Study of Acetolactate Synthase and its Mechanism of Inhibition by Sulfonylurea Active Ingredients: Amidosulfuron, Nicosulfuron, Cyclosulfuron – *In-silico Approach*

Received for publication, May 12, 2020, and in revised form, May 25, 2020

Mantu Chakraborty[1], Renuka Das[2], Sunil Kanti Mondal[2*]

[1]*Department of Microbiology, Vidyasagar University, Rangamati, Midnapore, West Bengal 721102.*

[2]*Department of Biotechnology, The University of Burdwan, Golapbag, Burdwan – 713104*

Abstract

Acetolactate synthase (ALS) or Acetohydroxyacid synthase (AHAS) catalyzes the first step in the synthesis of the branched-chain amino acids i.e., valine (2-amino-3-methylbutanoic acid), leucine (2-amino-4-methylpentanoic acid), and isoleucine (2-amino-3-methylpentanoic acid), in plants, bacteria, algae and fungi but not in humans. AHAS is the main target enzyme for sulfonylurea active ingredients; Amidosulfuron, Nicosulfuron and Cyclosulfuron those assist in lowering branched-chain amino acid synthesis through inhibition to form the complex of Lactyl-ThDP(TDL) to ALS with great practical importance. Amino acid composition, evolutionary and sequence analysis of the ALS protein from *Arabidopsis thaliana* and its homologous were systematically studied. Composition analysis reveals that ALS is a soluble protein. Moreover, the phylogenetic tree showed different clusters based on the source organism and multiple sequence alignment depicts conservative nature in amino acid residues. Furthermore, molecular docking has been conducted to study the interactions between ALS of *Arabidopsis thaliana* and TDL in presence/absence of the active ingredients of sulfonylurea herbicide groups. Molecular docking studies confirm active ingredients are effective to inhibit the binding of TDL to ALS. Our obtained results can be very useful to study specific protein interactions along with developing new herbicides using computational methods.

Key Words: Acetolactate synthase (ALS), Branched-chain amino acid synthesis (BCAA), Lactyl-ThDP (TDL), Herbicide, Molecular docking.

Introduction

Herbicides are chemicals which control or kill undesirable plants commonly known as weed killers. Potent low dose herbicides are known to target Acetohydroxyacid synthase (AHAS) or Acetolactate synthase (ALS), which makes this enzyme, as an important subject of research to weed scientists. Studies particularly relating to changes in amino acids Leucine, Valine, Isoleucine in plants that confer herbicide resistance, have remained the matter of interest for intensive studies.[1] The biosynthesis of neutral, nonpolar, hydrophobic branched-chain amino acids (BCAA) i.e., Leucine, Valine, Isoleucine, having an aliphatic side-chain with a branch (a central carbon atom bound to three or more carbon atoms) in plants, occurs through a series of parallel reactions, mediated by AHAS (E.C

Corresponding author
Email address: skmondal@biotech.buruniv.ac.in

4.1.3.18) or acetolactate synthase (ALS).[2][3]

Weed Science Society of America (WSSA) and Herbicide Resistance Action Committee (HRAC) divided herbicides according to their mode of action into 27 groups (https://www.intechopen.com/books/herbicides-physiology-of-action-and-safety/modes-of-action-of-different-classes-of-herbicides) like lipid synthesis inhibitor, branched or aromatic amino acid synthesis inhibitor, carotenoid synthesis inhibitor (refer **supplementary Table 1.**) etc.[4][5][6] Based on their mechanism of action herbicide possess an active chemical ingredient which is active in plants, affecting the plant tissue and inhibit many metabolic activities and ultimately leads to plants death.[7]

AHAS from *Arabidopsis thaliana* crystallize as a tetramer whereas AHAS from *Saccharomyces cerevisiae* acts as a dimer.[8] AHAS composed of two monomers and each monomer contains three domains α (85–269), β (281–458), γ (463–639). α domains and γ domains form subunit interface, β domains are faraway from them and play a minor role in stabilizing the dimer interface.[9] There is a C-terminal tail also called capping region (650–687) which is observable when herbicide bound to AHAS and the polypeptide segment is then called mobile loop (580–595). Two types of subunits present in AHAS, one for catalysis, the catalytic subunit that comprises the cofactor thiamine diphosphate (ThDP; also referred to as Thiamine phosphate, TPP) and the regulatory subunit significantly trigger the action of the catalytic subunit even with zero AHAS activity but can confers susceptivity to feedback inhibition by one or more of the branched-chain amino acids synthesis.[10]

Process of condensation of pyruvate to acetolactate, AHAS enzyme requires three co-factors such as Thiamine diphosphate (ThDP), divalent cation i.e., Mg2+, Flavine Adenosine Diphosphate (FAD) in complex with catalytic subunit of *Saccharomyces cerevisiae* as reported by Pang et al., in 2002.[8][9][11][12] This reaction is emphatically mediated by co-factor ThDP with the help of 2-ketobutyrate enzyme and produce 2-aceto-2-hydroxybutyrate.[2][13]

In the first step Thiamine diphosphate (ThDP) is protonated in C2 position and then ionizes to a reactive yield. Then this nucleophilic yield attracts a pyruvate to give Lactyl-ThDP (L-ThDP). At third step ThDP is decarboxylated in presence of pyruvate to give resonating Hydroxyethyl-ThDP/emine intermediates. Emine of Hydroxyethyl–ThDP undergoes separation of charge to give α-carbanion that reacts with another pyruvate or 2-keto acid and to generate product complex. Finally, product is released and ThDP regenerated.[14][15]

Sulfonylurea herbicides were discovered by Levitt and his colleagues, the most active herbicide for inhibition of AHAS enzyme.[16] Sulfonylurea forms complex with Acetolactate synthase (ALS) of *Arabidopsis thaliana* and inhibits binding of pyruvate to ALS by the way of competitive or uncompetitive inhibition.[17][18] It contains an aromatic ring attached to the sulphur atom by the sulfonyl urea bridge and a heterocyclic ring attached to the nitrogen atom. Aromatic ring is ortho substituted where the heterocyclic ring is meta substituted. Earlier studies have been carried out on the expression, purification, crystallization, structural

modelling of AHAS and describe how AHAS bound with some sulfonylurea herbicide but the mechanism of inhibition is not clear till date.[9][19][20][21]
In this paper, molecular docking has been conducted to study the interaction between AHAS of *Arabidopsis thaliana* and Amidosulfuron, Nicosulfuron and Cyclosulfamuron; selective active ingredients of sulfonylurea herbicide group to know their inhibition mechanism. Amino acid composition, evolutionary and conserved domain analysis of AHAS had also been evaluated. Our analysis will be beneficial for further development of new and effective herbicides in future.

2. Materials and method

2.1 Retrieval of the sequences

Initially we have collected amino acid sequence of ALS of *Arabidopsis thaliana* genome from Plant ensembl (http://plants.ensembl.org/index.html). Then we have searched for the homologous sequences using blastp (Basic Local Alignment Search Tool – protein to protein) 2.2.32 (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and ALS protein sequences of *Arabidopsis thaliana* as query.[22][23][24] On the basis of query coverage 70% and e-value 0.0 we have selected most related 53 sequences from different organisms (refer in **supplementary Table 2.**). FASTA format of those homologous ALS protein sequences were retrieved from NCBI (National Center for Biotechnology Information).

2.2 Analysis of the protein primary and secondary structure and prediction of the conserved domains

We know that Domains are responsible for function and interaction of particular protein and these are the conserved part of protein sequence. To identify domains present in these selected protein sequences, we have analysed the sequences by using conserved domain search tool (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).[25] The secondary structural characteristics/content was predicted through Self-Optimized Prediction Method with Alignment (SOPMA) tool. It is available free from http://npsaprabi.ibcp.fr/cgibin/npsa_automat.pl?page=/NPSA/npsa_sopma.html.[26]

2.3 Multiple Sequence Alignment and Phylogenetic Analysis

To contemplate the concise sequence alignment of the selected homologous ALS protein sequences from different organisms along with *Arabidopsis thaliana*, multiple sequence alignment was executed using Clustal Omega tool (www.ebi.ac.uk/tools/msa/clustalw2) with default parameter values and jalview display option was used to visualize along with sequence logo option for the same. A 1000-time boot strapped phylogenetic tree was generated following neighbour joining method based on the multiple sequence alignment of their protein sequences and displayed using Molecular Evolutionary Genetic Analysis (MEGA V6.06) program, which will depicts the evolutionary relationships among those selected organisms on the basis of the mutation acquired by AHAS.[22][27]

2.4 Physicochemical property

Physiochemical properties of selected proteins including molecular weight, theoretical pI, negative charge, positive charge, instability index, GRAVY (grand average hydropathy) etc. was calculated by using ProtParam tool available on Expert Protein Analysis System (ExPASy) proteomics server (http://web.expasy.org/protparam/).[25][28]

2.5 Statistical analysis

Individual amino acid frequencies of the selected ALS protein sequences were calculated using in-house PERL script. Some statistical analysis such as average, standard deviation and Z value of amino acids frequencies and groups on the basis of amino acids physicochemical properties for those ALS proteins were measured in MS Excel 2007.[29][30]

$$Z=(X_c-X_g)/\sqrt{((S_c^2+S_g^2)/(N_c+N_g)}$$

Here 'X_i' and 'S_i' represents the average and standard deviation of individual amino acids of AHAS from the selected organisms where 'i' represent c and g to denote Brassicales and Rosales order respectively and 'N' represents the sample size.[31]

2.6 Identification of interacting proteins of ALS

We have searched within STRING online resources at http://string-db.org/ to know the interacting proteins of ALS in *Arabidopsis thaliana.*[32] Ketoacid reductoisomerase, Dihydroxyacid dehydratase, 2-isopropylmalate synthase and 3-isopropylmalate dehydrogenase are procured as frequently interacting proteins of ALS in *Arabidopsis thaliana*.

2.7 Retrieval of AHAS (Arabidopsis thaliana) homologous sequences and their three dimensional structure

Again we have used blastp and protein sequence of ALS from *Arabidopsis thaliana* against PDB (Protein Data Bank) (http://www.rcsb.org) to search for the related sequences whose 3D structure is also available. From blastp result, protein sequences were collected based on query coverage 82% and e-value 8e-47 as threshold. ALS sequences in FASTA format from 6 related organisms including *Arabidopsis thaliana(ath), Cadida albicans SC 5314(cal), Saccharomyces cerevisiae(sce), Pseudomonas protegens(pfl), Basillus subtilis PY79(bsp)* and *Klebsiella pneumoniae(kpn*) were retrieved from NCBI and boot strapped (1000 times) phylogeny analysis of these six sequences was performed through MEGA V6.06 to know their closeness.[27] The PDB id's of the selected sequences are as follows 5WJ1(ath), 6DEK(cal), 1JSC(sce), 5AHK(pfl), 4RJJ(bsp) and 1OZF(kpn). Further more, structure alignment was also performed to see structural similarity between the closely related sequences using PyMOL (version 2.1.0).[33][34]

2.8 Molecular Docking study

Molecular docking is the computational study that involves interaction of two or more molecules (e.g., drug and protein or enzyme) with ligand (small molecules) and predicts the ligand binding site as well as ligand affinity of the targeted molecule (protein).[35]
For docking, at first 3D structure of active ingredients of sulphonylurea chemical family herbicides: Amidosulfuron (CID: 91777), Nicosulfuron (CID: 73281), Cyclosulfamuron (CID: 6451137) and Lacty-lThDP(L-ThDP/TDL were collected from PubChem (https://pubchem.ncbi.nlm.nih.gov/) and reserved the 3D structure as SDF file. That SDF file was converted to PDB file by using Auto dock and UCSF Chimera (an extensible software platform).[36] PDB file of ALS protein of *Arabidopsis thaliana* (PDB ID 5K6Q) was retrieved from RCSB PDB protein data bank (https://www.rcsb.org/) following molecular docking by using Autodock vina tool and Autodock (version 4.2.6) tool, forming a grid above the selected protein structure. The selected ligands were bind to targeted protein and this compound was then saved as PDB file. The best interaction between the ALS protein and active ingredients was displayed through LIGPLOT+ v.2.1 tool.[37] We have used 5K6Q in place of 5WJ1 as the later one is in a complex with other small molecule whereas the previous one is only the structure of *Arabidopsis thaliana*, acetohydroxyacid synthase catalytic subunit and their sequences are identical to each other along with their structural similarity judgment through PyMOL.

