Research Article **Open Access**

In-Silico Analysis, Structural Modelling and Phylogenetic Analysis of Acetohydroxyacid Synthase Gene of *Oryza sativa*

Ubaid Yaqoob[1]*, Tanushri Kaul[2] and Irshad Ahmad Nawchoo[1]

[1]*Plant Reproductive Biology, Genetic Diversity and Phytochemistry Research Laboratory, Department of Botany, University of Kashmir, Srinagar, Jammu and Kashmir, India*
[2]*Plant Molecular Biology Lab, International Centre for Genetic Engineering and Biotechnology, New Delhi, India*

Abstract

The Acetohydroxyacid synthase (EC 2.2.1.6) or Acetolactate synthase (ALS) belongs to a family of thiamine diphosphate (TPP) dependent enzymes which catalyzes the first reaction in the biosynthesis of essential amino acids - isoleucine, leucine and valine. Acetohydroxyacid synthase (AHAS) is present in plants, algae, fungi and bacteria and is found to be a vital target of multiple herbicides. We revealed the homology model of OsAHAS protein using the structure of *Arabidopsis thaliana* AHAS (PDB ID: 3E9Y) as template. The resulting model structure was refined by PROCHECK, ProSA, RMSD and Verify3D that indicated the model structure is reliable with 76% amino acid sequence identity with template. RMSD (1.75Å), Verify3D (86.02%), Z-score (-9.55) and Ramachandran plot analysis showed that conformations for 81.6% of amino acid residues are within the most favoured regions. The phylogenetic tree constructed revealed different clusters based on AHAS in respect of bacteria, fungi, algae and plants. The multiple sequence alignment of these AHAS protein sequences from different organisms showed conserved regions at different stretches with homology in amino acid residues. Through motif analysis, it was revealed that conserved AHAS domain are found in all AHAS proteins suggesting its possible role in cellular and metabolic functions.

Keywords: Acetohydroxyacid synthase; *Oryza sativa*; Herbicide; Amino acids

Introduction

The Acetohydroxyacid synthase (EC 2.2.1.6) or Acetolactate synthase (ALS), a plastid enzyme [1] which catalyzes the first reaction in the biosynthesis of branched-chain essential amino acids - isoleucine, leucine and valine [2-4] is the vital target of multiple herbicides. Acetohydroxyacid synthase (AHAS) belongs to a family of thiamine diphosphate (TPP) dependent enzymes present in plants, algae, fungi, and bacteria [5]. The ion cofactor is typically Mg^{2+} [6] which anchors TPP to AHAS. Flavin adenine dinucleotide (FAD) molecule, a third cofactor is also required by AHAS. Commercially available herbicides that inhibit AHAS include sulfonylureas (SU), imidazolinones (IMI), triazolopyrimidines (TP), pyrimidinyl-thiobenzoates (PTB) [also known as pyrimidinylsalicylic acids or pyrimidinyloxybenzoic acids] and sulfonyl-aminocarbonyl-triazolinones (SCT) [7,8]. Out of these the sulfonylureas and imidazolinones are the most significant, with the sulfonylureas being the leading group on an active ingredient basis. Due to amino acid starvation, AHAS inhibition leads to plant death [9]. The mammals lack the pathway for branched-chain amino acids biosynthesis and thus the ALS-inhibiting herbicides are thought to be non-toxic to them [10]. They are highly selective to plants and have a broad range of weed control activity [11-13]. The most common naturally occurring mutations are at amino acids Ala122 [14,15], Pro197 [16-18], Trp574 [14,16,19] and Ser653 [15,20]. Thus understanding its structural details would be a great revolution for engineering new herbicides, developing resistant crops and antimicrobial drugs.

Materials and Methods

Homology modelling and structural analysis: *Oryza sativa* AHAS (OsAHAS) sequence was retrieved by using NCBI database (http://www.ncbi.nlm.nih.gov). By searching the PDB of known protein structures, the homology modelling was performed with target sequence as the query [21]. The target sequence was searched for similar sequence using the BLAST (Basic Local Alignment Search Tool) [22] against Protein Database (http://www.rcsb.org). The BLAST results yielded X-ray structure of AHAS from *Arabidopsis thaliana* (AtAHAS) with 76% similarity to our target protein (OsAHAS). Using ClustalW

[23], all the sequences of AHAS were aligned to find out the similarity present among the sequences. 2D and 3D structure alignment was carried out using ClustalW [24] and MATRAS 1.2 [25] respectively. The sequences of the AHAS were further analysed for the presence of specific AHAS domains and motifs through motifscan (myhits.isb-sib.ch/cgi-bin/motif scan) and scan prosite (Prosite.expasy.nlm.nih. gov). Analysis of conserved motifs was done by MEME version 3.5.7 [26] using minimum and maximum motif width of 20 and 50 residues respectively and maximum number of 7 motifs, keeping rest of the considerations at default. Via Modeller 9.12 by comparative modelling of protein structure prediction, the theoretical structure of OsAHAS was generated.

The secondary structural features of the AHAS sequences of template (AtAHAS) and target (OsAHAS) were calculated using SOPMA. The physico-chemical properties of AHAS sequences like molecular weight, theoretical isoelectric point (pI), number of amino acids, total number of positive and negative residues, aliphatic index [27], grand average hydropathy (GRAVY) [28] extinction coefficient [29] and instability index [30] were evaluated by using Expasy's ProtParam server (http://us.expasy.org/tools/protparam.html) [31]. Using String software (http://string-db.org/) the interacting partners of AHAS and its co-expressed genes were predicted [32].

