Cell Research (2008) 18:566-576.
© 2008 IBCB, SIBS, CAS All rights reserved 1001-0602/08 $ 30.00
www.nature.com/cr

npg

ORIGINAL ARTICLE

AtTHIC, a gene involved in thiamine biosynthesis in *Arabidopsis thaliana*

Danyu Kong[1, 2, *], Yuxing Zhu[1, 2, *], Huilan Wu[1], Xudong Cheng[1, 2], Hui Liang[1], Hong-Qing Ling[1]

[1]*The State Key Laboratory of Plant Cell and Chromosome Engineering, National Center for Plant Gene Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Datun Road, Chaoyang District, Beijing 100101, China;* [2]*Graduate University of the Chinese Academy of Sciences, 19 Yuquan Road, Shijingshan District, Beijing 100049, China*

Thiamine (vitamin B_1) is an essential compound for organisms. It contains a pyrimidine ring structure and a thiazole ring structure. These two moieties of thiamine are synthesized independently and then coupled together. Here we report the molecular characterization of AtTHIC, which is involved in thiamine biosynthesis in *Arabidopsis*. AtTHIC is similar to *Escherichia coli* ThiC, which is involved in pyrimidine biosynthesis in prokaryotes. Heterologous expression of *AtTHIC* could functionally complement the *thiC* knock-out mutant of *E. coli*. Downregulation of *AtTHIC* expression by T-DNA insertion at its promoter region resulted in a drastic reduction of thiamine content in plants and the knock-down mutant *thic1* showed albino (white leaves) and lethal phenotypes under the normal culture conditions. The *thic1* mutant could be rescued by supplementation of thiamine and its defect functions could be complemented by expression of *AtTHIC* cDNA. Transient expression analysis revealed that the AtTHIC protein targets plastids and chloroplasts. *AtTHIC* was strongly expressed in leaves, flowers and siliques and the transcription of *AtTHIC* was downregulated by extrinsic thiamine. In conclusion, *AtTHIC* is a gene involved in pyrimidine synthesis in the thiamine biosynthesis pathway of *Arabidopsis*, and our results provide some new clues for elucidating the pathway of thiamine biosynthesis in plants.

***Keywords*:** *Arabidopsis, E. coli* ThiC, AtTHIC, thiamine, pyrimidine biosynthesis, vitamin B_1
Cell Research (2008) **18**:566-576. doi: 10.1038/cr.2008.35; published online 11 March 2008

Introduction

Thiamine (vitamin B_1) is an essential substance for organisms. Chemically, it consists of a pyrimidine (4-amino-5-hydroxymethylpyrimidine) and a thiazole (4-methyl-5-β-hydroxyethylthiazole) moiety. Free thiamine and its phosphoesters can interconvert each other in cells. The phosphoester thiamine pyrophosphate (TPP) functions as a coenzyme of some important enzymes in cellular metabolic pathways such as the citric acid cycle, glycolysis and the pentose phosphate cycle [1]. Most microorganisms and plants can utilize extrinsic thiamine efficiently to produce TPP and can synthesize thiamine *de novo*, but animals can accomplish thiamine pyrophosphorylation only. Therefore, thiamine is a dietary necessity for animals and human beings. In countries where rice is a staple food, thiamine deficiency is prevalent because the process of polishing rice hulls removes most thiamine from rice grains. Thiamine deficiency disturbs the central nervous and circulatory systems and causes beriberi disease [2, 3].

Extensive genetic and biochemical studies in microorganisms have revealed some of the steps involved in thiamine biosynthesis, especially in *Escherichia coli, Bacillus subtilis* and *Saccharomyces cerevisiae*. The thiazole and pyrimidine moieties of thiamine are synthesized independently. Then they are coupled to form thiamine phosphate (TP) [1, 4]. In *E. coli*, the thiazole moiety is formed from DXP (1-deoxy-D-xylulose-5-phosphate), tyrosine and cysteine [5-9]. Five gene products (ThiF, ThiS, ThiG, ThiH and ThiI) are involved in this step [10-14] (Figure 1). The pyrimidine unit is synthesized from AIR (5-aminoimidazole ribotide), and ThiC is re-

*These two authors contributed equally to this work.
Correspondence: Hong-Qing Ling
Tel/Fax: +86-10-64860377
E-mail: hqling@genetics.ac.cn
Received 22 November 2007; revised 14 January 2008; accepted 15 January 2008; published online 11 March 2008

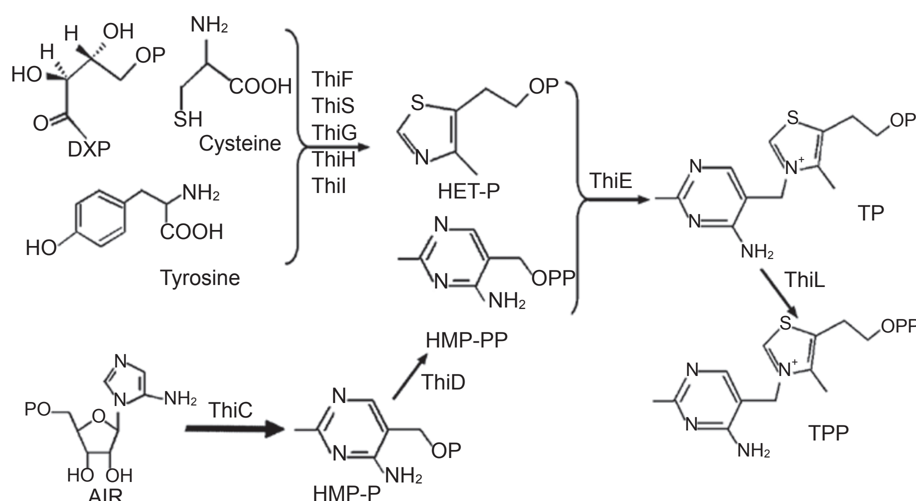

Figure 1 Outline of the biosynthesis of thiamine pyrophosphate (TPP) in *E. coli*. ThiF, ThiS, ThiG, ThiH and ThiI are involved in thiazole synthesis, whereas ThiC catalyses the synthesis of pyrimidine. DXP, 1-deoxy-D-xylulose-5-phosphate; AIR, 5-aminoimidazole ribotide; TP, thiamine phosphate; TPP, thiamine pyrophosphate; HMP, 4-amino-5-hydroxymethylpyrimidine; HET, 4-methyl-5-β-hudroxyethylthiazole.

quired for pyrimidine biosynthesis [14-17] (Figure 1). Subsequently, ThiE, a thiamine phosphate synthase, couples the two ring structures of pyrimidine and thiazole to form TP [9, 17-20] (Figure 1). Further, ThiL catalyses TP to form TPP [10, 21]. In yeast, the mechanism of thiamine biosynthesis is different from prokaryotes in the precursors and the related genes [3]. Pyrimidine biosynthesis is more complicated in yeast than in prokaryotes. Four genes (*THI5, THI11, THI12* and *THI13*) have been reported to be involved in this step [22, 23].

In plants, the biosynthesis of thiamine is still not clear although some mutants defective in thiamine biosynthesis have been reported [24-26]. In *Arabidopsis*, three types of thiamine-auxotroph mutants have been identified [24, 27]. The *py* mutant requires pyrimidine or thiamine supplementation for growth [24], and the *tz* mutant needs thiazole or thiamine for growth, whereas the mutants of *th-1, th-2* and *th-3* can be restored to normal growth only through supplementation of thiamine [27]. This suggests that plants may utilize a biosynthetic pathway that is similar to that of microorganisms. Two genes involved in thiamine biosynthesis have so far been cloned and characterized in *Arabidopsis* [28-30]. The first one is *THI1*, which is involved in thiazole synthesis. *THI1* was isolated by complementation of the *E.coli* mutant strain BW535, a mutant with defective DNA base excision repair pathways [28]. Further characterization showed that THI1 has two functions: it is involved in the thiazole synthesis of the thiamine biosynthesis pathway and in DNA repair [28, 29]. Also, THI1 possesses a signal peptide at its N terminus, which targets the THI1 protein to mitochondria and chloroplasts [31, 32]. The second is the *Arabidopsis thaliana HMPPK/TMPPase* gene (At1g22940) [30, 33]. The HMPPK/TMPPase protein is localized in the chloroplast compartment and has two enzymatic activities — namely as a TMP-PPase at the C terminus and a HMP-P kinase at the N terminus [30, 33]. The arrangement of the two functional domains in one protein results in the coordinated synthesis of the two enzymes for thiamine biosynthesis. More recently, Raschke *et al*. [34] described that *Arabidopsis* THIC is an iron-sulfur cluster protein and is involved in thiamine biosynthesis.

