Journal of Advanced Bioinformatics Applications and Research
ISSN 0976-2604.Online ISSN 2278 – 6007
Vol 3, Issue 2, 2012, pp 294-303
http://www.bipublication.com

MODELING OF ARABIDOPSIS AND SOYBEAN THIAMIN PYROPHOSPHO KINASE GENES

[1]G.V.Ramana, [2]Veera Bramhachari P and [1]Chaitanya K.V.*

[1]Department of Biotechnology, GITAM Institute of Technology, GITAM University, Visakhapatnam-530045
[2]Department of Biotechnology, Krishna University, Machilipatnam, INDIA
*Corresponding Author:viswanatha.chaitanya@gmail.com

[Received- 31/03/2012, Accepted- 29/04/2012]

ABSTRACT:

Thiamin also called, as Vitamin B1 is an important nutrient essential for the human metabolism. Thiamin pyrophosphate (TPP) is an important co-factor needed by the enzymes involved in maximum number of metabolic processes in most of the organisms. Synthesis of TPP in plants is controlled by the enzyme thiamin pyprophospho kinase (TP kinase). In Arabidopsis genome two genes have been identified for the TP kinase enzyme and one gene in soybean which is the richest source of proteins. However, not much information is available on the structure of the TP kinase proteins. In this study the Sequence differences in the amino acids of soybean and Arabidopsis TP kinase proteins, its structural properties have been discussed.

KEY WORDS: Thiamin, Arabidopsis, soybean, TPP kinase, modeling

I.INTRODUCTION:

Thiamin (vitamin B1) is an essential nutrient in human beings whose deficiency causes *Beri Beri*, disturbing the central nervous system and circulatory system by accumulating pyruvate and lactate in vessels [1]. Thiamin pyrophosphate (TPP) is an important co-factor needed by the enzymes of almost all biological systems involved in a number of metabolic processes including the production of acetyl-CoA, the tricarboxylic acid cycle, the pentose phosphate pathway/Calvin cycle, branched chain amino acid biosynthesis and isoprenoid biosynthesis [2]. TPP also acts as a carrier of activated aldehyde groups in decarboxylation and transketolation reactions [3]. It is a bicyclic compound composed of a pyrimidine ring that is covalently linked to a pyrophosphorylated thiazole ring. A pyrophosphate group is attached to the thiazole ring plays a key role in the integration of TPP

into the enzyme active site [4]. Deficiencies in TPP have been associated with disease states, including Wernicke-Korsakoff Syndrome and megablastic anemia [5]. In addition, TPP also functions as an efficient antioxidant protecting the cells against lipid peroxidation and free radical oxidation [6].

Activation of TPP in plants is controlled by the enzyme thiamin pyrophosphokinase (TPK; EC 2.7.6.2), requiring mg^{2+} and ATP for catalyzing the TPP formation through direct phosphorylation of free thiamin. In Arabidopsis, two genes have been reported for the TPK showing homology with many plants including Selaginella but the details of its gene structure and regulatory elements is not available. Soybean (*Glycine max*. L) is one of the economically important plants, popular for the protein and the edible oil. In the present study, we described the structure model of TPK gene from soybean and compared with the two genes of *Arabidopsis thaliana*.

II. MATERIALS AND METHODS:

Retrieving Sequence Information

The sequences of the Arabidopsis thaliana TP kinase1 and TP kinase 2 were retrieved from TAIR site with accession number. Sequences retrieved using Arabidopsis TP kinase 1&2 as reference from DFCI BLAST search where populated and their ORF' were translated. Sequences of plants such as lettuce, barley, rice, and wheat corresponding to accessions DY973873, TC252070, TC501862 and TC420894 respectively populated together. Peculiarly soybean and other plants listed above have unique sequence pattern for both the TP kinase types.

Phylogenetic analysis

For further identification of genes, corresponding to TP kinase 1 and 2 phylogenetic analysis was performed by constructing the UPGMA tree using MEGA. Input of the sequences was the Clustal w alignment of all the sequences to obtain the tree with bootstrap values at each branch length showing the evolutionary relationship between the sequences of input [7].

Boxshade Analysis

Box shade analysis server generates editable graphics or pretty boxes from an input of multiple alignment file. The amino acid pattern of TP kinase 1&2 genes in different plants with the help of various degrees of shading emphasizing conserved or similar residues was studied using box shade analysis server.

Secondary structure prediction

The secondary structure of Soybean TP kinase 1 and 2 genes were predicted using Psi pred server (http://bioinf.cs.ucl.ac.uk/psipred/), which will analyze the input protein sequence basing on PSI-BLAST [8].

Binding site, Heterogen and 3D structure prediction

3D ligand site was used for predicting the binding sites, heterogens and also for the prediction of 3D structures of the input protein sequences in FASTA format. Depending on the best predictions of binding sites which was a human method 3d ligand site functions; explained at CASP8. Ligands were grouped by single linkage clustering where clusters with most ligands are selected to form binding site. 3D structure was predicted using the phyre server. Predicted residues were colored in blue. 3D view powered by jmol applets will allowed spinning, changing color of residues or heterogens, labeling or space fill in various ranges from display options [9][10].