Model validation of OsAHAS: The model was evaluated on the basis of geometrical and stereo-chemical constraints using RAMPAGE server (http://mordred.bioc.cam.ac.uk/-rapper/rampage.php),

PROCHECK [33], Verify 3D [34] and ProSA-Web [35]. The model with the least number of residues in the disallowed region was selected for the further studies. The RMSD value between the template and target was calculated using MOE [36]. The best model structure was then compared with the template protein by superimposition using SuperPose Version 1.0 [37].

Phylogenetic analysis: Phylogenetic analysis of the sequences was done by Molecular Evolutionary Genetic Analysis (MEGA) software Version 4.1 [38] by using UPGMA method. Each node was tested using the bootstrap approach by taking 5,000 replicates.

Results and Discussion

Homology modelling and structural analysis: The *Oryza sativa* AHAS (OsAHAS) protein sequence consist of 644 amino acid residues. The query sequence from OsAHAS protein was selected for homology based searching of the template structure by the BLAST program against the structural database of PDB (http://www.rcsb.org) [30,31]. Sequences that showed maximum identity with high score and low e-value were aligned and the alignment was used to build a 3D model for OsAHAS. According to the result of BLAST search against PDB [39], three reference proteins (PDB: 3E9Y, 1YBH, 1NOH) represented a high level of sequence identity that is 76%, 75% and 41%, respectively. The homology search of AHAS revealed 76% sequence identity to *Arabidopsis thaliana* (PDB ID: 3E9Y) with an e-value of 0.0 and was selected for comparative modelling. Multiple sequence alignment of the AHAS sequences highlighted the sequence conservation of amino acid residues among different species (Supplementary File S1). Structurally conserved regions (SCRs) between model OsAHAS (target) and homologous proteins (PDB: 3E9Y, 1YBH, 1NOH) were determined by multiple sequence alignment (Figure 1). Structurally conserved regions (SCRs) between model OsAHAS and template (PDB: 3E9Y) were also determined (Figure 2). An extensive search

of the motifs and their positions was done by MEME software which identified several conserved motifs in the protein sequences of AHAS (Figure 3). Multilevel consensus sequences for the MEME defined motifs along with their e-values are shown in Figure 4.

The initial model of OsAHAS was built by homology modelling methods using Modeller 9.12. software [40]. In this study, predicted 3D structure of OsAHAS was generated and the N-terminal and C-terminal domains were identified (Figure 5). Each subunit consists of three domains - α, β, and γ, plus a C-terminal tail. In *Arabidopsis thaliana,* each subunit consists of three domains, α (residues 86-280), β (281-451), and γ (463-639), plus a C-terminal tail (646-668) that loops over the active site [41]. The secondary structural features of the *Arabidopsis thaliana* and OsAHAS sequences were calculated using SOPMA [42] with default parameters (Table 1). The AHAS protein is composed of 31.52% α-helices, 22.52% extended strands and 9.94% β-turn in rice. In case of *Arabidopsis thaliana,* the AHAS protein is composed of 33.22% α-helices, 23.63% extended strands and 9.76% β-turn. Thus the α-helices and the β-sheets cover comparatively larger portions of the rice and *Arabidopsis thaliana* AHAS enzymes. Similar results have been observed by McCourt et al. [41] in *Arabidopsis thaliana*. Several physico-chemical properties of AHAS sequences were calculated by using Expasy's ProtParam server [31]. The results are shown in Table 2. In developing buffer system for protein purification, the computed isoelectric point (pI) will be useful. The very high aliphatic index of the AHAS enzyme sequences indicate that these enzymes may be stable for a wide temperature range. The high extinction coefficient of enzyme in rice indicates the presence of more Cys, Trp and Tyr. The instability index value for the AHAS proteins were found to be ranging from 36.51 to 41.61 indicating the stable and instable nature of the *Arabidopsis thaliana* and rice AHAS protein respectively.

