



**Molecular characterization of the thi3 gene involved in thiamine biosynthesis in Zea mays: cDNA sequence and enzymatic and structural properties of the recombinant bifunctional protein with 4-amino-5-hydroxymethyl-2-methylpyrimidine (phosphate) kinase and thiamine monophosphate synthase activities**

Maria Rapala-Kozik, Mariusz Olczak, Katarzyna Ostrowska, Agata Starosta, Andrzej Kozik

► **To cite this version:**

Maria Rapala-Kozik, Mariusz Olczak, Katarzyna Ostrowska, Agata Starosta, Andrzej Kozik. Molecular characterization of the thi3 gene involved in thiamine biosynthesis in Zea mays: cDNA sequence and enzymatic and structural properties of the recombinant bifunctional protein with 4-amino-5-hydroxymethyl-2-methylpyrimidine (phosphate) kinase and thiamine monophosphate synthase activities. *Biochemical Journal*, 2007, 408 (2), pp.149-159. <10.1042/BJ20070677>. <hal-00478811>

**HAL Id: hal-00478811**

**<https://hal.science/hal-00478811v1>**

Submitted on 30 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

**Molecular characterization of *thi3* gene involved in thiamine biosynthesis in *Zea mays*: cDNA sequence and enzymatic and structural properties of recombinant bifunctional protein with 4-amino-5-hydroxymethyl-2-methylpyrimidine (phosphate) kinase and thiamine monophosphate synthase activities**

Maria RAPALA-KOZIK<sup>\*1</sup>, Mariusz OLCZAK<sup>†</sup>, Katarzyna OSTROWSKA<sup>‡</sup>, Agata STAROSTA<sup>\*</sup> and Andrzej KOZIK<sup>\*</sup>

<sup>\*</sup>Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian University, Gronostajowa 7, 30-387 Kraków, Poland, <sup>†</sup>Faculty of Biotechnology, University of Wrocław, Tamka 2, 50-138 Wrocław, Poland, and <sup>‡</sup>Faculty of Chemistry, Jagiellonian University, Ingardena 3, 30-060 Kraków, Poland

Short title: THI3 protein involved in thiamine biosynthesis in maize

---

Abbreviations used: CBD, chitin-binding domain; HET, 4-methyl-5-(2-hydroxyethyl)thiazole; HET-P, 4-methyl-5-(2-hydroxyethyl)thiazole phosphate; HMP, 4-amino-5-hydroxymethyl-2-methylpyrimidine; HMP-P, 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate; HMP-PP, 4-amino-5-hydroxymethyl-2-methylpyrimidine diphosphate; IPTG, isopropyl- $\beta$ -D-thiogalactopyranoside; RACE, rapid amplification of cDNA ends; RP, reverse phase; TBA, tetrabutylammonium hydrogensulphate; TDP, thiamine diphosphate; TMP, thiamine monophosphate.

<sup>1</sup>To whom correspondence should be addressed (e-mail rapala@mol.uj.edu.pl).

Sequence data from this article were deposited in the EMBL/DDBJ/GenBank database on December 5, 2005 under Accession Numbers AM16973 (GI:83423371) (cDNA) and CAJ45026 (protein).

---

A thiamine biosynthesis gene *thi3* of *Zea mays* has been identified through cloning and sequencing of cDNA and heterologous overexpression of encoded protein (THI3) in *Escherichia coli*. The recombinant THI3 protein was purified to homogeneity and showed to possess two essentially different enzymatic activities of 4-amino-5-hydroxymethyl-2-methylpyrimidine (phosphate) (HMP/HMP-phosphate) kinase and thiamine monophosphate (TMP) synthase. Both activities were characterized in terms of basic kinetic constants with interesting findings that TMP synthase is uncompetitively inhibited by excess of one of the substrates (HMP diphosphate) and ATP. A bioinformatic analysis of THI3 sequence suggested that these activities were located in two distinct, N-terminal kinase and C-terminal synthase, domains. Models of the overall folds of THI3 domains and the arrangements of active centre residues were obtained with the SWISS-MODEL protein modelling server, based on the known three-dimensional structures of *Salmonella typhimurium* HMP/HMP-phosphate kinase and *Bacillus subtilis* TMP synthase. The essential roles of Gln-98 and Met-134 residues for HMP kinase activity and of Ser-444 residue for TMP synthase activity were experimentally confirmed by site-directed mutagenesis.

Key words: cDNA cloning, enzyme kinetics, protein overexpression, thiamine phosphate diphosphorylase, site-directed mutagenesis, structural modelling.

## INTRODUCTION

Thiamine diphosphate (TDP) plays an essential role in all organisms as a cofactor of several universal enzymes involved in carbohydrate and amino acid metabolism, including pyruvate decarboxylase, pyruvate-,  $\alpha$ -ketoglutarate-, and branched fatty acid- dehydrogenases, transketolase, and  $\alpha$ -acetolactate synthase [1]. TDP is synthesized from exogenous thiamine taken with food (as the vitamin B<sub>1</sub>) by higher animals or taken up by microbial cells from the environment [2, 3], or from endogenous thiamine monophosphate (TMP) which is synthesized by many prokaryotic and eukaryotic microorganisms as well as higher plants. The last step of *de novo* TMP biosynthesis involves a condensation of two precursors, 4-amino-5-hydroxymethyl-2-methylpyrimidine diphosphate (HMP-PP) and 4-methyl-5-(2-hydroxyethyl)thiazole phosphate (HET-P) which are separately synthesized on the routes that significantly differ between various groups of thiamine-synthesizing organisms. Many microorganisms can additionally salvage HMP-PP and HET-P by phosphorylation of 4-amino-5-hydroxymethyl-2-methylpyrimidine (HMP) and 4-methyl-5-(2-hydroxyethyl)thiazole (HET) taken up from the medium (Scheme 1).

The pathways of bacterial thiamine biosynthesis are well recognized (for review see: [4]). Many of the enzymes which catalyze the specific steps of thiamine synthesis in model bacteria species such as *Escherichia coli*, *Bacillus subtilis* and *Salmonella typhimurium* have been obtained as recombinant proteins and their three-dimensional structures have recently been determined (for review see: [5]). In a sharp contrast, our knowledge of this process in eukaryotic cells is at a very preliminary stage, in terms of both the identification of individual biosynthetic reactions and the structural biology of enzymes involved. In baker's yeasts *Saccharomyces cerevisiae*, at least fifteen TDP- biosynthesis genes were identified (for review see: [6]) but products of only four genes, *THI80*, *THI6*, *THI4* and *THI20*, are known as purified, natural or recombinant proteins [7-10]. Kinetic characteristics are available for THI80 protein with thiamine diphosphokinase (EC 2.7.6.2) activity [11] and for THI6 bifunctional enzyme with TMP synthase (TMP diphosphorylase, EC 2.5.1.3) and HET kinase (EC 2.7.1.50) activities [8]. Three-dimensional structures were reported for thiamine diphosphokinase [7] and THI4 protein [9], involved in thiazole synthesis [12] and possibly playing some role in yeast tolerance against DNA damage [13]. In higher plants, a limited number of thiamine biosynthesis genes have been identified and sequenced, including several orthologues of yeast *THI4* thiazole-synthesizing gene, found in *Zea mays* [14], *Arabidopsis thaliana* [15] and *Alnus glutinosa* [16], and a family of orthologues of *Brassica napus bth1* gene [17] which probably codes for a bifunctional enzyme with TMP synthase and

HMP/HMP-P kinase (EC 2.7.1.49 / EC 2.7.4.7) activities. Any protein product of *bth1* gene was not isolated although its function in thiamine biosynthesis was drawn from the functional complementation of thiamine-requiring mutants of some bacteria [17]. Recently, a three-dimensional structure of *A. thaliana* orthologue of yeast THI4 protein was reported [18] but its enzymatic activity remains to be defined.

In this work we cloned and sequenced cDNA of a *Z. mays* gene designated *thi3* which is orthologous to *B. napus bth1* gene. A product of this gene was isolated as a recombinant active enzyme and its dual TMP synthase and HMP/HMP-P kinase activity was directly presented and kinetically characterized. Models for the structure of enzyme active centres are presented. This is the second plant thiamine-synthesizing enzyme which has been characterized on the purified protein level and the first one whose enzymatic properties were quantitatively analyzed.

## EXPERIMENTAL

### Cloning and sequencing of *thi3* cDNA from *Zea mays* leaves

The 10-day old leaves were grinded in liquid nitrogen to a fine powder using mortar and pestle followed by NucleoSpin filter homogenization (Macherey-Nagel, Germany). Total RNA from 100 mg of disrupted tissue was purified using NucleoSpin RNA II Kit (Macherey-Nagel). The quantity and integrity of the total RNA was checked by formaldehyde agarose gel electrophoresis.

To identify the *thi3* cDNA, a set of six degenerate primers based on the known eukaryotic and prokaryotic conservative sequences of BTH1-related proteins (sequences derived from EMBL database) were designed. The cDNA was synthesized using Superscript III reverse transcriptase (Invitrogen, USA) with poly-dT(20) primer. From nine reactions performed in the experiment, the pair of MO1-F primer and MO2-R primer (Table 1) gave a dominant PCR product of about 1000 bp. The product was purified from the agarose gel, cloned into the pCR4-TOPO vector, and sequenced. All PCR reactions were performed with Perpetual Opti Taq Polymerase (Eurx, Gdansk, Poland).

The 5' and 3' RACE (rapid amplification of cDNA ends) was performed according to Frohman [19]. Based on the partial sequence of *thi3*, a set of 5'RACE reverse primers and 3'RACE forward primers was designed (Table 1). The synthesis of cDNA from 4 µg of the total RNA was performed using the ZTH51 primer in 5'RACE and adaptor-oligo-dT(17) primers in 3'RACE. After 5'RACE RT reaction, primers, nucleotides and salts were removed by ultrafiltration on Microcon-100 columns (Millipore, USA), using sterile water as dilution liquid. The purified cDNA was tailed using terminal transferase (Promega, USA) in the presence of 0.25 mM dATP at 37°C for 10 min. The second strand of cDNA was synthesized using adaptor-dT(17) primer. After adding of adaptor and ZTH52 primers, the first 5'RACE reaction was performed, followed by "semi-nested" PCR using ZTH53 primer and adaptor primer. The discrete product was purified and cloned into the pCR4-TOPO sequencing vector. The 3'RACE product was amplified after adding of ZTH31 primer and adaptor primer to 3'RT cDNA. Then ZTH32 primer and adaptor primer were used in "semi-nested" PCR. The specific product was purified and cloned into the pCR4-TOPO vector. All DNAs were analyzed by sequencing (Sequencing Laboratory, Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland).

