Journal of Advanced Bioinformatics Applications and Research ISSN 0976-2604
Vol 2, Issue 2, June-2011, pp 167-172 http://www.bipublication.com

CHARACTERIZATION OF HYDROXYMETHYL PYRIMIDINE MONOPHOSPHATE KINASE FOR THIAMIN SYNTHESIS IN WHEAT

Sneha C, Narayan Agarwal, Ramya B, G.V. Ramana, K.V. Chaitanya
Department of Biotechnology, GIT, GITAM University, Visakhapatnam- 530045 (INDIA)

ABSTRACT
Thiamin (Vitamin B1) plays an important role in maintaining human health by aiding in the metabolism of carbohydrate, fats and amino acids, deficiency of which causes beriberi, and Parkinson's, Alzheimer's and Huntington's which have no permanent cure. Thiamin is synthesized by plants, bacteria and yeast. In plants, synthesis of thiamine involves Hydroxymethyl pyrimidine monophosphate kinase (HMP-P kinase, EC 2.7.1.49) and thiamin pyrophospho kinase (TPK, EC 2.7.6.2), which is rate limiting step in this process. The present work involves the identification of gene for the HMP-P kinase and TP kinase enzymes in wheat and other cereals using *Arabidopsis* genome, phylogenetic analysis of the gene, its differences in protein and secondary structure prediction.

Key words: Thiamin, HMP-P Kinase, TP kinase, Cereals, Wheat, *Arabidopsis.*

I. INTRODUCTION:
Vitamins are natural substances found in the plants and animals and are considered as essential nutrients for human beings. Thiamin is one of the eight water soluble vitamin B-complex members whose coordinated function provides health living conditions for human beings. Thiamin is synthesized by bacteria, plants and fungi. Humans must obtain thiamin in their diet. Deficiency of thiamin causes disease beriberi, a potentially lethal disturbance of the central nervous and circulatory systems due to accumulation of pyruvic acid and lactic acid and also causes other neurodegenerative diseases such as Parkinson's, Alzheimer's and Huntington's which have no permanent cure [1]. In plants, thiamin is synthesized in the leaves and is transported to the root system, where it promotes root development. Synthesis of thiamine confers systemic acquired resistance on susceptible plants, leading to rapid counterattack against pathogens [2]. Thiamin also acts as a cofactor of enzymes in key cellular metabolic pathways such as citric acid cycle, glycolysis, acetyl-CoA synthesis, the tricarboxylic acid cycle, anaerobic ethanolic fermentation, the oxidative pentose phosphate pathway, the Calvin cycle, the branched–chain amino acid pathway, and plant pigment biosynthesis and pentose phosphate pathway [3].

Synthesis of thiamin in plants involves the independent syntheses of two substituted thiazole and pyrimidine compounds, 4-methyl-5-(2-hydroxyethyl) thiazole phosphate (HET-P) and 4-amino-5-hydroxymethyl-2-methylpyrimidine diphosphate, which are coupled to form thiamine monophosphate (TMP). The first step involves a complex chemical rearrangement of 5-aminoimidazole ribonucleotide (AIR) to HMP-P which is catalyzed by HMP-P synthase [4]. It requires S-adenosylmethionine (SAM) and reduced nicotinamide. HET-P synthase catalyzes the formation of thiazole moiety from Nicotinamide Adenine Dinucleotide (NAD), glycine and a yet to be identified sulphur donor. Phosphorylation of HMP-P to HMP-PP and condensation of HET-P and HMP-PP to form TMP is catalyzed by a bifunctional enzyme in plants. This enzyme possesses the activities of both HMP-Kinase and TMP-PPase. The TMP formed due to the condensation of HET-P and HMP-PP is dephosphorylated to thiamine.

Thiamine is then pyrophosphorylated to thiamine diphosphate which is catalyzed by Thiamine pyrokinase (TPK) (Figure 1). Thiamine diphosphate thus formed binds to metabolite binding domain within certain messenger RNA's called riboswitches [5]. This leads to allosteric rearrangement of the messenger RNA structures that result in modulation of gene expression and protein production.