3. Results and Discussions

3.1 Conserved Domain Analysis

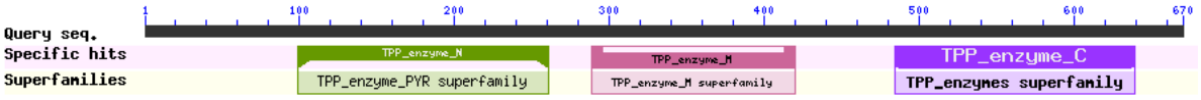


Fig1. Conserved domain structure of ALS protein from *Arabidopsis thaliana* obtained from NCBI Conserved Domain Search – NIH.

A number of enzymes require thiamine pyrophosphate (TPP) (vitamin B1) as a cofactor. It has been shown that some of these enzymes are structurally related. Some of them are TPP_enzyme_N, TPP_enzyme_M and TPP_enzyme_C. The N-terminal TPP binding domain of TPP enzymes represents the alpha(α) subunit, or the N-terminal region. The M-terminal TPP binding domain of TPP enzymes represents the gamma(γ) subunit, or the M-terminal region. Nearly every member of the C-terminal TPP binding domain of TPP enzymes is the beta(β) subunit, or else the C-terminal region. Aligning query sequence of TPP_enzyme_N, TPP_enzyme_M and TPP_enzyme_C are approximately ranges from 100-250, 290-420, 485-640 respectively. Conserved domain structure of ALS protein of every considerable organisms (fig.1) are quite similar to each other (for complete graphical representation refer **supplementary** (fig.1).

3.2 Multiple Sequence Alignment (MSA)

Multiple sequence alignment (MSA) methods refer to a series of algorithmic solution for aligning more than two biological sequences (DNA, RNA or Protein) at a time considering evolutionary events (mutations, insertions, deletions etc.).

Multiple sequence alignment was performed to study amino acid conservation at different site of selected AHAS sequences. There are three conserved domains present in AHAS protein sequence of *Arabidopsis thaliana* N-terminal domain(α), M-terminal domain(γ) and C-terminal domain(β) and sequences from selected organisms i.e., TPP_enzyme_N, TPP_enzyme_M, TPP_enzyme_C. Sequence alignments have revealed conservation of interacting amino acids. Conserved domain of amino acids are present from position within alignment 140 to 302 (fig 2a.), 334 to 466 (fig 2b.) and 530 to 686 (fig2c). In the jalview representation corresponding height of the different amino acid residues in sequence logo further the colour conservation at the different sites, unveil a few variations in the ALS encoding gene, suggesting the level of conservation of amino acids.

Fully conserved amino acids (site number in sequence alignment) are as follows: G(140), D(142), L(144), A(147), E(149), L(148), A(150), G(152), V(153), V(156), F(157), A(158), Y(159), P(160), G(161), G(162), E(166), I(167), H(168), Q(169), A(170), L(171), T(172), R(173), G(200), G(203), L(217), M(241), G(243), A(253), A(255), E(264), I(272), R(287), H(300), H(301), H(302), L(334), Q(336), R(339), I(440), I(442), G(448), A(468), K(449), K(451), K(462), A(464), G(530), G(532), H(534), Q(535), M(536), W(537), A(539), Q(540), G(555), G(557), L(602), P(603), L(614), Y(625), H(631), S(635), L(683) respectively. In contrast to the fully conserved amino acid sites, the non- or semi-conserved sites equip with more beneficial information for evolutionary study. The amino acid pattern present in the non/semi conserve sites yield quite useful facts on the possible places of changes or mutations occurred in these sequences during evolution, bring out differences in the grouping of the identified proteins in different clades, as evidenced in the phylogenetic analyses (fig. 3.).

Less conserved site in sequence alignment indicates nearly similar pattern within organisms in orders.[9] From this multiple sequence alignment and jalview representation along with sequence logo, we are able to understand about the effect of mutation acquired in the ALS encoding gene and reflection on phylogeny. However many of the mutations are silent mutation, *Tarenaya hassleriana, Arabis alpina. Trema orientale, Parasponia andersonii* belongs to Brassicales and Rosales respectively and most non conserved site within ALS domains of these organisms are mutated and their effect of mutations are notable on phylogeny for these organisms (e.g., pointed as no.; i, ii, v, vi, viii) besides a few with hardly any impact over phylogenetic tree (e.g., pointed as no.; iii, iv, vii, ix, x).

A number of mutations have been mentioned below supporting the analysis carried out with jalview (For detailed record refer **supplementary Table 3. A., 3. B., 3. C.**).

ISSN **2689-6389 (Print)**
*ISSN***2687-7939 (Online)**

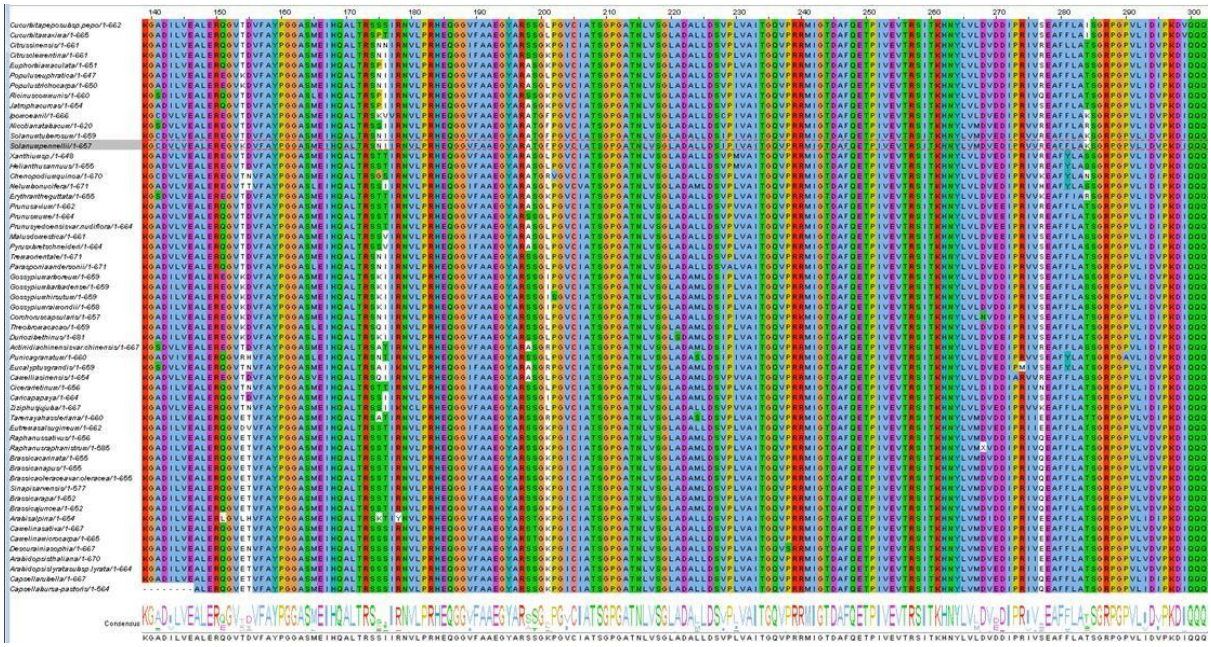


Fig 2A**.** JalView representation with sequence logo of multiple sequence alignment using Clustal Omega of the ALS sequences from selected organisms for the domain TPP_Enzyme_N

From the above jalview representation (Fig2A**.**) we have observed many non-conserved amino acids present in this conserved domain. Depending upon that we have detected a number of mutable sites as follows with positive impression over the phylogenetic tree besides some with no impact:

i. At site 151 in aligned condition is occupied by amino acids Glutamine, Glutamic acid and Leucine. Brassicales, Rosales, Myrtales, Cucurbitales are possess Glutamine. *Populas euphratica, Populus tricocarpa* belongs to Malpighiales order, Solanales, Asterales, Proteales, Laminales, Mavales, Ericales possess amino acid Glutamic acid and only *Arabis alpina* belongs to Brassicales which possess Lysine at this site. Glutamic acid and glutamine both are polar and hydrophilic whereas Lysine is nonpolar, hydrophobic and Glutamine and Lysine both are neutral whereas Glutamic acid is acidic. Therefore, at this site mutation has occurred.

Effect of mutation observed in Brassicales (*Arabis alpina*), Malpighiales (*Populas euphratica, Populas triocarpa*) on phylogenetic tree.

ii. Likewise, site 178 is almost occupied by amino acid Arginine with some differences for *Arabis alpina, Ziziphus jujube, Eucalyptus grandis* which belongs to Brassicales, Rosales, Myrtales respectively. *Arabis alpina* (Brassicales) possess amino acid Tyrosine while *Ziziphus jujuba* and *Eucalyptus grandis* possess Lysine. Arginine and Lysine are basic whereas Tyrosine is neutral. Lysine and Tyrosine are hydrophobic in nature but Arginine is hydrophilic. Therefore, mutation occurred in this site. Effect of mutation on phylogenetic tree observed in *Ziziphus jujuba* (Rosales) and *Arabis alpina* (Brassicales).

iii. Site 221 is occupied by Alanine and only *Durio zibenthus* belongs to malvales and possess Serine amino acid at this site. Alanine and Serine both are neutral but Alanine is nonpolar, hydrophobic whereas Serine is polar, hydrophilic in nature. Even while mutation occurs at this site but no phylogenetic impact was observed.

iv. Similarly, site 274 is almost occupied by Arginine. *Eucalyptus grandis* (Myrtales) possess Methionine. Arginine is basic, polar and hydrophilic whereas Methionine is neutral, nonpolar hydrophobic hence mutation occurs yet no phylogeny impact observed.

Just as in conserved domain TPP_Enzyme_N, kind of similarity noticed for the mutable sites from the above representation (Fig 2B) too.

v. Site 344 is almost occupied by amino acid Serine. *Xanthium sp., Helianthus annus* (Asterales), *Chenopodium quinoa* (Caryophyllales), *Actinidia chinensis ver chinensis* (Ericales), *Capsellarubella* (Brassicales), *Capsellabursa-pastoris* (Brassicales) occupied by amino acid Alanine.

Serine and Alanine both are neutral other than Serine is polar, hydrophilic but Alanine is nonpolar and hydrophobic. Considering this, mutation occurs at this site and phylogenetic impact is observed on Brassicales order.

vi. Site 346 is hold up with amino acid Lysine, Arginine, Methionine. *Punica granatum (Myrtales), Eutrema salsugineum, Raphanus sativas, Raphanus raphanistrum, Brassica rapa,*

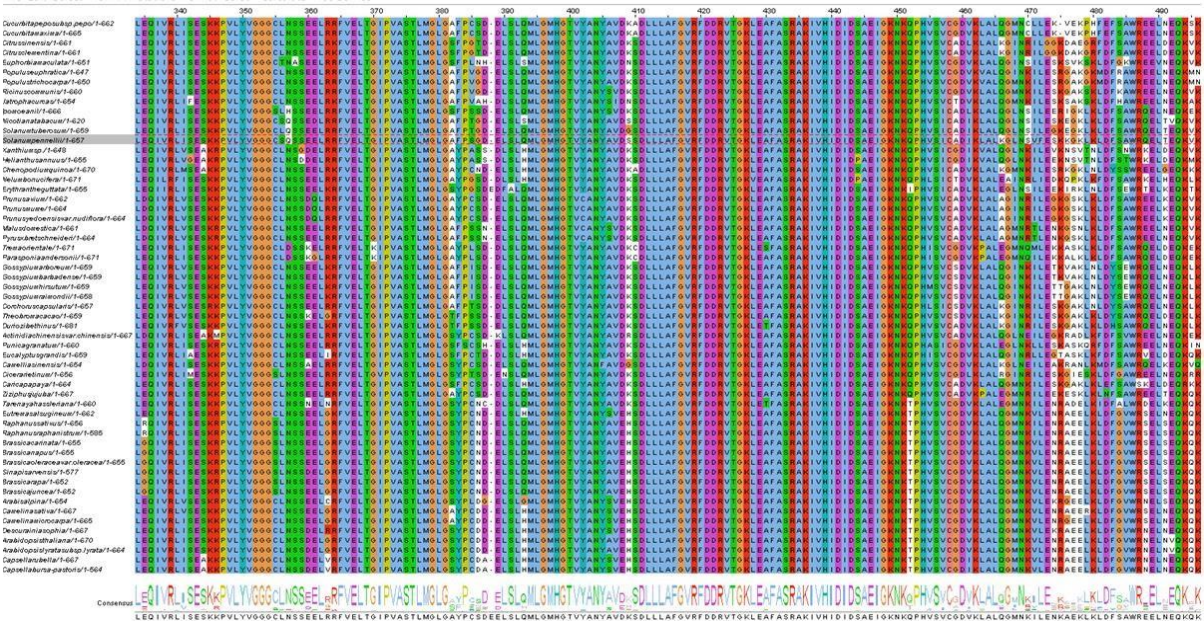


Fig2b**.** JalView representation with sequence logo of Multiple Sequence Alignment using Clustal Omega of the ALS sequences from selected organisms for the domain TPP_Enzyme_M.