Quarternary structure prediction

The quaternary structure of TP kinase 1&2 protein in soybean was predicted using protein interfaces and surfaces tool (PISA). Using this, the Probable Quaternary Structure (PQS) of a protein with different group of assemblies that can make a crystal can be known by the sets, which represent the solutions, which are observed on the top of the table. Top most values in assemblies table is at most appropriate in contrast to other values. MM size infers number of

Macromolecular Monomeric units in that particular assembly, which represent an oligomeric or multimeric state. Formula indicates an assembly chemical composition which denotes the number of different monomeric units and their types. Stable column imply the stability of an assembly which may or may not dissociate in solution. Solvation free energy (ΔG^{int}) calculated by a difference in solvation energies of isolated and assembly structures indicates free energy gain during the formation of assembly in Kcal/M. Free energy of dissociation (ΔG^{diss}) represent free energy difference between associated and dissociated states. Assemblies with $\Delta G^{diss}>0$ are thermodynamically more stable because positive values includes external energy utility in dissociation of assembly.

Model evaluation using PROSA

Z-score values indicate overall quality of a model which can find be determined using PROSA web server. Local model quality plot will plot the energies of knowledge based on y-axis and sequence positions on x-axis. Positive values in the graph are erroneous structures, which imply the model quality. Plots with single residues are with more fluctuations and are of limited value. Hence, average energy values of a model for each 40 residues (i+39) will be calculated which will be assigned to central residue of the entire fragment at position i+19 (thick line graph). In the background another thin line which will represent window 10 i.e. 10 residues considered as a single fragment [11].

Surface calculation

Inorder to perform the calculations of molecular mechanics such as calculation of solvent accessibility surface area VEGA ZZ was used. After starting VEGA ZZ in the calculate option of tool bar Surface can be calculated. For that loading the pdb structure is preliminary step. Then using the surface option we can calculate the SASA in either the dot method or mesh and solid method where the radius is 1.4. Result will be displayed in the console of VEGA ZZ depicting SAS of the input molecule [12].

Energy prediction

Electrostatic and desolvation interaction free energy between two proteins estimation in the units of kcal/mol is feasible with the Free contact server available at http://structure.pitt.edu/servers/fastcontact/ by submitting the PDB format of the sequence as the input. The output contains the desolvation free energy and electrostatic free energy but also residue contact free energy which evaluates van der waals interaction using CHARMm [13].

III.RESULTS ANALYSIS AND DISCUSSION:

UPGMA Tree Analysis

UPGMA tree constructed using MEGA was used to analyze the evolutionary relationship between these plants. Arabidopsis TP kinase1& 2 clustered together in a single clade with a branch length of 0.0281which further clustered with soybean possessing a branch length difference of 0.0281. Lettuce and rice forms the second cluster with a branch length of 0.1817. Barley and wheat segregated with a unique branch length of 0.0199. Cluster 1& 2 segregates further with a branch length difference of 0.0111. Further clustering between the remaining two groups with a branch length difference of 0.1773 observed. The phylogenetic analysis indicates that the evolution time of both TP kinase1 and 2 genes is almost same and a very high degree of similarity exists between the TP kinase genes of Arabodopsis and soybean resulted in the segregation of soybean TP kinase genes with Arabidopsis (Figure1). Segregation of lettuce and rice separately is another interesting observation. Barley and wheat were segregated separately (Figure 1).

Box shade Analysis

Box shade analysis reveals the amino acid pattern between the TP kinase 1and 2 genes of all the plants considered for this study which was almost unique. There is a wide range of similarity existing among these sequences particularly in amino acids of 1-8, 22-38, 60-70 (Figure 2). This

study brought the points to be emphasized into lime light, which allow us to concentrate on the Arabidopsis TP kinase 1&2 along with soybean TP kinases, which forms the first cluster for further predictions and analysis.

Secondary structure prediction

PSI PRED server predictions revealed that the Arabidopsis secondary structures have 6 helices and soybean 5 helices. They maintained same pattern in strands with 16 strands each for Arabidopsis structures and 15 for soybean. In the case of coils also Arabidopsis tp kinase 1 & 2 have 21 coils whereas soybean have 20 (Table 1). There is a subtle difference exist between the secondary structures of Arabidopsis tp kinase 1 & 2. First helical structure in both the structures is similar i.e., from 39 to 44. However, the second helix is from 55 in tp kinase 1 of Arabidopsis whereas 57 in tp kinase 2 and in contrast the difference in culmination of helixes in both structures is just a base pair. Similarly third helix also has one residue difference in inception in contrast for two structures. Remaining helixes are in same position for both structures. In the case of coils also such subtle difference can be observed in coils numbered 5,6,10,13,17 and 18 where only single residue difference is observed either at inception or culmination of the coils.5,6,7,8,11,12,13 and 14 strands of Arabidopsis tp kinase 1& 2 have subtle variations. In contrast to above two structures soybean tpkinase structure have no similarity in helixes, strands or in coils. Soybean helixes range from 40-46, 56-63, 94-102,118-129 and 150-162. Strands of soybean TP Kinase ranges from 27-32, 49-54, 84-87,106-109,139-145,168-172,175-180,185-190,198-203,211-213,217-219,230-238,241-246 and 250-256. Whereas coils of this structure have an observable high range of 1-26,33-39,47-48,55,64-83,88-93,103-105,110-117,130-138,146-149,163-167,173-174,181-184,191-197,204-210,214-216,220-229,239-240,247-249 and 257-259 respectively[14] (Table 1).