Here, we report the genetic and molecular characteristics of *AtTHIC*, which is required for thiamine biosynthesis and is putatively involved in the synthesis of the pyrimidine moiety of thiamine in *A. thaliana*. The results presented in this work provide an important clue for understanding the biosynthetic pathway of thiamine in plants.

Results

Isolation of AtTHIC

Early reports showed that *E. coli* ThiC is involved in pyrimidine biosynthesis. ThiC (which is also known as ThiA, a homolog of ThiC in *B. subtilis*) represents a unique gene product that has so far been identified in the synthesis of pyrimidine in prokaryotes. Sequence similarity searches using a BLAST algorithm in the *Arabidopsis* database revealed that the AGI code At2g29630 showed 60% similarity to *E. coli* ThiC at the amino-acid

level. Therefore, we have named At2g29630 as *AtTHIC*. It encodes a protein that contains 644 amino acids and possesses a conserved domain from amino-acid sequence 164 to 587, which is known as the THIC domain in the NCBI (National Center for Biotechnology Information) database, based on its ubiquity in ThiC proteins across different species (Figure 2). Phylogenic analysis using the known ThiC proteins from microorganisms (*Azotobacter vinelandii* ThiC, *B. subtilis* ThiC, *E. coli* ThiC and *Salmonella typhimurium* ThiC) and the putative ThiC proteins from plants (AAG49550 from *Poa secunda*, NP001050897 from rice, and NC008467 and NC008475 from black cottonwood) on the NCBI database, as well as using rice Os12g0270200 (which shows some sequence similarity to AtTHIC), revealed that AtTHIC of *A. thaliana* is more similar to the putative THICs from plants than the ThiCs from microorganisms. They were separately clustered into two different groups. Although rice Os12g0270200 shows some sequence similarity to AtTHIC, it was not clustered into the ThiC group (Figure 3).

Functional complementation of the thiC mutant of E. coli with AtTHIC

Considering the high sequence similarity to *E. coli* ThiC at the protein level, AtTHIC may have similar functions to *E. coli* ThiC in the thiamine biosynthesis of *A. thaliana*. To verify this hypothesis, we isolated the cod-

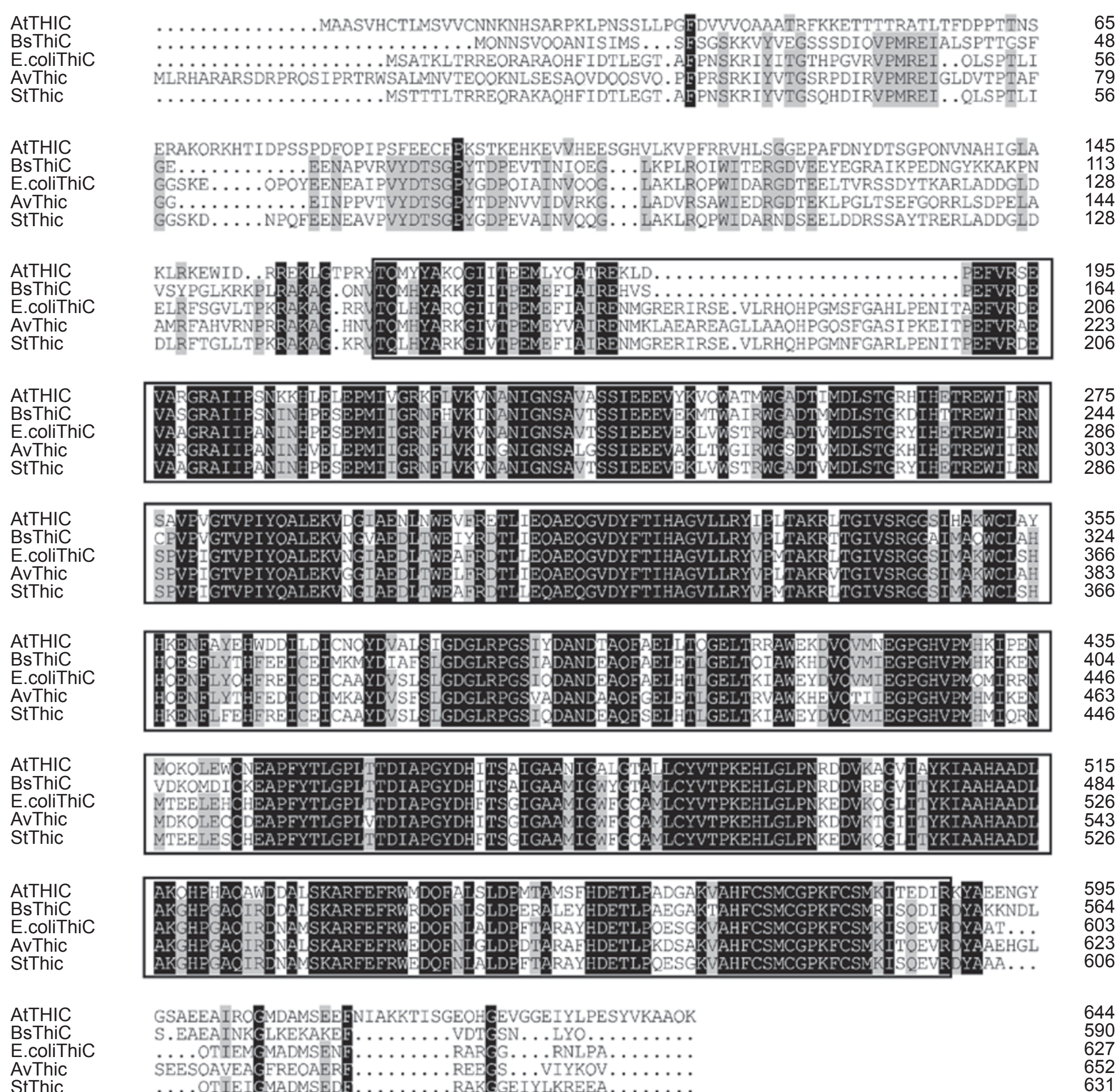

Figure 2 The alignment of AtTHIC with the known ThiCs from microorganisms. The sequence alignment was performed using the ClustalW program. The boxed sequences are the conserved THIC domain. *Av, Azotobacter vinelandii; Bs, Bacillus subtilis; E. coli, Escherichia coli; St, Salmonella typhimurium*.