Brassica juncea, Brassica carinata, Sinapis arvensis, Arabis alpine (Brassicales), Mulus domestica, Pyrus X bretsctneideri (Rosales) filled by amino acid Arginine. *Actinidia chinensis ver. chinensis* belongs to Ericales order and occupied by amino acid Methionine. Lysine and Arginine both are basic, polar and hydrophilic bow out Methionine as neutral, nonpolar and hydrophobic. Hence mutation occurs at this site and *Mulus domestica and Pyrus X bretschneideri* cut off from other plants which belongs to Rosales.

vii. In contrast, at site 405 it is almost occupied by amino acid Tyrosine. *Camelina sinesis* belongs to Ericales and filled up with Phenylalanine. Tyrosine is neutral, polar, hydrophilic in other hand Phenylalanine is neutral, nonpolar, hydrophobic. Although, mutation occurred at this site but no phylogeny impact is notable.

By the same token in the above case (fig2c.) of conserved domain TPP_Enzyme_C some inequalities are noted in the event of mutation at different mutable sites.

viii. Site 570 of alignment is almost occupied by Alanine and Serine. Only Brassicales order (except *Tarenaya orientale, Arabis alpina*) possess with amino acid Serine. Where Alanine is neutral, nonpolar, hydrophobic in nature, Serine is neutral, polar and hydrophilic. That is why, mutation occurs with evident of phylogenic impact in Brassicales order.

ix. Site 572 is occupied by amino acid Alanine. Only Solanales order possess with amino acid Glycine. Alanine and Glycine both are neutral, nonpolar but alanine is hydrophobic and glycine is hydrophilic. Therefore, mutation occurs at this site with no positive phylogeny impact.

Fig 2c. JalView representation with sequence logo of Multiple Sequence Alignment using Clustal Omega of the ALS sequences from selected organisms for the domain TPP_Enzyme_C.

x. Site 678 is almost reserved with amino acid Proline. Cucurbitales order possess with amino acid Glutamate. Proline is neutral, nonpolar, hydrophobic but Glutamine is acidic, polar and hydrophilic. Hence, mutation occurs at this site but no phylogeny impact is observed.

These above mentioned outcomes support the evolutionary relationship of ALS protein from selected organisms as obtained in phylogenetic tree.

3.3 Phylogenetic analysis

Phylogeny is the evolutionary development of an organ or of a kind of organism. Perusal of Biological evolutionary relationship between two or more organisms shared common ancestors in the past referred to as phylogenetic analysis.

Phylogenetic tree (Fig3) on the basis of AHAS sequences across the selected organisms were categorized into fourteen different clades which supports taxonomical order such as Brassicales, Rosales, malvales, malpighiales, Solanales, Asterales, Laminales, Ericales, Caryophyllales, Cucurbitales, Myrtales, Fabales, Sapindales according to their phylogenetic tree. The twelve different colors in taxonomical order indicate members of the corresponding clades.

Cucurbetales and Fabales are close to each other and Cucurbetales are also related to Malvales. Brassicales are close to Fabales whereas *Carica papaya* situated in Brassicales close to Malvales. *Helianthus annus* belongs to Asterales and *Xanthiam sp.* is closely related to Asterales. Therefore, sequence of ALS from *Xanthiam sp.* is similar to *Helianthus annus*. Solanales, Caryophyllales and Laminales orders are closely associated to each other. Rosales are situated close to Proteales and Solanales. *Ricinus communis, Jatropha curcas, Populus euphratica, Populus trichocarpa* are belong to Malpighiales but *Ricinus communis* and *Jatropha curcas* are more close to *Euphorbia maculate* than, *Populus euphratica, Populus trichocarpa with* Sapindales.

This phylogenetic tree indicates that Brassicales order is the most common ancestor among the selected organisms.

3.4 Analysis of physicochemical properties

Next we have analyzed the physicochemical properties of ALS protein from *Arabidopsis thaliana* and other selected organisms to determine the molecular characteristics. The ALS protein comprised of an average of 600-670 amino acid residues with an average molecular weight of ~65-72 k Da. From physicochemical analysis, obtained grand average hydropathy (GRAVY) score is negative. Higher (positive) GRAVY score indicates that the protein is membrane bound protein where as negative GRAVY score suggests that the protein is soluble in nature. Hence, ALS from *Arabidopsis thaliana* and selected organisms is soluble in nature.[8]

Protein structure and function is contributed by its compositional/constituent amino acids.[20][28] Each of the 20 most common amino acids, with their specific chemical characteristics, contributes their unique role to determine a protein structure as well as its function. Significant positive Z-values were obtained for amino acids glutamate, methionine, glycine and isoleucine, indicates that these amino acids are present with higher frequencies within ALS from Brassicales, whereas, significant negative Z-values for amino acids serine, alanine and proline indicates the presence to a greater extent of these amino acids within ALS from Rosales (**Table$_1$.**). Non-polar aliphatic side groups like methionine, glycine and isoleucine are hydrophobic ("water fearing") in nature. In aqueous solutions, the protein which are made of these amino acids, fold into their characteristic three-dimensional shape, to protect these hydrophobic side chains in the protein interior.

***ISSN* 2689-6389 (Print)**
***ISSN*2687-7939 (Online)**

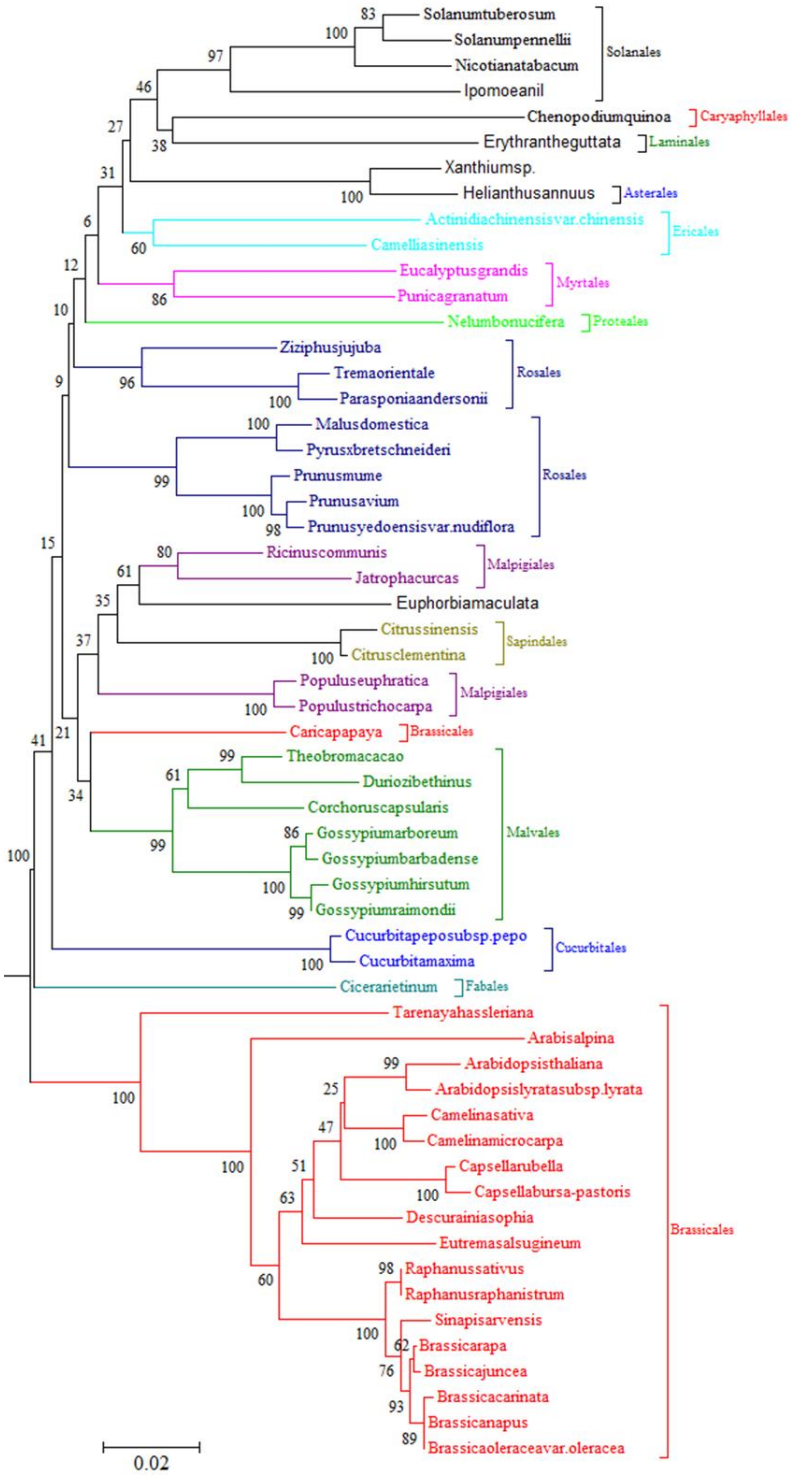


Fig. 3. Phylogenetic relationship of ALS of different organisms on the basis of AHAS sequences using the Neighbour Joining (NJ) method. Amino acid sequences were aligned with Clustal Omega tool (www.ebi.ac.uk/tools/msa/clustalw2) and the phylogenetic tree was constructed using MEGA V6.06 software (http://www.megasoftware.net/). The branching order is validated by 1000 steps of bootstrap replicates. The bootstrap values are shown at the nodes. The number in each node indicates the confidence value of that branch after bootstrapping the phylogenetic tree.

3.5 Statistical analysis

Table 1. Z – scores of different amino acids and their physic-chemical groups of ALS enzyme in two different orders i.e., Brassicales and Rosales. Z- value greater than +1.96 and less than -1.96 ($p<0.05$) are significant**.**

Amino acid	**Z- score**	**Amino acid**	**Z- score**
Negative	**5.827**	**Non polar aliphatic**	-1.575
Glutamate	**6.62**	Proline	**-10.907**
Aspartate	0.353	Methionine	**6.353**
Positive	**3.554**	Glycine	**5.679**
Lysine	0.38	Alanine	**-5.047**
Arginine	0.994	Valine	-1.479
Histidine	1.603	Isoleucine	**3.547**
Aromatic	0	Leucine	-1.616
Tyrosine	-1.058	Polar Uncharged	**-2.129**
Tryptophan	1.806	Glutamate	**2.735**
Phenylalanine	0.34	Asparagine	-1.585
		Serine	**-2.403**
		Threonine	-0.025
		Cysteine	-1.516

Table 2. Secondary structure content and responsible region within corresponding organisms to provide different secondary structure, domain architecture, active sites, binding sites etc. obtained from SOPMA for the selected ALS protein sequences

Organisms	**Alpha helix**	**Extended strand**	**Beta turn**	**Random coil**
Brassicales				
Arabidopsis thaliana	29.40%	20.75%	7.31%	42.54%
Arabidopsis lyrata subsp. lyrata	29.07%	20.33%	7.23%	43.37%
Camelina sativa	29.69%	19.04%	6.75%	44.53%
Raphanus sativus	30.79%	19.21%	7.32%	42.68%

Brassica napus	29.92%	19.08%	6.87%	44.12%
Brassica carinata	29.92%	18.78%	6.87%	44.43%
Brassica oleracea var. oleracea	29.92%	19.08%	5.95%	45.04%
Brassica rapa	31.29%	18.10%	6.29%	44.33%
Brassica juncea	31.44%	18.56%	6.90%	43.10%
Capsella rubella	30.28%	19.04%	6.75%	43.93%
Arabis alpina	29.97%	20.03%	6.57%	43.43%
Raphanus raphanistrum	33.33%	18.63%	8.03%	40.00%
Tarenaya hassleriana	31.36%	17.73%	6.36%	44.55%
Sinapis arvensis	33.97%	19.93%	8.49%	37.61%
Capsella bursa-pastoris	33.69%	19.68%	8.87%	37.77%
Carica papaya	32.53%	18.83%	6.63%	42.02%
Rosales				
Prunus mume	33.58%	17.77%	7.23%	41.42%
Ziziphus jujuba	30.28%	19.49%	7.80%	42.43%
Trema orientale	29.66%	18.63%	7.15%	44.56%
Prunus avium	29.15%	18.88%	7.10%	44.86%
Prunus yedoensis var. nudiflora	33.73%	17.47%	6.93%	41.87%
Malus domestica	29.80%	18.76%	7.11%	44.33%
Pyrus x bretschneideri	30.12%	18.67%	6.78%	44.43%
Parasponia andersonii	29.21%	19.23%	7.00%	44.56%

Brassicales and Rosales are consequences similar nature in the secondary structural characteristics within ALS among the selected groups. The percentages of the alpha(α) helix and the random coils are highest in the ALS of both the groups (**Table$_2$.**). Proteins those are in absolute dearth of well-defined secondary structure, generally exhibits Random coil structures as an alternate. Adjacent residues through the peptide bonds is the atmost alliance among the amino acids those occupy Random coil structures.