### Expression and purification of THI3 protein in *E. coli*

The overexpression and purification of THI3 protein containing intein (with the chitin-binding domain, CBD) attached at either the N- or C-terminus of the protein was performed according to the published method [20]. For this purpose, the *thi3* cDNA was cloned into the pTYB1 vector (New England Biolabs, USA) utilizing the *NdeI* and *SapI* restriction sites, and into pTYB11 vector utilizing the *SapI* and *XhoI* restriction sites, resulting in pMO1 and pMO2 plasmids, respectively. The N-tagged version of recombinant THI3 protein lacked N-terminal 21-amino acid sequence. Recombinant proteins containing the CBD tag were overexpressed after induction with 0.3 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) in ER2566 *E. coli* cells at 16°C for 16 h. The cells were suspended in 20 mM Tris/HCl buffer, pH 8.5, containing 500 mM NaCl and sonicated. After centrifugation at 40 000 x g for 20 min, the recombinant proteins containing the CBD tag were purified according to the IMPACT-CN expression system (New England Biolabs) from soluble fraction using 50 mM dithiothreitol to induce the tag cleavage.

The protein samples were dialyzed against 20 mM Tris/HCl buffer, pH 7.5, containing 150 mM NaCl and subjected to the fine purification on Superdex 200 HR 10/50 high performance size-exclusion column (Amersham, England) eluted with the same buffer. The purified protein samples were stored in aliquots at -20°C. Under these storage conditions the enzyme was stable for at least 3 months.

Protein samples from various preparation steps were analyzed by SDS-PAGE in the Laemmli system [21] with 10% resolution gel. The protein bands were either visualized by Coomassie Brilliant Blue R-250 staining or subjected to Western blotting on nitrocellulose membranes, using primary polyclonal anti-CBD antibodies (New England Biolabs) and secondary anti-rabbit antibody-horse-radish peroxidase conjugate (Promega).

### Site-directed mutagenesis

The QuickChange Multi Site-Directed Mutagenesis Kit (Stratagene, USA) and the pTyb11 bacterial expression plasmid with *thi3* insert as a template was used for mutagenesis. Initially, two double mutants were produced: in the first mutant Gln-98 was replaced by Leu and Gln-373 by Leu, in the second Met-134 was replaced by Lys and Thr-472 by Asp. The single mutants Q98L, M134K, Q373L and T472D were constructed by digestion of the double mutants and the wild-type Tyb11-*thi3* by *SacI* and *XhoI* restriction enzymes and appropriate ligation. Additionally, a single mutant in which Ser-444 was replaced by Ala was designed.

Table 1 contains primers used in site-directed mutagenesis. All seven mutagenesis constructs were sequenced.

### **Synthesis of HMP and its phosphate- and diphosphate esters**

HMP synthesis was performed through the intermediary syntheses of 4-amino-5-aminomethyl-2-methylpyrimidine [22] and 4-amino-5-bromomethyl-2-methylpyrimidine [23]. HMP-P and HMP-PP were obtained from HMP [24].

### **Enzyme assays**

For TMP synthase assay, the reaction mixture contained, unless otherwise specified, 15  $\mu\text{M}$  HMP-PP, 15  $\mu\text{M}$  HET-P and 10 mM  $\text{MgCl}_2$  in 50 mM Tris-HCl buffer, pH 8.0 and 5  $\mu\text{g}$  of the pure enzyme in the final volume of 100  $\mu\text{l}$ . After incubation at 37°C for 60 min, the reaction was stopped with 16  $\mu\text{l}$  of 10% metaphosphoric acid and the samples were centrifuged at 10 000  $\times g$  for 10 min, diluted 3 times with 0.1 M potassium phosphate and analyzed for TMP content by reverse-phase HPLC (RP-HPLC) with postcolumn derivatization and fluorimetric detection [25]. A Merck LiChrosphere 100 RP-18 (5  $\mu\text{m}$ ) column (250 mm  $\times$  4 mm) was used, with a two-solvent system including 0.15 M ammonium citrate, pH 4.2 (solvent A) and 0.1 M formic acid, containing 0.4% diethylamine, pH 3.2 (solvent B). A gradient elution (0–90% B in 16 min) at a flow rate of 1 ml min<sup>-1</sup> was applied. The fluorogenic derivatization was performed with 0.0025% sodium hexacyanoferrate (III) in 2.25% NaOH, pumped through a peristaltic pump with a flow rate of 0.8 ml min<sup>-1</sup>. The fluorescence emission at 430 nm excited at 365 nm was monitored.

The production of TMP from HMP or HMP-P due to the consecutive action of HMP/HMP-P kinase and TMP synthase were determined in analogous samples, with HMP-PP replaced by these substrates (15  $\mu\text{M}$ ) and with the 10 mM ATP supplement.

For a direct HMP kinase assay, the reaction mixtures contained 20  $\mu\text{M}$  HMP, 1 mM ATP, 10 mM  $\text{MgCl}_2$ , 50 mM Tris-HCl buffer pH 8.0 and 5  $\mu\text{g}$  enzyme in the final volume of 100  $\mu\text{l}$ . After 60 min-incubation at 37°C, the reaction was stopped and the sample, diluted 3 times with 0.1 M potassium phosphate, was analyzed for HMP-P content by RP-HPLC on a Supelcosil LC-18 column (250 mm  $\times$  4.6 mm), eluted with a ternary solvent system composed



of 4 mM tetrabutylammonium hydrogensulphate (TBA) in water (solvent A), 4 mM TBA in methanol (solvent B) and 4 mM TBA in 50 mM potassium phosphate buffer pH 6.6 (solvent C). The following time program was applied: at t=0 min: 10%B, 30%C, at t=14 min: 50%B, 50%C, at t=16 min: 10%B, 80%C, at t=18 min: 10%B, 30%C, at t=35min: 10%B, 30%C. The absorbance of the eluate at 245 nm was monitored.

### Kinetic data analysis

The plots of the initial reaction velocity *versus* substrate concentrations were analyzed for various kinetic parameters as outlined by Bisswanger [26]. For hyperbolic kinetics, the apparent  $K_m$  and  $V_{max}$  parameters were estimated by fitting the classic Michaelis-Menten equation (untransformed) with the use of GraphPad Prism software. In the cases where the reaction velocity reached a maximum and decreased thereafter, indicative of the inhibition by the excess of substrate, the following equation was fitted:

$$v = (V_{max}[S]) / (K_m + [S] + [S]^2/K_{iu})$$

where  $v$  is the reaction velocity,  $[S]$  – the substrate concentration, and  $K_{iu}$  – the apparent substrate (uncompetitive) inhibition constant.  $K_{iu}$  was first estimated from the abscissa intercept of the high-concentration part of the Dixon plot,  $1/v$  *versus*  $[S]$ . The value obtained was used as a constant during the fitting of the above equation by the GraphPad Prism to estimate the remaining  $K_m$  and  $V_{max}$  parameters. The uncompetitive inhibition of TMP synthase by ATP was analyzed from the plots of the reaction velocities *versus* HET-P concentrations at various ATP concentrations which should follow the equation:

$$v = (V_{max}[S]) / (K_m + [S] + [S][I]/K_{iu})$$

where  $[I]$  is the inhibitor concentration, and  $K_{iu}$  – the apparent uncompetitive inhibition constant. The Hanes-Woolf transformations,  $[S]/v$  *versus*  $[S]$ , were used to classify the type of inhibition, and their slopes, determined by linear regression, were re-plotted *versus*  $[I]$  to estimate  $K_{iu}$  as the abscissa intercept of this secondary plot. In addition, all three parameters,  $K_{iu}$ ,  $K_m$  and  $V_{max}$ , of the above equation were also estimated by direct non-linear regression fits to the untransformed data.

## Bioinformatic analyses

Searching for protein sequence similarities was performed with the BLAST (v. 2.2.16) program [27] available on the server of National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov>). The multiple sequence alignments were generated with the use of CLUSTAL W (v. 1.82) [28] from the ExPASy Proteomics Server of Swiss Institute of Bioinformatics (<http://www.expasy.org>) and visualized with JalView 2.3 [29] multiple alignment editor (<http://www.jalview.org>). Putative chloroplast transit peptides were found by the ChloroP (v. 1.1) program [30] at the server of Center for Biological Sequence Analysis of the Technical University of Denmark (<http://www.cbs.dtu.dk/services/ChloroP>). The THI3 protein sequence was also analyzed using the Conserved Domains Search service [31] on the NCBI server. Models of three dimensional structures of THI3 protein domains, based on the known crystal structures of *S. typhimurium* HMP-P kinase (PDB code 1JXH, [32]) and *B. subtilis* TMP synthase (PDB code 2TPS, [33]), were generated using the SWISS-MODEL automated homology modelling server [34] of Swiss Institute of Bioinformatics (<http://swissmodel.expasy.org>). The structures were visualized using the Swiss-PdbViewer and PyMOL programs available from the ExPASy Server.