Binding site and Heterogen prediction

Among the predicted binding sites, 53, 87, 89, 111, 112, 117 and 120 are common in both the structures of Arabidopsis coding amino acid asparagines. Only binding site 90 found to be excess in Arabidopsis tp kinase 1 coding amino acid serine with a contact of 4, average distance 0.21 whose j s divergence is 0.72. Residues 87 and 89 are with a contact of 9 whose j s divergence is 0.77 and 0.76 respectively (Table2). In resemblance to residue 90, 111 and 120 residues have unique contact of 4 with varied average distances of 0.49 and 0.18 respectively. Their j s divergence was found to be 0.25 and 0.45. Another peculiarity we can found in residues 111 and 120 were code by histidine and lysine unlike to remaining amino acids whose contact in both the structures is 4. Only two base pairs are with zero average distance in both the Arabidopsis tp kinase structures, which are none other than 53 and 87.In Arabidopsis tp kniase 2 residues 87 and 89 are with similitude contact of 8. Contact of 53,112 and 117 are 6, 11 and 10 respectively. J S Divergence of 87 and 89 are 0.76. 117 residues have 0.77 scoring in this parameter. Base pair 111 acquired least J.S.D of 21 whereas 112 stand next to it with 0.35 values. Residue 111 has highest avg. distance in the structure with a value of 0.39.All the residues of soybean tp kinase had unique amino acid coding by asparagines [15]. Residues 88 and 90 have contact of 8 with a j s divergence of 0.76 and 0.75 respectively. Average distance is least i.e. 0.00 in 54 and 88 whereas in 118 it is 0.08 where J S Divergence is 0.77 for 118 and 0.74 for residue 54 (Table 2).

Heterogens TPP,MG and AMP are in common among all the three structures. ATP and ZN are the heterogens present in Arabidopsis tp kinase 2 and soybean tp kinase, which were not present in the Arabidopsis tp kinase1. SAH is the only heterogen peculiarly obsereved in Arabidopsis tp kinase 1. In all the structures except for MG for all remaining heterogens count is 1. For

Arabidopsis tp kinase 1 and soybean count for MG is similar i.e. is 8. For Arabidopsis tp kinase 2 it is 9 (Table 3). Our prediction of binding sites revealed that the heterogens TPP, MG and AMP are omnipresent in three sequences with similarity in the count between the Arabidopsis TP kinase and soybean TP kinase. The heterogen prediction revealed that Arabidopsis and soybean have some functional similarity with respect to TP kinase genes. SAH is the heterogen, which is specific to Arabidopsis TP kinase 2. Another significant feature of heterogens is ATP and Zn are absent in Arabidopsis TPkinase1 and are present in other three proteins.

3D Structures

Analysis of the 3D structures is much easier when compared to the analysis of binding site prediction or heterogen prediction. It was powered by jmol view with easy color differentiation between various heterogens and unique color expression of binding sites enhancing easy identification in complex structures of Arabidopsis Tp kinase 1, Arabidopsis Tp kinase 2, and Soybean Tp kinase (Figure 3a, 3b and 3c).

Binding sites indicted with a blue color along with labeling. Heterogens were differentiated in multiple colors. Zn indicated by gray color with a round ball like structure. Whereas Heterogens SAH, TPP, MG, AMP and ATP were denoted by the colors sky-blue, red, green, pink and sky blue respectively. If the spacefill format activated, it will be more feasible to distinguish between metal ions and binding sites [10]. When there is a high demand for differentiation space fill is activated, otherwise labeling is enabled which show amino acids coding the accessions of binding sites.

Quaternary Structure prediction:

All the assembly prediction values for the three proteins observed to be similar. PQS set values are 1, 2 and 3 representing three most probable assemblies among which 1st set is appropriate with mmsize of 4 and formula $A_4a_4b_{12}c_4d_2e_2$. There are six different monomeric units with four types of assemblies in A, a and c, 12 in b and two in d and e. Here id is 1 and biomolecule R350 is 2 representing two annotations for assembly. The structure is thermodynamically stable with 5.2 dissociation free energy which is positive value. ΔG^{int} value is -275.6 Kcal/M, which means that the dissociation energy is less than the association energy of this assembly. All the above stated statements will imply for Tp kinase1, tp kinase 2 and soybean tp kinase. All the structures are thermodynamically stable with different sequid's of 36.842, 37.218 and 38.346 respectively (Table 4). Thus, we can say that there is subtle difference in quaternary structures among the three proteins. Further clarification in this arena can be observed in PXS calculations.

