPK/TMPPase gene, which is involved in TMP synthesis and catalyzes the sequential phosphorylation of HMP/HMP-P (Kim et al. 1998, Ajaiji et al. 2007b). Ajaiji et al. (2007a) reported that thiamine pyrophosphokinase was required for the conversion of free thiamine to the enzyme cofactor TPP in plants because lack of expression of AtTPK1 and AtTPK2, which encoded thiamine pyrophosphokinase, resulted in a seedling-lethal phenotype which was able to be rescued by spraying with TPP, but not by free thiamine and TMP. Additionally, Belanger et al. (1995) reported that thi1 encoding the functional homolog of the yeast THI4 gene was involved in the synthesis of the thiazole moiety in maize, and Rapala-Kozik et al. (2007) demonstrated that maize thi3, like thi1 of A. thaliana and bth1 of Brassica napus, encoded a product possessing HMP phosphate kinase and TMP synthase enzymatic activities, functioning in the biosynthesis of thiamine.

As shown above, a defect in the py locus, which is responsible for pyrimidine biosynthesis, reveals a block in the pathway to pyrimidine synthesis. Recently, the AtTHIC gene of Arabidopsis was identified by sequence similarity to the THIC gene of E. coli, and reported to be required for Arabidopsis HMP-P biosynthesis (Raschke et al. 2007, Kong et al. 2008). AtTHIC encoded a putative Fe–S cluster protein and was post-transcriptionally suppressed by TPP through a riboswitch (a metabolite-sensing gene control element that is located in non-coding regions of mRNA, where it selectively binds its target compound and subsequently modulates gene expression) in the 3′-untranslated region (UTR) of precursor mRNA (Bocobza et al. 2007, Wachter et al. 2007). However, the detailed functions and expression regulation of THIC genes have still not been illustrated in plants.

In tomato, a pyrimidine moiety synthesis mutant was found by Langridge and Brock (1961). Clayberg et al. (1966) termed this mutant thiamineless (tl). In this work, we isolated its corresponding gene, which encodes an amino acid sequence homologous to EcTHIC of E. coli and AtTHIC of A. thaliana. We demonstrated that it was a chloroplastic protein and required Fe for maintenance of its enzyme activity. Further, we showed that expression of the gene was regulated at the transcriptional and post-transcriptional levels by multiple factors, such as light, riboswitch-mediated feedback inhibition by TPP, and Fe status.

**Results**

**The tomato thiamineless (tl) mutant is caused by inactivation of LeTHIC expression**

The tomato tl mutant is a spontaneous mutant, exhibiting a chlorotic and albino phenotype under culture conditions without thiamine supply (Langridge and Brock 1961). As described previously, it grew extremely slowly, and formed a few leaves which were small and pale yellow with few developed chloroplasts in mesophyll cells under normal culture conditions (Fig. 1A, B, upper part). With time, the leaves withered and died. Further, the root growth of the tl mutant was characterized after being grown for 1 month in a hydroponics system (Fig. 1C). Total root length and average root volume of the tl mutant were 161.48 cm and 0.18 cm³, respectively, significantly less than those (260 cm and 0.28 cm³) of the wild type (Fig. 1D, E). Additionally, tl plants also showed an impaired lateral root development, the root tips per tl plant were only one-third of those of the wild type (Fig. 1F). The defective functions in the tl mutant could be completely rescued by spraying with thiamine (Fig. 1A, B, lower part). With HPLC, we measured the thiamine content of the tl mutant and its wild type. The thiamine content in leaves of the wild type was about 28.1 µg 100 g FW⁻¹, while only 2.4 µg 100 g FW⁻¹ was determined in the mutant leaves (Fig. 1G). These results indicate that thiamine synthesis in the tl mutant was defective and the phenotypes of the tl mutant were obviously caused by thiamine deficiency.