Using String software, the AHAS interacting partners as well as its coexpression genes were predicted in both rice and *Arabidopsis*
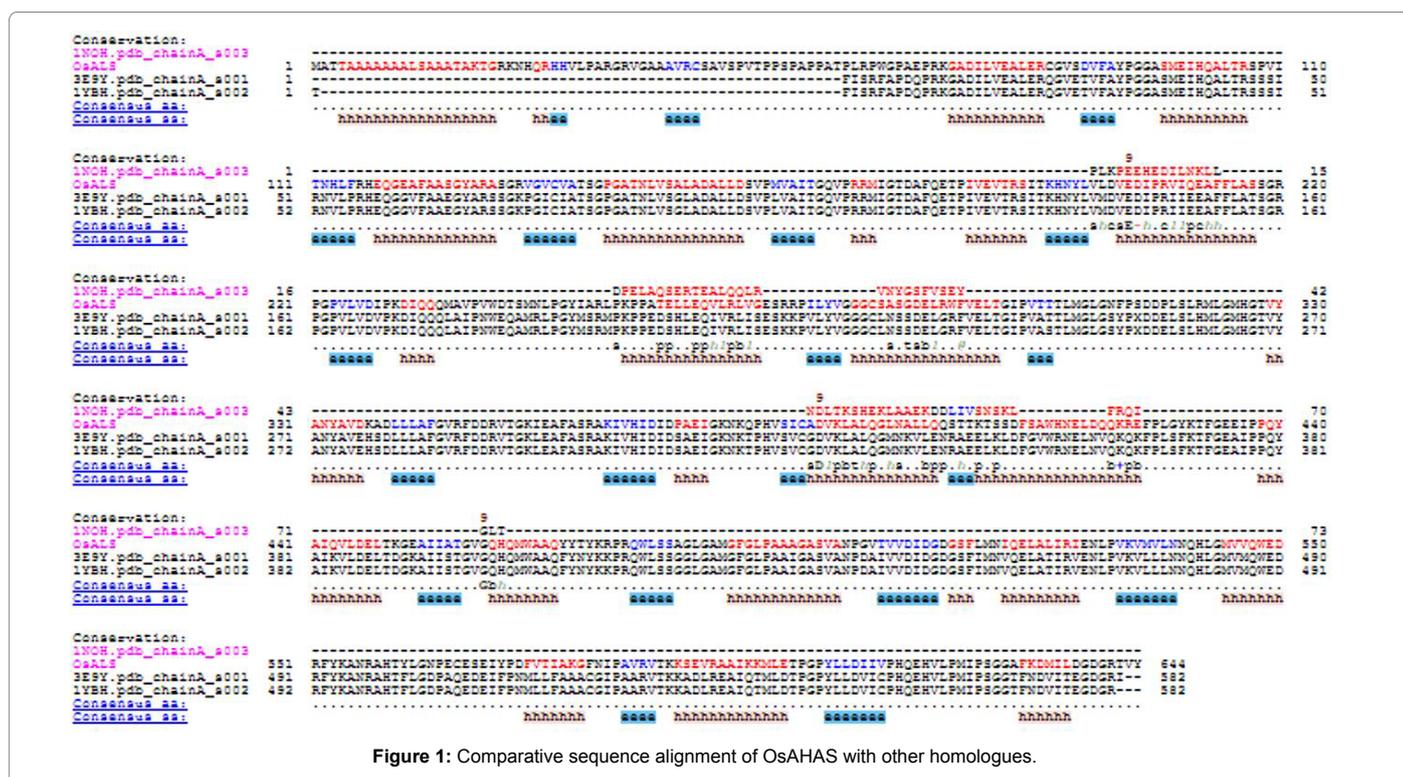
Figure 1: Comparative sequence alignment of OsAHAS with other homologues.
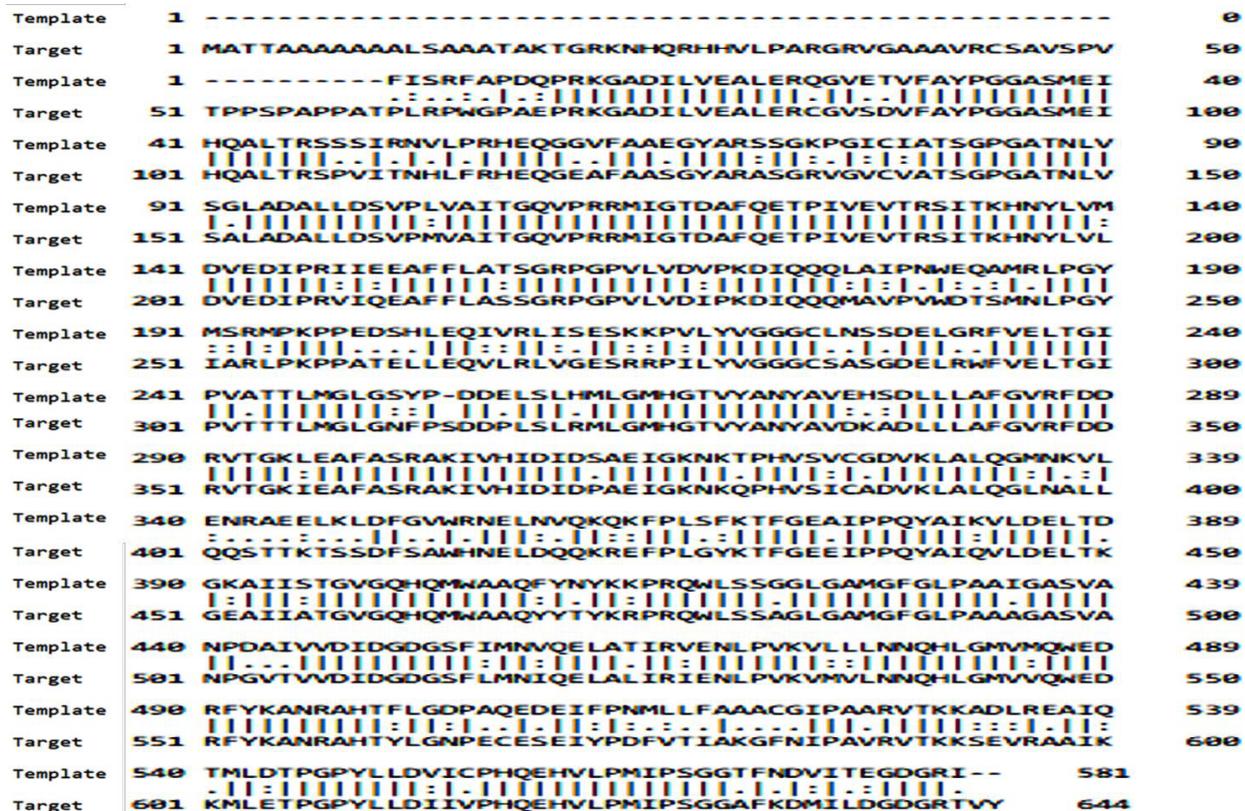
Figure 2: Comparative sequence alignment of OsAHAS (target) and AtAHAS (template) using superpose.
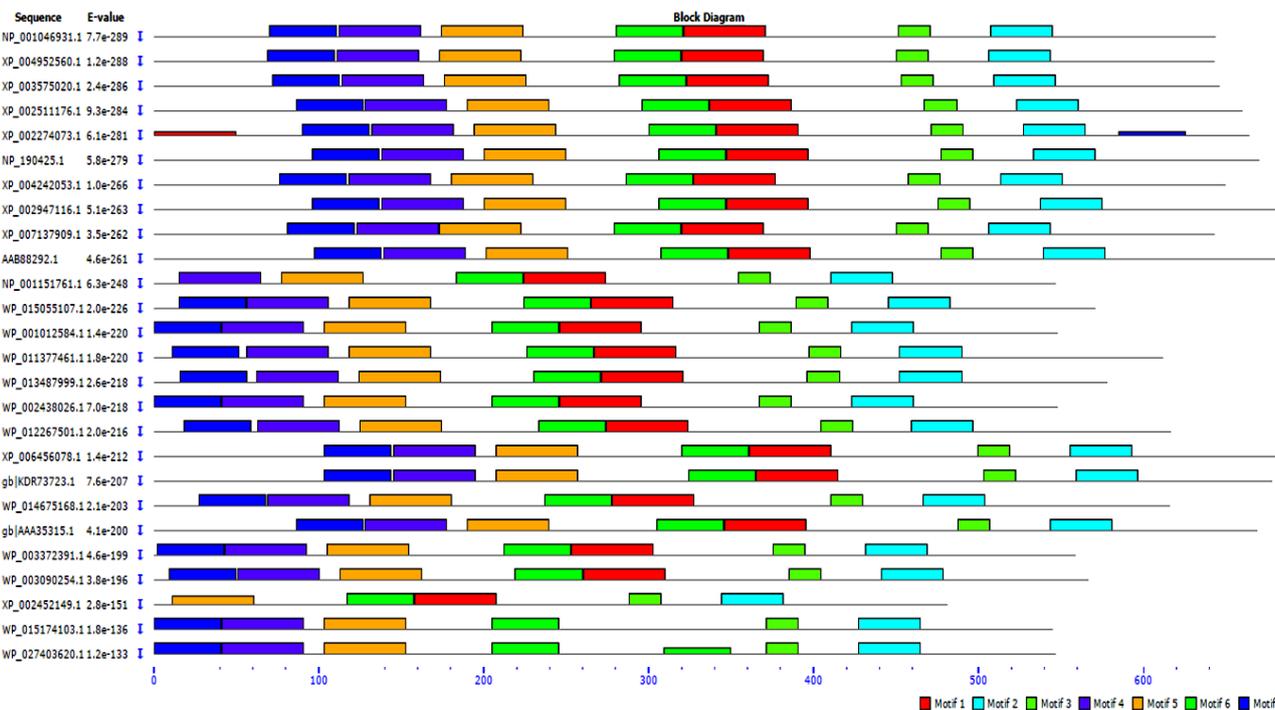
Figure 3: Block diagram of multilevel consensus sequences for the MEME defined motifs of AHAS proteins: Seven motifs were obtained by MEME software. Different motifs are indicated by different filled boxes with numbers 1 to 7.