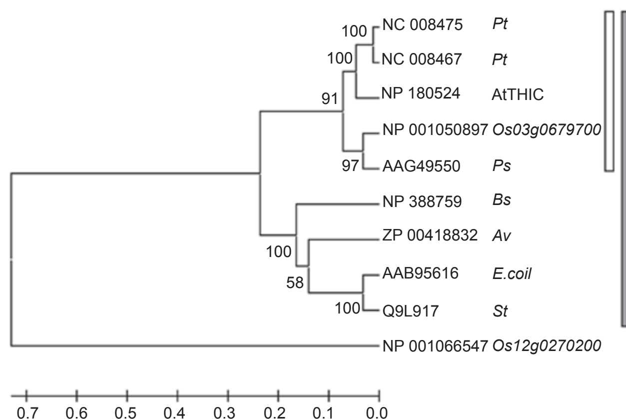

Figure 3 Dendrogram showing amino-acid sequence similarity among the well characterized THICs from microorganisms and their homologs found by a BLAST search of the NCBI database. Multiple sequence alignment was performed using the ClustalW program (http://www.ebi.ac.uk/clustalw/) and transformed into a dendrogram using MEGA version 4 (http://www.megasoftware.net/). The gray box labels the proteins that contain the THIC domain, the white box labels the putative THICs of plants. *At, Arabidopsis thaliana; Av, Azotobacter vinelandii; Bs, Bacillus subtilis; E. coli, Escherichia coli; Os, Oryza sativa; Ps, Poa secunda; Pt, Populus trichocarpa; Sc, Saccharomyces cerevisiae; St, Salmonella typhimurium*. NC 008467 and NC 008475 are the linkage group numbers of the genome sequences of *Populus trichocarpa*.

ing sequence of *AtTHIC* using RT-PCR, cloned into an expression vector of *E. coli* and introduced into the *thiC* knock-out mutant strain KG6953 of *E. coli* for functional complementation. The results are presented in Figure 4. On minimal medium without thiamine supply, the *thiC* mutant of *E. coli* expressing *AtTHIC* grew as well as the wild type (Figure 4A, left panel), whereas the mutant strain transformed with the plasmid without *AtTHIC* barely grew (Figure 4A, right panel). On the minimal medium supplemented with thiamine, both strains grew well (Figure 4B). The results reveal that the expression of *AtTHIC* can functionally complement the *thiC* mutant of *E. coli*, suggesting that *AtTHIC* should have a similar function to *thiC* in pyrimidine biosynthesis.

Characterization and functional complementation of thic1

To characterize the biological functions of AtTHIC, we obtained a T-DNA insertion mutant of *AtTHIC* (Salk_011114, hereafter called *thic1*) from the *Arabidopsis* Biological Resource Center (ABRC). According to ABRC's description, the T-DNA was inserted 403 bp upstream of the predicted translation start codon. To confirm this, we sequenced the whole region and found that the T-DNA was actually inserted 474 bp upstream of the first ATG of *AtTHIC* (Figure 5A). The homozygous insertion mutant of *AtTHIC* was selected and termed *thic1*. The *thic1* mutant exhibited white leaves (Figure 6A, right) and died at an early growth stage under normal culture conditions, whereas it grew in a normal manner when thiamine was supplied in the culture medium (Figure 6B). This indicates that *thic1* is a thiamine-auxotroph mutant. Northern blot analysis showed that the *AtTHIC* transcript was not detected in *thic1* (Figure 5B). To further confirm whether *AtTHIC* transcription in the *thic1* mutant is downregulated or knocked out, more sensitive analysis was performed using RT-PCR. Weak *AtTHIC* transcription activity was detected when the PCR reaction with 30 cycles (Figure 5C). These results indicate that *AtTHIC* transcription in the *thic1* mutant is downregulated. Furthermore, we cloned the *AtTHIC* coding sequence into a T-plasmid under the control of the CaMV35S promoter (hereafter termed 35S), and introduced this into the genome of the *thic1* mutant through *Agrobaterium*-mediated transformation. The transgenic *thic1* plants showed normal growth on the thiamine-free MS plate and in soil (Figure 6C and 6D). *AtTHIC* expression in the transgenic lines was also confirmed using northern blot hybridization (Figure 7A, lanes 3-5). These results demonstrate that the downregulation of *AtTHIC* in

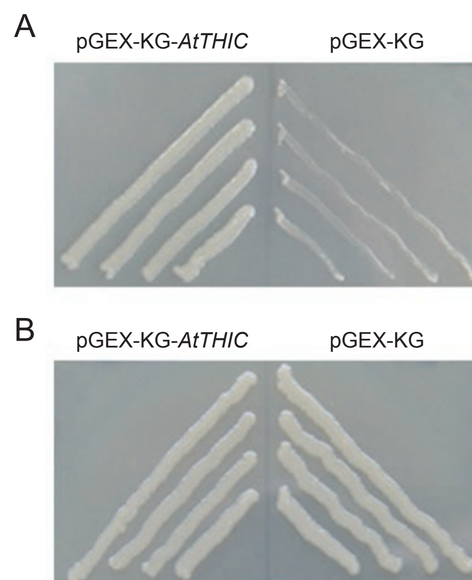

Figure 4 Functional complementation of the *E. coli thiC* mutant strain KG6953 with *AtTHIC*. Cells were grown on minimal medium (M9) plates without thiamine **(A)** and with 50 mg/L thiamine **(B)**. The mutant strain transformed with the plasmid pGEX-KG was used as a negative control. The cells were grown on M9 plates at 37°C for 24 h.

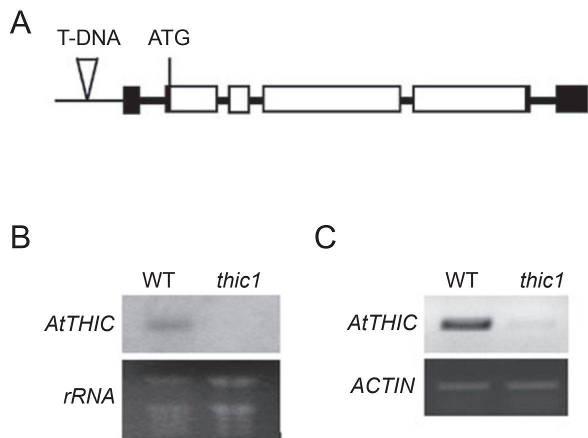

Figure 5 Molecular characterization of the *thic1* mutant. **(A)** Scheme of the T-DNA integration site in *thic1*. The triangle shows the T-DNA insertion position (at 474 bp upstream of the start codon ATG). The thick lines depict introns, boxes mean extrons (white boxes indicate the coding region, and the black boxes are the 5′- and 3′-UTRs) and the thin line depicts the promoter of *AtTHIC*. **(B)** Northern blot analysis of *AtTHIC* expression in wild-type and *thic1* shoots. **(C)** RT-PCR analysis of *AtTHIC* expression in the wild type and the *thic1* mutant.

thic1 is responsible for the albino and lethal phenotypes under normal culture conditions.

Bearing in mind that an exogenous supply of thiamine can rescue *thic1* allowing normal growth, AtTHIC might be involved in thiamine synthesis. To confirm this, we measured the total thiamine content in the *thic1* mutant, the wild type and the complement lines (*thic1:AtTHIC*) grown on a thiamine-free medium using high-pressure liquid chromatography (HPLC) analysis. Consistent with these phenotypes, the thiamine content in the *thic1* mutant was significantly lower than in the wild type (approximately one-third of the wild type), whereas the thiamine content in the *thic1:AtTHIC* lines was the same as in the wild type (Figure 7B). In addition, we introduced *AtTHIC* under the control of the 35S promoter into the wild type to generate plants that overexpressed *AtTHIC* (Ox*AtTHIC*). Three homogenous Ox*AtTHIC* lines were analysed. They all overexpressed *AtTHIC* (Figure 7A, lanes 6-8), showed normal growth like the wild type, and had higher content of thiamine than the wild type (Figure 7B).