Previous genetic analysis revealed that the tl mutant was a recessive, single-gene mutation located between markers GP164 and GP79 on the short arm of tomato chromosome 6.
Using an F2 segregating population derived by a cross between LA0758/C0 (a tl mutant derived from LA0758/+ by selfing) and Lycopersicon pennellii (a wild tomato species), and polymorphic markers from Tomato-EXPEN 2000 (SGN; http://www.sgn.cornell.edu/), the tl locus was further mapped between the markers GP164 and SSR48, 3.0 cM from SSR48 (Supplementary Fig. S1). In this region, a unigene marker T1188 showed a high sequence similarity to the putative thiamine synthase gene AtTHIC (Raschke et al. 2007, Kong et al. 2008). Due to the lack of polymorphism of T1188, a new polymorphic marker, TG1, directly upstream of T1188 was developed using reverse PCR amplification (Fig. 2A, upper part). A segregation test showed that TG1 was co-segregated with the albino phenotype of the tl mutant. Thus, we speculated that the gene residing in T1188 might be the gene...
responsible for the \( tl \) mutant. Therefore, a bacterial artificial chromosome (BAC) library was constructed using the genomic DNA of the wild type and screened by PCR with \( \text{LeTHIC} \)-T primers designed from the sequence of T1188 (Supplementary Table S1). A single BAC clone was isolated and sequenced. An open reading frame (ORF) encoding a 651 amino acid protein was predicted, and it was considered as the candidate gene for the \( tl \) mutant. The predicted amino acid sequence of the ORF is highly similar to the sequence of \( \text{AtTHIC} \); therefore, we termed it \( \text{LeTHIC} \). Using the same primers as for BAC library screening, a tomato leaf cDNA library was screened, and several positive clones were obtained. Comparison of the cDNA sequence with the genomic sequence revealed that the \( \text{LeTHIC} \) gene was 5,256 bp and composed of seven exons and six introns (Fig. 2A, upper part). Northern blot analysis showed that \( \text{LeTHIC} \) was expressed abundantly in various tissues (such as leaf, root, stem, cotyledon, flower and fruit) of the wild type, but not in such tissues of the \( tl \) mutant (Fig. 2B), suggesting that lack of \( \text{LeTHIC} \) expression in the mutant may be the reason for the \( tl \) phenotype.

To test this hypothesis, the full-length coding region of \( \text{LeTHIC} \) under the control of the cauliflower mosaic virus (CaMV) 35S promoter was introduced into the \( tl \) mutant by \( \text{Agrobacterium tumefaciens} \)-mediated transformation following the protocol described by Ling et al. (1998). Several positive transformants were obtained and displayed normal growth under culture conditions without thiamine supply (Fig. 3A). HPLC analysis showed that the transgenic plants accumulated thiamine like the wild type (Fig. 3B, upper part). Reverse transcription-PCR (RT–PCR) analysis revealed that transcription of \( \text{LeTHIC} \) was restored in the transgenic plants (Fig. 3B, lower part). The fact that expression of the wild-type \( \text{LeTHIC} \) gene in the \( tl \) mutant was able to complement the mutant to wild type suggests that the \( tl \) mutant of tomato is due to inactivation of \( \text{LeTHIC} \) expression.

To detect why \( \text{LeTHIC} \) is not expressed in \( tl \) plants, the genomic sequence of the \( \text{LeTHIC} \) gene from the \( tl \) mutant and its wild type was amplified by PCR and compared. A narrow region (about 160 bp indicated by a red line above the \( tl \) genomic DNA schematic representation in the lower part of Fig. 2A) in its...
LeTHIC could not be amplified in the tl mutant, while the sequences of the remaining regions were identical to their counterparts in wild-type LeTHIC. Southern blot analysis displayed different hybridization signals between the mutant and wild type when probed with the 160 bp segment amplified from wild-type LeTHIC (Fig. 2C). These results hint that an insertion of a large unknown nucleotide fragment in the 5'-UTR of LeTHIC occurred in the tl genome, and consequently disturbed LeTHIC expression.

LeTHIC is an evolutionarily conserved HMP-P synthase functioning in pyrimidine moiety synthesis of thiamine

An NCBI database search identified hundreds of LeTHIC homologs distributed in archaeal, bacterial and plant genomes. These THICs shared high similarity at both the nucleotide and amino acid level. LeTHIC displayed 50.7% sequence identity with E.coliTHIC and 83% with AtTHIC. Phylogenetic analysis based on sequence alignment of THIC proteins revealed that all the THICs clustered into a clear monophyletic group, and LeTHIC was closest to AtTHIC on the phylogenetic tree (Fig. 4A), indicating that the THICs may originate from a common ancient gene.

Thus far, the bacterial THIC was demonstrated to be a HMP-P synthase involved in the biosynthesis of the pyrimidine moiety of thiamine by reconstitution of the enzyme-catalyzed reaction (Chatterjee et al. 2008, Martinez-Gomez and Downs 2008). Based on the high sequence similarity to the THIC of E.coli, we speculated that LeTHIC may have the same functions as E.coliTHIC in plants. To test it, the coding region of LeTHIC was amplified by PCR and cloned into the expression vector pGEX-KG (Yam et al. 2000) under the control of the T7 promoter. The construct was transformed into THIC-deficient strains THIC3178::Tn10Kan (Nichols et al. 1998) and THIC765::Kan (Baba et al. 2006) of E.coli. The growth of transformed cells was observed after 36 h incubation on M9 and M9 + thiamine medium. All strains grew well when supplied with exogenous thiamine (Fig. 4B, left). Under culture conditions without thiamine supplementation, only the strains carrying LeTHIC grew, but those carrying the empty vector did not (Fig. 4B, right). These results clearly indicate that LeTHIC, like THIC of E.coli, is a HMP-P synthase involved in synthesis of the pyrimidine moiety of thiamine.