Figure 4: Multilevel consensus sequences for the MEME defined motifs and their e values.

Figure 5: Cartoon structure of OsAHAS.

Predicted Functional Partners:

	Neighborhood	Gene Fusion	Cooccurrence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
4328319 ketol-acid reductoisomerase, chloroplast precursor, putative, expressed (578 aa)	•		•	•		•	•		0.998
OsJ_02845 ketol-acid reductoisomerase, chloroplast precursor, putative, expressed (581 aa)	•		•	•		•	•		0.997
4329938 ACT domain containing protein, expressed (484 aa)	•		•	•		•	•		0.996
4333898 threonine dehydratase biosynthetic, chloroplast precursor, putative, expressed (602 aa)	•		•	•		•	•		0.994
4333613 dehydrogenase, putative, expressed; Catalyzes the oxidation of 3-carboxy-2-hydroxy-4- methylpen [...] (408 aa)	•		•	•		•	•		0.990
OsJ_21322 ACT domain containing protein, expressed (476 aa)	•		•			•	•		0.984
4351460 2-isopropylmalate synthase B, putative, expressed (635 aa)	•		•			•	•		0.975
4346318 dihydroxy-acid dehydratase, putative, expressed (594 aa)	•		•	•		•	•		0.939
4336624 serine racemase, putative, expressed; Catalyzes the synthesis of D-serine from L-serine. Has de [...] (339 aa)	•			•		•	•		0.905
OsJ_32373 aminotransferase domain containing protein, putative, expressed (381 aa)	•		•	•		•	•		0.901

Predicted Functional Partners:

	Neighborhood	Gene Fusion	Cooccurrence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
AT3G58610 ketol-acid reductoisomerase (591 aa)	•		•	•		•	•		0.998
OMR1 L-O-methylthreonine resistant 1; Catalyzes the formation of alpha-ketobutyrate from threonine i [...] (592 aa)	•		•			•	•		0.996
VAT1 VALINE-TOLERANT 1; Regulatory subunit of acetohydroxy-acid synthase. Involved in the feed-back [...] (477 aa)	•		•			•	•		0.995
AT2G31810 ACT domain-containing small subunit of acetolactate synthase protein; Regulatory subunit of ace [...] (491 aa)	•		•			•	•		0.994
IMS1 2-isopropylmalate synthase 1; Catalyzes the condensation of the acetyl group of acetyl-CoA with [...] (631 aa)	•		•			•	•		0.991
IMD3 isopropylmalate dehydrogenase 3; Catalyzes the oxidation of 3-carboxy-2-hydroxy-4- methylpentan [...] (404 aa)	•					•	•		0.990
MAML-4 2-isopropylmalate synthase; Catalyzes the condensation of the acetyl group of acetyl-CoA with 3 [...] (631 aa)	•					•	•		0.989
IMD2 isopropylmalate dehydrogenase 2; Catalyzes the oxidation of 3-carboxy-2-hydroxy-4- methylpentan [...] (405 aa)	•					•	•		0.988
MAM1 methylthioalkylmalate synthase 1; Determines the side chain length of aliphatic glucosinolate s [...] (506 aa)	•		•				•		0.968
AT3G23940 dihydroxy-acid dehydratase (608 aa)	•			•			•		0.956