Interest in thiamine homeostasis is rapidly expanding because abnormalities in thiamine-dependent enzymes and diminished metabolism leads to several neurodegenerative disorders [6]. Diminished thiamine-dependent processes and abnormal thiamin metabolism accompany neurodegeneration in Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's disease, Wernicke–Korsakoff Syndrome, progressive supranuclear palsy (PSP) and the adult-onset neurodegenerative diseases [7, 8]. In addition, abnormalities of thiamine-metabolic enzymes contribute to thiamine-responsive maple syrup urine disease[9], Leigh's disease [10], sudden infant death syndrome [11], cerebellar degeneration [12], thiamine-responsive anemia [13]. The molecular and cellular basis of the reductions in thiamine-dependent enzymes is unknown [14]. One of the possible ways to control the reductions in thiamin level is the healthy diet or increased thiamin production in the staple food crops. In India, three fourth of the population consumes wheat as the staple food which is one of the sources of the thiamin in humans (1.882 mg^{-1} 100 g). India is one of the leading wheat producing and wheat consuming countries in the world. It is cultivated as a spring as well as winter cereal and is also cultivated in many developing countries. The objective of the present investigation is to identify and characterize the genes for the thiamin synthesis in cereals especially wheat.

II. MATERIALS AND METHODS:

Source sequences

The nucleotide sequence of HMP-P kinase (Acc. No. AT1G22940) of Arabidopsis was obtained from the site www. tair.org and were used as the source sequences for obtaining the required sequences from wheat and other cereal plants. Using these genes as query, nucleotide BLAST was performed .

Identification of genes for HMP-P kinase

Nucleotide BLAST was performed for the identification of the ESTs homologous to the enzyme in the site http://compbio.dfci.harvard.edu/cgi_bin/tgi/Blast/index.cgi by selecting wheat, maize, rice, sugar cane, barley plants respectively and the ESTs with high homology was pooled. Inorder to determine the homology at the protein level, these ESTs were translated in the EXPASY proteomics server and their homology at the protein levels were calculated and the EST showing high homology was shortlisted for the phylogenetic analysis.

Phylogenetic analysis

For further confirmation of the genes corresponding to HMP-P kinase, the proteins of these genes were aligned using multiple sequence alignment and the alignment product was subjected to phylogenetic tree construction using UPGMA. The HMP and TPK genes of wheat, separated along with Arabidopsis were considered for further studies of comparative homology and amino acid differences.

Comparision of Arabidopsis and wheat HMP-P kinase proteins

The open reading frames of these genes were compared for their homology and amino acid differences in protein using BOX SHADE SERVER (http://www.ch.embnet.org/ software/ BOX_form.html).

Prediction of secondary structure

The probable secondary structure was predicted for the HMP-P Kinase gene of Arabidopsis and

wheat using PSIPRED Protein Structure Prediction Server http://bioinf.cs.ucl.ac.uk/psipred/. .

III.RESULTS AND DISCUSSION:

Thiamin biosynthetic pathway in plants is known to feature many similarities to both plants and yeast pathways and many differences with bacteria [15]. In all the three cases, the thiazole and pyramidine moieties are synthesized in separate branches of the pathway and coupled to form thiamin phosphate, the active form of the co-factor. The metabolic pathway for the production of thiamin has been identified with little information regarding their genes controlling their activities. In the present study the genes for the enzyme HMP-P kinase has been identified from the databases using their Arabidopsis sequence as query and the ESTS with high homology from the cereals wheat, maize, rice, sugar cane, barley were selected for the study (Table 1 and 2). In wheat, HMP gene of Arabidopsis on nucleotide BLAST has given two ESTs with accession numbers TC409723 and TC394520 which is covering the query sequence from 100 bases to 1680 bases (Figure 2). These two ESTs are having an overlap of 200 bases. These ESTs were studied for the similarity and there was 92% identity and 3% gaps between these two ESTs at the regions of 191-417 bases and 877-1103 bases. The overlapping regions were removed to give the wheat HMP-P kinase gene. The obtained HMP gene of wheat is 2204 bases long and the open reading frame is 728 bases starting from 970 to 1698 bp. The protein of the Wheat HMP-P kinase is has shown avery high homology with barley (89%), maize and rice (79%), Arabidopsis (80%) HMP-P kinase genes (Data not shown), which confirms that the predicted protein belongs to wheat.