## RESULTS

### Cloning and sequence analysis of maize *thi3* cDNA

The amplification with a pair of degenerate MO1-F forward and MO2-R reverse primers (Table 1) resulted in a discrete, specific RT-PCR product, which was cloned into a sequencing vector. After *E. coli* transformation, 10 clones were randomly chosen for sequencing. All clones contained the same 997-nucleotide DNA fragment. Based on this sequence, 5'RACE and 3'RACE were performed. In both reactions, nested PCR gave dominant products, which were cloned into a sequencing vector. For each RACE, four randomly chosen colonies were picked up and a plasmid purified from each clone was sequenced. All 3'RACE sequences were identical. From clones received after 5'RACE we choose one with the longest region upstream of the putative ATG start codon. This *thi3* cDNA clone (Accession No. AM16973) had a total length of 1947 bp and contained a 48 bp untranslated 5' region, a 1653 bp-long open reading frame and a 246 bp untranslated 3' region with the stop codon and a polyA tail. A putative open reading frame-encoded protein (THI3, Accession No. CAJ45026) consisted of 551 amino acid residues and its predicted molecular mass and isoelectric point were 57.8 kDa and 8.51, respectively.

### Purification and basic molecular characterization of recombinant maize THI3 protein, overproduced in *E. coli*

Of several versions of THI3 recombinant protein tested, the protein lacking N-terminal 21-amino acid fragment and expressed in *E. coli* with 55-kDa intein as a tag protein attached to the N-terminus gave a satisfactory level of expression, especially when bacterial cultures were propagated at temperature lowered to 16°C. Under these conditions the majority of recombined fusion THI3 protein was found in the soluble fraction.

The SDS-PAGE record of a typical THI3 purification run is presented in Figure 1. After the affinity chromatography on a chitin-containing column followed by the Superdex-200 high performance gel filtration, a homogeneous protein was obtained in a typical yield of about 3 mg from 4 l of bacterial culture. The protein apparent monomer molecular mass was 55 kDa as determined by SDS-PAGE. Upon the Superdex-gel filtration the protein eluted in a volume corresponding to an apparent molecular mass of 95 kDa (results not shown), suggesting that under non-denaturing conditions the protein existed as a dimer.

### Characterization of enzymatic activities of THI3 protein

As shown in Table 2, the recombinant maize THI3 protein could synthesize TMP from HMP-PP and HET-P substrates. This TMP synthase activity required the presence of magnesium ions in the reaction mixture. For this reaction, HMP-PP could not be substituted by HMP or HMP-P, nor HET-P by HET. The TMP synthase activity was inhibited by ATP, but in the presence of ATP TMP could be formed from HET-P plus HMP or HMP-P, clearly indicating that the enzyme possessed additional HMP kinase and HMP-P kinase activities. At the equivalent ATP concentration, the rate of TMP formation from HMP/HMP-P plus HET -P substrates, *i.e.* due to the sequential action of HMP/HMP-P kinase and TMP synthase, was *ca.* 30% of that from HMP-PP plus HET-P. On the other hand, no HET kinase activity could be assigned to THI3 protein as TMP was not formed from unphosphorylated thiazole in combination with any pyrimidine in the presence of ATP.

The conventional kinetic characteristics of maize TMP synthase is illustrated in Figure 2. A dependence of initial reaction rate on HET-P concentration was hyperbolic but the dependence on HMP-PP concentration showed a pattern indicating for a strong inhibition by excess substrate (Figure 2A). The rate was maximal at 15-20  $\mu$ M HMP-PP and then gradually dropped. The inhibition of TMP synthase by ATP was analyzed against varying HET-P concentration at a constant HMP-PP concentration. An example analysis using Hanes-Woolf plots is shown in Figure 2B. These plots, prepared for various ATP concentrations, had a common intercept with the rate axis, suggesting that the inhibition is “uncompetitive” [26]. The kinetic characteristics of HMP kinase activity was obtained through direct determinations of HMP phosphorylation. The dependences on HMP or ATP concentration at the fixed concentration of the second substrate were hyperbolic (Figure 3). The values of estimated kinetic constants of maize HMP kinase and TMP synthase are collected in Table 3.

### Site-directed mutagenesis of THI3 protein at the TMP synthase and HMP/HMP-P kinase active centres

The active sites of THI3 protein were probed by substitution of several amino acid residues which were selected based on the literature data on bacterial TMP synthases and HMP kinases, as well as on the model of THI3 structure (see Discussion). The rates of TMP synthase and HMP kinase reactions in the extracts from cells transformed with mutant plasmids are presented in Table 4. Substitutions of Leu for Gln-98 and Lys for Met-134 were detrimental to HMP kinase activity while TMP synthase activity was fully preserved. A substitution of Ala for Ser-444 represents an opposite situation of a complete elimination of

TMP synthase activity without any significant effect on the second activity. Two further single mutants, Q373L and T472D, exhibited full TMP synthase activity and moderately reduced (by 30% and 50%, respectively) HMP kinase activity. A double mutant Q98L+Q373L differed from the Q98L mutant only in a detectable but still small (*ca.* 10%) HMP kinase activity. In contrast, another mutant with double amino acid substitution both in the TMP synthase and HMP/HMP-P kinase active centres, M134K+T472D, showed anomalous behaviour, presenting too high HMP kinase activity, *ca.* 25% of that of the wild-type protein.

## DISCUSSION

The *Z. mays* gene which was identified in this work through cloning of its cDNA and characterization of its product in the form of pure, heterologously expressed recombinant protein is hereby designated *thi3* as the third discovered gene of this species involved in thiamine biosynthesis. Searching through available protein databases revealed several orthologues of *Z. mays* THI3 in four higher plant species, including *Oryza sativa*, *Medicago truncatula*, *Arabidopsis thaliana* and *Brassica napus*. A multiple sequence alignment of plant THI3 orthologues (one from each species) is shown in Figure 4. Amino acid identity with *Z. mays* THI3 varies from 77% for rice ABA96049 protein to 63% for *Medicago* ABE87727 protein. In the *Arabidopsis* genome, three homologous sequences, AAM 91567 (Figure 4), BT020417 and AF000657 were identified. The first two differ by only one amino acid substitution; the third shows one short insertion and one deletion and may be the product of alternative splicing of the same gene. This gene exists in a single copy and maps to the locus AGT1G22950 on chromosome 1. Similarly, four homologous sequences, ABA96049 (Figure 4), EAZ19897, BAF29362 and ABG21903, were found in the rice genome. The first two probably originate by alternative splicing from one gene at the locus Os12g0192500 on chromosome 12, and two others are their truncated versions.

Like all other members of this plant protein family, THI3 contains a putative N-terminal chloroplast transit peptide [35]. All THI3 orthologues were described as having two putative conserved domains: an N-terminal domain with high sequence similarity to many bacterial HMP/HMP-P kinases, and a C-terminal domain homologous to bacterial TMP synthases. As analyzed with the use of NCBI CDD Search Server, THI3 shows an analogous two-domain structure, with the N-terminal kinase domain comprising residues 59-302 and the C-terminal synthase domain including residues 335-546.

In spite of the presence of putative HMP/HMP-P kinase and TMP synthase-encoding domains in cDNAs of several plant species, their protein products have not yet been obtained in isolated, pure form for which the suggested enzymatic activities could be directly demonstrated. However, the *B. napus bthi1* cDNA was isolated based on its ability to complement an *E. coli* mutant deficient in TMP synthase and was additionally shown to complement some HMP kinase-deficient mutants [17]. Our maize THI3 protein, obtained by heterologous expression in *E. coli*, is the first pure plant bifunctional enzyme with confirmed HMP/HMP-P kinase and TMP synthase activity.

The recombinant THI3 was homogeneous as determined by SDS-PAGE and apparently existed in solution as a dimer, similarly to bacterial HMP kinase and TMP

synthase which form dimers in crystals [32, 33]. The enzymatic properties of recombinant THI3 were readily detectable, although the determined specific activities (Table 3) were much lower than those reported for corresponding bacterial enzymes. For example, the specific activity of THI3 HMP kinase for HMP as the substrate ( $20 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ ) was one-eighth of the lowest value reported for a bacterial HMP kinase, ThiD from *E. coli* [36]. Similarly, the specific activities of a bacterial TMP synthase, ranging from  $65 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$  for the enzyme isolated from *E. coli* [37] to  $700 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$  for ThiC from *B. subtilis* [38], exceed at least eight times the value for the plant bifunctional enzyme ( $8.5 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ ). Whether these comparisons indicate for an intrinsically lower turnover of the bifunctional plant enzyme or for some conformational imperfections of its recombinant form, remains to be established. The kinetic characterization of maize THI3 protein revealed two interesting properties of the TMP synthase activity of this enzyme: the substrate inhibition by excess of HMP-PP and uncompetitive inhibition by ATP. The uncompetitive inhibition by ATP, ADP and inorganic diphosphate was also reported for *E. coli* TMP synthase [37].

In an attempt to identify specific amino acid residues which form the active centres on THI3 molecule we took advantage of a significant sequence similarity of THI3 with *S. typhimurium* HMP/HMP-P kinase (Code P55882) and *B. subtilis* TMP synthase (P39594) whose crystal structures have been reported [32, 33]. As presented in Table 5, most of amino-acid residues that occur in the active sites of bacterial enzymes are conserved in the maize enzyme except that Thr-96 and Gln-98 replaced Val and Glu, respectively, in HMP/HMP-P kinase domain and Val-523 replaced Ile in TMP synthase domain. The sequence identity of *S. typhimurium* HMP/HMP-P kinase and *B. subtilis* TMP synthase with the maize THI3 protein, ca. 39% as estimated by BLAST, was sufficient to use the three-dimensional structures of those bacterial enzymes as templates for modelling the THI3 structure. For this purpose, we used the SWISS-MODEL automatic homology-based modelling server of the Swiss Institute of Bioinformatics. The predicted overall structures of THI3 domains were very similar to those of the reference enzymes (Figure 5). Thus, the model of HMP kinase domain revealed a ribokinase-like fold with a core structure composed of a central eight strand  $\beta$ -sheet covered by two layers of five and three  $\alpha$ -helices [32]. The modelled TMP synthase domain had the triosephosphate isomerase fold, containing a core barrel composed of eight  $\alpha$ -helices and eight parallel  $\beta$ -strands [33, 39]. More importantly, the predicted arrangements of active-site residues within the three-dimensional protein structure basically match those in the reference bacterial proteins. In the HMP kinase domain of THI3 protein (Figure 5A), eight residues:

Gly-65, Ala-72, Asp-77, Thr-96, Gln-98, Met-134, Val-161 and Cys-268, may be involved in HMP binding. Gln-98 probably forms a direct hydrogen bond to 4-amino group of HMP and the other residues may interact with the pyrimidine ring indirectly, through fixed water molecules, or by making some hydrophobic contacts [32]. Three residues, Lys-230, Thr-266 and Lys-292 may contribute to a putative ATP-binding site. At least three residues in TMP synthase domain of THI3 (Figure 5B), Thr-470, Thr-472 and Lys-473, are shifted from the locations which the corresponding residues of *B. subtilis* protein occupy to make direct hydrogen bonds with thiazole phosphate group (Thr-156 and Thr-158) and van der Waals contacts with thiazole ethyl chain as well as electrostatic interactions with HMP-PP diphosphate group (Lys-159) [39]. A new location of Lys-473 may provide favourable interactions with thiazole phosphate group. Positions of other active centre residues of plant and bacterial proteins nearly match. Residues Ser-524, Gly-500 and Val-523 may constitute the phosphate-binding site, Arg-375, Lys-377, Asn-405, Asp-406, Asp-425 and Ser-444 may fix the diphosphate group and Gln-373 may interact with pyrimidine moiety, possibly *via* two hydrogen bonds to the ring N-3 and the 4-amino group [39].