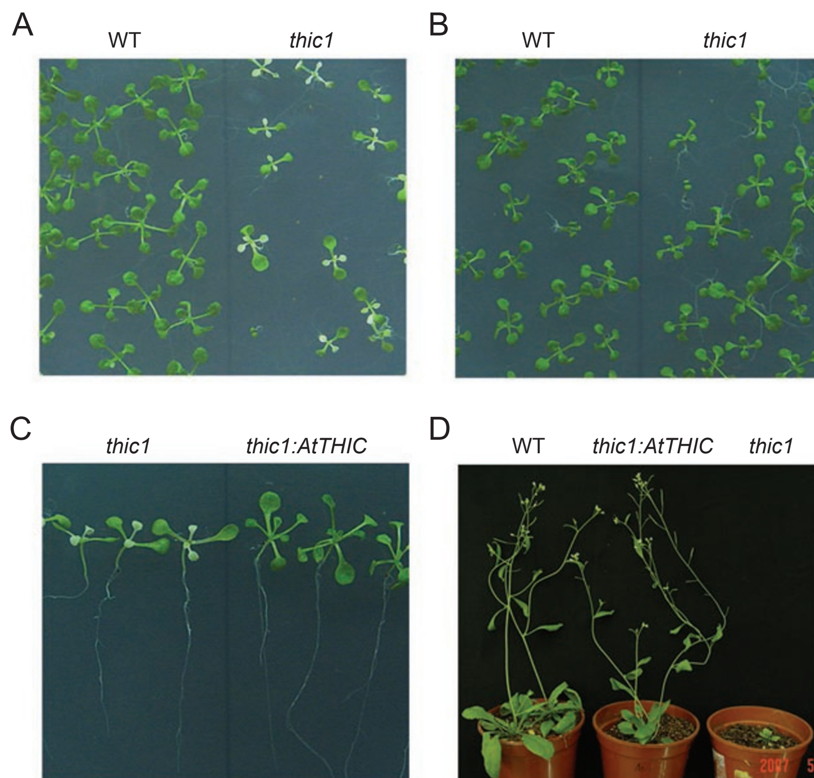

Figure 6 Phenotypic characterization of the *AtTHIC*-knockdown mutant *thic1* and the *thic1:AtTHIC* complementation lines. **(A)** Phenotype of *thic1* on MS plates without thiamine. **(B)** Phenotype of the *thic1* mutant on MS plates with 50 mg/L thiamine. **(C)** Phenotypes of the *thic1* mutant and the *thic1:AtTHIC* line on MS plates without thiamine. **(D)** Phenotypes of the *thic1* mutant, the *thic1:AtTHIC* line and the wild type in soil.

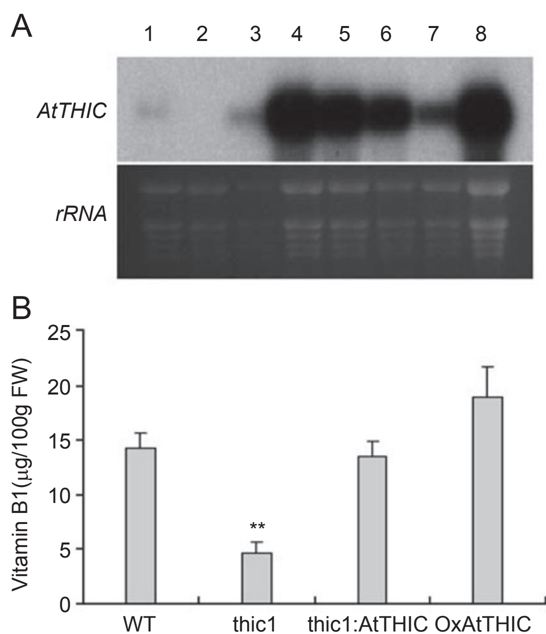

Figure 7 *AtTHIC* expression and thiamine levels in the T-DNA insertion mutant *thic1*, the complementation lines *thic1:AtTHIC* and the *AtTHIC*-overexpression lines Ox*AtTHIC*. **(A)** Northern blot analysis of *AtTHIC* expression in the wild type (lane 1), the *thic1* mutant (lane 2), the *thic1:AtTHIC* lines (*AtTHIC* expressed in the mutant background, lanes 3-5), and the Ox*AtTHIC* lines (*AtTHIC* expressed in the wild-type background, lanes 6-8). **(B)** Total thiamine contents in the wild type, the *thic1* mutant, and the *thic1:AtTHIC* and Ox*AtTHIC* lines were measured using HPLC. The data shown are the mean values of three biological replicates and the error bars represent the standard deviation. ** indicates that there is significant difference (P<0.01) in comparison with the wild type.

Subcellular localization of AtTHIC

The sequence analysis of AtTHIC by TargetP revealed that AtTHIC contains a signal peptide, implying that AtTHIC may be targeted to organelle(s). To test this hypothesis, we fused the coding sequence of *AtTHIC* in-frame to the 5′-end of *GFP* and placed it under the control of a 35S promoter to generate an AtTHIC-GFP fusion protein (GFP was fused at the C terminus of AtTHIC) expression plasmid (35S::*AtTHIC-GFP*). The plasmid was introduced into the epidermis cells of onion by bombardment for transient expression. After culture for 16-24 h, GFP fluorescence was detected using a confocal microscope (Olympus, Japan). Abundant green fluorescence spots appeared in the cells that were bombarded with the 35S::*AtTHIC-GFP* plasmid, whereas the green fluorescence was distributed across the whole cytoplasm of the cells that were transformed with the 35S::*GFP* plasmid (Figure 8A). In order to further identify in which organelle the AtTHIC-GFP protein is localized, we introduced the 35S::*AtTHIC-GFP* construct into cowpea leaf protoplasts using PEG transformation. The fluorescence signals of the AtTHIC-GFP fusion protein were observed in chloroplasts, whereas green fluorescence was detected across the whole cytoplasm of the cells that were transformed with the 35S::*GFP* plasmid (Figure 8B). To identify whether AtTHIC also targets mitochondria, we performed Mito Tracker Red staining, but no merged image of GFP and Mito Tracker Red was observed (data not shown). These results clearly show that the AtTHIC protein targets plastids and chloroplasts, but not mitochondria. This is consistent with the description that the thiamine synthesis of plants occurs in plastids [35, 36].

Expression profiles of AtTHIC

To investigate the expression patterns of *AtTHIC*, total RNAs were extracted from roots, leaves, flowers and siliques, and analysed using northern blot hybridization. *AtTHIC* was expressed strongly in leaves, flowers and siliques, but only with a trace expression in roots (Figure 9A). Raschke *et al*. [34] recently reported that the

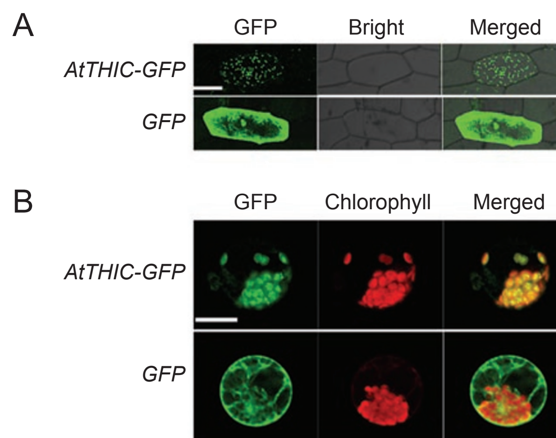

Figure 8 Transient expression analysis of *AtTHIC* in onion epidermis cells and in cowpea protoplasts. **(A)** *AtTHIC-GFP* (a fusion gene of *AtTHIC* and *GFP*) and *GFP* (control), both under the control of a 35S promoter, were transiently expressed in onion epidermis cells introduced by bombardment. The left panel shows GFP fluorescence, the middle panel shows the transmission view and the right panel shows the merged images. **(B)** *AtTHIC-GFP* and *GFP* expressed in cowpea protoplasts introduced using PEG transformation. The left panel shows GFP fluorescence, the middle panel shows chlorophyll autofluorescence and the right panel shows the merged images. The GFP signal is localized in chloroplasts in the cells that were transformed with 35S::*AtTHIC-GFP*, whereas the GFP is spread through the whole cytoplasm in the cells that were transformed with 35S::*GFP*. Scale bar = 50 μm.