Phylogenic trees are reconstructed to elucidate the functional relationship within living cells [16]. To identify the relationship between these genes these sequences were subjected to phylogenetic analysis by UPGMA (Figure 3). In our analysis Wheat HMP-Pkinase gene was placed in the same clade along with sugar cane EST SCRFL3008D06 and Arabidopsis HMP-P kinase gene along with barley. To study the amino acid pattern in these genes box shade analysis has been done (Figure 4) in which the amino acids are homologous in wheat, sugar cane EST SCRFL3008D06 and Arabidopsis, whereas Barley and EST TC149008 of sugarcane had the common pattern of amino acids in their proteins.

Secondary structure of a protein will play primordial role in bioinformatics for the purpose of analysis and research in drug development. In the present study, the secondary structure for the HMP proteins of wheat were predicted using PSIPRED prediction server and compared with the secondary structure of Arabidopsis. The secondary structure of the wheat HMP-P kinase protein (Figure 5) showed differences with Arabidopsis with respect to the amino acids which forms strands, coils and helices (Table 2). There are five binding sites in this protein which are occupied by the amino acids PHE at 34, GLN at 65, ARG at 67 and LYS at 69 and ASN at 97 respectively[17]. In this protein there are three heterogens among which MG is metallic, whereas FMN and NAD are non- metallic (Table 4).

IV.CONCLUSIONS:

In this study the probable sequence of the HMP-P kinase gene and its protein structure has been elucidated using the Arabidopsis gene as the model. There is a high rate of similarity between these two genes which has been identified through protein secondary structure prediction analysis. Such studies are useful in identification of the genes from the databases and their isolation and expression analysis. Further studies are in progress in studying the expression of this gene in cereals and comparative analysis with

dicotyledonous plants to identify the plants with probability for the manipulation of thiamin metabolism.

V. REFERENCES:

1. Chatterjee A, Jurgenson CT, Schroeder FC, Ealick SE, Begley TP: Biosynthesis of the Thiamin Thiazole in Eukaryotes: The Conversion of NAD to and Advanced Intermediate. *J Am Chem Soc* 2007, 129:126-134.
2. Ahn IP, Kim S, Lee Y-H: Vitamin B1 functions as an activator of plant disease resistance. *Plant Physiology* 2005, 138: 1505–1515.
3. Maria R-K, Ewa K, Katarzyna O: Modulation of thiamine metabolism in Zea mays seedlings
under conditions of abiotic stress. *Journal of Experimental Botany* 2008, 56: 4133–4143.
4. Raschke M, Bu¨rkle L, Mu¨ller N, Nunes-Nesi A, Fernie AR,Arigoni D, Amrhein N, Fitzpatrick TB: Vitamin B1biosynthesis in plants requires the essential iron sulfur cluster protein, THIC. *Proceedings of the National Academy of Sciences USA* 2007, 104: 19637–19642.
5. Ajjawi I, Rodriguez Milla MA, Cushman J, Shintani DK: Thiamin pyrophosphokinase is required for thiamin cofactor activation in Arabidopsis. *Plant Molecular Biology* 2007, 65: 151–162.
6. Gibson G E, Park LC, Zhang H., Sorbi S, Calingasan NY 1999. Oxidative stress and a key metabolic enzyme in Alzheimer brains, cultured cells, and an animal model of chronic oxidative deficits. *Ann New York Acad Sci* 1999, 893: 79–94.
7. Albers DS, Augood SJ, Park LC, Browne SE, Martin DM., Adamson J, Hutton M., Standaert DG, Vonsattel JP, Gibson GE, Beal MF: 2000. Frontal lobe dysfunction in progressive supranuclear palsy: evidence for oxidative stress and mitochondrial impairment. *J Neurochem* 2000, 74: 878–881.
8. Cooper AJ, Sheu KF, Burke JR, Strittmatter WJ, Gentile VPG, Blass JP: Pathogenesis of inclusion bodies in (CAG)n/Qn-expansion diseases with special reference to the role of tissue transglutaminase and to selective vulnerability. *J Neurochem* 1999, 72: 889–899.
9. Elsas LJ, Danner DJ: 1982. The role of thiamine in maple syrup urine disease. *Ann New York Acad Sci* 1982, 378: 404–421.
10. Pincus JH: Subacute necrotizing encephalomyelopathy (Leigh's disease): a consideration of clinical features and etiology. *Dev Med Child Neuro* 1972, 14: 87–101.
11. Read DJ: The aetiology of the sudden infant death syndrome: current ideas on breathing and sleep and possible links to deranged thiamine neurochemistry. *Aust New Zealand J Med* 1978, 8: 322–336.
12. Poloni M, Patrini C: Cerebrospinal fluid, plasma and erythrocyte thaimine levels in ataxias of degenerative origin. *Ital J Neurol Sci* 1984, 4: 141–143.
13. Haworth C, Evans DI, Mitra J, Wickramasinghe SN: Thiamine responsive anaemia: a study of two further cases. *Br J Haematol* 1982, 50: 549–561.
14. Gibson GE, Park LC, Zhang H, Sorbi S, Calingasan NY: Oxidative stress and a key metabolic enzyme in Alzheimer brains,cultured cells, and an animal model of chronic oxidative deficits. *Ann New York Acad Sci* 1999, 893: 79–94.
15. Mirinda-Rio's J, Claudia M, Hermene GT, Aracelida V, Sergio EN, Jaime M,Mario SN: Expression of Thiamin Biosynthetic Genes (*thiCOGE*) and Production of Symbiotic Terminal Oxidase *cbb*3 in *Rhizobium etli*. *Journal of Bacteriology* 1997, 179: 6887-6893.
16. Rajeswari P, Apparao A: Normalized distance matrix method for construction of phylogenetic trees using new compressor-DNABIT compress. *Journal of Advanced Bioinformatics Applications and Research* 2011, 2: 89-97.
17. Wass MN, Kelley LA, Sternberg MJ: 3DLigandSite: predicting ligand-binding sites using similar structures.*NAR* 2010, 38: W469-73.