LeTHIC is a chloroplastic protein

To determine the possible site(s) of pyrimidine heterocycle biosynthesis in plants, the subcellular localization of LeTHIC was investigated. The protein sequence of LeTHIC was first subjected to computer-based localization prediction. An eight amino acid long chloroplast targeting sequence was identified in the N-terminus by both PSORT (Nakai and Horton 1999) and TargetP programs (Emanuelsson et al. 2000). To confirm the prediction, a fusion plasmid 35S:LeTHIC-GFP and the control vector 35S:GFP were transiently expressed in Arabidopsis protoplasts. After 18 h of culture, the subcellular location of the LeTHIC–green fluorescent protein (GFP) chimeric protein in cells was determined via fluorescence microscopy. Green fluorescence was detected throughout the cell transformed with the control vector 35S:GFP (Fig. 5A), while the signal of green fluorescence in the cell transformed with 35S:LeTHIC-GFP plasmid was distinguished exclusively in chloroplasts (Fig. 5B). Further, a strong green fluorescence signal was observed in chloroplast stroma in enlarged images of chloroplasts (Fig. 5C, D), hinting that LeTHIC was not equally distributed in chloroplasts, but...
mainly localized at a special zone of the chloroplast stroma. The chloroplastic localization of LeTHIC protein suggests that the synthesis of the pyrimidine moiety of thiamine in plants occurs in chloroplasts.

**TPP-riboswitch is involved in controlling LeTHIC expression at the post-transcriptional level**

According to previous research, a TPP-dependent alternative splicing of mRNA exists in *E. coli* and *A. thaliana* (Winkler et al. 2002, Bocobza et al. 2007, Wachter et al. 2007). Through screening a leaf cDNA library of tomato and RACE (rapid amplification of cDNA ends) analysis, three different transcripts of *LeTHIC* were obtained and named as L (long), M (medium) and S (short) variants based on their length. Three forms of *LeTHIC* mRNA shared an identical 5’-UTR and internal regions, and displayed varying 3’-UTR lengths. The common 5’-UTR was composed of two exons (66/16 bp) separated by a 1,523 bp intron, and the 3,071 bp long internal region consisted of four introns and five exons which encode a 651 amino acid protein (Fig. 2A). The 3’-UTR lengths of the L, M and S forms of *LeTHIC* mRNA were 861, 701 and 149 bp, respectively. A 160 bp large intron was present in the 3’-UTR of the L form (Fig. 6A, B). A potential TPP-riboswitch motif was identified in the 3’-UTR of the L variant by riboswitch explorer (RibEx) (Fig. 6C). It shared high sequence similarity (87% nucleotide sequence identity) with the TPP-riboswitch Eco1 of *E. coli* (Sudarsan et al. 2003) and Ath1 of *A. thaliana* (Wachter et al. 2007), which have equivalent function in mediating the TPP-dependent alternative splicing of mRNA (Winkler et al. 2002, Bocobza et al. 2007, Wachter et al. 2007). According to the results obtained in *E. coli* and *A. thaliana*, the L variant of *LeTHIC* mRNA will further split to form the M or S variant dependent on the thiamine status in cells. The M variant is predicted to be an unstable form because the distance between its stop codon and last intron–exon boundary is 89 bp (once this distance exceeds 50 nucleotides, nonsense-mediated mRNA decay will be activated; Maquat 2002). This inference was then confirmed with a half-life test in yeast. Full-size M and S variants of *LeTHIC* were separately cloned into the yeast expression vector pYES2.0 (Invitrogen) and then transformed into yeast cells which were grown in YPD medium. After induction by galactose for 9 h, actinomycin D (50 μg ml\(^{-1}\) final concentration) was added to block transcription. M type mRNA was unstable and completely disappeared 2 h after adding actinomycin D, whereas the S type mRNA remained relatively unchanged even at 4 h after the treatment (Supplementary Fig. S2).

Due to the presence of the TPP-riboswitch in the 3’-UTR of the L variant, TPP was regarded as a potential regulation element. Therefore, multiplex RT–PCR analysis with specific primers for the three transcripts of *LeTHIC* (Supplementary Table S1) was used to determine the impact of increased TPP concentrations on *LeTHIC* expression. When sprayed with thiamine at a concentration of 0.05, 0.1 and 1.0 mM, the total amount of *LeTHIC* transcripts decreased to ~63, ~47 and ~39% of that measured in seedlings sprayed with only water (as a control). The S and L types of *LeTHIC* transcripts exhibited a decreased abundance with increased thiamine concentration, whereas the M variant was dramatically elevated (>40% at concentrations of 0.05 and 0.1 mM, and >130% at a concentration of 1.0 mM) (Fig. 7A, upper part). Further, the protein accumulation of *LeTHIC* was determined by Western blot analysis. It showed that *LeTHIC* accumulation is consistent with the abundance of S type mRNA and gradually decreased with increasing thiamine concentration (Fig. 7A, lower part). Additionally, the...
time-dependent effects of thiamine treatment on LeTHIC transcripts were assessed with seedlings sprayed with 0.05 mM thiamine. The total mRNA of LeTHIC was reduced to 63 and 45% of that measured in the control at 4 and 28 h after thiamine application, respectively. The reduction of the total mRNA is obviously caused by the reduced amounts of the S and L types, while the M type of LeTHIC was increased (about 40%) (Fig. 7B, upper part). The reduction of the S form mRNA was more dramatic (about 40% at 4 h, 55% at 28 h) than that of the L type (about 18% at 4 h, 22% at 28 h). Immunoblot analysis revealed that the significant reduction of LeTHIC protein accumulation was only observed at 28 h after the supply of thiamine (Fig. 7B, lower part).