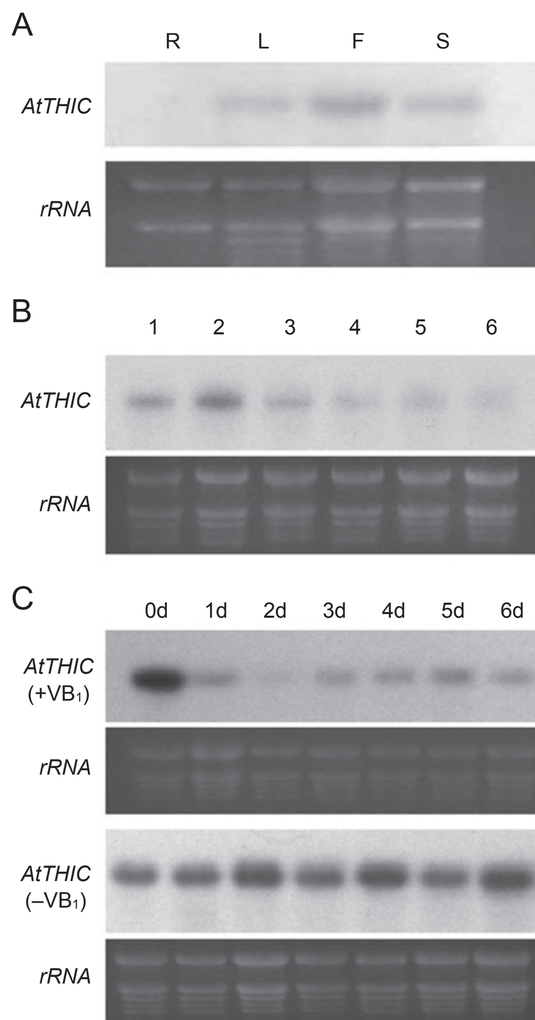

Figure 9 Expression profile analysis of *AtTHIC*. **(A)** Northern blot analysis of *AtTHIC* expression in different plant tissues. F, flowers; L, leaves; R, roots; S, siliques. **(B)** Northern blot analysis of *AtTHIC* expression under culture conditions with different concentrations of thiamine. 1 to 6 indicate thiamine concentrations of 0, 0.5, 1.0, 10, 100 and 1 000 mg/L, respectively. **(C)** A time course experiment of *AtTHIC* expression under culture conditions with 100 mg/L (upper part) and without (lower part) thiamine supplementation as analysed using northern blotting. 0 d to 6 d indicate the number of days after treatment with 100 mg/L thiamine.

expression of *AtTHIC* is regulated by light. This might explain why *AtTHIC* expression is weak in roots.

Considering that AtTHIC is involved in thiamine biosynthesis, we analysed the effect of extrinsic thiamine supplementation on *AtTHIC* expression. One-week old seedlings were transferred to MS plates supplied with different thiamine concentrations (0, 0.5, 1.0, 10, 100 and 1 000 mg/L) and were grown for 1 day. Then, the total RNAs were extracted and analysed using northern blot analysis. The results showed that the expression of *AtTHIC* was significantly repressed when more than 1 mg/L of thiamine was added to the culture medium and that the degree of repression was proportional to the concentrations of thiamine supplied (Figure 9B). Further, we performed a time course experiment to investigate the effect of thiamine supplementation on *AtTHIC* expression at different times. The seedlings were transferred on MS medium with 100 mg/L thiamine and were grown for up to 6 days. The shoot samples were collected from the first to the sixth day after the treatment and their total RNAs were analysed (Figure 9C). As described above, the expression of *AtTHIC* was clearly downregulated by the supplementation of 100 mg/L thiamine on the first day and its expression intensity was persistently kept at a low level during the whole treatment period (6 days) compared with the control (Figure 9C).

To characterize the *AtTHIC* expression patterns in plant tissues, the *AtTHIC* promoter was amplified and fused with β-glucuronidase gene (*GUS*). The transgenic lines expressing *GUS* driven by *AtTHIC* promoter were generated via *Agrobacterium*-mediated transformation using the Columbia ecotype of *A. thaliana*. The GUS activity in the transgenic lines was assayed by histochemical analysis at different developmental stages (Figure 10). At the stage of germination, the most GUS activity was at the root tips and in the jointed section between the hypocotyl and root (Figure 10A-10D). At the stage of two to three true leaves, GUS expression was observed in the leaves, cotyledons and vascular bundles of hypocotyls, and very weakly in the roots (Figure 10E). In the flowers, GUS staining was visible in the sepals, filaments and pistil tips, but not in the petals (Figure 10F and 10G). GUS expression also occurred in the siliques, with strong GUS staining in the jointed region between the silique and silique stem, and in the tips of siliques (Figure 10H). These results are consistent with the northern blot results presented above.

Discussion

In this work, we have characterized *AtTHIC*, which encodes an amino-acid sequence that is very similar to the pyrimidine synthetase ThiC in *E. coli*. It is a single-copy gene in the *Arabidopsis* genome and is localized on chromosome 2 of *Arabidopsis*. Downregulation of *AtTHIC* transcription by T-DNA insertion in the *AtTHIC* promoter region in the *thic1* mutant resulted in a significant decrease in the thiamine level (Figure 7B), and the mutant plants exhibited typical thiamine-deficient phenotypes (albino and lethal [24]), which could be restored to

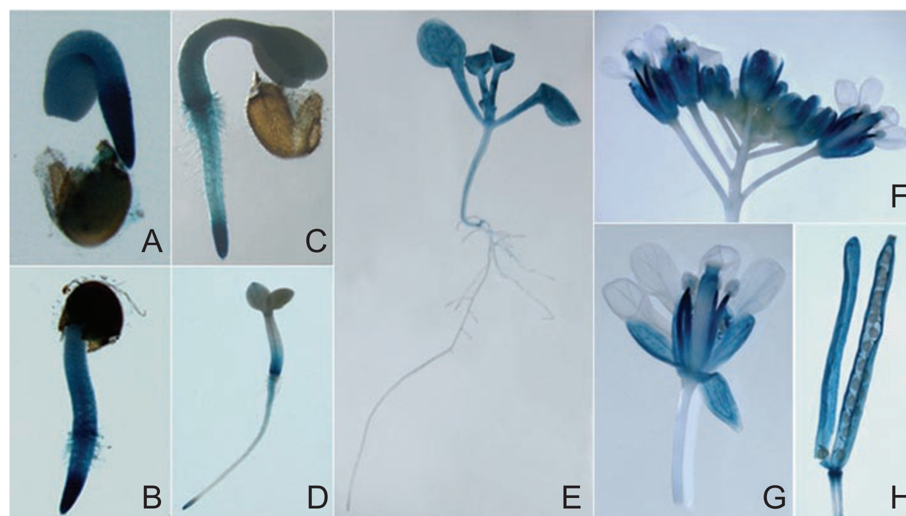

Figure 10 Histochemical analysis of GUS activity driven by the *AtTHIC* promoter in *Arabidopsis*. **(A-D)** GUS activity at the different stages of germination. **(E)** GUS activity assay at the stage of two true leaves. **(F** to **H)** GUS staining in flowers and siliques.

normal through extrinsic thiamine supplementation (Figure 6). Additionally, the *AtTHIC* overexpressing plants (Ox*AtTHIC*) had higher levels of thiamine (Figure 7B). These data support that AtTHIC is involved in the thiamine biosynthesis of *Arabidopsis*.