Model evaluation

First plot was Z score plot with number of residues plotted on x axis and z score on y axis. From the 1st plot of all the three models we can say that these models are similar to NMR models which was indicated by a black spot on the NMR region. Z score of Arabidopsis tp kinase1, Arabidopsis tp kinase 2 and Soybean tp kinase are -6.66, -5.14 and -5.9 respectively (Figure 4a, 4b & 4c). PROSA web server results displayed Z score values in this plot, which contains all Z score values determined experimentally. Z score plots elucidated that Arabidopsis and soybean TP kinases possess model quality (Figure 5a, 5b &5c). In this plot, sequence position was plotted against knowledge-based energy. Window 10 and 40 indicated in the inset represents number of residues considered for plotting their average energies. Window 10 graphs with more fluctuations given trivial importance hence shaded in light color. To the mean for 40 residues in Arabidopsis tp kinase 1 we can find average energy as greater than -1 with reference to knowledge based energy predictions. For tp kinase 2 minimum free energy is in between 0 to -1, precisely more nearer to -1. Soybean tp kinase minimum energy levels are also clearly found to

be greater than -1. Plots with more positive values found to be residues with erroneous structures. Thin lined graphs are with more fluctuations showing highest energies than windows 40 graphs. Proteins with minimum free energy are stable with fewer fluctuations and need to be considered for further molecular mechanical studies.

VEGA ZZ

To display analyze and manage 3D structures of the molecules some features were highly useful in VEGAZZ. The most important features are (1) file format conversion (with assignment of the atom types and atomic charges), (2) surface calculation and (3) trajectory analysis [16].

We need to input the secondary structure in pdb format as input for calculation of Solvent Accessibility Surface Area and minimum free energy. Total solvent accessibility was calculated for the entire molecule by fixing hydrogen's and charge. Soybean tp kinase, Arabidopsis tp kinase 1 and Arabidopsis tp kinase 2 SASA is calculated to be 97.936 nm2, 100.590 nm^2 and 107.276 nm^2 respectively (Table 5). Solvent accessibility of soybean TP kinase is high when compared to that of Arabidopsis.

Energy predictions

For predicting the energy in the proteins of Arabidopsis and soybean TP kinases, electrostatic, desolvation and van der waal energies were analyzed using CHARMm. Electro static energies of Arabidopsis tp kinase 1, Arabidopsis tp kinase2 and soybean tp kinase were 1670.3506 kcal/mol, 1793.2414 kcal/mol and 1588.13562 kcal/mol respectively (Table 6). De solavtion free energies are -127.277693 kcal/mol, -202.408025 kcal/mol and -221.759403 kcal/mol respectively for Arabidopsis tp kinase 1 & 2 and for soybean tp kinase. Vander waals energy is almost similar to all the three structures ranging from 4.01 to 4.08 E+69 kcal/mol (Table 6). The energy predictions for Arabidopsis and soybean TPkinases suggest that they have independent free energies and the levels of free energy in Arabidopsis TPkinase 2 were high when compared to that of other two proteins.

In the present study, we have analyzed the TP kinase proteins of Arabidopsis and soybean, which has very less difference with respect to structure, and sequence properties, which hypothesize that they have similar function, which is yet to be studied. Further studies are in progress to identify the functional similarities in TP kinase genes of Arabidopsis and soybean.

V.REFERENCES:

1. Begley TP, Downs DM, Ealick SE, McLafferty FW, Van Loon AP, Taylor S, Campobasso N, Chiu HJ, Kinsland C, Reddick JJ, Xi J: 1999. Thiamin biosynthesis in prokaryotes. *Arch. Microbiol.* 1999, 171: 293–300.
2. Tunc-Ozdemir M, Miller G, Song L, Kim J, Sodek A, Koussevitzky S, Misra AN, Mittler R, Shintani D: Thiamin confers enhanced tolerance to oxidative stress in Arabidopsis. *Plant Physiol.* 2009, 151: 421–432.
3. Ajjawi I, Rodriguez Milla MA, Cushman J, Shintani DK: 2007. Thiamin pyrophosphokinase is required for thiamin cofactor activation in Arabidopsis. *Plant Mol. Biol.* 2007, 65: 151–162.
4. Roux C, Delfourne E, Bastide J: 1996. Activation and inhibition of maize acetohydroxyacid synthase by thiamine analogues. *Plant Physiol. Biochem.* 1996, 34: 293–299.
5. Timm DE, Liu J, Baker LJ, Harris RA: Crystal structure of thiamin pyrophosphokinase. *J. Mol. Biol.* 2001, 310:195-204.
6. Jung IL, Kim IG: 2003. Thiamin protects against paraquat-induced damage: scavenging activity of reactive oxygen species. *Environ. Toxicol. Pharmacol.* 2003, 15: 19–26.
7. Peterson N, Stecher G, Nei M, Kumar S: MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol. Biol. Evol. 2011, doi: 10.1093/molbev/msr121.
8. Jones DT: Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* 1999, 292: 195-202.
9. Wass MN, Sternberg MJ: 2009. Prediction of ligand binding sites using homologous structures and conservation at CASP8. *Proteins*, 2009, 9:147-51.
10. Wass MN, Kelley LA, Sternberg MJ: 3DLigandSite: Predicting ligand-binding sites using similar structures. *Nucl. Acid Res.* 2010, 38 Suppl:W469-73.
11. Wiederstein M, Sippl MJ: ProSA-web: Interactive web service for the recognition of errors in