We performed a preliminary experimental verification of the above models of the active centres of *Z. mays* HMP kinase/TMP synthase by mutations at selected residues. We exchanged two residues, Gln-98 and Met-134, at the HMP kinase site and three residues, Gln-373, Ser-444 and Thr-472 at the TMP synthase site. The enzymatic properties of Q98L, M134K and S444A mutants exactly followed the above structural predictions. Two first exclusively express full TMP synthase activity, and the last has only the HMP kinase activity. The last type mutation (S130A) had been extensively exploited during the structural studies of the *B. subtilis* TMP synthase to prove the essential role of Ser-130 residue in the catalytic mechanism [33, 39]. A full TMP synthase activity of T472D mutant is consistent with the predicted escape of Thr-472 from the bound ligand. A lack of effect of Q373L mutation is more difficult to explain, although the model of TMP synthase active site shows a possible rotation of amide group which may prevent the formation of hydrogen bonds to aminopyrimidine moiety. Unfortunately, there is no data available on the effects of similar mutations in the bacterial enzyme. Two double-mutants of THI3 protein, Q98L+Q373D and M134K+T472D, were additionally studied, with a single residue substituted at each active centre. As expected, both exhibited full TMP synthase activity and decreased HMP kinase activities, although the HMP kinase of M134K+T472D mutant exceeds the predictions from the single mutant data. A systematic study of the mutations of all residues which were



preliminarily classified in this work as occurring at active centres of THI3 protein is currently carried out in our laboratory on the level of purified mutant proteins.

Several features of recombinant bifunctional enzyme of *Zea mays* described in this paper implicate putative biological functions of the natural product of *thi3* gene. *In vivo*, the full-length THI3 polypeptide would be targeted to chloroplasts, within which the entire thiamine biosynthesis pathway is known to occur [14]. The matured two-domain enzyme with HMP/HMP-P and TMP synthase activities could catalyze two consecutive steps of thiamine biosynthesis that lead from HMP-P to TMP. An additional ability of THI3 to phosphorylate HMP in the presence of ATP does not appear to be important *in vivo* as no mechanisms by which HMP could enter chloroplasts have ever been discussed. This reaction is well known in various microorganisms as a salvage pathway. Bacterial TMP synthases were often suspected to play a regulatory function, primarily because of their susceptibility to inhibition by ATP, ADP and inorganic diphosphate [37]. The unique fusion of HMP-P kinase and TMP synthase activities into one molecule of higher plant protein such as THI3 may result in a more selective response to ATP as the TMP synthesis is quenched by HMP-PP (which is formed in the presence of ATP on the same enzyme and is immediately consumed for TMP formation). It may be that ATP levels regulate the yield of entire thiamine synthesis in higher plants through the inhibitory action on the bifunctional enzyme HMP/HMP-P kinase / TMP synthase. The THI3 orthologues occur in several higher plant species, strongly suggesting a specific role of the combination of two activities into one protein for the thiamine biosynthesis in higher plants.

## **Acknowledgements**

This work was partly supported by the grant No. 2P04C 017 27 from the Ministry of Science and Higher Education, Poland (to M.R.-K.).

## REFERENCES

1. Friedrich, W. (1987) Thiamin (Vitamin B<sub>1</sub>, aneurin). In *Hanbuch der Vitamine*, pp. 240-258, Urban and Schwarzenberg, München, Wien, Baltimore
2. Yamada, K. and Kawasaki, T. (1980) Properties of the thiamine transport system in *Escherichia coli*. *J. Bacteriol.* **141**, 254-261
3. Iwashima, A., Nosaka, K., Nishimura, H. and Enjo, F. (1997) Thiamin transporters in yeast. *Methods Enzymol.* **279**, 109-117
4. Begley, T. P., Downs, D. M., Ealick, S. E., McLafferty, F. W., Van Loon, A. P., Taylor, S., Campobasso, N., Chiu, H. J., Kinsland, C., Reddick, J. J. and Xi, J. (1999) Thiamin biosynthesis in prokaryotes. *Arch. Microbiol.* **171**, 293-300
5. Settembre, E., Begley, T. P. and Ealick, S. E. (2003) Structural biology of enzymes of the thiamin biosynthesis pathway. *Curr. Opin. Struct. Biol.* **13**, 739-747
6. Nosaka, K. (2006) Recent progress in understanding thiamin biosynthesis and its genetic regulation in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **72**, 30-40
7. Baker, L. J., Dorocke, J. A., Harris, R. A. and Timm, D. E. (2001) The crystal structure of yeast thiamin pyrophosphokinase. *Structure* **9**, 539-546
8. Kawasaki, Y. (1993) Copurification of hydroxyethylthiazole kinase and thiamine-phosphate pyrophosphorylase of *Saccharomyces cerevisiae*: characterization of hydroxyethylthiazole kinase as a bifunctional enzyme in the thiamine biosynthetic pathway. *J. Bacteriol.* **175**, 5153-5158
9. Jurgenson, C. T., Chatterjee, A., Begley, T. P. and Ealick, S. E. (2006) Structural insights into the function of the thiamine biosynthetic enzyme Thi4 from *Saccharomyces cerevisiae*. *Biochemistry* **45**, 11061-11070
10. Haas, A. L., Laun, N. P. and Begley, T. P. (2005) Thi20, a remarkable enzyme from *Saccharomyces cerevisiae* with dual thiamin biosynthetic and degradation activities. *Bioorg. Chem.* **33**, 338-344
11. Voskoboyev, A. I. and Ostrovsky, Y. M. (1982) Thiamin pyrophosphokinase: structure, properties and role in thiamin metabolism. *Ann. N. Y. Acad. Sci.* **378**, 161-176
12. Dorrestein, P. C., Huili Zhai, H., Taylor, S. V., McLafferty, F. W. and Begley, T. P. (2004) The biosynthesis of the thiazole phosphate moiety of thiamin (vitamin B<sub>1</sub>): the early steps catalyzed by thiazole synthase. *J. Am. Chem. Soc.* **126**, 3091-3096

13. Machado, C. R., Praekelt, U. M., de Oliveira, R. C., Barbosa, A. C. C., Byrne, K. L., Meacock, P. A. and Menck, C. F. M. (1997) Dual role for the yeast *THI4* gene in thiamine biosynthesis and DNA damage tolerance. *J. Mol. Biol.* **273**, 114-121
14. Belanger, F. C., Leustek, T., Chu, B. and Kriz, A. L. (1995) Evidence for the thiamine biosynthetic pathway in higher-plant plastids and their developmental regulation. *Plant Mol. Biol.* **29**, 809-821
15. Machado, C. R., Costa de Oliveira, R. L., Boiteux, S., Praekelt, U. M., Meacock, P. A. and Menck, C. F. M. (1996) *Thi1*, a thiamine biosynthetic gene in *Arabidopsis thaliana*, complements bacterial defects in DNA repair. *Plant Mol. Biol.* **31**, 585-593
16. Ribeiro, A., Praekelt, U. , Akkermans, A. D. L., Meacock, P. A., van Kammen, A., Bisseling, T. and Pawlowski, K. (1996) Identification of *agthi1*, whose product is involved in biosynthesis of the thiamine precursor thiazole, in actinorhizal nodules of *Alnus glutinosa*. *Plant J.* **10**, 361-368
17. Kim, Y. S., Nosaka, K., Downs, D. M., Kwak, J. M., Park, D., Chung, I. K. and Nam, H. G. (1998) A *Brassica* cDNA clone encoding a bifunctional hydroxymethylpyrimidine kinase/thiamin-phosphate pyrophosphorylase involved in thiamin biosynthesis. *Plant Mol. Biol.* **37**, 955-966
18. Godoi, P. H., Galhardo, R. S., Luche, D. D., Van Sluys, M. A., Menck, C. F. and Oliva, G. (2006) Structure of the thiazole biosynthetic enzyme THI1 from *Arabidopsis thaliana*. *J. Biol. Chem.* **281**, 30957-30966
19. Frohman, M. A. (1993) Rapid amplification of complementary DNA ends for generation of full-length complementary DNAs: Thermal RACE. *Methods Enzymol.* **218**, 340-356
20. Chong, S., Williams, K. S., Wotkowicz, C. and Xu, M.-Q. (1998) Modulation of protein splicing of the *Saccharomyces cerevisiae* vacuolar membrane ATP-ase intein. *J. Biol. Chem.* **273**, 10567-10577
21. Laemmli, U. K. (1970) Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
22. DiBella, E. P. and Hennessey, D. J. (1961) Isomeric S-methylthiamins. *J. Am. Chem. Soc.* **83**, 2017-2021
23. Morimoto, H., Hayashi, N., Naka, T. and Kato, S. (1973) Bildungsmechanismus von 4-Amino-5-aminomethyl-2-methyl-pyrimidin aus 3-Athoxy-2-(diathoxymethyl)propionitril und Acetamidin. *Chem. Ber.* **106**, 893-901