Thiamine is a compound containing thiazole and pyrimidine moieties. Early research has suggested that the thiamine biosynthesis of plants is similar to that of microorganisms [24, 36]. In *E. coli*, thiamine is synthesized by coupling HET-P (4-methyl-5-β-hydroxyethylthiazole phosphate) with HMP-PP (4-amino-5-hydroxymethyl-pyrimidine pyrophosphate) catalysed by ThiE (a thiamine phosphate synthase) [15, 18, 20, 21] (Figure 1). HET-P (thiazole moiety) and HMP-PP (pyrimidine moiety) are synthesized separately under the influence of different genes [24, 27]. *E.coli ThiC* is a key gene involved in the process of HMP-P synthesis from 5-aminoimidazole ribotide (AIR) to 4-amino-5-hydroxymethylpyrimidine phosphate (HMP-P). Based on the high sequence similarity (60%) and its functional complementation with *E.coli* ThiC, AtTHIC may have a similar function to *E.coli* ThiC, stimulating HMP-P synthesis in *A. thaliana*. The *AtTHIC*-knockdown mutant *thic1* can be rescued by supplying thiamine in culture solution, but not by supplying HET (4-methyl-5-β-hydroxyethylthiazole, data not shown), which further suggests that AtTHIC is involved in the pyrimidine but not the thiazole synthesis of thiamine in *A. thaliana*.

Thiamine is an essential trace nutrient for plant growth and development. The biosynthesis of thiamine should be tightly regulated in the tissues where thiamine is required. Northern blot analysis showed that *AtTHIC* was strongly expressed in leaves, flowers and siliques, but very weakly expressed in roots (Figure 9A). Furthermore, the transcription intensity of *AtTHIC* was negatively regulated by exogenous supplementation of thiamine (Figure 9B and 9C). Early research indicates that the TPP-binding riboswitches are involved in controlling thiamine biosynthesis in prokaryotes and eukaryotes [37-40]. The TPP-binding RNA domains have been found in the 5′-UTR, intron and 3′-UTR of mRNAs in different organisms [41]. A TPP-binding element has been identified in the 3′-UTR of the *Arabidopsis* EST sequence AC005496.3 [41], which is an *AtTHIC* mRNA. Owing to the location of the TPP-binding element immediately upstream of the polyA tail [41], it is believed that the riboswitch may regulate the processing and stability of *AtTHIC* mRNA. Considering the result that additional extrinsic thiamine had a negative effect on the steady state mRNA level of *AtTHIC*, it is reasonable to assume that the TPP-binding riboswitch functions in the stability of *AtTHIC* mRNA.

In conclusion, *AtTHIC* is clearly involved in thiamine biosynthesis in *A. thaliana*. It putatively encodes an enzyme that functions in the pyrimidine synthesis in the thiamine biosynthesis pathway in *A. thaliana*. The high sequence and function similarities of AtTHIC with *E. coli* ThiC indicate that pyrimidine biosynthesis in *Arabidopsis* is similar to that in *E. coli*. Furthermore, overexpression of *AtTHIC* can increase thiamine levels in shoots. These findings are useful for the further understanding and manipulation of thiamine biosynthesis in plants.

Materials and Methods

AtTHIC cloning and sequence analysis

We searched the complete *Arabidopsis* genome sequence us-

ing a BLAST algorithm with the amino-acid sequence of *E. coli* ThiC. AGI gene code At2g29630, which showed high similarity to *E. coli* ThiC, was identified from http://mips.gsf.de/proj/plant/jsf/athal/searchjsp/index.jsp and termed *AtTHIC*. Based on the predicted DNA sequences of *AtTHIC* from the database, 3′ and 5′ -primers for *AtTHIC* (5′-AAG CTT ATG GCT GCT TCA GTA CAC TG-3′, 5′-GTC GAC TTA TTT CTG AGC AGC TTT GAC-3′) were designed and the coding sequence was amplified from total RNAs with these primers using RT-PCR. The DNA fragment was purified using a GFX gel purification kit (Pharmacia, Sweden), and was subsequently cloned into the pGEM T-easy vector (Promega, USA) and sequenced. Sequence alignment was performed using the ClustalW program (http://www.ebi.ac.uk/clustalw/) and phylogenic analysis was done using MEGA version 3.1 (http://www.megasoftware.net/).

Cell lines and media

An *E. coli* DH10B strain grown in LB medium was used for normal bacterial transformation. An *E. coli thiC* mutant strain (KG6953) was obtained from CGSC (The Coli Genetic Stock Center, MCDB Department, Yale University) to be grown in M9 medium [42] with 50 mg/L thiamine and was used for the functional complementation experiment. An *Agrobacterium tumefaciens* strain (GV3101) was grown in YEB medium [42] and was used for *Arabidopsis* transformation.

Plant materials and growth conditions

The Columbia ecotype of *A. thaliana* and the T-DNA insertion mutant *thic1* (Salk_011114, ABRC) were used in this experiment. Unless otherwise stated, an MS medium (Sigma, USA) without organic components, supplemented with 3% sucrose and 1% agar, and at a pH of 5.8 was applied for growing the wild-type and transgenic lines under sterile conditions at 23 °C with a 16-h light period.

Identification of the T-DNA insertion mutation line of AtTHIC

Putative mutants were screened using an electronic BLAST search of the available populations of sequence-indexed *Arabidopsis* T-DNA-insertion mutants using the *Arabidopsis* genomic sequence of *AtTHIC* (At2g29630). Salk_011114 from the ABRC was found to be a putative T-DNA insertion mutant of *AtTHIC*. The homozygous lines were identified using PCR with the primers LBb1 (5′-GCG TGG ACC GCT TGC TGC AAC T-3′), LP (5′-TGC GTC AAT GGT TGT TGT TAC-3′) and RP (5′-CAG CCA TAG CTG GAG ACA AAC-3′), which were designed using the SIGnAL T-DNA Verification Primer Design Tool. Because the homozygous lines are lethal under normal culture conditions, multiplication of the mutant line was carried out under culture conditions that had an exogenous thiamine supply (0.5 mg/L thiamine in MS medium and 100 mg/L thiamine in solution for irrigating mutants in soil).

Plasmid construction and transformation

For the functional complementation of the thiamine-auxotroph strain KG6953 of *E. coli*, the coding sequence of *AtTHIC* was amplified from total RNAs by RT-PCR using the 5′-CCA TGG ATG GCT GCT TCA GTA CAC TG-3′ and 5′-AAG CTT TTA TTT CTG AGC AGC TTT GAC-3′ primers. After verification by sequencing, the cDNA sequence of *AtTHIC* was cloned into the pGEX-KG vector at the *Nco*I/*Hind*III sites to generate the pGEX-KG-*AtTHIC* expression vector. Both plasmids (pGEX-KG-*AtTHIC* and pGEX-KG) were then introduced into the *E. coli thiC* mutant strain KG6953 and plated on an LB solid plate. The positive clones harboring pGEX-KG-*AtTHIC* and pGEX-KG were separately streaked on an M9 minimal plate supplemented with 15 mg/L IPTG (isopropyl-β-D-thiogalactopyranoside), 100 mg/L ampicillin and 50 mg/L thiamine, and on the M9 plate only with IPTG and ampicillin. After cultivation at 37 °C for 24 h, the plates were then photographed.

Transgenic *thic1* plants expressing functional AtTHIC were generated by transformation with a T-plasmid containing the coding sequence of *AtTHIC*, which was cloned at the *Nco*I/*Sal*I cut sites of pCAMBIA1300 under the control of a 2× 35S promoter (CAMBIA, Australia). The coding sequence of *AtTHIC* was amplified from total RNAs by RT-PCR using the 5′-AAG CTT ATG GCT GCT TCA GTA CAC TG-3′ and 5′-GTC GAC TTA TTT CTG AGC AGC TTT GAC-3′ primers. To generate transgenic plants expressing the *AtTHIC* promoter-*GUS*, a 1.5-kb genomic fragment upstream of the *AtTHIC* transcription start condon was amplified by PCR using the primers 5′-GGT ACC CAC ACA CGA AAT GAT CAA AG-3′ and 5′-CCA TGG AGC TGG AGA CAA ACG AAA AT-3′, and cloned at the *Nco*I/*Kpn*I sites of pJIT166 (http://www.pgreen.ac.uk) to generate the *AtTHIC* promoter-*GUS* plasmid. Then the expression cassette of the *AtTHIC* promoter –*GUS* was cut out from pJIT166 by *Xho*I/*Kpn*I and cloned into pBINPLUS [43]. The constructs were introduced into the GV3101 strain of *A. tumefaciens* for *Arabidopsis* transformation according to the floral dip protocol [44]. Seeds obtained from the primary transformants were germinated on MS plates containing 50 mg/L kanamycin for transformant selection. The homozygous transgenic lines were selected in the T3 generation and used for analysis.