**Effect of light on LeTHIC expression**

Considering that LeTHIC was mainly expressed in leaf and its product was localized in chloroplasts, we speculated that light may have an effect on the expression of LeTHIC. To test the hypothesis, a time course experiment of LeTHIC expression was performed and LeTHIC transcripts were quantified by multiplex RT–PCR. All the three transcript variants of LeTHIC were detectable in light (09:00–21:00 h) and dark (21:00–09:00 h) periods, and the S type was most abundant at all time points investigated (Fig. 8A). The three transcripts of LeTHIC displayed a gradual increase over the whole light period (from the morning at 09:00 h to the evening at 21:00 h), and reached their peak at 21:00 h. During the dark period, the abundance of the three variants of LeTHIC mRNA decreased gradually with extension of the dark period. At the end of the dark period, the amount of S type mRNA was only half of that measured at 21:00 h (beginning of the dark period), and the M variant almost disappeared (Fig. 8A). At the protein level, LeTHIC accumulation was basically similar to that of its transcript abundance, increased with extension of the light period, reached its peak

---

**Fig. 5** Subcellular localization of LeTHIC–GFP fusion protein in Arabidopsis protoplasts. The construct of the full-length encoding sequence of LeTHIC fused with GFP was transformed and transiently expressed in Arabidopsis protoplasts. All images were made with a confocal scanner (Zeiss LSM 710 NLO). GFP fluorescence (green) was excited with an argon laser (488 nm) and detected at 515–520 nm. Chl autofluorescence (red) was detected simultaneously at 650–670 nm. The sizes of the scale bars are shown directly in the images. (A) Expression of the pJIT163-hGFP vector as a control. (B–D) Expression of the LeTHIC–GFP fusion protein.

---

**Tomato LeTHIC is an Fe-requiring HMP-P synthase**
Fig. 6 (A) LeTHIC has three transcript variants, named L, M and S. The 5'-UTRs, exons, introns and 3'-UTRs are represented by boxes shaded green, black, white and blue, respectively. The TPP-riboswitch located in the 3'-UTR of L is indicated by a 2-D structure. (B) Genomic DNA sequence contexts of TPP-riboswitches in LeTHIC, which is the same as the statement of Wachter et al. (2007). TAA marked with a black box is the stop codon of the LeTHIC ORF; GT and AG labeled by black boxes are the splice sites of the type M transcript variant. The 3'-UTR of the type S transcript variant is displayed in blue letters; the aptamer sequence is highlighted by a yellow line. (C) Comparison of the thiamine-riboswitch from LeTHIC with its homologs. Leth1 is thiamine-riboswitch identified in the 3'-UTR of LeTHIC. Ath1 and Eco1 are well documented thiamine-riboswitches identified in the AtTHIC (Wachter et al. 2007) and E. coli THICEFSGH operon (Sudarsan et al. 2003). Arrows and a circle above the alignments represent structures of base-paired stems and a stem-loop, respectively. Sequences that form stems are highly conserved (shaded gray).
at 18:00 h (9 h after the beginning of the light period), and then reduced gradually to the lowest level at 09:00 h (at the end of the dark period) (Fig. 8B). These findings clearly indicated that the abundance of LeTHIC expression was dependent on light.

**LeTHIC is an Fe-requiring protein for thiamine synthesis**

Eleven THIC protein sequences (see Fig. 4A) that represented Bacteria, Archaea and Eukarya were aligned. A consensus CX,CX,CXC motif for coordinating 4Fe–4S clusters was identified and located in the C-terminus of these THICs (Fig. 9A). According to the research of Raschke et al. (2007), an N-terminally truncated peptide lacking the first 71 amino acids of LeTHIC was produced in *E. coli* with a C-terminal glutathione S-transferase (GST) tag. The purified precipitate showed the characteristic oxygen sensitivity of Fe–S cluster proteins, it quickly turned brown upon exposure to air, and showed an absorption maximum at 410 nm and a shoulder at 320 nm (Fig. 9B). Thus, LeTHIC, like E.cTHIC and AtTHIC (Raschke et al. 2007, Chatterjee et al. 2008, Martinez-Gomez and Down 2008), was an Fe–S cluster protein.