24. Brown, G. M. (1970) Preparation of the mono- and pyrophosphate esters of 2-methyl-4-amino-5-hydroxymethylpyrimidine for thiamine biosynthesis. *Methods Enzymol.* **18A**, 162-164
25. Lee, B. L., Ong, H. Y. and Ong, C. N. (1991) Determination of thiamine and its phosphate esters by gradient-elution high-performance liquid chromatography. *J. Chromatogr.* **567**, 71-80
26. Bisswanger, H. (2002) *Enzyme kinetics. Principles and methods.* Wiley-VCH, Weinheim
27. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402
28. Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680
29. Clamp, H., Cuff, J., Searle, S. M. and Barton, G. J. (2004) The Jalview alignment editor. *Bioinformatics* **20**, 426-427
30. Emanuelsson, O., Nielsen, H. and von Heijne, G. (1999) ChloroP, a neutral network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* **8**, 978-984
31. Marchler-Bauer, A. and Bryant, S. H. (2004) CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* **32**, 327-331
32. Cheng, G., Bennett, E. M., Begley, T. P. and Ealick, S.E. (2002) Crystal structure of 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate kinase from *Salmonella typhimurium* at 2.3 Å resolution. *Structure* **10**, 225-235
33. Chiu, H. J., Reddick, J. J., Begley, T. P. and Ealick, S. E. (1999) Crystal structure of thiamin phosphate synthase from *Bacillus subtilis* at 1.25 Å resolution. *Biochemistry* **38**, 6460-6470
34. Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* **31**, 3381-3385
35. Wallace, T. P. and Howe, C. J. (1993) Plant organellar targeting sequences In *Plant Molecular Biology Labfax* (Croy, R. R. D., ed.), pp. 287-292, Bios Scientific Publishers, Oxford

36. Reddick, J. J., Kinsland, C., Nicewonger, R., Christian, T., Downs, D. M., Winkler, M. E. and Begley, T. P. (1998) Overexpression, purification and characterization of two pyrimidine kinases involved in the biosynthesis of thiamin: 4-amino-5-hydroxymethyl-2-methylpyrimidine kinase and 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate kinase. *Tetrahedron* **54**, 15983-15991
37. Kawasaki, T. (1979) Thiamine phosphate pyrophosphorylase. *Methods Enzymol.* **62**, 69-73
38. Zhang, Y., Taylor, S. V., Chiu, H. J. and Begley, T. P. (1997) Characterization of the *Bacillus subtilis* *thisC* operon involved in thiamine biosynthesis. *J. Bacteriol.* **179**, 3030-3035
39. Peapus, D. H., Chiu, H. J., Campobasso, N., Reddick, J. J., Begley, T. P., and Ealick, S. E. (2001) Structural characterization of the enzyme-substrate, enzyme-intermediate, and enzyme-product complexes of thiamin phosphate synthase. *Biochemistry* **40**, 10103-10114

## TABLES

**Table 1 Primers used in this study**

Name of primer	Sequence	Description
MO1-F	TCGTGAAGACTGGGATGCT(TC)CC	Degenerate, forward primer
MO2-R	GTAGTC(CT)GCACCATCTTTCCA	Degenerate, reverse primer
ZTH51	GCTTCTTTCACATTTGGAGTGAC	5'RACE RT primer
ZTH52	GTAGCGAGAGTAGATGGTCC	First 5'RACE primer
ZTH53	GCTGGTGGACACCATGACC	"Nested" 5'RACE primer
ZTH31	GCGTCCATGTTGGCCAATC	First 3'RACE primer
ZTH32	CATACCAGCATGGGAAGTAC	"Nested" 3'RACE primer
Adaptor	GACTCGAGTCGACATCG	Adaptor 3' and 5'RACE primer
ZTH1-F	GGTGGTCATATGACGTCCGTACCGCTAC	Forward primer for PCR of <i>thi3</i> cloned into pTYB1 vector
ZTH1-R	GGTGGTTGCTCTTCCGCAAGACCTTGAGTGAGCATTCTTC	Reverse primer for PCR of <i>thi3</i> cloned into pTYB1 vector
ZTH11-F	GGTGGTTGCTCTTCCAACGCCACCCTTGGCTTCCGCA	Forward primer for PCR of <i>thi3</i> cloned into pTYB11 vector
ZTH11-R	GGTGGTCTCGAGCTAAGACCTTGAGTGAGCATTC	Reverse primer for PCR of <i>thi3</i> cloned into pTYB11 vector
MutQ98L	GCTGTCACCGCGCTGAACACCGTCGG	Mutagenetic primer, replaces Gln-98 to Leu-98
MutQ373L	GGCGCTACCATTGTCCTACTGAGAGAAAAAGACGCC	Mutagenetic primer, replaces Gln-373 to Leu-373
MutM134K	GGTGAAAACAGGGAAGCTCCCTTCAGCTGG	Mutagenetic primer, replaces Met-134 to Lys-134
MutT472D	GTGTCTTCCCGACCACGGACAAGGCAAACAATCCTACC	Mutagenetic primer, replaces Thr-472 to Asp-472
MutS444A	GGAAAAATCATCGGCGTCGCATGTAAGACCCTGG	Mutagenetic primer, replaces Ser-444 to Ala-444

**Table 2 Enzymatic activity of maize THI3 protein: the rates of TMP formation from various combination of substrates and cofactors**

Concentrations in the reaction mixture were: 15  $\mu$ M HMP, HMP-P, HMP-PP, HET-P or HET, 10 mM ATP and 10 mM  $MgCl_2$ . Samples containing 5  $\mu$ g protein/100  $\mu$ l in 50 mM Tris-HCl buffer pH 8.0 were incubated for 60 min at 37°C and then analyzed for TMP content by RP-HPLC with fluorogenic postcolumn derivatization. Values are means  $\pm$  standard deviations from at least 3 independent kinetic experiments.

Activity tested	Substrates and cofactors	Specific activity (nmol TMP/min·mg protein)
TMP synthase	HMP-PP, HET-P	$0.52 \pm 0.05$
	HMP-PP, HET-P, $Mg^{2+}$	$2.46 \pm 0.12$
	HMP-PP, HET-P, ATP	$0.05 \pm 0.02$
	HMP-PP, HET-P, $Mg^{2+}$ , ATP	$0.83 \pm 0.11$
HMP kinase /TMP synthase	HMP, HET-P, ATP, $Mg^{2+}$	$0.22 \pm 0.03$
HMP-P kinase /TMP synthase	HMP-P, HET-P, ATP, $Mg^{2+}$	$0.24 \pm 0.02$
HET kinase /TMP synthase	HMP-PP, HET, ATP, $Mg^{2+}$	0



**Table 3 Kinetic constants of TMP synthase and HMP kinase activities of THI3 protein**

The apparent kinetic parameters were determined at the second substrate concentration of 15  $\mu\text{M}$  (HMP-PP), 1.5  $\mu\text{M}$  (HET-P), 0.5 mM (HMP) or 5 mM (ATP).  $K_m$  – Michaelis constant;  $K_{iu}^{\text{HMP-PP}}$  – substrate (HMP-PP) uncompetitive inhibition constants for TMP synthase,  $K_{iu}^{\text{ATP}}$  – the constant for uncompetitive inhibition of TMP-synthase by ATP (determined against variable concentration of HET-P as the substrate),  $V_{\text{max}}$  - maximal velocity recalculated per 1 mg protein. The applied kinetic equations with these parameters are specified under Methods. Values are means  $\pm$  standard deviations; the number of independent kinetic experiments are given in parentheses.

Activity	Parameter	Parameter value
TMP synthase	$K_m^{\text{HMP-PP}}$ ( $\mu\text{M}$ )	$12.8 \pm 6.9$ (4)
	$K_{iu}^{\text{HMP-PP}}$ ( $\mu\text{M}$ )	$14.8 \pm 6.3$ (4)
	$V_{\text{max}}^{\text{HMP-PP}}$ (nmol TP/min·mg protein)	$8.5 \pm 3.9$ (4)
	$K_m^{\text{HET-P}}$ ( $\mu\text{M}$ )	$1.79 \pm 0.63$ (8)
	$V_{\text{max}}^{\text{HET-P}}$ (nmol TP/min·mg protein)	$2.66 \pm 0.22$ (8)
	$K_{iu}^{\text{ATP}}$ (mM)	$3.4 \pm 0.9$ (3)
HMP kinase	$K_m^{\text{HMP}}$ ( $\mu\text{M}$ )	$6.58 \pm 1.55$ (4)
	$V_{\text{max}}^{\text{HMP}}$ (nmol HMP-P/min·mg protein)	$20.1 \pm 1.90$ (4)
	$K_m^{\text{ATP}}$ ( $\mu\text{M}$ )	$7.24 \pm 1.80$ (7)
	$V_{\text{max}}^{\text{ATP}}$ (nmol HMP-P/min·mg protein)	$17.7 \pm 0.34$ (7)

**Table 4 Enzymatic activity of THI3 mutants**

Activity was determined in the lysates from *E. coli* cells transformed with plasmids containing THI3 sequence with indicated amino acid substitutions. The expression levels of all mutants were similar to that of wild-type protein as judged from immunochemical analysis using CBD-specific antibodies (results not shown). The activity were expressed relatively to that determined in the wild-type protein samples. Values given are ranges from at least three independent determinations.

Mutant	Relative activity (%)	
	TMP synthase	HMP kinase
Q98L	104-124	0-2
M134K	97-111	0-5
Q373L	96-113	64-75
S444A	0-1	80-91
T472D	101-126	46-54
Q98L+Q373L	93-105	7-13
M134K+T472D	83-97	19-32

**Table 5 Active centre residues of *S. typhimurium* HMP kinase, *B. subtilis* TMP synthase and the corresponding domains of *Z. mays* THI3 bifunctional enzyme**

For bacterial enzymes, the residues found in the closest neighbourhood of the bound ligands in the crystal structures [32, 33, 39] are listed, except the residues of putative ATP-binding site of *S. typhimurium* HMP kinase, modelled through the homology to other members of ribokinase family [32]. The corresponding putative active centre residues of *Zea mays* THI3 domains were identified by sequence alignment with CLUSTAL W program and structural modelling with the use of the SWISS-MODEL server.