Determination of thiamine content

Seeds were germinated and grown on MS medium without thiamine for 3 weeks. Then the seedlings were collected 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 in seedlings were performed according to the protocol (GB/T 5009.84-2003 Determination of thiamine (vitamin B1) in foods) described in the National Standard of the People's Republic of China (http://www.cssn.net.cn). Briefly, approximately 5-7 g (fresh weight) of seedlings were homogenized in 0.1 M HCl and digested by boiling for 30 min in an autoclave at 121 °C. Subsequently, protein and amylum in the solution were removed by treatment with protease and amylase at 37 °C overnight. After purification by filtration, the thiamine in the solution (10 ml) was oxidized to a thiazole pyrimide pigment by adding 5 ml reaction buffer (0.25 g/L $K_3Fe(CN)_6$ and 10 g/L NaOH), which was then fractionated with a Diamonsil^{TM} C_{18} column (5 μm, 250 mm × 4.6 mm) (Dikma, USA) on an LC-10ATvp HPLC (Shimadzu, Japan) and measured using an RF-10AXL fluorescence detector (excitation: 375 nm, emission: 435 nm) according to the manufacturer's instructions. The thiamine content in plants was calculated based on a standard curve of thiamine.

GUS activity assay

Five independent transgenic lines were germinated and grown on MS medium as described above. Seedlings at different devel-

opmental stages were collected for histochemical assays of GUS activity according to a previously described procedure [45]. To investigate GUS activity in generative tissues, flowers and siliques were harvested from the plants growing in a growth chamber.

Northern blot analysis

Total RNAs were extracted with Trizol reagent (Invitrogen, USA) and 10 μg RNAs for each sample were used for northern blotting. RNA samples were denatured and electrophoresed on a 1.2% 3-(*N*-morpholino)-propane-sulfonic acid/formaldehyde/agarose gel before transfer to a nylon Hybond-N^+ membrane according to the manufacturer's instructions (Amersham Pharmacia Biotech, USA). The DNA probe, a 300-bp fragment of *AtTHIC* cDNA amplified by PCR with the 5′-CTA AGG AAG GAG TGG ATT G-3′ and 5′-GAC GAC CAG TTG AGA GAT-3′ primers, was labeled with ~30 μC ^{32}P-dCPT using the Prime-A-Gene Labeling System according to the manufacturer's instructions (Promega, USA). Hybridization and membrane washing were performed according to a previously described protocol [46]. The membrane was exposed to a Fujifilm (BAS-SR2025, Tokyo, Japan) in an X-ray imaging plate under –70 °C for 1-3 days for radioautography.

RT-PCR analysis

Total RNAs were extracted with Trizol reagent. After eliminating genomic DNA contamination according to a previously described method [47], the first–strand cDNA was synthesized from 2 μg total RNAs with M-MLV reverse transcriptase (Invitrogen) and oligo $(dT)_{17}$ as a primer according to the manufacturer's instructions. The content of cDNAs in the reverse transcription mixtures were standardized by amplifying *ACTIN* transcripts using the 5′-GAT TCG CTG GAG ATG ATC TG-3′ and 5′-TCA GGA GCA ATA CGG AGC CT-3′ primers. To analyse *AtTHIC* expression, the 5′-CTA AGG AAG GAG TGG ATT G-3′ and 5′-GAC GAC CAG TTG AGA GAT-3′ primers were applied. PCR amplification was conducted as follows: 95 °C for 3 min, 30 cycles of 94 °C for 30 s, 55 °C for 30s, 72 °C for 30 s and a final extension at 72 °C for 5 min. The PCR products were separated by electrophoresis on 1.5% (w/v) agarose gel with ethidium bromide and documented using a CCD camera.

Subcellular localization of AtTHIC

To investigate the subcellular location of AtTHIC, the coding sequence of *AtTHIC* was amplified from *A. thaliana* cDNA by PCR using the 5′-AAG CTT ATG GCT GCT TCA GTA CAC TG-3′ and 5′-GTC GAC TTT CTG AGC AGC TTT GAC ATA G-3′ primers. The PCR product was cloned into the pGEM T-easy vector (Promega, USA). After verification by sequencing, the coding sequence of *AtTHIC* was excised by *Nco*I and *Hind*III and cloned into the pJIT163-*hGFP* vector [48] to generate the AtTHIC-GFP fusion protein (GFP at C-terminus of AtTHIC) expression plasmid pJIT163-*AtTHIC-hGFP*.

For transient expression in onion epidermis cells, 5 μg of the pJIT163-*AtTHIC-hGFP* plasmid were used to bombard the onion epidermis cells according to a described procedure [49]. For transient expression in cowpea protoplasts, 10~15 μg of the purified pJIT163-*AtTHIC-hGFP* plasmid were used to transform cowpea mesophyll protoplasts by PEG transformation according to a described procedure [50]. The GFP fluorescence in onion epidermis cells and in cowpea protoplasts was detected using a Confocal Laser Scanning Microscope (Olympus, FV500, Japan). The GFP fluorescence was observed at an excitation wavelength of 488 nm and an emission wavelength of 506-538 nm, and the autofluorescence of chloroplasts was observed at an excitation wavelength of 488 nm and an emission wavelength of 664-696 nm.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant numbers 30530460 and 30521001), the Ministry of Science and Technology of China (grant numbers 2005cb20904 and 2006AA10A105) and the Chinese Academy of Sciences (grant number KSCX2-YW-N-001).

References

1 Begley TP. The biosynthesis and degradation of thiamine (vitamin B_1). *Nat Prod Rep* 1996; **13**:177-185.
2 Burdick R. Thiamine. In: Howe-Grant M, ed. Kirk-Othmer encyclopedia of chemical technology, Wiley Press: New York 1998; **25**:152-171.
3 Nosaka K. Recent progress in understanding thiamine biosynthesis and its genetic regulation in *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 2006; **72**:30-40.
4 Spenser ID, White RL. Biosynthesis of vitamin B_1 (thiamine): An instance of biochemical diversity. *Angew Chem Int Ed Engl* 1997; **36**:1032-1046.
5 Estramareix B, Therisod M. Tyrosine as a factor of biosynthesis of the thiazole moiety of thiamine in *Escherichia coli*. *Biochim Biophys Acta* 1972; **273**:275-282.
6 Bellion E, Kirkley DH, Faust JR. The biosynthesis of the thiazole moiety of thiamine in *Salmonella typhimurium*. *BBA-Gen Subjects* 1976; **437**:229-237.
7 White RH, Rudolph FB. The origin of the nitrogen atom in the thiazole ring of thiamine in *Escherichia coli*. *Biochim Biophys Acta* 1978; **542**:340-347.
8 DeMoll E, Shive W. Determination of the metabolic origin of the sulfur atom in thiamine of *Escherichia coli* by mass spectrometry. *Biochem Biophys Res Commun* 1985; **132**:217-222.
9 Tazuya K, Yamada K, Nakamura K, Kumaoka H. The origin of the sulfur atom of thiamine. *Biochim Biophys Acta* 1987; **924**:210-215.
10 Webb E, Downs D. Characterization of *thiL*, Encoding thiamine-monophosphate Kinase, in *Salmonella typhimurium*. *J Biol Chem* 1997; **272**:15702-15707.
11 Kinsland C, Taylor SV, Kelleher NL, McLafferty FW, Begley TP. Overexpression of recombinant proteins with a C-terminal thiocarboxylate: Implications for protein semisynthesis and thiamine biosynthesis. *Protein Sci* 1998; **7**:1839-1842.
12 Taylor SV. Thiamine biosynthesis in *E. coli*: biosynthesis of the thiazole moiety. PhD thesis. Cornell University, Ithaca, NY, 1998.
13 Taylor SV, Kelleher NL, Kinsland C, *et al*. Thiamine biosynthesis in *Escherichia coli*. Identification of this thiocarboxylate as the immediate sulfur donor in the thiazole formation. *J Biol Chem* 1998; **273**:16555-16560.