Further, a series of physiological tests were performed to investigate the impact of Fe on thiamine synthesis in vivo. Seedlings of tomato were grown in full Hoagland solution (Hoagland and Arnon 1950) with different concentrations of macro- and microelements (Supplementary Table S2) for 2 weeks, and then the leaves were harvested for determination of their thiamine content. As shown in Fig. 9C, Fe deficiency resulted in a significant fall in the thiamine content of leaves (21 \( \mu \)g 100 g FW\(^{-1} \) at 1 \( \mu \)M Fe, 17 \( \mu \)g 100 g FW\(^{-1} \) at 0 \( \mu \)M Fe) compared with the control (38.7 \( \mu \)g 100 g FW\(^{-1} \) of thiamine). At 10 \( \mu \)M Fe which was considered as a sufficient concentration for plant growth, the thiamine content of plants was measured and found to be at the same level as the control. To test whether the decreased thiamine content of plants grown in the solution with a low concentration (1 \( \mu \)M) and without Fe supply is a specific response to Fe deficiency or a general stress reaction, the thiamine contents were also determined in plants grown in the culture solution without Zn or Mn supply. The results revealed that Zn- or Mn-deficiency stress did not affect the synthesis of thiamine significantly; and the same thiamine level was measured as in the control (Fig. 9C). Using the same samples, the riboflavin content was measured as a control and found to be independent of the Fe concentration (Supplementary Fig. S3). These results clearly reveal that the low thiamine content in leaves of plants grown under the conditions with 1 \( \mu \)M or without Fe supply is a specific reaction to Fe deficiency, indicating that Fe is required for thiamine synthesis in plants.

As shown above, LeTHIC is an Fe–S protein involved in the synthesis of the pyrimidine moiety of thiamine. To test whether the reduced thiamine synthesis in the plants grown under Fe deficiency conditions is due to an affect on LeTHIC expression...
or other reasons, LeTHIC expression was analyzed at the transcriptional and protein levels with plants grown under conditions with different Fe supply via Northern and Western blot analysis. The expression intensity of LeTHIC at both the transcriptional (S form mRNA) and protein levels was clearly increased under Fe deficiency (1 or 0 μM Fe) compared with Fe sufficiency (10 and 100 μM Fe), whereas no effect on LeTHIC expression were observed under deficiency of Zn or Mn (Fig. 9D, E).

Discussion

In contrast to convenient genetic and biochemical assays in microorganisms, there are more difficulties in characterizing the metabolic and catabolic pathways in higher plants due to lack of selectable markers. Thus, stable auxotrophs are instrumental in identifying the components of various metabolic and catabolic pathways. Therefore, tl mutants are valuable in exploring the biosynthetic pathway of thiamine in plants. In this work, we characterized the tl mutant of tomato. Our results showed that the insertion of a large unknown DNA fragment in the 5' UTR of LeTHIC (encoding a HMP-P synthase) inactivated its expression in the tl mutant, consequently blocking thiamine biosynthesis. Expression of wild-type LeTHIC in tl plants complemented the mutant to wild type (Fig. 3). These findings definitively demonstrate that LeTHIC is the gene responsible for the tl mutant of tomato.

The LeTHIC product was localized in the chloroplasts of leaf cells. Expression of this protein in THIC-deficient strains of E. coli restored normal growth. Previously, E. cTHIC had been documented to catalyze the condensation of AIR and AdoMet to generate HMP-P (Chatterjee et al. 2008, Martinez-Gomez and Downs 2008), thus LeTHIC was believed to play the same role in tomato. Previous studies showed that thi1 of A. thaliana (Papini-Terzi et al. 2003) and thi1-1 and thi1-2 of maize (Belanger et al. 1995), which were responsible for thiazole ring synthesis, shared high similarity with S. cerevisiae THI4 (Nosaka 2006, Chatterjee et al. 2007). The Arabidopsis thi1 and maize Thi3 genes which encode TMPP/HMPPKase, involved in the formation of TMP and phosphorylation of HMP/HMPP, were homologous to the E. coli counterpart ThiE/ThiD (Kim et al. 1998, Ajjawi et al. 2007b, Rapala-Kozik et al. 2007). These findings suggest that plant combines a bacteria-like synthesis of the pyrimidine moiety and TMP in the chloroplast with a yeast-like synthesis of the thiazole moiety in the mitochondria and chloroplast.