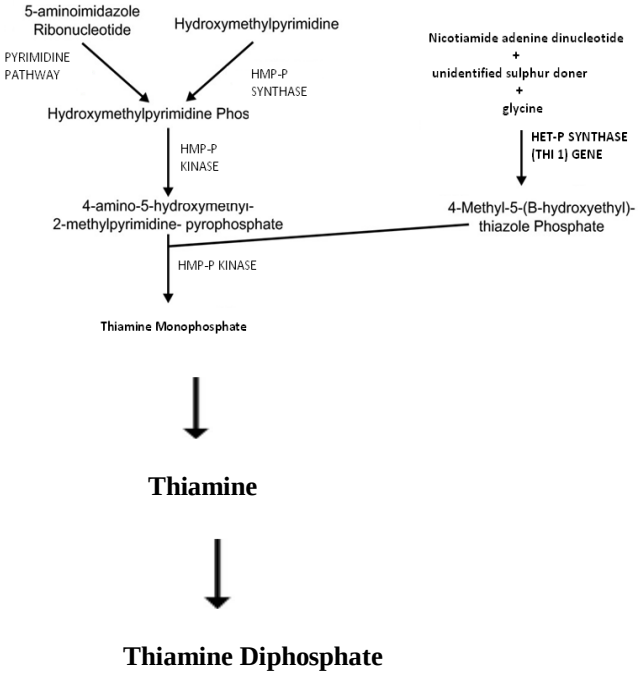

Figure 1: Thiamin biosynthesis pathway in plants

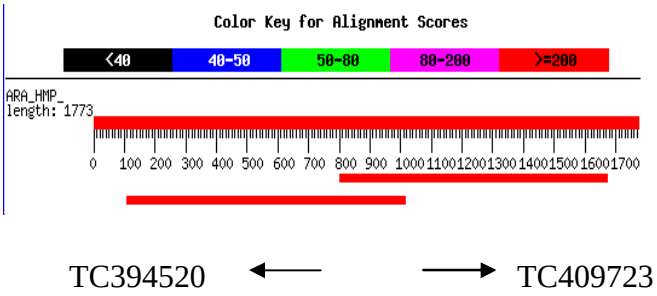

Figure2: Distribution of EST TC409723 and TC394520 of wheat against the query *Arabiodpsis*