Figure 6: AHAS interacting partners as well as its coexpression genes predicted by STRING. (A) Rice (B) Arabidopsis thaliana (C) The key to the putat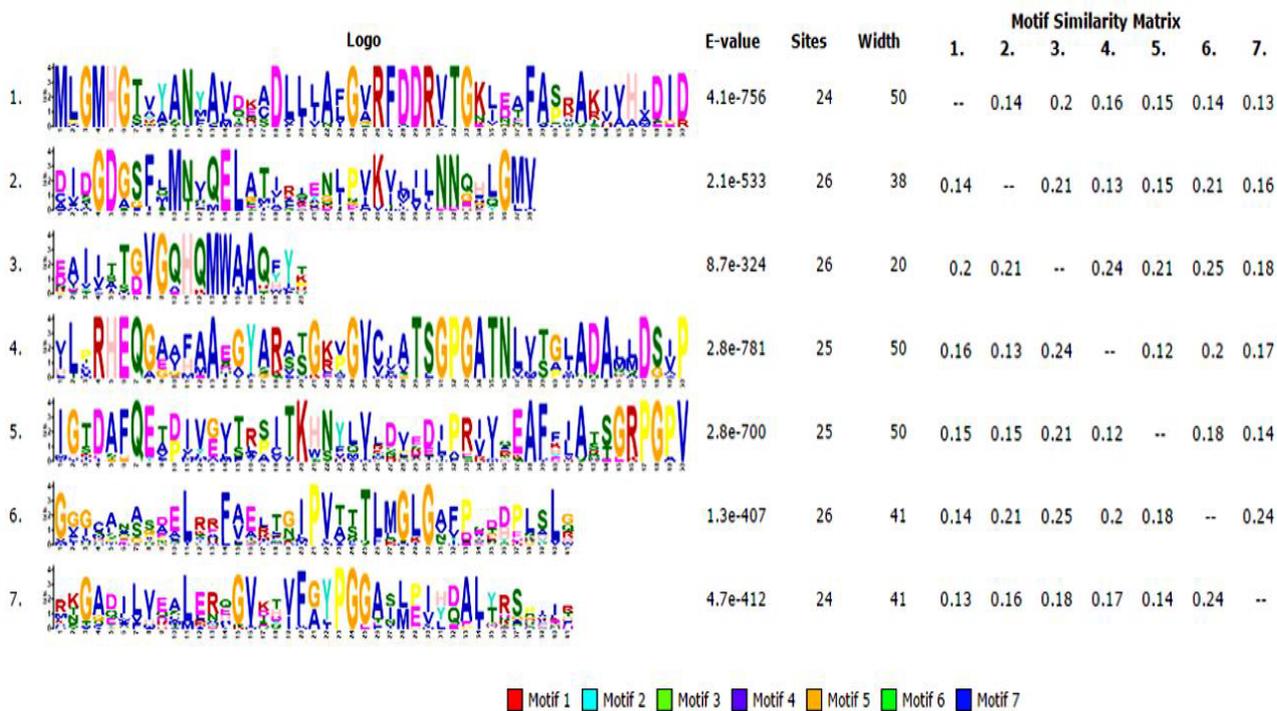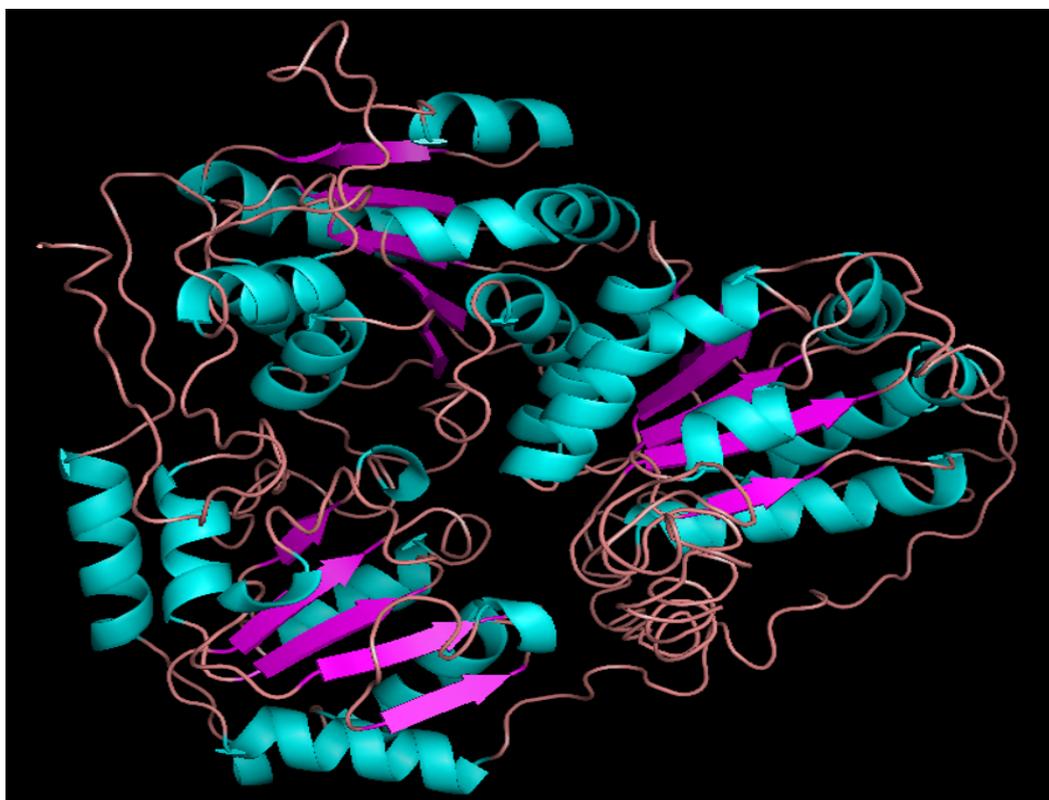ive interacting partners for OsAHAS gene is listed. (D) The key to the putative interacting partners of AtAHAS gene is listed.