Negative regulation of THIC expression mediated by a TPP-riboswitch was reported by Bocobza et al. (2007) and Wachter et al. (2007). In the established model, the TPP-riboswitch present in the 3'-UTR of the plant THIC gene controls the formation of transcripts with different 3'-UTR.
lengths through binding of TPP, which then affected mRNA accumulation and protein production (Wachter et al. 2007). The sequence analysis and experimental data of this work support this model. *LeTHIC* has three (L, M and S) mRNA variants. The M and S mRNA types are spliced from the L variant, and the M type of *LeTHIC* mRNA is an unstable type (Supplementary Fig. S2). With increased concentration or extended time of thiamine treatment, the abundance of both total *LeTHIC* mRNAs and its S type transcript decreased, but the M variant increased markedly (Fig. 7A and B, upper parts). These results indicate that more of the L variant of *LeTHIC* mRNA was processed to yield the unstable M variant under thiamine excess conditions, resulting in reduction of S variant mRNA, which is stable and translates into protein for its function.

From an evolutionary perspective, *THIC* genes are functionally conserved. Like its bacterial and plant homologs, *E.cTHIC* and *AtTHIC* (Raschke et al. 2007, Chatterjee et al. 2008, Martinez-Gomez and Downs 2008), *LeTHIC* contains a conserved CX2CX4C motif in the C-terminus, and the protein precipitate displayed oxygen sensitivity (a typical feature of Fe–S cluster proteins). We also showed that thiamine synthesis in tomato is dependent on the Fe status because the plants grown under Fe deficiency conditions revealed a significantly lower
thiamine content than that of the control, but this was not observed in the plants grown under the conditions with Zn or Mn deficiency (Fig. 9C). As a control, we also measured riboflavin content, and showed that the synthesis of riboflavin was not affected by Fe deficiency (Supplementary Fig. S3). Based on the results, we conclude that Fe is required for thiamine synthesis. Expression analysis revealed that the mRNA and protein product of LeTHIC accumulated significantly under Fe-deficient conditions (Fig. 9D, E). The up-regulated expression of LeTHIC at the mRNA and protein levels with decreased thiamine content of leaves under Fe-deficient conditions can be explained through a feedback regulation mechanism, i.e. Fe is required to maintain LeTHIC activity, deficiency of Fe results in reduction of its enzyme activity and thiamine content, which, in turn, enhanced the expression of LeTHIC. This is the first time that the impact of Fe on plant thiamine synthesis has been observed and quantified through in vivo experimentation. In E. coli, iron starvation diminishing the enzyme activity of an Fe–S cluster-containing protein, 6-phosphogluconate dehydratase, has already been reported (Outten et al. 2004).

In this study, we further investigated the effects of light on LeTHIC expression. At the mRNA level, the abundance of LeTHIC transcription was enhanced by light. The protein accumulation of LeTHIC is well correlated with the abundance of S type mRNA of LeTHIC with an exception at the time point of 21:00 h. The protein accumulation at this time point is a little less than that at the time point of 18:00 h, but the abundance of S type mRNA is obviously higher at 21:00 h than at 18:00 h (Fig. 8). The miscorrelation can be explained due to activation of the feedback regulation mechanism of thiamine synthesis resulting in the reduction of LeTHIC accumulation.

In conclusion, LeTHIC encoding a pyrimidine synthase (an Fe–S cluster protein) acts as a key gene in tomato thiamine synthesis. Inactive transcription of LeTHIC via insertion of an unknown DNA fragment in the 5′-UTR region in the tl mutant resulted in defects in thiamine synthesis and blocked plant growth and development. The expression of LeTHIC was subject to multiple regulatory mechanisms, including light regulation on a transcriptional level, negative regulation by TPP on the post-transcriptional level, and both expression and enzyme functional regulation by Fe status. The knowledge from this research provides new insight for understanding the biosynthetic pathway of thiamine, as well as shedding light on the design of plants engineered for elevating thiamine levels to improve human and animal nutrition.

Materials and Methods

**Genetic materials and growth conditions**

Seeds of the tomato line LA0758+/− (heterozygous at the tl locus) were obtained from the Tomato Genetics Stock Center in Davis (CA, USA). Seedlings were grown under standard greenhouse conditions (25°C day/21°C night) with additional artificial light (12 h photoperiod). The tl mutant, a lethal mutant under normal culture condition, was maintained via the heterozygotes (LA0758+/−). Thus, the tl mutant (LA0758−/−) and wild type (LA0758+/+) were isolated by selfing of LA0758 (+/−). THIC-deficient E. coli strains THIC3178::Tn10Kan (Nichols et al. 1998) and THIC765::Kan (Baba et al. 2006) were kindly provided by the E. coli Stock Center at Yale University.

**Root measurements**

The wild type (LA0758+/+) and tl mutant (LA0758−/−) plants grew in full Hoagland solution (Supplementary Table S2) for 1 month. Subsequently, the roots were washed, screened by an Epson Expression 10000XL scanner at 300 dpi, and analyzed for total root length, average root volume and the number of root tips with WINRhizo software (Regent Instruments) (Watt et al. 2005, Watt et al. 2008).