14 Begley TP, Diana MD, Steven EE, *et al*. Thiamine biosynthesis in prokaryotes. *Arch Microbiol* 1999; **171**:293-300.
15 Costello C. Mechanistic studies on thiamine biosynthesis in *E.coli* and thiamine degradation in *B. thiaminolyticus*. PhD thesis. Cornell University, Ithaca, NY, 1996.
16 Zhang Y, Begley TP. Cloning, sequencing and regulation of *thiA*, a thiamine biosynthesis gene from *Bacillus subtilis*. *Gene* 1997; **198**:73-82.
17 Zhang Y, Taylor SV, Chiu HJ, Begley TP. Characterization of the *Bacillus subtilis thiC* operon involved in thiamine biosynthesis. *J Bacteriol* 1997; **179**:3030-3035.
18 Vander Horn PB, Backstrom AD, Stewart V, Begley TP. Structural genes for thiamine biosynthetic enzymes (*thiCEFGH*) in *Escherichia coli* K-12. *J Bacteriol* 1993; **175**:982-992.
19 Backstrom AD, McMordie RAS, Begley TP. Biosynthesis of thiamine I: The function of the thiE gene product. *J Am Chem Soc* 1995; **117**:2351-2352.
20 Kelleher NL, Taylor SV, Grannis D, *et al*. Efficient sequence analysis of the six gene products (7-74 kDa) from the *Escherichia coli* thiamine biosynthetic operon by tandem high-resolution mass spectrometry. *Protein Sci* 1998; **7**:1796-1801.
21 Webb E, Febres F, Downs DM. Thiamine pyrophosphate (TPP) negatively regulates transcription of some *thi* genes of *Salmonella typhimurium*. *J Bacteriol* 1996; **178**:2533-2538.
22 Rodríguez-Navarro S, Llorente B, Rodríguez-Manzaneque MT, *et al*. Functional analysis of yeast gene families involved in metabolism of vitamin B_1. *Yeast* 2002; **19**:1261-1276.
23 Wightman R, Meacock PA. The *THI5* gene family of *Saccharomyces cerevisiae*: distribution of homologues among the hemiascomycetes and functional redundancy in the aerobic biosynthesis of thiamine from pyridoxine. *Microbiology* 2003; **149**:1447-1460.
24 Li SL, Rédei GP. Thiamine mutants of the crucifer, *Arabidopsis*. *Biochem Gene* 1969; **3**:163-170.
25 Kumar S, Sharma SB. Mutations in three of the genes determining thiamine biosynthesis in *Pisum sativum*. *MGG* 1986; **204**:473-476.
26 Proebsting WM, Maggard SP, Guo WW. The relationship of thiamine to the Alt locus of *Pisum sativum L*. *J Plant Physiol* 1990; **136**:231-235.
27 Koornneef M, Hanhart CJ. *Arabidopsis* Information Service 18. 1981; http://www.*Arabidopsis*.org/ais/1981/koorn-1981-aactn.html.
28 Machado CR, Costa de Oliveira RL, Boiteux S, Praekelt UM, Meacock PA, Menck CFM. Thi1, a thiamine biosynthetic gene in *Arabidopsis thaliana*, complements bacterial defects in DNA repair. *Plant Mol Biol* 1996; **31**:585-593.
29 Ribeiro A, Praekelt U, Akkermans ADL, *et al*. Identification of *agthi1*, whose product is involved in biosynthesis of the thiamine precursor thiazole, in actinorhizal nodules of *Alnus glutinosa*. *Plant J* 1996; **10**:361-368.
30 Ajjawi I, Tsegaye Y, Shintani D. Determination of the genetic, molecular, and biochemical basis of the *Arabidopsis thaliana* thiamine auxotroph *th1*. *Arch Biochem Biophys* 2007; **459**:107-114.
31 Chabregas SM, Luche DD, Farias LP, *et al*. Dual targeting properties of the N-terminal signal sequence of *Arabidopsis thaliana* THI1 protein to mitochondria and chloroplasts. *Plant Mol Biol* 2001; **46**:639-650.
32 Chabregas SM, Luche DD, van Sluys MA, Menck CF, Silva-Filho MC. Differential usage of two in-frame translational start codons regulates subcellular localization of *Arabidopsis thaliana* THI1. *J Cell Sci* 2002; **116**:285-291.
33 Kim YS, Nosaka K, Diana MD, *et al*. A Brassica cDNA clone encoding a bifunctional hydroxymethylpyrimidine kinase/thiamin-phosphate pyrophosphorylase involved in thiamine biosynthesis. *Plant Mol Biol* 1998; **37**:955-966.
34 Raschke M, Burkle L, Muller N, *et al.* Vitamin B1 biosynthesis in plants requires the essential iron sulfur cluster protein, THIC. *Proc Natl Acad Sci USA* 2007; **104**:19637-19642
35 Julliard J, Douce R. Biosynthesis of the thiazole moiety of thiamine (vitamin B_1) in higher plant chloroplasts. *Proc Natl Acad Sci USA* 1991; **88**:2042-2045.
36 Faith CB, Thomas L, Boyang C, Alan LK. Evidence for the thiamine biosynthetic pathway in higher-plant plastids and its developmental regulation. *Plant Mol Biol* 1995; **29**:809-821.
37 Miranda-Ríos J, Navarro M, Soberón M. A conserved RNA structure (*thi* box) is involved in regulation of thiamine biosynthetic gene expression in bacteria. *Proc Natl Acad Sci USA* 2001; **98**:9736-9741.
38 Mironov AS, Gusarov I, Rafikov R, *et al*. Sensing small molecules by nascent RNA: A mechanism to control transcription in bacteria. *Cell* 2002; **111**:747-756.
39 Winkler W, Nahvi A, Breaker RR. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* 2002; **419**:952-956.
40 Lai EC. RNA sensors and riboswitches: self-regulating messages. *Curr Biol* 2003; **13**:285-291.
41 Sudarsan N, Barrick JE, Breaker RR. Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA* 2003; **9**:644-647.
42 Sambrook J, Russell DW, eds. Moleular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, 2001.
43 Van Engelen FA, Molthoff JW, Conner AJ, Nap JP, Pereira A, Stiekema WJ. pBINPLUS: an improved plant transformation vector based on pBIN19. *Transgenic Res* 1995; **4**:288-290.
44 Clough SJ, Bent AF. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J* 1998; **16**:735-743.
45 Jefferson RA, Kavanagh TA, Bevan MW. Assaying chimeric genes in plants: The GUS gene fusion system. *Plant Mol Biol Rep* 1987; **5**:387-405.
46 Church GM, Gilbert W. Genomic Sequencing. *Proc Natl Acad Sci USA* 1984; **81**:1991-1995.
47 Li L, Cheng X, Ling HQ. Isolation and characterization of Fe(III)-chelate reductase gene *LeFRO1* in tomato. *Plant Mol Biol* 2004; **54**:125-136.
48 Liu X, Liu X, Zhang J, *et al*. Molecular and functional characterization of sulfiredoxin homologs from higher plants. *Cell Res* 2006; **16**:287-296.
49 Weigel D, Glazebrook J, eds. *Arabidopsis*: a Laboratory Manual. New York: Cold Spring Harbor Laboratory Press, 2002.
50 Shah K, Russinova E, Gadella TW Jr, Willemse J, de Vries SC. The *Arabidopsis* kinase-associated protein phosphatase controls internalization of the somatic embryogenesis receptor kinase 1. *Genes Dev* 2002; **16**:1707-1720.