three-dimensional structures of proteins. *Nucl. Acids Res.* 2007, 35: W407-W410.

12. Champ PC, Camacho CJ: FastContact: A free energy scoring tool for protein–protein complex structures *Nucl. Acids Res*. 2007, 35: W556-W560 .

13. Camacho CJ, Zhang C: FastContact: Rapid estimate of contact and binding free energies. *Bioinformatics* 2005, 21: 2534-2536.

14. Capra J, Singh M: Characterization and prediction of residues determining protein functional. *Bioinformatics* 2008, 24: 1473-1480.

15. Bryson K, McGuffin LJ, Marsden RL, Ward JJ, Sodhi JS, Jones DT: Protein structure prediction servers at University College London. *Nucl. Acids Res.* 2005, 33: W36-38.

16. Pedretti A, Villa L, Vistoli G: VEGA: A Versatile program to convert, handle and visualize Molecular structure on windows-based PCs. *J. Mol. Graph.* 2002, 21: 47-49 .

INDENTS FOR FIGURES AND TABLES:

Figure 1: phylogenetic analysis of TP kinase genes

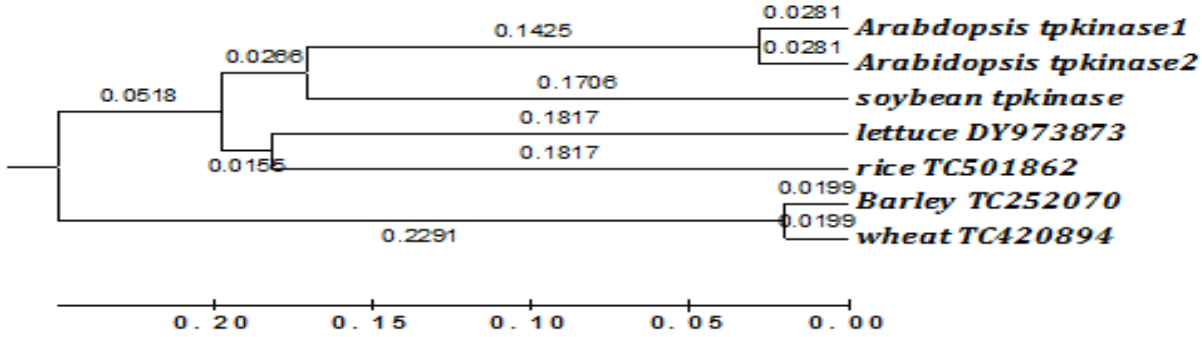


Figure 2: Box shade analysis of Tp kinase proteins

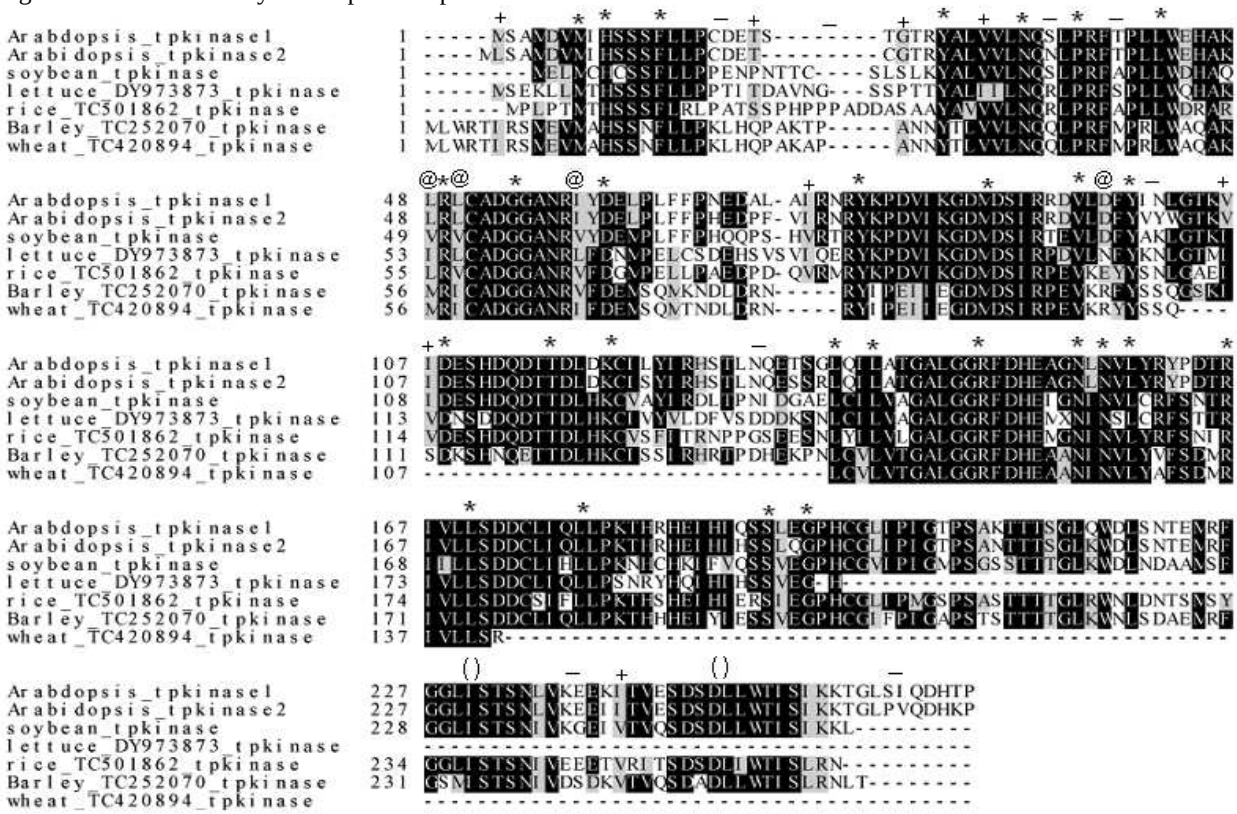
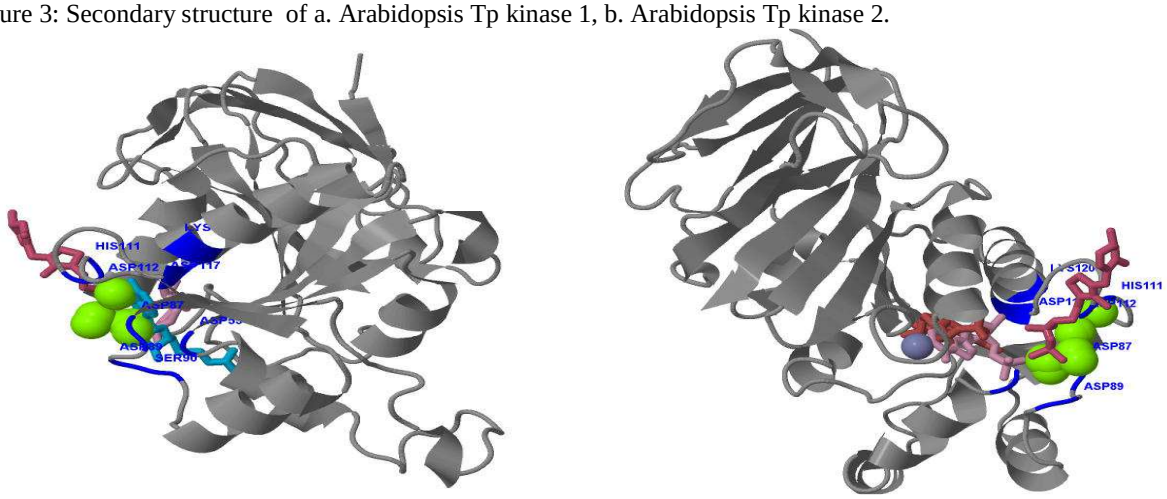


Figure 3: Secondary structure of a. Arabidopsis Tp kinase 1, b. Arabidopsis Tp kinase 2.



c. Soybean Tp kinase. Blue residues are binding sites with labeling, Remaining colored objects and wire frames are heterogens

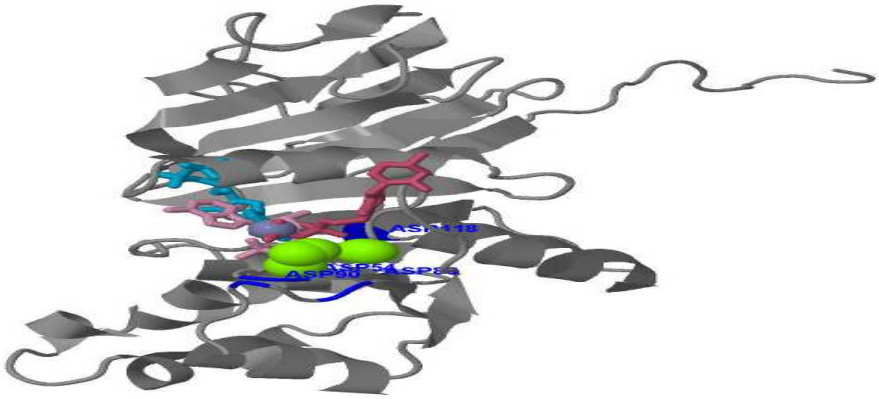


Figure 4: PROSA Model evaluation of a: Arabidopsis Tp kinase 1 b. Arabidopsis Tp kinase 2. c Soybean Tp kinase.