**Genetic mapping and isolation of the candidate gene**

To map the tl locus, a segregating F2 population was constructed by crossing LA0758−/− with L. pennellii (a wild tomato species). The leaves of 3-week-old F2 seedlings were harvested and the genomic DNA was extracted. Genetic analysis and fine mapping of the tl locus were performed as described by Ling et al. (1996) using the molecular markers from the map of Tomato-EXPEN 2000 (SGN; http://www.sgn.cornell.edu/). Map distances in centiMorgans were calculated according to Kosambi (1944).

For isolation of the candidate gene, the genomic BAC library of tomato (cultivar Moneymaker) was constructed by partial digestion with HindIII and subcloned into a pIndigoBAC-5 vector (EPICENTRE Biotechnologies). A PCR-based library screening (Cheng et al. 2004) was performed with LeTHIC-T primers (Supplementary Table S1). The BAC clone was sequenced and subjected to gene prediction in SoftBerry (http://www.softberry.com/all.htm). A leaf cDNA library of tomato cultivar Moneymaker was constructed with the SMART cDNA Library Construction Kit (Clontech) and screened using the same method and primers as for the BAC library screening for isolation of LeTHIC transcripts. The positive colonies were isolated and their cDNAs were sequenced. The complete sequences of the 3′-UTR of clones were obtained by 3′ RACE analysis (Full RACE Core Set, TAKARA) with Race3 primers (Supplementary Table S1).

**Complementation assays**

The full-length coding region of LeTHIC was amplified from a cDNA clone of LeTHIC by PCR with LeTHIC-CA primers (Supplementary Table S1), and subcloned into the binary vector pBINPLUS (van Engelen et al. 1995) at the KpnI and SalI restriction enzyme sites under the control of the 35S promoter. The construct was then introduced into the strain GV3101 of A. tumefaciens. Transformation and regeneration of the tl mutant were done as described by Ling et al. (1998).
The regenerated plants which contained the transgene (confirmed by PCR and Southern blot analysis) were grown in a greenhouse and used for further analysis.

For the functional complementation assay of *E. coli* THIC-deficient strains, the coding region of *LeTHIC* was amplified with *LeTHIC*-CA primers (Supplementary Table S1) and subcloned into the bacterial expression vector pGEX-KG at the EcoRI and Xhol restriction enzyme sites under the control of the T7 promoter. After sequence construction, the construct was transformed into the *E. coli* THIC-deficient strains THIC3178::Tn10Kan (Nichols et al. 1998) and THIC765::Kan (Baba et al. 2006), and an empty vector pGEX-KG (Yam et al. 2000) was transformed as a negative control. Complementation of the auxotrophy was subsequently assayed by cultivation of the transformed cells at 37°C on an agar plate with M9 minimal medium containing 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG), 100 μg ml⁻¹ ampicillin and 100 μg ml⁻¹ kanamycin with or without 10 μM thiamine.

**Measurement of thiamine content**

Fresh leaves (8 g per sample) were harvested and sent to the Analysis Department of the Beijing Research Institute for Nutritional Resources (Beijing, China) to determine the thiamine content. The extraction and measurement of total thiamine were done following the protocol described by Kong et al. (2008). Briefly, 8 g (FW) of leaves were homogenized in 0.1 M HCl and digested by boiling for 30 min in an autoclave at 121°C. Subsequently, protein and starch in the solution were removed by treatment with protease and amylase at 37°C overnight. After purification by filtration, thiamine in the solution (10 ml) was oxidized to thiochrome by adding 5 ml of the reaction mixture [0.25 g l⁻¹ K₂Fe(CN)₆ and 10 g l⁻¹ NaOH], and then separated on a Diamonsil TMC18 column (5 μm, 250 mm × 4.6 mm) (Dikma) using LC-10ATvp HPLC (Shimadzu) and measured using a RF-10AXL fluorescence detector (excitation 375 nm, emission 435 nm) according to the manufacturer's instructions. The thiamine content of plants was calculated based on a standard curve of thiamine.

**Sequence analysis**

The homologous sequence search was performed by Blast programs on the NCBI database and the SGN unigene database on the SGN web site (the threshold value was designed as an E-value cut-off <1e⁻10). The intron–exon origination of *LeTHIC* was determined by Spidey (NCBI). Sequence identity and similarity were calculated by the Sequence Manipulation Suite (SMS) (http://bio.dfci.harvard.edu/Tools/SMS/). The riboswitch was predicted by RibEx (http://www.ibt.unam.mx/biocomputo/ribex.html). Sequence alignment was performed with PROBCONS (http://probcns.stanford.edu/about.html) and the phylogenetic tree was constructed with the program MEGA3.0 (Kumar et al. 2004). Further, the alignment file was used to generate the sequence logo by WebLogo (http://www.bio.cam.ac.uk/cgi-bin/seqlogo/logo.cgi) (Crooks et al. 2004). Targeting signal prediction of subcellular protein localization was performed with iPSORT in http://hc.ims.u-tokyo.ac.jp/ipSORT/index.html.