3.6 Role of interacting proteins

By using STRING software, interacting partners of ALS and its co- expression genes were identified in *Arabidopsis thaliana* (fig 4**.**) Some protein including Ketoacid reductoisomerase, Dihydroxyacid dehydratase, 2-isopropylmalate synthase and 3-isopropylmalate dehydrogenase are found to be common interacting proteins of ALS in *Arabidopsis thaliana*.In the Isoleucine biosynthesis pathway initially Threonine dehydratase converts Threonine to 2-ketobutyrate and in the 2nd step Acetohydroxyacid synthase converts 2-ketobutyrate to 2-acetohydroxybutyrate in case of Isoleucine and for Valine and Leucine it kick off the

ISSN **2689-6389 (Print)**
*ISSN***2687-7939 (Online)**

biosynthesis by forming Acetolactate from Pyruvate. Ketoacid reductoisomerase then produces 2,3-Dihydroxy-3-methylvalerate in the 3rd step for Isoleucine and Dihydroxy isovalerate in the 2nd step for Valine and Leucine biosynthesis. In the 4th step of Isoleucine, Dihydroxyacid dehydratase forms 2-Oxo-3-methylvalerate and 2-Oxoisovalerate in the 3rd step of Valine and Leucine. Aminotransferase produce Isoleucine and Valine in the final step of Isoleucine and Valine biosynthesis respectively. Now Isopropylmalate synthase forms 3-Carboxy-3-hydroxyisocaproate for synthesis of Leucine in the 5th step. Then 3-carboxy-2-hydroxyisocaproate and 2-Oxoisocaproate are produced by Isopropylmalate isomerase and Isopropylmalate dehydrogenase in the 6th and 7th step respectively. Finally Aminotransferase converts 2-Oxoisocaproate to Leucine as a final product.

These all proteins are involved in branched-chain amino acid synthesis pathway for biosynthesis of Valine, Isoleucine, and Leucine conserved in plants, fungi and algae. In plants, the syntheses of these amino acids arise by a series of analogous reactions, mediated by AHAS or ALS.[8][9][12] Unlike most bacterial and fungal AHAS that is sensitive only to Valine, a characteristic of plant that AHAS is sensitive to each of the three branched-chain amino acids is conferred by the regulatory subunits. Further determination of the 3-D structure of the plant AHAS regulatory subunits, and catalytic plus regulatory subunits, could substantially accommodate in assimilation the cross-talk between the subunits, and inevitably, how herbicide resistance mutations in the catalytic subunit affect the enzyme activity and its sensitivity.[38]

3.7 Sequence and Structural similarity among homologous sequences within PDB

Further more, the phylogenetic tree based on the retrieved homologous sequences whose structure is available along with the query sequence depicts their closeness. This phylogenetic tree (**FIG. 5.**) exhibits a clear outline of ALS into two clusters. Phylogenetic tree which we obtained resembles that sequence of ALS from *Arabidopsis thaliana* indistinguishable to the sequence of ALS from *Saccharomyces Cerevisiae* that belongs to Fungi.

3.8 Docking result

We have used UCSF Chimera and Autodock vina to perform the molecular docking that is the study of molecular interaction between the receptor and ligand molecule (refer **supplementary Table 4.**). Molecular docking studies through UCSF Chimera and Autodock vina offer 10 best predictions of interaction and for each and every possible interaction it will give a binding energy value which indicates about the significance of interaction. More negative binding energy indicates more favourable interaction between the receptor and ligand molecule. Then Ligplot + v.2.1 is used to obtained the interaction sites.[36]

3.8.1 Interaction between Acetolactate synthase (ALS) and Lactyl-ThDP (TDL)

Acetolactate synthase (ALS) of *Arabidopsis thaliana* form complex with Amidosulfuron,

ISSN 2689-6389 (Print)
*ISSN*2687-7939 (Online)

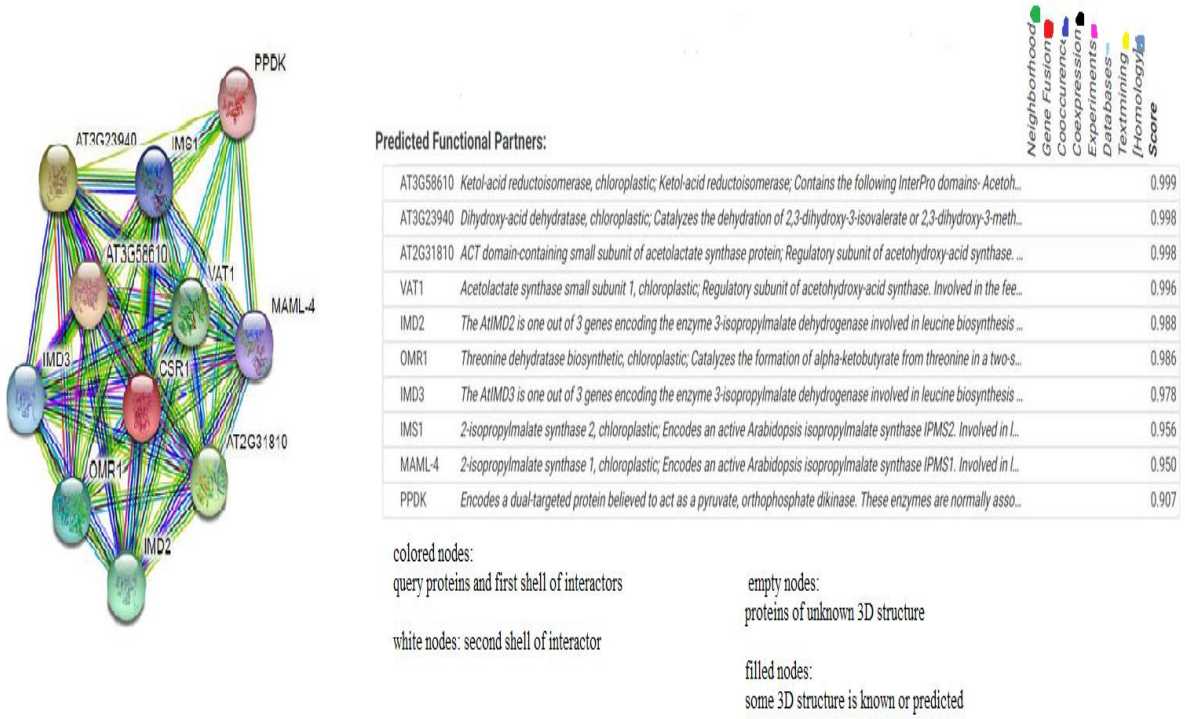


Fig 4. AHAS interacting partners as well as its co expression genes predicted by STRING software (https://string-db.org/cgi/network.pl?taskId=I7kzqYzL745x)

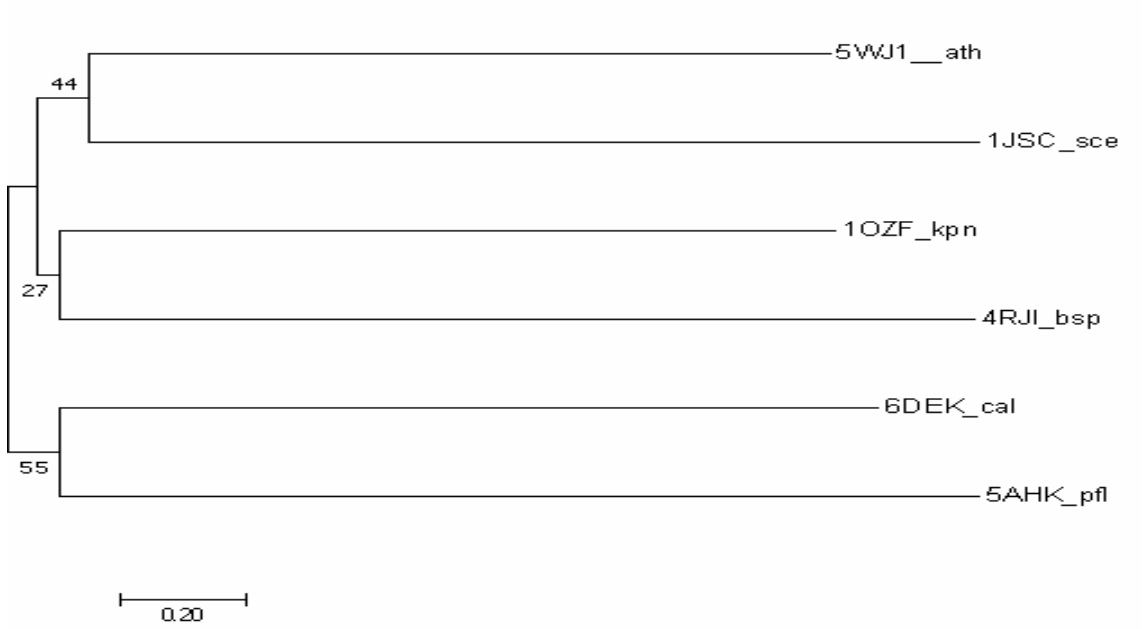


Fig. 5. 1000 times bootstrapped phylogenetic tree obtained through MEGA V 6.06 software based on the 6 homologous sequences whose structure is available in PDB.

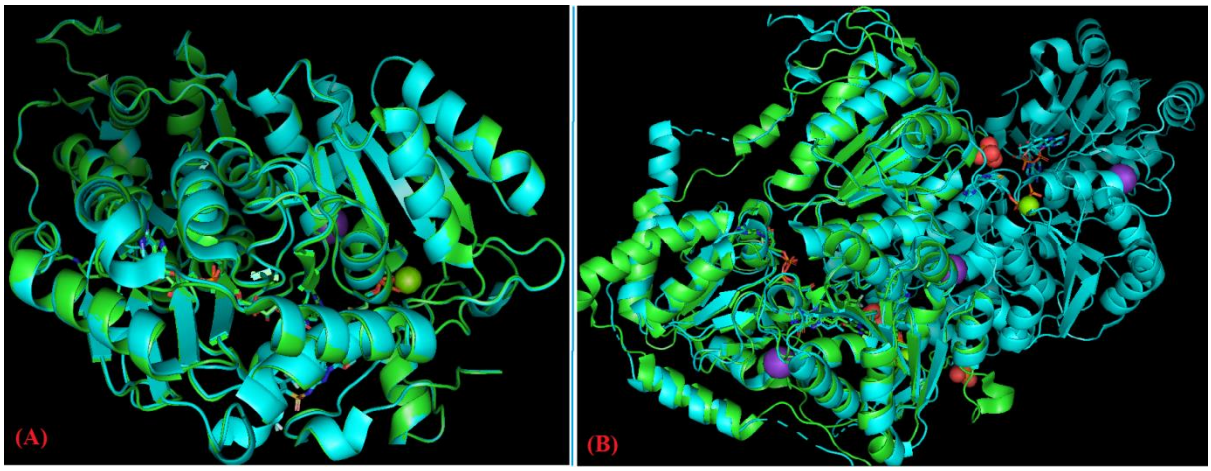


Fig. 6. Structure alignment between **A.** 5K6Q(Green) and 5WJ1(Cyans) with RMSD 0.227; and **B.** 5WJ1(Green) and 1JSC(Cyans) with RMSD 0.824.

NOTE: Moreover, RMSD (Root-Mean-Square-Deviation) values obtained from the structure alignment between 5K6Q and 5WJ1 through PyMOL is 0.227 (Fig.6a.) and between 5WJ1 and 1JSC it is 0.824 (Fig.6b.) which are very low and indicates about their structural similarities.