Figure 7: The plot for OsAHAS designed by Rampage.

thaliana (Figure 6). Some proteins such as ketol-acid reductoisomerase, dihydroxyacid dehydratase, 2-isopropylmalate synthase and 3-isopropylmalate dehydrogenase are found to be common interacting partners of AHAS in both rice and *Arabidopsis thaliana*. These proteins are involved in the BCAA synthesis pathway for the biosynthesis of amino acids which is conserved in prokaryotes, algae, fungi and plants.

Validation of OsAHAS structure: RAMPAGE server and PROCHECK generated model revealed that 81.6% residues are falling in the most favoured region, 11.7% residues in allowed region, and 6.7% residues in outlier region of the Ramachandran plot (Figure 7). ProSA-Web analysis of the model revealed a Z-score value of target protein. The Z-score value of the target model OsAHAS (-9.55) is located within the space of proteins determined by NMR and X-ray crystallography. This Z-score value is close to the value of template 3E9Y (-11.49) which suggested that the obtained model was reliable and very close to experimentally determined structures (Figure 8a). Verify3D showed a score greater than 0.2 in 86.02% of the residues that corresponded to the quality of the OsAHAS model that was acceptable and reliable. The value of RMSD indicates the degree to which the two three dimensional structures are similar. The lesser the value, the more similar the structures are. The Cα RMSD and backbone RSMD deviation for the OsAHAS model and the AtAHAS template were 1.03Å, and 1.10 Å, respectively and overall RMSD was 1.75 Å. Thus, the OsAHAS model generated by Modeller 9.12 was confirmed to be reliable and accurate. The superimposition of the template and the model structure is shown in Figure 8b. It shows that the helix and the sheet regions of the template and model structure superimposed in a better way and a large deviation can be observed mainly in loop regions. It is reported that the loop region is the main region where the accuracy of a model protein structure deviates from the templates [43].

Phylogenetic analysis: The phylogenetic analysis of AHAS across the selected organisms showed a delineation of AHAS into four clusters - fungi, algae, bacteria and plants with some variations. Phylogenetic tree results outline the development of AHAS in *Arabidopsis thaliana*, *Aphanizomenon flos-aquae*, *Agaricus bisporus*, *Brachypodium distachyon*, *Bacillus cellulosilyticus*, *Bacillus thuringiensis*, *Clostridium botulinum*, *Chlamydomonas reinhardtii*, *Escherichia hermannii*, *E. coli*, *Galerina marginata*, *Microcystis aeruginosa*, *Oryza sativa*, *Oscillatoria nigro-viridis*, *Phaseolus vulgaris*, *Ricinus communis*, *Setaria italica*, *Sorghum bicolor*, *Solanum lycopersicum*, *Schizosaccharomyces pombe*, *Streptococcus ratti*, *Synechococcus elongatus*, *Streptomyces hygroscopicus*, *Vitis vinifera*, *Volvox carteri* and *Zea mays*. Many of these exhibited orthologous and paralogous relations with each other (Figure 9). The algae *A. flos-aquae* and *O. nigro-viridis* differs from others. Similarly, the bacteria *C. botulinum* differs from others. However, *B. distachyon* showed highest sequence similarity to OsAHAS. The results indicate that AHAS protein gene family is conserved and has evolved from bacteria and algae.