Active centre	Ligand	<i>S. typhimurium</i>	<i>B. subtilis</i>	<i>Z. mays</i>
HMP kinase	HMP	Gly-11		Gly-65
		Ala-18		Ala-72
		Asp-23		Asp-77
		Val-42		Thr-96
		Glu-44		Gln-98
		Met-80		Met-134
		Val-107		Val-161
		Cys-213		Cys-268
	ATP	Lys-176		Lys-230
		Asp-187		Asp-242
		Arg-202		Arg-257
		Thr-211		Thr-266
		Lys-237		Lys-292
TMP synthase	HMP-PP		Gln-57	Gln-373
			Arg-59	Arg-375
			Lys-61	Lys-377
			Asn-92	Asn-405
			Ser-130	Ser-444
			Lys-159	Lys-473
	HET-P		Thr-156	Thr-470
			Thr-158	Thr-472
			Gly-188	Gly-500
			Ile-208	Val-523
	Mg <sup>2+</sup>		Ser-209	Ser-524
			Asp-93	Asp-406
			Asp-112	Asp-425

## LEGENDS TO SCHEMES AND FIGURES

### Scheme 1 Proposed scheme for thiamine and TDP biosynthesis in *Eukaryota*

The biosynthesis pathways may differ in some details between yeasts and higher plants, as exemplified by different substrates (underlined) used by plants for the synthesis of thiazole moiety (HET-P). The sequenced *genes* and purified proteins, unequivocally assigned to thiamine biosynthesis in baker's yeast and higher plants (underlined) as specified. The symbols of proteins whose three-dimensional structures were reported are bolded. A – *Arabidopsis thaliana*, Z – maize, b – *Brassica napus*, ag – *Alnus glutinosa*.

### Figure 1 SDS-PAGE patterns from a representative purification of maize THI3 protein overproduced in *E. coli*

The electrophoresis was carried out in the Laemmli system [21] with 10% resolution gel. Protein bands were visualized by Coomassie Brilliant Blue R-250 staining. Left lane: molecular mass standards, lane 1: transformed cell lysate before IPTG induction; lane 2: transformed cell lysate after IPTG induction for 16 h at 16°C; lane 3: sample after chitin-affinity chromatography; lane 4: sample after Superdex-200 gel filtration.

### Figure 2 Kinetic characteristics of TMP synthase activity of THI3 protein

(A) Dependences of TMP-formation rate on concentrations of substrates. (B) Hanes-Woolf plots of the dependences of TMP-formation rates on HET-P concentration in the presence of ATP. In A, the line drawn through data points for the HMP-PP concentration dependence was generated with  $K_{iu}=24\ \mu\text{M}$  found from Dixon plot, and  $K_m=6.35\ \mu\text{M}$  and  $V_{max}=5.05\ \text{nmol TMP per min per mg protein}$ , estimated by fitting the equation for the uncompetitive inhibition by excess substrate. For the HET-P concentration dependence, the classic hyperbola with  $K_m=1.74\ \mu\text{M}$  and  $V_{max}=2.52\ \text{nmol TMP per min per mg protein}$ , was fitted. In B, from the Hanes-Woolf plot obtained in the absence of ATP, the  $K_m$  and  $V_{max}$  values for HET-P were estimated as  $2.55\ \mu\text{M}$  and  $2.88\ \text{nmol TMP per min per mg protein}$ , respectively, and the  $K_{iu}$  for ATP ( $4.3\ \text{mM}$  in this example) was obtained from the abscissa intercept of the secondary plot of the slope *versus* the ATP concentration.

### Figure 3 Kinetic characteristics of HMP kinase activity of THI3 protein

A dependence of HMP-P formation rate on concentrations of HMP and ATP substrates. The production of HMP-P was directly quantified by RP-HPLC with spectrophotometric

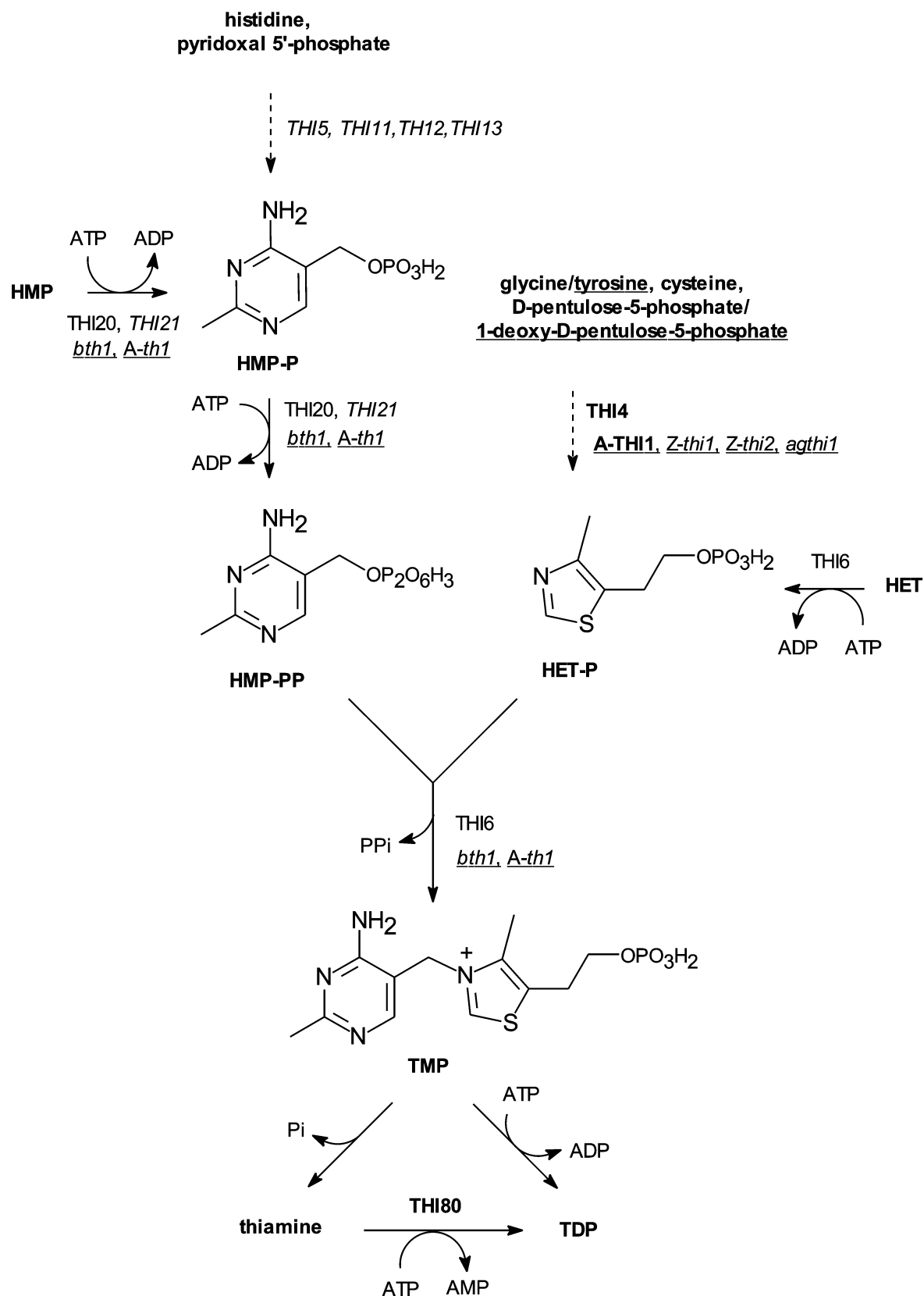
detection. Ordinary hyperbolic fits are shown with  $K_m$  and  $V_{max}$  values of 7.46  $\mu$ M and 20.5 nmol HMP-P per min per mg protein, respectively, for HMP concentration dependence, and of 6.27  $\mu$ M and 17.4 nmol HMP-P per min per mg protein, respectively, for the dependence on ATP concentration.

**Figure 4 Multiple alignment of the amino acid sequences of THI3 and its orthologues occurring in various plant species**