Nicosulfuron, Cyclosulfuron and inhibits binding of TDL (Lactyl-ThDP) to ALS. This inhibition is either competitive or uncompetitive inhibitions.[17][18]

At first we have used Acetolactate synthase as a receptor molecule and Lactyl-Thdp as a ligand. From these molecular docking studies we have obtained the interacting site molecules of ALS with TDL and binding affinity between ALS and TDL (Fig.7a.). After that we have framed an interaction between ALS and one of the considered active ingredients to form a complex of ALS and active ingredient. Next we analyzed the molecular interaction between ALS-active ingredient complex and TDL (Fig7b.). From the successive molecular docking studies we came to know how the binding affinity decreases when sulfonylurea herbicides binds to ALS and inhibit binding of TDL to ALS.

3.8.2 Interaction between ALS and Amidosulfuron a commonly used sulfonylurea herbicide

Here ALS acts as a receptor protein and Amidosulfuron (Unk1) acts as a ligand . From this molecular interaction (Fig 8a) it is clear that Amidosulfuron binds to ALS by hydrogen bond and hydrophobic bond. The responsible amino acids from ALS to form hydrogen bond are Arg246, Arg279 and for hydrophobic interactions are as follows: Gly245, Tyr276, Asp397, Leu183, Lys220 and co factor NHE 703 of chain A (Fig8b). The binding affinity for this interaction is -6.3 kcal/mol.[39]

ISSN 2689-6389 (Print)
*ISSN*2687-7939 (Online)

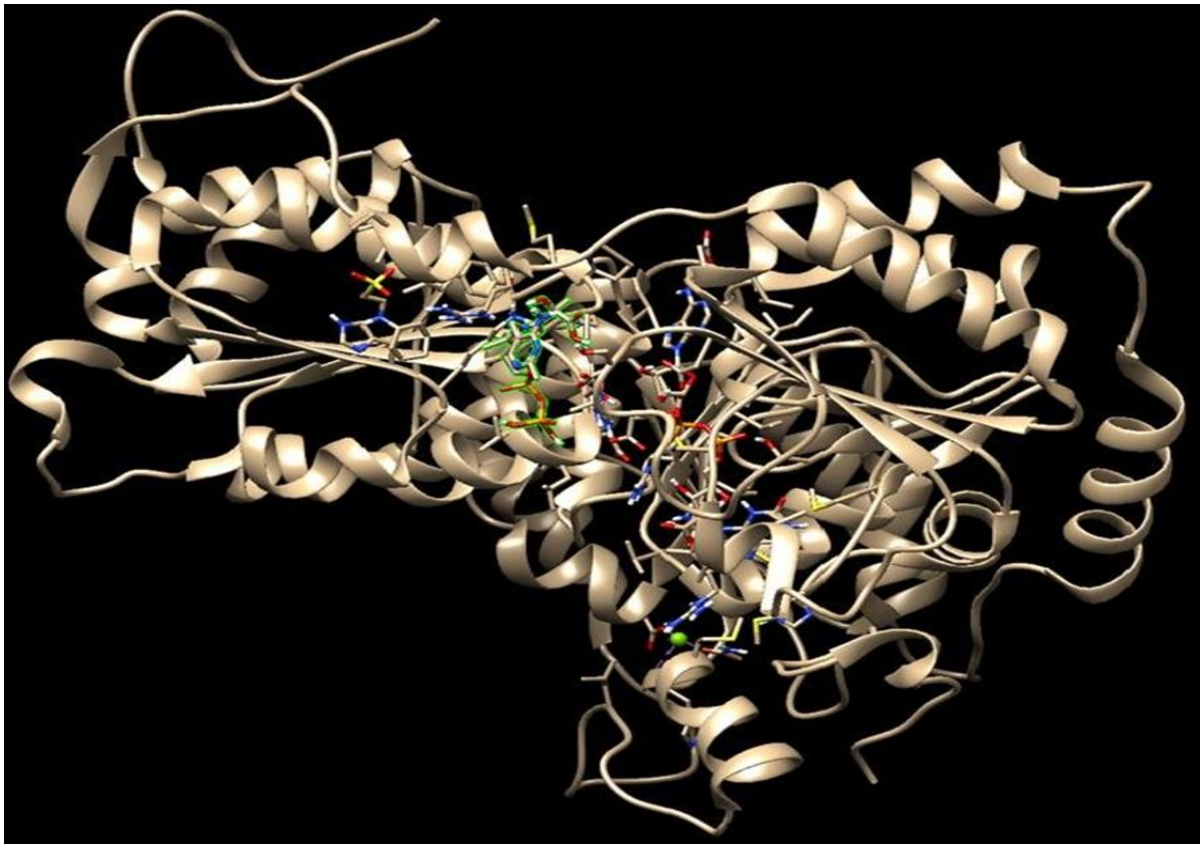
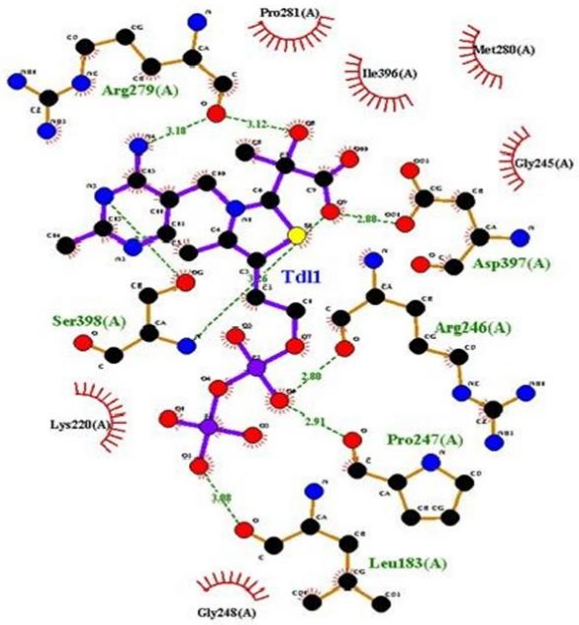


Fig.7a. Docked structure of ALS and TDL by using Autodock vina



ALS+TDL

Fig.7b**.** Interaction site of ALS and TDL by using LigPlot+ v.2.1. where green coloured dashed lines indicate hydrogen bond and red sun like symbol indicates hydrophobic bonds**.**

***ISSN* 2689-6389 (Print)**
***ISSN*2687-7939 (Online)**

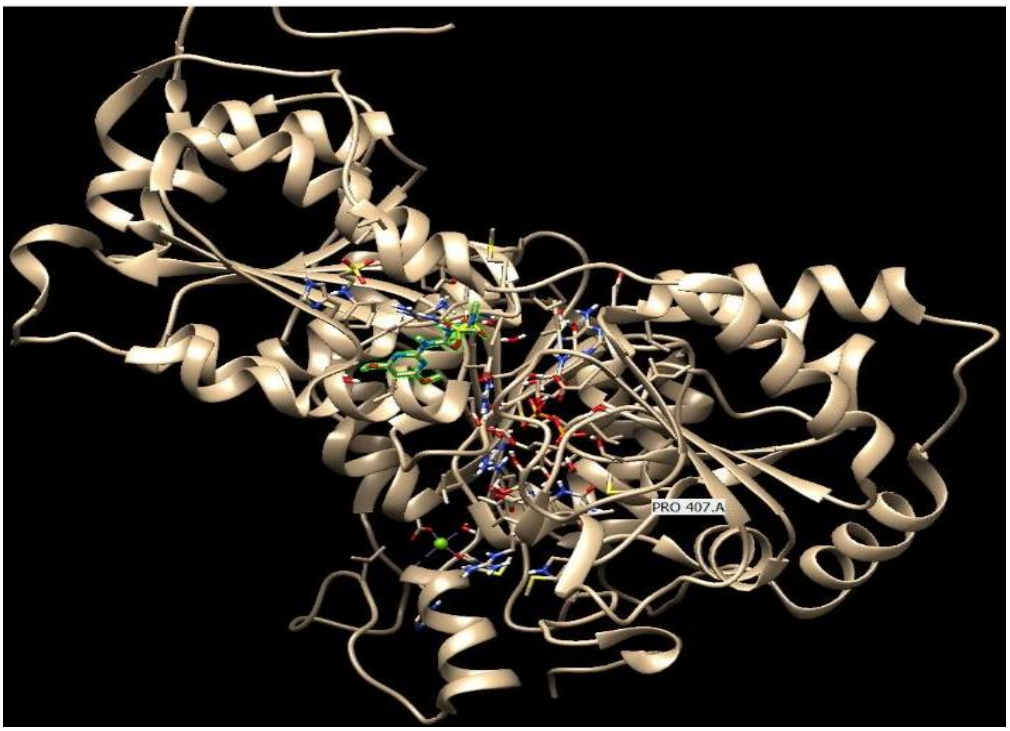


Fig 8a**.** Docked structure of Amidosulfuron and ALS using by Autodock vina

3.8.3 Interaction between TDL and ALS -Amidosulfuron complex

In this molecular docking study (Fig 9a) we can observe when TDL interacts with ALS – Amidosulfuron (ASM) complex, Arg246, Arg279 from ALS is responsible to forms hydrogen bond and Gly245, Tyr276, Asp397, Leu183, Lys220 and co-factor NHE703 forms hydrophobic bond with TDL (Fig9b). The binding energy for this interaction is -6.1 kcal/mol which is less than the binding energy between ALS and TDL i.e., -7.0 kcal/mol and more over these interaction sites are exactly same to ALS- Amidosulfuron interaction. Therefore, we have concluded that Amidosulfuron inhibits TDL by competitive inhibition.[39]

3.8.4 Interaction between ALS and Nicosulfuron (Nsf)

In this molecular docking study (Fig9A) we can observe when TDL interacts with ALS – Amidosulfuron (ASM) complex, Arg246, Arg279 from ALS is responsible to forms hydrogen bond and Gly245, Tyr276, Asp397, Leu183, Lys220 and co-factor NHE703 forms hydrophobic bond with TDL (Fig9B). The binding energy for this interaction is -6.1 kcal/mol which is less than the binding energy between ALS and TDL i.e., -7.0 kcal/mol and moreover these interaction sites are exactly same to ALS- Amidosulfuron interaction. Therefore, we have concluded that Amidosulfuron inhibits TDL by competitive inhibition.[39]

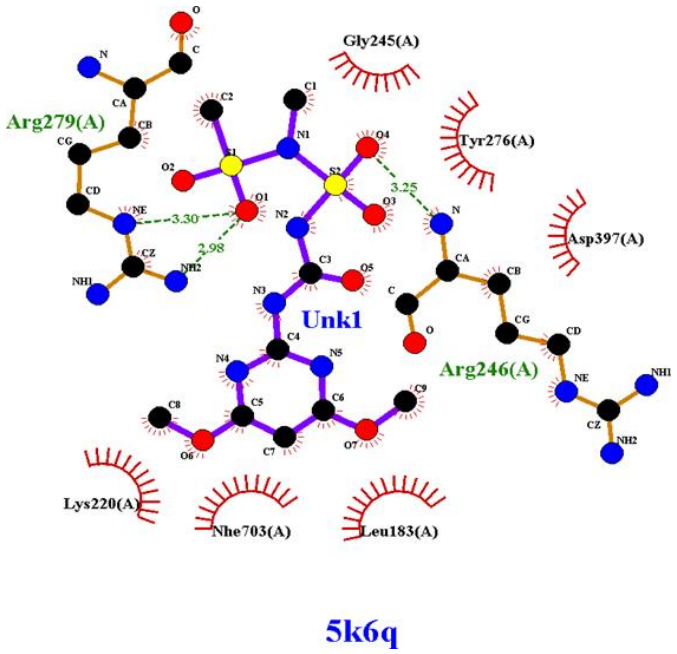


Fig.8B. Interaction site of ALS and Amidosulfuron (in picture denoted by Unk1) by using LigPlot+ v.2.1. where green colored dashed lines indicate hydrogen bond and red sun like symbol indicate hydrophobic bonds.