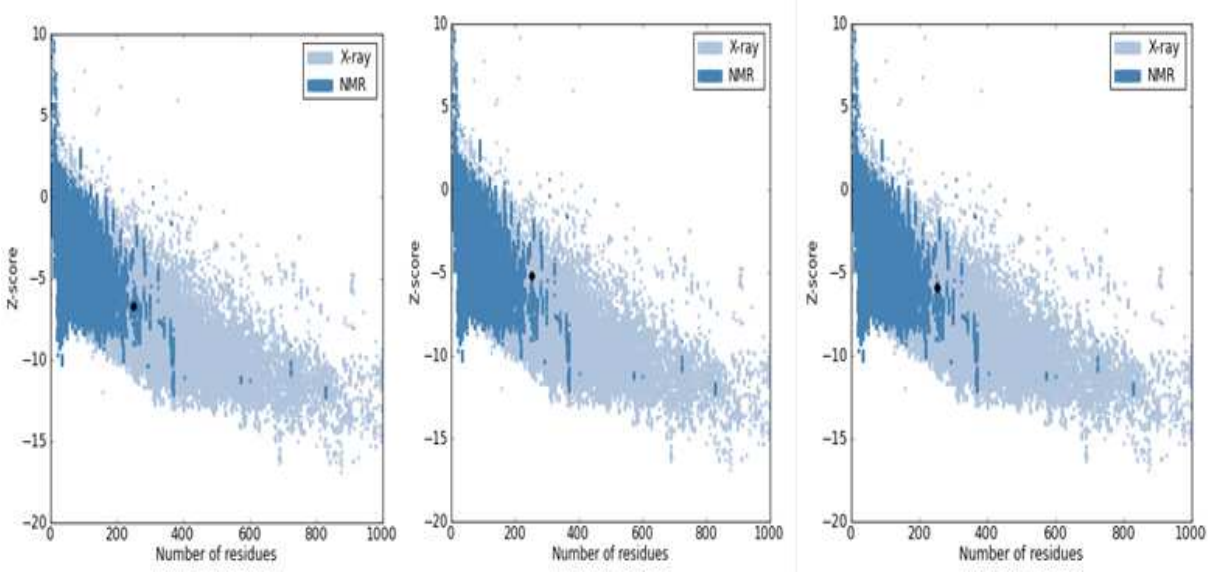


Figure 5: PROSA model evaluation with knowledge based energy predictions.
a: Arabidopsis Tp kinase 1,b. Arabidopsis Tp kinase 2. C. Soybean Tp kinase.

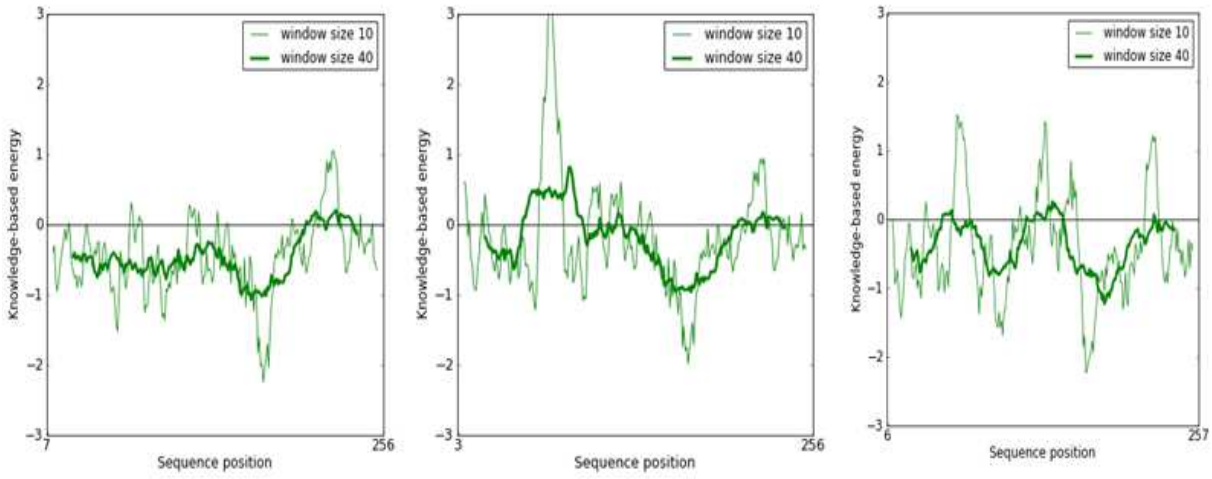


Table 1: Secondary structure Prediction table of Arabidopsis and Spybean TP kinase