**Southern and Northern blot analysis**

Genomic DNA extraction was performed as described (Fulton et al. 1995). For Southern blot analysis, approximately 10 μg of genomic DNA was separately digested with HindIII and NdeI. Gel separation, blotting of DNA fragments onto Hybond-N⁺ membrane (Amersham) and hybridization were performed as described by Ling et al. (2002). The probes were synthesized by PCR amplification with *LeTHIC*-SB primers (Supplementary Table S1), and labeled with [³²P]dCTP by Prime-a-Gene System (Promega).

For Northern analysis, total RNA isolation and purification were carried out as described by Li et al. (2004). Northern blot analysis was performed following the protocol of Yuan et al. (2008). About 15 μg of total RNAs were loaded per lane; the coding region of *LeTHIC* was amplified by PCR with *LeTHIC*-T primers (Supplementary Table S1) and used as a probe.

**Immunoblot analysis of LeTHIC**

The coding sequence of *LeTHIC* (72–651 amino acids) was amplified by *LeTHIC*-WB primers (Supplementary Table S1) and subcloned into the bacterial expression vector pGEX-4T-1 (Amersham Biosciences) at the EcoRI and SalI restriction enzyme sites. It was expressed in bacteria and purified as an antigen. The polyclonal antibody against the peptide (72–651 amino acids) of *LeTHIC* was generated at the Animal Center of the Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, and used for the immunoblot analysis. Total proteins of tomato were extracted from leaves following the protocol described by Connolly et al. (2003). The proteins (20 μg per sample) were separated by SDS–PAGE and transferred to polyvinylidene fluoride membranes by electroblotting. The procedure for Western blot analysis was that of Yuan et al. (2008).

**RT–PCR analysis**

The total mRNA extracted from different tomato seedling tissues was treated with DNase 1 (Promega) to eliminate genomic DNA, and then 2 μg per sample was used for reverse transcription with an M-MLV RT kit (Promega) according to the manufacturer’s instructions. The reactions were diluted 6-fold for RT–PCR analysis as described by Li et al. (2004). PCR conditions were 30 s at 94°C, 30 s at 55°C, 60 s at 72°C, 24 cycles. The elongation factor *LeEF* was used as a positive control.

**Multiplex RT–PCR analysis**

The expression patterns of *LeTHIC* were assayed using the GenomeLab GeXP Analysis System Multiplex (Beckman) (Yuan et al. 2008). Primer design and multiplex optimization was performed using the GeXP Express Profiler and primer design module (Beckman; http://www.beckman.com).
The primer pairs utilized in multiplex RT–PCR assay are listed in Supplementary Table S1. Each reverse primer is chimeric, with the 5’ end containing a 19 nucleotide universal priming sequence (GTAGCAGCTCAGTATAGGA) and the 3’ end containing the gene-specific sequence. Each forward primer is chimeric, with the 5’ end containing a second 20 nucleotide universal forward priming sequence (AGGGTACCTAGTAAATA) and the 3’ end containing the gene-specific sequence. The expression patterns of multiple genes were examined using the GenomeLab GeXP Analysis System Multiplex RT–PCR assay (Beckman) according to the protocol described by Chen et al. (2007). The fragment analysis was performed with a module of the GeXP Genetic Analysis System, which linked the gene expression profile to the corresponding peak area. The gene expression data acquired from multiplex RT–PCR were normalized by dividing the peak area of each gene by the peak area of the housekeeping gene LeEF.

**Subcellular localization analysis**

The construct expressing the fusion protein LeTHIC–GFP was generated by subcloning the LeTHIC coding sequence amplified with LeTHIC-SL primers (Supplementary Table S1) into the plant expression vector pIT163-hGFP tagged with GFP. The construct and the empty vector pIT163-hGFP (as negative control) were transiently expressed in Arabidopsis mesophyll protoplasts. The Arabidopsis protoplast preparation and transformation were done as described by Shah et al. (2002). Microscopic analysis of the transformed cells was carried out with a confocal scanner (Zeiss LSM 710 NLO). The generation of transmission micrographs for visualization of non-fluorescent protoplast structures was achieved using the manufacturer’s filter settings. GFP fluorescence was excited with an argon laser (488 nm) and detected at 515–520 nm. Chl autofluorescence (red) was detected simultaneously at 650–670 nm.

**Supplementary data**

Supplementary data are available at PCP online.

**Funding**

This work was supported by the Ministry of Science and Technology of China [grant No. 2011CB100304]; the Chinese Academy of Sciences [grant No. KSCX2-YW-N-056]; the HarvestPlus Challenge Program [Agreement #8236] and the Ministry of Agriculture of China for transgenic research [grant No. 2008ZX08009-003].

**References**


Haagland, D.R. and Arnon, D.J. (1950) The water-culture method for growing plants without soil. In California Agricultural Experimental...
Tomato LeTHIC is an Fe-requiring HMP-P synthase

Plant Mol. Biol. 37: 955–966.