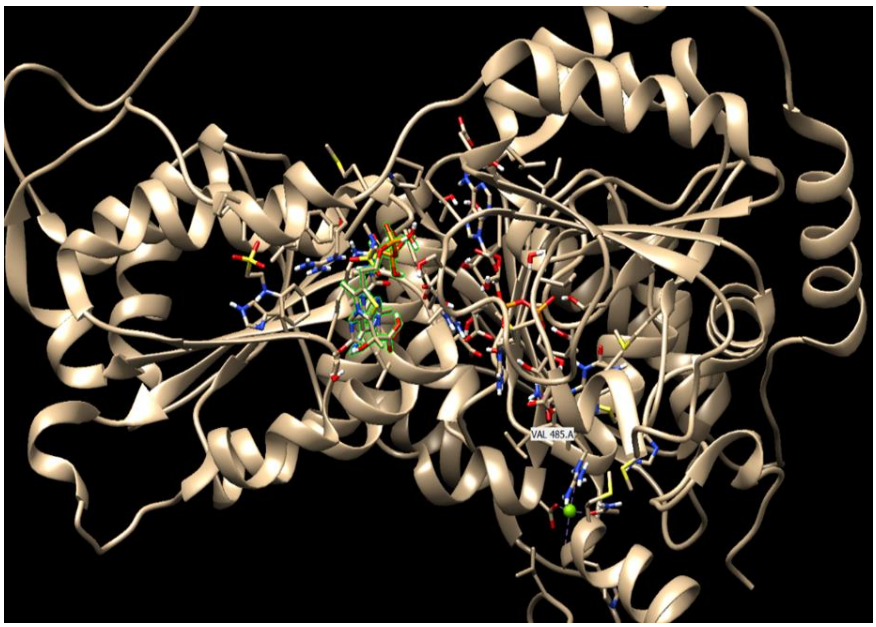


Fig 9A. docked structure of tdl binds with als -amidosulfuron complex

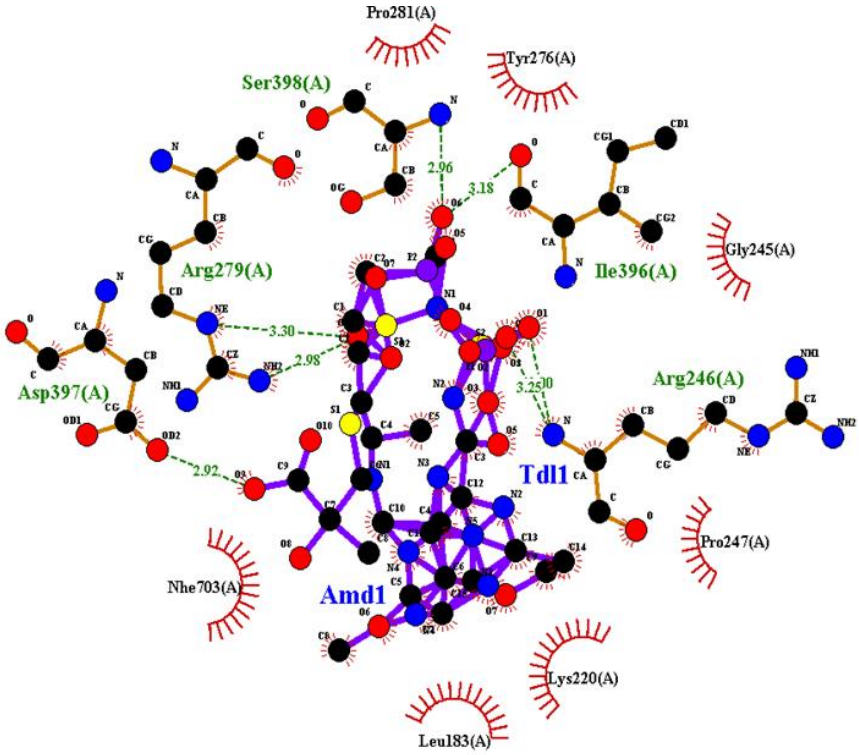


Fig 9B**.** Interaction sites of Amidosulfuron-ALS complex and TDL visualized through LigPlot+v.2.1.

3.8.4 Interaction between ALS and Nicosulfuron (Nsf)

Nicosulfuron interacts with ALS through hydrogen and hydrophobic bond formation (Fig. 10A.). Thr662, Ser454 of ALS are the participating amino acids for hydrogen bond formation and His646, Gly664, Leu684, Asp665, Glu663, Lys450, Phe452, Val355, Glu645

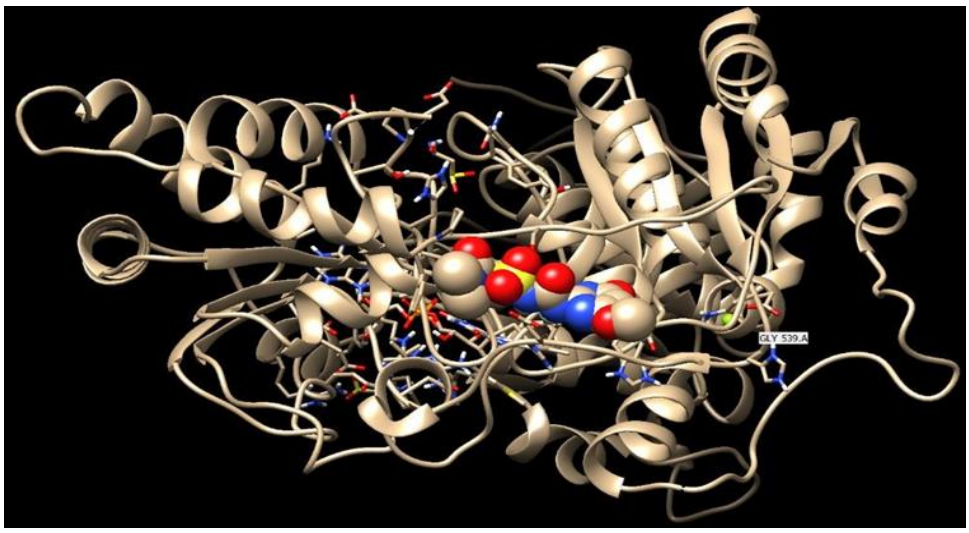


Fig 10 A**.** Docked structure of ALS and Nicosulfuron, a sulfonylurea herbicide active ingredients

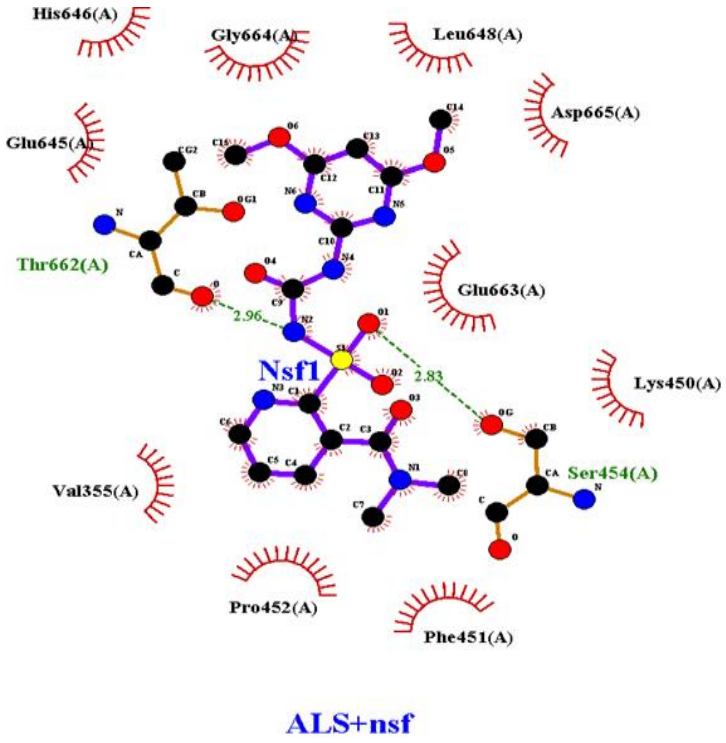


Fig 10B. Interaction site between ALS and Nicosulfuron visualized through LigPlot+ v.2.1are responsible for hydrophobic interaction with Nsf1 (Fig.10B). The binding energy for this interaction is -7.55Kcal/mole.

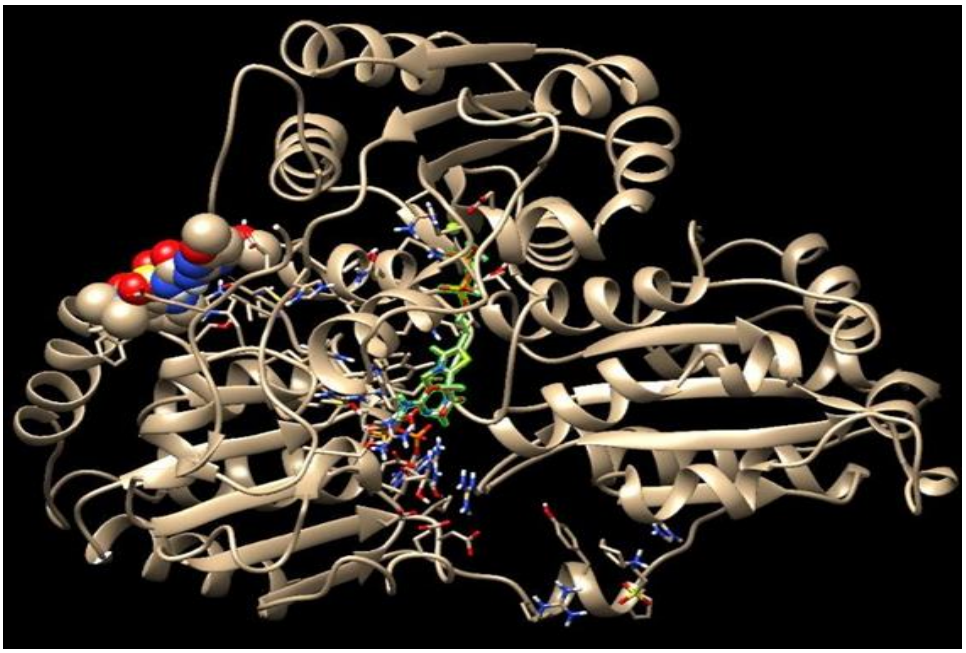


Fig. 11A. Docked structure of TDL and ALS- Nicosulfuron complex by using Autodock vina

3.8.5 Interaction between TDL and ALS -Nicosulfuron complex

TDL interacts with ALS-NSF complex (Fig 11A.) by hydrogen bond in Ser454, Thr662 and hydrophobic bond in His646, Gly664, Leu648, Asp665, Glu645, Glu663, Val355, Pro452, Lys450, Phe451 (Fig 11B**.**). Binding affinity for this interaction is -2.48Kcal/mole which is much lower than normal binding affinity -7.00 Kcal/mole of TDL- ALS. Binding sites are different from ALS-TDL binding site (Fig.7A.) and decrease the binding affinity that is why we have concluded that this inhibition is uncompetitive inhibition.

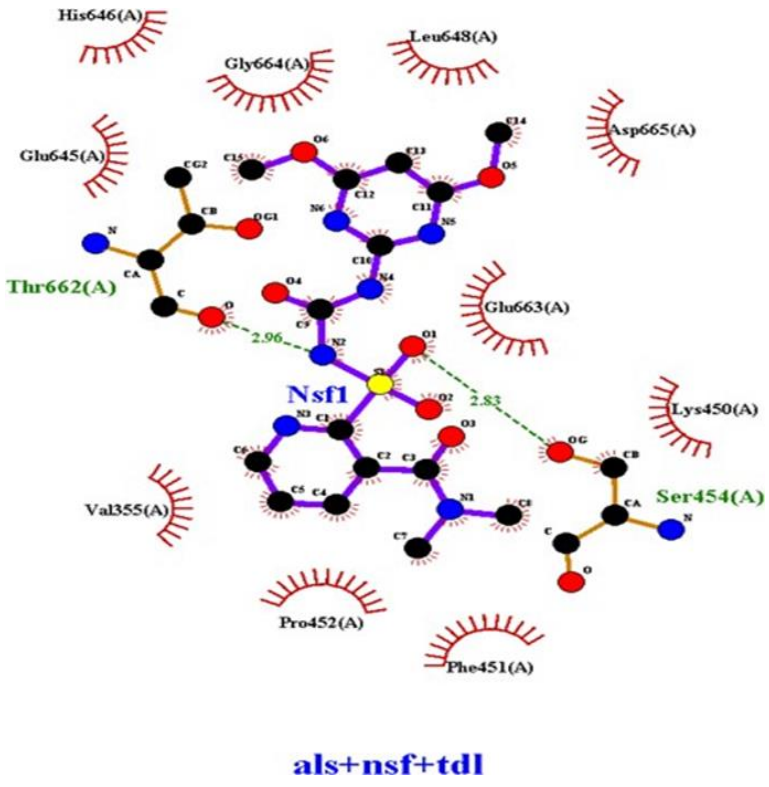


Fig11B. Interaction sites of TDL and ALS -Nicosulfuron complex obtained by Ligplot+ v.2.1