Conclusions

The homology model of OsAHAS protein was revealed using the structure of *Arabidopsis thaliana* AHAS (PDB ID: 3E9Y) as template. The resulting model structure was refined by PROCHECK, ProSA, RMSD and Verify3D that indicated the model structure is reliable with 76% amino acid sequence identity with template. The multiple sequence alignment of these AHAS protein sequences from different organisms showed conserved regions at different stretches with homology in amino acid residues. Through motif analysis, it was revealed that conserved AHAS domain are found in all AHAS proteins suggesting its possible role in cellular and metabolic functions.
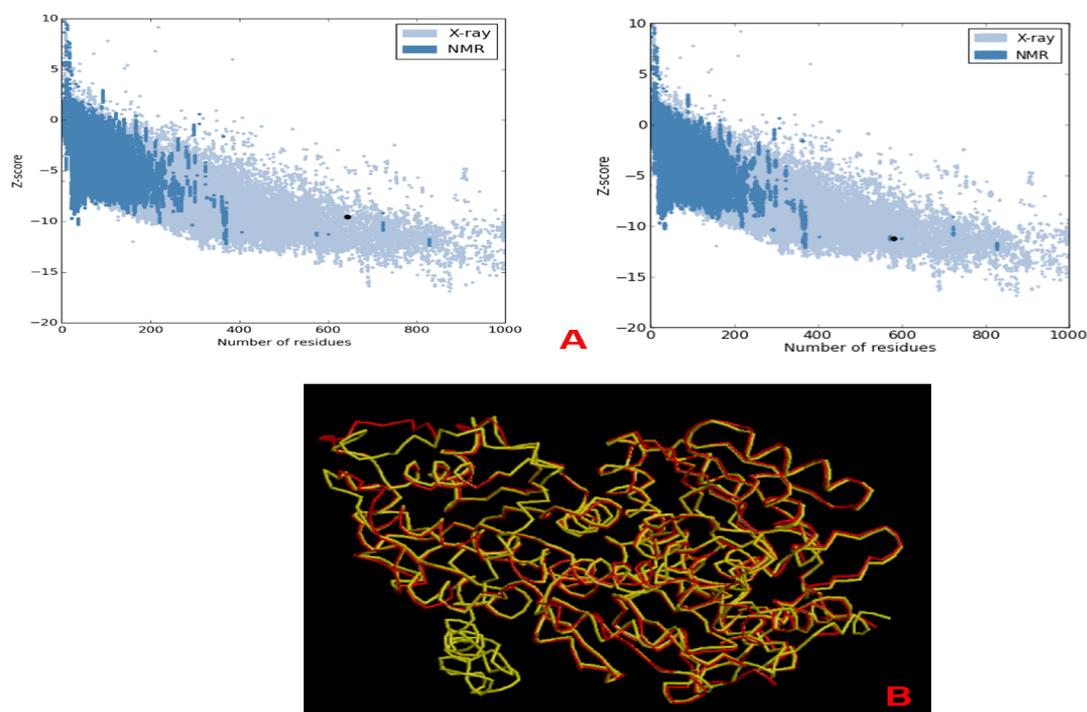

Figure 8: (A) Validation of OsAHAS by ProSA tool. The Z-score value OsAHAS (target) and AtAHAS (template) protein were determined by NMR (represented in dark blue colour) and X-ray (represented in light blue colour). The two black dots represent Z-score value of target and the template. (B) Superposition of OsAHAS (target) and AtAHAS template (PDBID: 3E9Y) shown in yellow and red colour respectively.
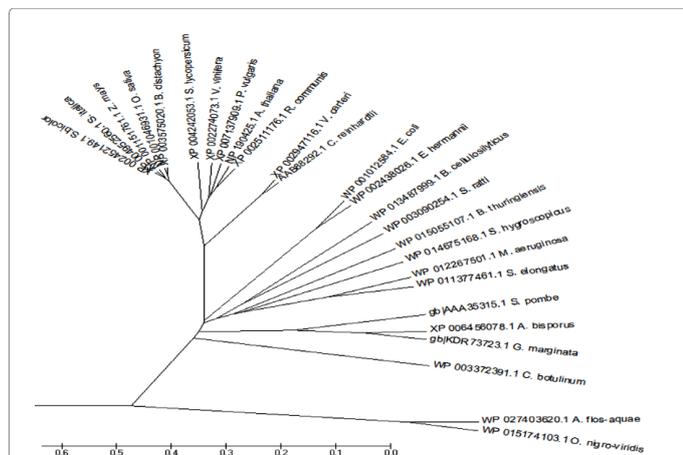
Figure 9: Phylogenetic tree constructed by UPGMA method of MEGA version 4.1 showing similarity of OsAHAS with monocots, dicots, fungi and bacteria.

Secondary structure element	OsAHAS	AtAHAS
Alpha helix	31.52%	33.22%
3_{10} helix	0.00%	0.00%
Pi helix	0.00%	0.00%
Beta bridge	0.00%	0.00%
Extended strand	22.52%	23.63%
Beta turn	9.94%	9.76%
Bend region	0.00%	0.00%
Random coil	36.02%	33.39%
Ambiguous states	0.00%	0.00%
Other states	0.00%	0.00%

Table 1: Details of the calculated secondary structure elements by SOPMA.

Properties	OsAHAS	AtAHAS
Molecular weight	69392.7	63739.1
Theoretical pI	6.48	5.47
Number of amino acids	644	584
-R	64	69
+R	60	54
Aliphatic index	92.13	93.68
Grand average of hydropathicity (GRAVY)	-0.027	-0.066
Extinction coefficients (M^{-1} cm^{-1})	64205	48735
Instability index	41.61	36.51

Table 2: Physiochemical, structural and sequence properties of the AHAS protein sequences.

Acknowledgements

Conflict of Interest

We declare that we have no conflict of interest.

References

1. Chaudhry O (2009) Herbicide-Resistance and Weed-Resistance Management.

2. Ray TB (1984) Site of action of chlorsulfuron: inhibition of valine and isoleucine biosynthesis of plants. Plant Physiol 75: 827-831.

3. Chipman D, Barak Z, Schloss JV (1998) Biosynthesis of 2-aceto-2-hydroxy acid; acetolactate synthases and acetohydroxyacid synthases. Biochim Biophys Acta 1385: 401-419.

4. Duggleby RG, McCourt JA, Guddat LW (2008) Structure and mechanism of inhibition of plant acetohydroxyacid synthase. Plant Physiol Biochem 46: 309-324.

5. Gonzalo J, Verónica J, Jordi VF, Xavier PR, Eduardo D, et al. (2010) Computational study on the carboligation reaction of acetohidroxyacid synthase: New approach on the role of the HEThDP- intermediate. Proteins: Structure Function Bioinfor 8: 774-788.

6. Umbarger HE, Brown B (1958) Isoleucine and valine metabolism in Escherichia coli. VIII. The formation of acetolactate. J Biol Chem 233: 1156-1160.