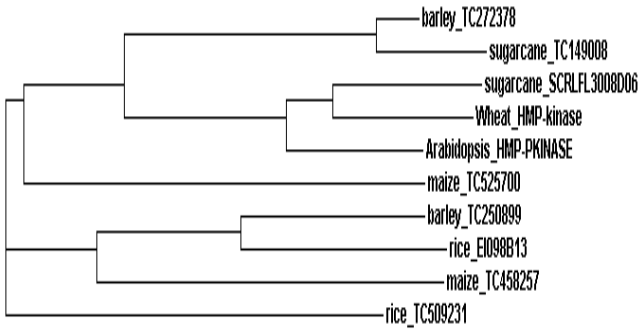

Figure 3: Phylogenetic tree of HMP-Pkinase genes

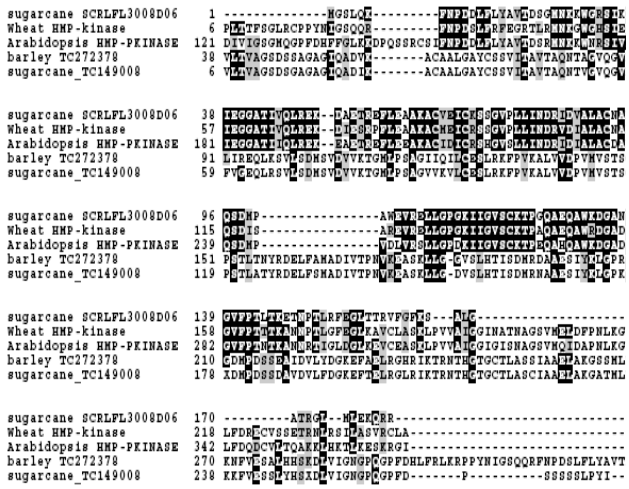

Figure 4: Box shade analysis of HMP-Pkinase proteins of *Arabidopsis* clade

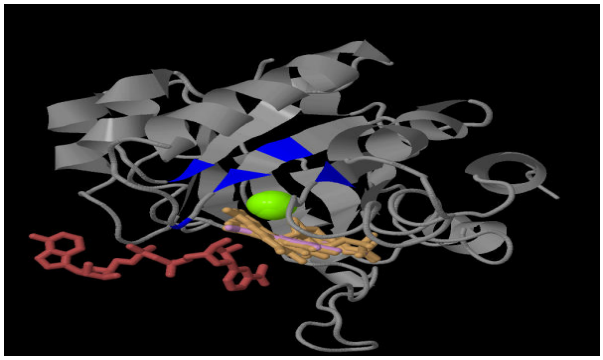

Figure 5: Secondary structure of the wheat HMP-P kinase. BLUE-Residues of binding sites; LIGHT GREEN-Heterogens; OTHER WIRE FRAMES:Non metallic heterogens

S.No	PLANT	EST ACCESSION No.	IDENTITY (%)
1.	Wheat	TC 409723	67
		TC394520	67
2.	Rice	TC509231	68
		EI098B13	70
3.	Maize	TC458257	67
		TC525700	68
4.	Sugar cane	TC149008	66
		CA226622	67
5.	Barley	TC272378	67
		TC250899	67

Table 1: Homology of the ESTs with HMP gene of Arabidopsis

S.No.	STRUCTURE	AMINO ACIDS (in no.)
1.	HELIX	**Wheat-** 2-5,8-13, 46-53,62-74,97-103,118-131 **Arabidopsis**- 5-15,40-48,61-68,76-88,135-146,151- 168,223-232-240-241,247-264,274-279,294-300, 314-322, 347-356, 373-376,398-413
2.	STRAND	**Wheat**- 34-39,56-61,88-90,109-112,137-140 **Arabidopsis**- 23-25, 29-30, 55-57,93-96,109-113,116-121,296-210,237-239,268-271,306-309,327-330,361-364,385-391
3.	COIL	**Wheat**-1,6-7,14-33,40-45,54-55,75-87,91-96,104-108,113-117,132-136,141-142 **Arabidopsis**-1-4,16-22,26-28,31-39,47-54,58-60,69-75,89-92,97-108,114-115,122-134,147-150,167-205,211-222,233-236,242-246,265-267,272-273,280-293,301-305,310-313,323-326,331-346,357-360,365-371,377-384,392-396,414-416.

Table2: Secondary structure of Wheat HMP-P kinase

S.NO.	RESIDUE	POSITION	CONTACT	AV DISTANCE	JS DIVERGENCE
1.	PHE	34	14	0.25	0.73
2.	GLN	65	13	0.20	0.86
3.	ARG	67	17	0.00	0.85
4.	LYS	69	5	0.01	0.81
5.	ASN	097	16	0.52	0.80

Table 3: Predicted binding sites of wheat HMP-P kinase

S.NO	HETEROGEN	COUNT
1.	FMN	4
2	MG	1
3.	NAD	1

Table 4: Predicted heterogens of wheat HMP-P kinase