.3.8.6 Interaction between ALS- Cyclosulfamuron

Cyclosulfamuron binds to ALS (Fig12A.) by hydrogen bond and hydrophobic bond. Leu183, Thr219, Lys220, Tyr276, Met280, Gly245, Pro281 make hydrophobic bond to ALS and Arg246 and Arg279 make hydrogen bond to ALS (Fig12B.). The binding affinity for this interaction is -7.7 kcal/mol

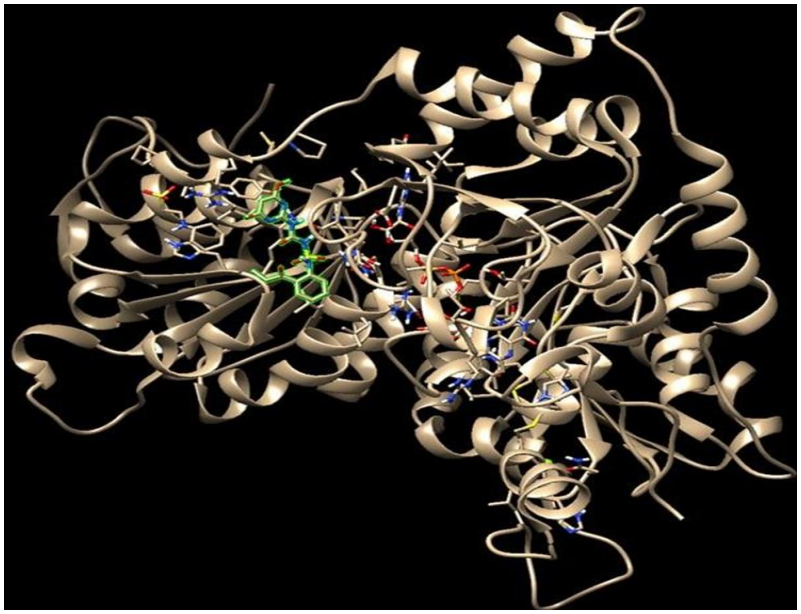


Fig. 12A**.** Docked structure of ALS and Cyclosulfamuron (CSM) interaction by using Autodock vina

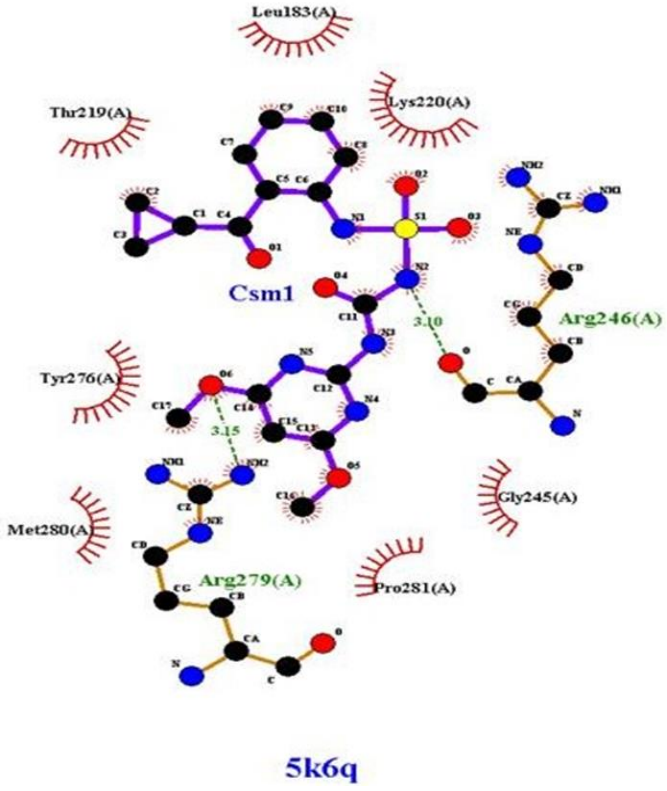


Fig. 12 B**.** Interaction sites of ALS and cyclosulfamuron

Cyclosulfamuron binds to ALS (Fig12A**.**) by hydrogen bond and hydrophobic bond. Leu183, Thr219, Lys220, Tyr276, Met280, Gly245, Pro281 make hydrophobic bond to ALS and Arg246 and Arg279 make hydrogen bond to ALS (Fig12B.). The binding affinity for this interaction is -7.7 kcal/mol.

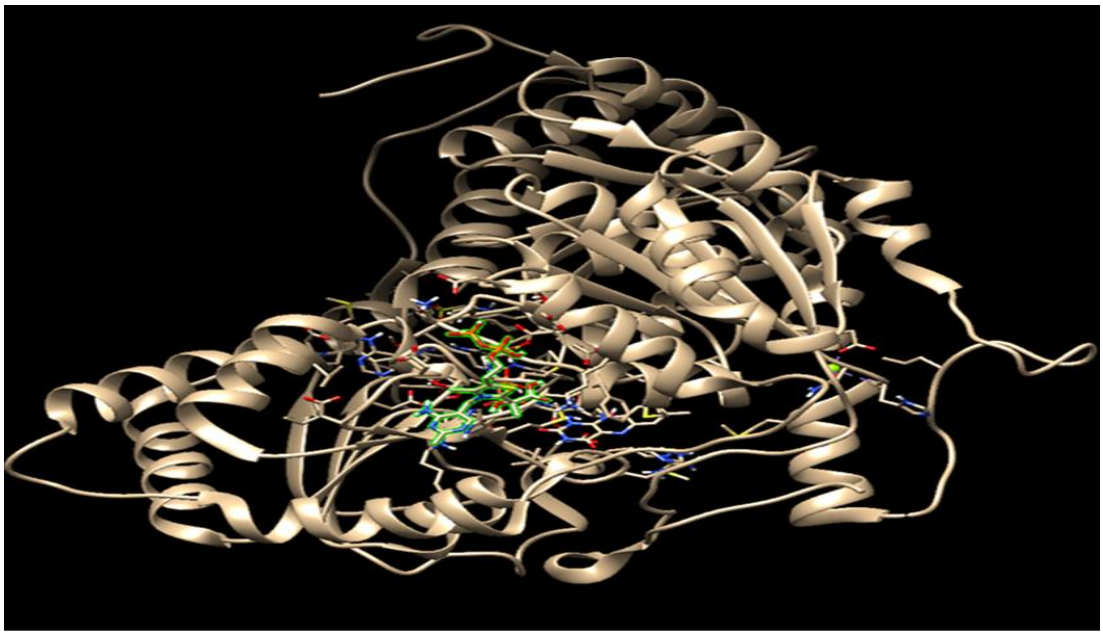


Fig. 13a**.** Docked Structure of TDL and ALS-Cyclosulfamuron complex by using Autodock vina

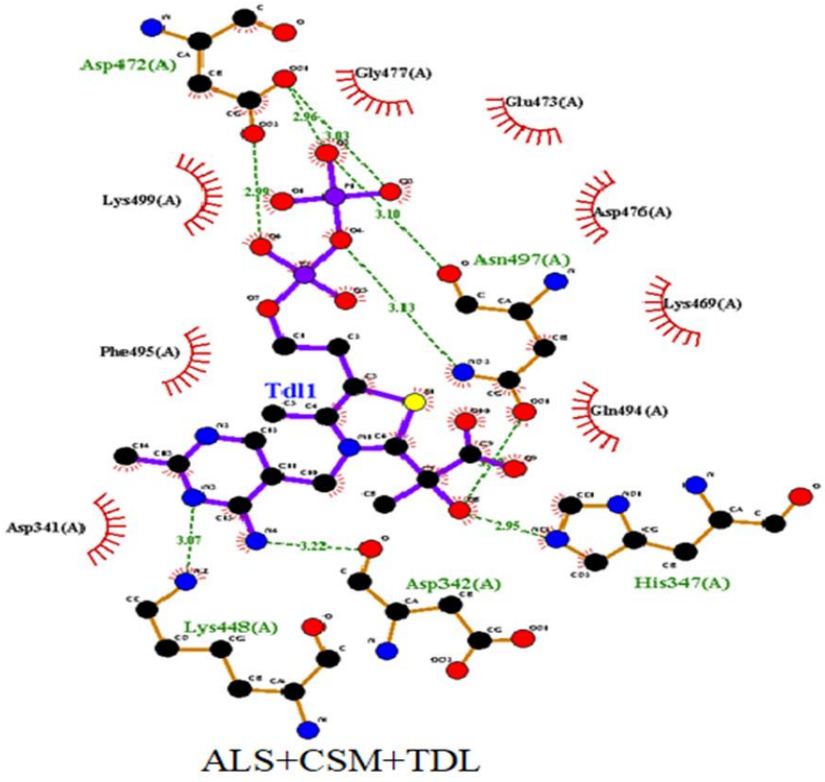


Fig.13b**.** Interaction site of TDL and ALS-Cyclosulfamuron

3.8.7 Interaction between TDL and ALS-Cyclosulfamuron complex

TDL interacts with ALS – Cyclosulfamuron complex (**FIG. 13. A.**) by hydrogen bond in Asp472, Asn497, His347, Asp342, Lys448 and by hydrophobic bond with Gly477, Glu473,

ISSN **2689-6389 (Print)**
*ISSN***2687-7939 (Online)**

Asp476, Lys469, Gln494, Lys499, Phe495, Asp341 (**FIG. 13. B.**) and binding affinity is -6.5 kcal/mol lower than ALS-TDL binding affinity. Interacting sites are different from the interacting sites of ALS and TDL (**FIG. 7. A.**). As for decreased binding affinity, we have concluded that this inhibition is uncompetitive inhibition.[12]

From docking result we can understand that binding of ALS to TDL is inhibited by either competitive or uncompetitive inhibition. After getting interacting site of ALS protein we have searched for the respective positional character in other organisms in jalview representation of multiple sequence alignment and from this analysis it is cleared that these interacting sites are totally occupied by same character irrespective of the organisms.Therefore, interacting site of active ingredients to ALS may be same for the other selected organisms too.

4. Conclusion

Acetolactate synthase (ALS) or Acetohydroxyacid synthase (AHAS) [E.C. 2.2.1.6] catalyses first step of branched-chain amino acid synthesis i.e., Valine, Isoleucine, Leucine and it is the prime target of the active ingredients to stop the growth of weeds. Multiple sequence alignment of these ALS protein sequences from different organisms showed conserved region with homology in amino acid residue at different site. Thereafter, phylogenetic tree constructed based on protein sequence of acetolactate synthase of *Arabidopsis thaliana* and it's homolog from different organisms revealed 14 cluster based on their taxonomical order such as Brassicales, Rosales, Malvales, Asterales, Malpighiales, Laminales, Proteales, Solanales, Cucurbitales, Myrtals, Fables, Ericales, Sapindales and Caryaphyllales. Moreover, from phylogenetic analysis we may conclude that Brassicales order is the common ancestor of the rest selected orders of organisms. The jalview representation of the multiple sequence alignment delineate that some of the sites within the conserved domain are not occupied by identical character. Further analysis resembles those apparently non-conserve sites might retain their conservation (also known as silent mutation) based on the similar physicochemical properties of the present amino acids in those sites and it has proper reflection in the phylogenetic tree with some exceptions and we can understand about mutation acquired at non-conserved site.

In this study we have also find out AHAS interacting partners as well as its coexpression genes were predicted in *Arabidopsis thaliana* from STRING resources. Some protein including ketoacid reductoisomerase, dihydroxyacid dehydratase, 2-isopropylmalate synthase and 3-isopropylmalate dehydrogenase are involved in branched-chain amino acid synthesis pathway of plants, fungi and algae are found to be common interacting proteins of ALS in *Arabidopsis thaliana*.

From compositional and physicochemical study we came to know that ALS protein comprises 600-670 amino acid residues and have an average molecular weight ~65-72 k Da and negative GRAVY score indicates it as a soluble protein.

The ALS enzyme is inhibited by Amidosulfuron, Nicosulfuron and Cyclosulfamuron are most common member of sulfonylurea herbicide family. Interaction between these active ingredients of sulfonylurea herbicide and ALS from *Arabidopsis thaliana* (PDB ID 5K6Q) shows that the formation of the complex inhibits the binding of TDL (Lactyl-ThDP) to ALS properly and results into the perturbation of the side chain biosynthesis.