7. Cobb AH, Reade JPH (2010) Herbicides and Plant Physiology. 2nd edn. Wiley-Blackwell, Oxford, United Kingdom.

8. Yu Q, Powles SB (2014) Resistance to AHAS inhibitor herbicides: current understanding. Pest Manag Sci.

9. Shaner DL (1991) The imidazolinone herbicides. In: Shaner DL, O'Connor SL (eds.), Ann Arbor: Lewis, pp: 129-138.

10. Singh BK, Shaner DL (1995) Biosynthesis of branched chain amino acids: from test tube to field. Plant Cell 7: 935-944.

11. Levirr G, Ploeg HL, Weigel RC, Fitzgerald DJ (1981) 2-Chloro-N-[4- methoxy -6-methyl-1,3,5-triazine-2-yl] aminocarbonyl benezenesulfonamide, a new herbicide. J Agric Food Chem 29: 416-424.

12. Whitcomb CE (1999) An introduction to ALS-inhibiting herbicides. Toxic Ind Health 15: 232-240.

13. Tan S, Evans R, Singh B (2006) Herbicidal inhibitors of amino acid biosynthesis and herbicide tolerant crops. Amino Acids 30: 195-204.

14. Bernasconi P, Woodworth AR, Rosen BA, Subramanian MV (1995) A naturally occurring point mutation confers broad range tolerance to herbicides that target acetolactate synthase. J Biol Chem 270: 17381-17385.

15. Tranel PJ, Wright TR (2002) Resistance to ALS-inhibiting herbicides: what have we learned? Weed Sci 50: 700-712.

16. Lee KY, Townsend J, Tepperman J, Black M, Chui CF, et al. (1988) The molecular basis of sulfonylurea herbicide resistance in tobacco. EMBO J 7: 1241-1248.

17. Mourad G, Williams D, King J (1995) A double mutant allele, csr1-4, of Arabidopsis thaliana encodes an acetolactate synthase with altered kinetics. Planta 196: 64-68.

18. Sibony M, Michel A, Haas HU, Rubin B, Hurle K (2001) Sulfometuronresistant Amaranthus retroflexus: cross resistance and molecular basis for resistance to acetolactate synthase-inhibiting herbicides. Weed Res 41: 509-522.

19. Hattori J, Brown D, Mourad G, Labbe H, Ouellet T, et al. (1995) An acetohydroxy acid synthase mutant reveals a single site incolved in multiple herbicide resistance. Mol Gen Genet 246: 419-425.

20. Hattori J, Rutledge R, Labbe H, Brown D, Sunohara G, et al. (1992) Multiple resistance to sulfonylureas and imidazolinones conferred by and acetohydroxyacid synthase gene with separate mutations for selective resistance. Mol Gen Genet 232: 167-173.

21. Westbrook J, Feng Z, Jain S, Bhat TN, Thanki N, et al. (2002) The Protein Data Bank: unifying the archive. Nucleic Acids Res 30: 245-248.

22. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402.

23. Thompson JD, Higgins DG, Gibson TJ (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.

24. Thompson JD, Gibson TJ, Higgins DG (2002). Multiple sequence alignment using Clustal W and Clustal X. Curr Protoc Bioinformatics Chapter 2: Unit 2.3.

25. Kawabata T (2003) MATRAS: a program for protein 3D structure comparison. Nucleic Acids Res 31: 3367-3369.

26. Bailey TL, Williams N, Misleh C, Li WW (2006) MEME: discovering and analysing DNA and protein sequence motifs. Nucleic Acids Res 34: W369-W373.

27. Ikai AJ (1980) Thermostability and aliphatic index of globular proteins. J Biochem 88: 1895-1898.

28. Kyte J, Doolottle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105-132.

29. Gill SC, Von Hippel PH (1989) Extinction coefficient. Anal Biochem 182: 319-328.

30. Guruprasad K, Reddy BVP, Pandit MW (1990) Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Prot Eng 4: 155-164.

31. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, et al. (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed), The proteomics protocols handbook. Humana Press, Totowa, pp: 571-607.

32. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, et al. (2013) STRING V9.1: Protein-protein interaction networks, with increased coverage and integration. Nuc Acid Res 41: 1.

33. Morris AL, MacArthur MW, Hutchinson EG, Thornton JM (1992) Stereochemical quality of protein structure coordinates. Proteins 12: 345-364.

34. Eisenberg D, Luthy R, Bowie JU (1997) VERIFY3D: Assessment of protein models with three-dimensional profiles. Methods Enzymol 277: 396-404.

35. Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35: W407-W410.

36. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, et al. (2004) UCSF Chimera - a visualization system for exploratory research and analysis. J Comput Chem 25: 1605-1612.

37. Maiti R, Domselaar GHV, Zhang H, Wishart DS (2004) SuperPose: a simple server for sophisticated structural superposition. Nucleic Acids Res 32: 590-594.

38. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4, Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. Mol Biol Evol 24: 1596-1599.

39. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26: 283-291.

40. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, et al. (2006) Comparative protein structure modeling with Modeller. Curr Protoc Bioinformatics Chapter 5: Unit 5.6.

41. McCourt JA, Pang SS, King-Scott J, Guddat LW, Duggleby RG (2006) Herbicide binding sites revealed in the structure of plant acetohydroxyacid synthase. Proc Natl Acad Sci USA 103: 569-573.

42. Geourjon C, Deleage G (1995) SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput Appl Biosci 11: 681-684.

43. Fiser A, Feig M, Brooks CL, Sali A (2002) Evolution and physics in comparative protein structure modeling. Acc Chem Res 35: 413-421.