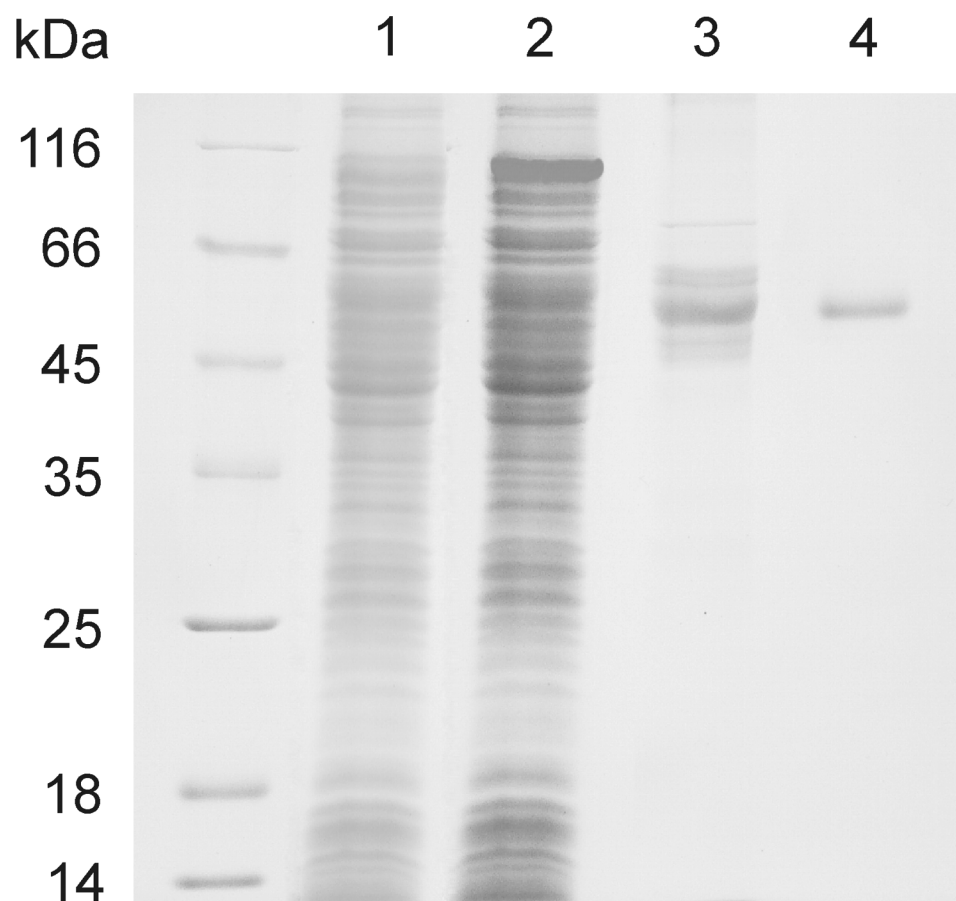
The accession codes of the sequences compared were emb/CAJ45026 (*Zea mays*, this work), gb/AAM91567 (*Arabidopsis thaliana*), gb/AAC31298 (*Brassica napus*, BTH1), gb/ABE87727 (*Medicago truncatula*) and gb/ABA96049 (*Oryza sativa*). The alignment was performed using the CLUSTALW program and the result visualized in JalView. Black, dark grey and light grey highlights mark the positions in the primary structures where all 5, 4 of 5 and 3 of 5 residues are identical, respectively. The box at the N-terminus of *Zea mays* THI3 protein includes the putative 36-amino acid residue chloroplast transit peptide predicted by ChloroP program [25].

**Figure 5 The models of overall structures and arrangements of putative active centre residues of THI3 HMP/HMP-P kinase (A) and TMP synthase (B) domains, overlaid on the corresponding features of three-dimensional structures of *S. typhimurium* HMP/HMP-P kinase [32] and *B. subtilis* TMP synthase [33]**

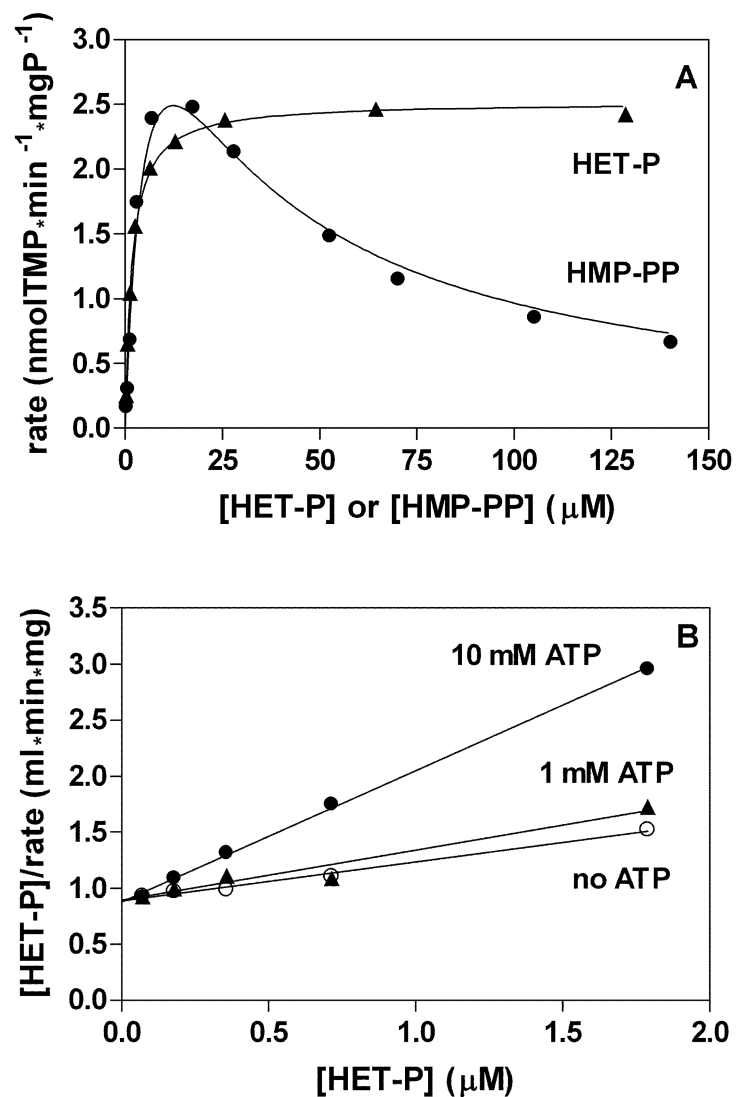
The models of THI3 domains were obtained with the SWISS-MODEL server and all structures were visualized using the Swiss-PdbViewer and PyMOL programs. The ribbon diagrams of the overall architectures of THI3 domains (in violet) and the corresponding bacterial enzymes (in orange) are shown in background. Stick representations of the active site residues in the bacterial proteins and corresponding putative active site residues of THI3 domains are marked with similar colours, lighter for THI3. Additionally, the THI3 residues are indicated by solid arrows and labelled with normal font, and the bacterial protein residues with dashed arrows and italic font. Dark grey stick models of substrates, HMP in A, and HMP-PP and HET-P in B, are shown in the positions they occupy in the crystal structures of bacterial enzymes.