Protein	Description	Accession Number
ARABIDOPSIS TP KINASE 1	HELIX	39-44, 55-61, 71-76, 93-101,117-128,149-161.
	STRAND	6-8, 26-31, 48-53,83-86,105-108,138-142,167-170,174-180,184 189,197-202,210-218,229-232,234-237,240-246,249-255.
	COIL	1-5,9-25,32-38,45-47,54,62-70,77-82,87-92,109-116,129-137, 143-148,162-166,171-173,181-183,190-196,203-209,219-228, 233,238-239,247-248,256-267.
ARABIDOPSIS TP KINASE 2	HELIX	39-44, 57-60,72-76,93-101,117-128,149-161.
	STRAND	6-8, 26-31,48-53,83-86,102-108,137-142,167-171,175-180,184-189,197-202,210-212,216-218,231-232,235-237,240-246,249-255.
	COIL	1-5,9-25,32-38,45-47,54-56,61-71,77-82,87-92,109-116,129-136,143-148,162-166,171-174,181-183,190-196,203-209,213-215,219-230,233-234,238-239,247-248,256-267.
SOYBEAN TP KINASE 1	HELIX	40-46,56-63,94-102,118-129,150-162.
	STRAND	27-32,49-54,84-87,106-109,139-145,168-172,175-180,185-190,198-203,211-213,217-219,230-238,241-246,250-256.
	COIL	1-26,33-39,47-48,55,64-83,88-93,103-105,110-117,130-138,146-149,163-167,173-174,181-184,191-197,204-210,214-

Table 2: Binding site prediction for the TP kinase proteins of Arabidopsis and soybean

Plant Protein	Residue	Amino acid		Contact	Av distance	Js Divergence
	53	ASP	6	0.00		0.76
Arabidopsis TPkinase1	87	ASP	9	0.00		0.77
	89	ASP	9	0.06		0.76
	90	SER	4	0.21		0.72
	111	HIS	4	0.49		0.25
	112	ASP	11	0.05		0.37
	117	ASP	10	0.08		0.77
	120	LYS	4	0.18		0.45
Arabidopsis TP Kinase2	53	ASP	6	0.00	0.73	
	87	ASP	8	0.00	0.76	
	89	ASP	8	0.07	0.76	
	111	HIS	4	0.39	0.21	
	112	ASP	11	0.05	0.35	
	117	ASP	10	0.12	0.77	
	120	LYS	4	0.18	0.46	
SOYBEAN TPKINASE	54	ASP	5	0.00	0.74	
	88	ASP	8	0.00	0.76	
	90	ASP	8	0.06	0.75	
	118	ASP	9	0.08	0.77	

Tables 3: Heterogen predictions for TP kinase proteins of Arabidopsis and soybean

Plant Protein	Heterogen	Count	Source Structures
Arabidopsis TP K1	TPP	1	3ihk_A
	MG	8	2f17_A,2hh9_B,2g9z_A,3ihk_A
	AMP	1	2f17_A
	SAH	1	1l3i_B
Arabidopsis TP K2	TPP	1	3ihk_A
	MG	9	2f17_A,2hh9_B,2g9z_A,3ihk_A, 2rhd_A
	AMP	1	2f17_A
	ATP	1	1mjh_B ZN 11ww1_A
	ZN	1	1ww1_A
Soybean TP Kinase	TPP	1	3ihk_A
	MG	8	2f17_A,2hh9_B,2g9z_A,3ihk_A
	AMP	1	2f17_A ZN 12wt9_B
	ZN	1	2wt9_B
	ATP	1	1mjh_B

Table 4: Quatenary structure analysis of Arabidopsis and soybean TP kinase proteins

Pqs set	mmsize	formula	id	Bio molR350	Stable	Buried surface area sq. Å	ΔG^{int}, kcal/mol	ΔG^{diss}, kcal/mol
1	4	$A_4a_4b_{12}c_4d_2e_2$	1	2	yes	16680	-275.6	5.2
2	2	$A_2a_2b_6c_2de$	2	1	yes	7300	-133.6	17.2
3	2	$A_2a_2b_{10}c_2d_2$	3	-	yes	6660	-186.4	5.8
4	1	Aabcd	4	-	yes	1400	-21.0	0.0

Table 5: Solvent Accessibility Calculations of Arabidopsis and soybean TP kinase proteins

STRUCTURE	SASA
Arabidosis tp kinase1	97.936 nm^2
Arabidopsis tp kinase 2	100.590 nm^2
Soybean tp kinase	107.276 nm^2

Table 6: Predicted Energies for the TP kinase proteins of Arabidopsis and soybean

Structure name	Electro static free energy kcal/mol	De-solvation free energy kcal/mol	Van der waals kcal/mol
Arabidopsis TPkinase1	1670.3	-127.2	4.080914E+69
Arabidopsis TP kinase2	1793.2	-202.4	4.050961E+69
SoybeanTP kinase	1588.1	-221.7	4.013326E+69