**Scheme 1**

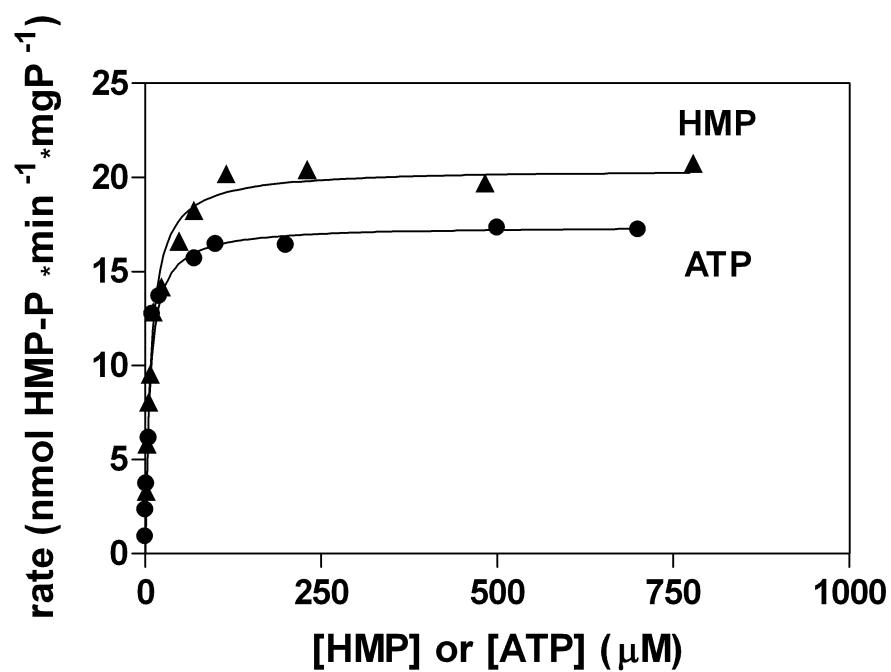


**Figure 1**



**Figure 2**

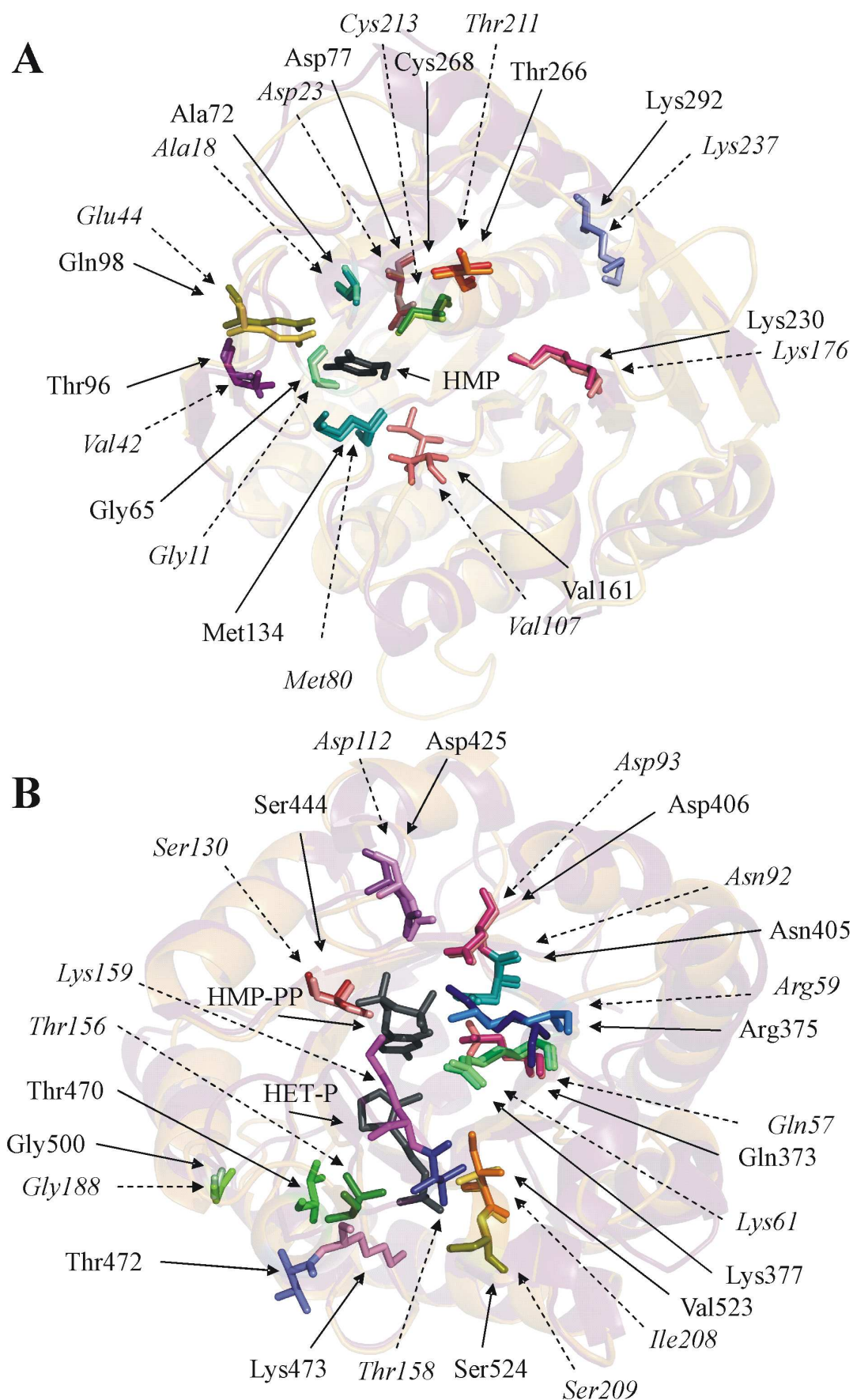




**Figure 3**

<i>Zea/1-551</i>	1	MTSVPLPPLPAALAQYRPSAAATLGFRTSPSSAPGVSAASSSTRARRFSVIAAS - - - - -	51
<i>Arabidopsis/1-522</i>	1	-----MNSLGGIRSWPANWRSTTASMTTT - - -	24
<i>Brassica/1-523</i>	1	-----MQSLGGIRSWPATWRTTASMTTTT - -	26
<i>Medicago/1-578</i>	1	MMMSCVFNVDVDFSSHYSPLHLHLHLHAHSQITPPLLSKRAINRFWNCNSDSINHQTPRF	58
<i>Oryza/1-548</i>	1	MAAAPQQSVHPSL - - - -PSSTSTLRLLISS-SPRRPPPPPPRRARRYNRLAASASA - -	50
<i>Zea/1-551</i>	52	-----AREMPWPHVLTVAGSDSSAGAGIQADIKACAALGAYCS	89
<i>Arabidopsis/1-522</i>	25	-----ESVRKVPQVLTVAGSDSGAGAGIQADLKVCAARGVYCA	62
<i>Brassica/1-523</i>	27	-----ESVRKVAQVLTVAGSDSGAGAGIQADIKVCAARGVYCA	64
<i>Medicago/1-578</i>	59	VKMQSNITSNSIPTPIHTFDSSELNKIPHVLTVAGSDSGAGAGIQADLKTCSARRVYCS	116
<i>Oryza/1-548</i>	51	-----AREMPWPHVLTVAGSDSGGGAGIQADIKACAALGAYCS	88
<i>Zea/1-551</i>	90	SAITAVTAQNTVGVQGIHAVPEKFVGEQLRSVLSDMSSVDVVKTGMLPSAGVVKVLCS	147
<i>Arabidopsis/1-522</i>	63	SVITAVTAQNTRGVQSVHLLPPEFISEQLKSVLSDFEFDDVVKTGMLPSTEIVEVLLQN	120
<i>Brassica/1-523</i>	65	SVKTAVKAKNTRAVQSVHLLPPDSVSEQLKSVLSDFEFDDVVKTGMLPSPEIVEVLLQN	122
<i>Medicago/1-578</i>	117	TVITAVTAQNTLGVQGVNIIPHDFVQHQLNSVLS DINVDVVKTGMLPSLSVLKVLCS	174
<i>Oryza/1-548</i>	89	SVTAVTAQNTAGVQGIHVVPPEFIREQLNSVLSDMSSVDVVKTGMLPSIGVVRVLCES	146
<i>Zea/1-551</i>	148	LRKFPVKALVVDPMVSTSGDTLSGPSTLATYRDELFSMADIVTPNVKEASKLLGDVS	205
<i>Arabidopsis/1-522</i>	121	LSDFPVRALVVDPMVSTSGHVLGSSILSIFRERLLPIADITPNVKEASALLDGFR	178
<i>Brassica/1-523</i>	123	LSEYPVRALVVDPMVSTSGHVLGSSILSIFRERLLPIADITPNVKEASALLGGVR	180
<i>Medicago/1-578</i>	175	LRKFPVKALVVDPMISTSGDILAGPSVLDGFRRELLPMADIVTPNVKEASALLGDLP	232
<i>Oryza/1-548</i>	147	LKRFVKALVVDPMVSTSGDTLSESSSTLSVYRDEL FAMADIVTPNVKEASRLGGVS	204
<i>Zea/1-551</i>	206	LHTISDMRNAAESIYKLGPKYVLVKGDDMPDSSDAIDVLFDGKEFTELRGLRIKTRNT	263
<i>Arabidopsis/1-522</i>	179	IETVAEMRSAAKSLHEMGPRFVLVKGDDL PDSSDSVDVYFDGKEFHRLSPRIATRNT	236
<i>Brassica/1-523</i>	181	IQTVAEMRSAAKSLHQMGRFVLVKGDDL PDSSDSVDVYFDGNEFHRLSPRIATRNT	238
<i>Medicago/1-578</i>	233	IKTVSDMRATAAKLIHDLGPRSVLVKGDDL PNSSAIDIFYDGQEFYELSSPRVNTRNT	290
<i>Oryza/1-548</i>	205	LRTVSDMRNAAESIYKFGPKHVLVKGDDMLESSDATDVFFDGKEFTELHAHRIKTHNT	262
<i>Zea/1-551</i>	264	HGTGCTLASCIAAELAKGATMLHAVQAAKKFVESALYHSDKLVIGNGPQGPFDFHFL	321
<i>Arabidopsis/1-522</i>	237	HGTGCTLASCIAAELAKGSSMLSAVKVAKRFVDNALDYSKDIVIGSGMQGPFDFHFL	294
<i>Brassica/1-523</i>	239	HGTGCTLASCIAAELAKGSSNMLSAVKVAKRFVDSALNYSKDIVIGSGMQGPFDFHFL	296
<i>Medicago/1-578</i>	291	HGTGCTMASCIAAELAKGSSMLSAVKIAKRFVEAAL EYSRDLLIGNGVQGPFDHLLAL	348
<i>Oryza/1-548</i>	263	HGTGCTLASCIASELAKGATMLHAVQVAKNFVESALHHSKDLVVGNGPQGPFDFHFL	320
<i>Zea/1-551</i>	322	KSPSYKMGSLQKFNPDDLFLYAVTDSGMNKKWGRS IKDAVKAAIEGGATIVQLREKDA	379
<i>Arabidopsis/1-522</i>	295	KK-DPQSSRCSIFNPDDLFLYAVTDSRMNKKWNRSIVDALKAAIEGGATIIQLREKEA	351
<i>Brassica/1-523</i>	297	KK-DPQSYRQSTFKPDDLFLYAVTDSRMNKKWNRSIVDAVKAAIEGGATIIQLREKEA	352
<i>Medicago/1-578</i>	349	KNINQSSYRQDRFNPDNDFLYAVTDSGMNKKWGRSIAEAVKAAVEGGATIVQLREKDA	406
<i>Oryza/1-548</i>	321	KCPPYNVGSQPSFKPDQLFLYAVTDSGMNKKWGRS IKEAVQAAIEGGATIVQLREKDS	378
<i>Zea/1-551</i>	380	ETREFLEAAKACVEICKSLGVPLL INDRV DVALACDAGGVHVQSDIPAEVVRGLLGP	437
<i>Arabidopsis/1-522</i>	352	ETREFLEAAKACIDICRSHGVSL INDRIDIALACDADGVHVQSDMPVDLVRSLG	409
<i>Brassica/1-523</i>	353	ETREFLEAAKSCVDICRNGVCL INDRFDIAIALDADGVHVQSDMPVDLVRSLG	410
<i>Medicago/1-578</i>	407	ETKDFVDAKVKCLKICRSYGVPLL INDRIDVALACDADGVHVQSDMPARLARTILGP	464
<i>Oryza/1-548</i>	379	ETREFLEAAKACMEICKSSGVPLL INDRV DIALACNADGVHVQQLDMSAHEVRELLGP	436
<i>Zea/1-551</i>	438	GKIIIGVSCKTLAQAEQAWKDGADYIGCGGVFPTTTKANNPTLGFEGRLTVCLASKLPV	495
<i>Arabidopsis/1-522</i>	410	DKIIIGVSCKTPEQAHQAWKDGADYIGSGGVFPTNTKANNRTIGLDGLKEVCASKLPV	467
<i>Brassica/1-523</i>	411	DKIIIGVSCKTQEQAHQAWKDGADYIGSGGVFPTNTKANNRTIGLDGLREVCKASKLPV	468
<i>Medicago/1-578</i>	465	EKIIIGVSCKTPEHAHQAWLDGADYIGSGGVYPTNTKENNRTIGLDGLKEVCASKLPV	522
<i>Oryza/1-548</i>	437	GKIIIGVSCKTPAQAAQAWNDGADYIGCGGVFPTSTKANNPTLGFEGRLTVCLASKLPV	494
<i>Zea/1-551</i>	496	VAIGGINAGNAGSVMELGLPNLKGAVVVSALFDRERVAAETRNLRSILMKNAHSRS	551
<i>Arabidopsis/1-522</i>	468	VAIGGIGISNAGSVMQIDAPNLKGAVVVSALFDQDCVL TQAKKLHKTLESKRGI -	522
<i>Brassica/1-523</i>	469	VAIGGIGISNAESVMRIGEPNLKGAVVVSALFDQECVL TQAKKLHKTLESKREH -	523
<i>Medicago/1-578</i>	523	VAIGGIGMSNARAVMELGVPSLKGAVVVSALFDRECILTETRNMQAVISEAALLTQ	578
<i>Oryza/1-548</i>	495	VAIGGINASNAGSVMELGLPNLKGAVVVSALFDRPSVVAETRNMKSILNTNTSRT - -	548

Figure 4



**Figure 5**