From molecular docking analysis it is clear that, this inhibition is either competitive or uncompetitive inhibition. Moreover, after identifying the interacting site of protein, we have search for the respective positional character in other organism through jalview and from this study it is cleared that these interacting sites are totally occupied by same amino acids irrespective of the organisms.

With the development of bioinformatic approaches, the scopes of understanding fundamental processes have increased. Here, in this study, bioinformatic tools were used with the objective to analyse the interactions between acetolactate synthase from *Arabidopsis thaliana* and active ingredients of sulfonylurea herbicide groups. The current work done by various *in silico* tools had produced many informative results. Our obtained outcomes will be a valuable sources of enlightenment to the weed biologists and beneficial to research specific protein interactions along with developing new herbicides in near future using computational methods.

Acknowledgements

All authors acknowledge Mr. Madhab Kumar Sen, PhD student, Department of Agroecology and Crop Production, Czech University of life sciences, Prague, for his constant and valuable help and guidance for this entire work. The authors are thankful to the Department of Biotechnology, The University of Burdwan, West Bengal for providing necessary infrastructural support.

Author contribution statement

SKM conceived and designed the experiments; SKM, MC and RD Performed the experiments; Analyzed, interpreted the data and wrote the manuscript.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing interest statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

[1] Vats, S. Herbicides: History, Classification and Genetic Manipulation of Plants for Herbicide Resistance, Springer, 15, 153-192 (2015) doi:10.1007/978-3-319-09132-7_3

[2] Duggleby, RG., Pang, SS. Acetohydroxyacid synthase. Biochemistry and Molecular Biology Journal 33: 1–36 Evidence that it is not a flavoprotein, Journal of Biological Chemistry, 243, 3740-3741 (2000) doi:10.1111/j.1742-4658.2012.08505.x

[3] Saari, LL., Cotterman, JC., Thill, DC. Resistance to acetolactate synthase-inhibiting herbicides, in: S.B. Powles, J.A.M. Holtum (Eds.), Herbicide Resistance in Plants: Biology and Biochemistry, Lewis Publishers, Boca Raton, FL, 83–139 (1994) doi:10.1177/074823379901500120

[4] Retzinger, EJ., Mallory-Smith, C. Classification of herbicides by site of action for weed resistance management strategies, Weed Technology, 11, 384–39. doi:10.1614/WS-D-11-00206.1 (1997)

[5] Schmidt, RR. Classification of herbicides according to mode of action. Bayer Ag, Leverkusen, 8 (1998) doi:10.1007/978-3-319-09132-7_3

[6] Shariq, I. Sherwani., Ibrahim, A. Arif., and Haseeb, A. Khan. Modes of Action of Different Classes of Herbicides, Intech Open, 176-177 (2015) doi:10.5772/61779

[7] Lowery, RF. Granular formulations and application, in: McWhorter CG, Gebhardt MR (eds)Methods of applying herbicides, WSSA monograph 4, Weed Science Society America, *Champaign*, 165–176.*358* (1987)doi:10.1007/978-3-319-09132-7_3

[8] McCourt, JA., Pang, SS., King-Scott, J., Duggleby, RG., Guddat, LW. Herbicide binding sites revealed in the structure of plant acetohydroxyacid synthase, Proceedings of the National Academy of Sciences USA, 103, 569–573 (2006) doi:10.1073/pnas.0508701103

[9] Pang, SS., Guddat, LW., Duggleby, RG. Molecular basis of sulfonylurea herbicide inhibition of acetohydroxyacid synthase, Journal of Biological Chemistry, 278, 7639–7644 (2003) doi.10.1073/pnas.0508701103

[10] Yu, Q., Han, H., Vila-Aiub, MM. and Powels SB. AHAS herbicide resistance endowing mutations: effect on AHAS functionality and plant growth, J Exp Bot. 2010 Sep; 61(14): 3925–3934(2010) doi:10.1093/jxb/erq205

[11] Störmer, FC. The pH 6 acetolactate-forming enzyme from Aerobacter aerogenes. II, Journal of Biochemistry, 243, 3740-3741 (1968) doi:10.1111/j.1432-1033.1969.tb00682.x

[12] McCourt, JA., Pang, SS., Duggleby, RG., Guddat, LW. Elucidating the specificity of binding of sulfonylurea herbicides to acetohydroxyacid synthase, Biochemistry, 44, 2330–2338 (2005) doi:10.1021/bi047980a

[13] Chipmana, D., Ze'evBaraka., John, V. Schlossb. Biosynthesis of 2-aceto-2-hydroxy acids: acetolactate synthases and acetohydroxyacid synthases, 18-23 (1998) doi:10.1016/s0167-4838(98)00083-1

[14] Breslow, R. On the mechanism of thiamine action. IV. Evidence from studies on model systems. Journal of the American Chemical Society, 80, 3719-3726 (1958) doi:10.1021/ja01494a070

[15] Umbarger, HE., Brown, B. Isoleucine and valine metabolism in Escherichia coli. VIII. The formation of acetolactate, Journal of Biological Chemistry, 233,1156–1160 (1958) doi:10.1007/BF00272245

[16] Sauers, RF., and Levitt, G. Sulfonylureas: New high potency herbicides, in Pesticide Synthesis Through Rational Approaches (Magee, P. S., Kohn, G. K., and Menn, J. J., Eds.) American Chemical Society, Washington, DC, 21-28 (1984) doi:10.1021/bk-1992-0504.ch002

***ISSN* 2689-6389 (Print)**
***ISSN*2687-7939 (Online)**

[17] LaRossa, RA. and Schloss, JV. The sulfonylurea herbicide sulfometuron methyl is an extremely potent and selective inhibitor of acetolactate synthase in *Salmonella typhimurium*. J Biol Chem, 25: 8753 – 8757 (1984)

[18] Xing, RY. and Whitman, WB. Sulfometuron methyl-sensitive and -resistant acetolactate synthases of the archaebacteria Methanococcus spp, Journal of Bacteriology 169(10):4486-92 (1987) doi:10.1128/jb.169.10.4486-4492.1987

[19] Muller, YA., Schulz, GE. Structure of the thiamine-and flavin dependent enzyme pyruvate oxidase, Science, 259, 965–967 (1993) doi:10.1126/science.8438155

[20] Hasson, MS., Muscate, A., McLeish, MJ., Polovnikova, LS., Gerlt, JA., Kenyon, GL., Petsko, GA., Ringe, D. The crystal structure of benzoylformate decarboxylase at 1.6A° resolution: diversity of catalytic residues in thiamine diphosphate-dependent enzymes, Biochemistry, 37, 9918–9930 (1998) doi:10.1021/bi973047e

[21] Caines, ME., Elkins, JM., Hewitson, KS., Schofield, CJ. Crystal structure and mechanistic implications of N2-(2-carboxyethyl) arginine synthase, the first enzyme in the clavulanic acid biosynthesis pathway, Journal of Biological Chemistry, 279, 5685–5692 (2004) doi:10.1074/jbc.M310803200

[22] Mondal, S.K. and Sen, M.K. An in-silico characterization of Sry-related HMG box C (SOXC) in humans and mouse, Meta Gene (2019) doi:10.1016/j.mgene.2018.12.012

[23] Mondal, SK., Kundu, S., Das, R. and Roy, S. Analysis of phylogeny and codon usage bias and relationship of GC content, amino acid composition with expression of the structural nif genes, J. Biomol. Struct. Dyn., pp. 2-4 (2015) doi:10.1080/07391102.2015.1087334

[24] Altschul, SF., Gish, W., Miller, W., Myers, EW. and Lipman DJ. Basic local alignment search tool, J Mol Biol, 215(3):403-10 (1990) doi:10.1016/S0022-2836(05)80360-2

[25] Mondal, SK. and Roy, S. Genome-wide sequential, evolutionary, organizational and expression analyses of phenylpropanoid biosynthesis associated MYB domain transcription factors in Arabidopsis, J. Biomol. Struct. Dyn., 36, pp. 1580-1592 (2017) doi:10.1080/07391102.2017.1329099

[26] Geourjon, C. and Deléage, G. SOPM: a self-optimized method for protein secondary structure prediction, *Protein Engineering, Design and Selection*, 7(2), 157–164 (1994) doi:10.1093/protein/7.2.157

[27] Tamura, K.,Stecher, G., Peterson, D., Filipski, A. andKumar, S. MEGA 6: Molecular Evolutionary Genetics Analysis Version 6.0, Molecular Biology and Evolution, 30(12), 2725-2729 (2013) doi: 10.1093/molbev/mst197

[28] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, MR., Appel, RD., Bairoch, A. Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed.), The proteomics protocols handbook. Humana Press, Totowa, 571-607 (2005) doi:10.1385/1-59259-890-0:571

[29] Mondal, SK Study of phoBR in Escherichia coli and phoPR in Bacillus: computational approach Int. J. Adv. Res. Biol. Sci., pp. 28-31 (2015) doi:10.3389/fmicb.2018.02678

[30] Begum, Y. and Mondal, SK. Comprehensive study of the genes involved in chlorophyll synthesis and degradation pathways in some monocot and dicot plant species, Journal of Biomolecular Structure and Dynamics (2020) doi.org/10.1080/07391102.2020.1748717

***ISSN* 2689-6389 (Print)**
***ISSN*2687-7939 (Online)**

[31] Mondal, SK., Shit, S., Kundu, S. A comparative computational study of the 'rbcL' gene in plants and in the three prokaryotic families – archaea, cyanobacteria, proteobacteria, Indian Journal of Biotechnology, 12, 58-66 (2013) doi:10.1080/07391102.2015.1087334

[32] Szklarczyk, D., Morris, JH., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, NT., Roth, A., Bork, P., Jensen, LJ., von, Mering C. (2016) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible, Epub Nucleic Acids Res, (D1):D362-D368 (2017) doi: 10.1093/nar/gkw937

[33] DeLano, WL. The PyMOL Molecular Graphics System, DeLano Scientific LLC, San Carlos, CA, PLOS ONE (2002) doi:10.1371/journal.pone.0036770

[34] Mondal, SK. and Sen, M. Loss of phosphatase activity in PTEN (phosphatase and tensin homolog deleted on chromosome ten) results in endometrial carcinoma in humans: An in-silico study Heliyon, 6(1), e03106 (2020b) doi:10.1016/j.heliyon.2019.e03106

[35] Mondal, SK Das, S. and Sen, MK Study of α-Amylase Based on their Compositional Parameters of Its Gene Along with Its Protein Structure, Microbial Fermentation and Enzyme Technology, edited by Hrudayanath Thatoi, Pradeep K. Das Mohapatra, Sonali Mohapatra, Keshab C. Mondal, CRC Press, 29-Apr-2020 (2020a) doi:10.1201/9780429061257

[36] Pettersen, SF., Goddard, TD., Huang, CC., Couch, GS., Greenblatt, DM., Meng, EC., & Ferrin, TE. UCSF Chimera - A visualization system for exploratory research and analysis, Journal of Computational Chemistry, 25,1605–1612 (2004) doi:10.1002/jcc.20084

[37] Wallace, AC., Laskowski, RA., Thornton, JM. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering, Design and Selection*, Volume 8, Issue 2 127-134 (1995) doi:10.1093/protein/8.2.127.

[38] Lee, YT., Duggleby, RG. Identification of the regulatory subunit of *Arabidopsis thaliana* acetohydroxyacid synthase and reconstitution with its catalytic subunit, Europe PMC, 29-33 (2001) doi:10.1021/bi002775q

[39] Ahan, TW., Kim, DW., Choi, JD. Inhibition of acetohydroxyacid synthase by sulfonylureas and imidazolinones. Korean Biochem Journal, 25, 636–641 (1992) doi:10.1007/s00726-005-0297-3

***ISSN* 2689-6389 (Print)**
***ISSN*2687-7939 (Online)**

Supplementary Table 1. Different groups of Herbicides, Mode of action, site of action and their chemical family.(https://www.intechopen.com/books/herbicides-physiology-of-action-and-safety/modes-of-action-of-different-classes-of-herbicides)(Schmidt, RR.1998)

Group	Mode of Action	Site of Action	Chemical Family
1	Lipid-Synthesis Inhibitors	ACCase Inhibitor	Aryloxyphenoxypropionate (FOPs), Cyclohexanedione (DIMs), Phenylpyrazolin (DENs)
2	Amino-Acid Synthesis Inhibitors	ALS Inhibitors	Imidazolinones, pyrimidinylthiobenzoates, sulfonylaminocarbonyltriazolinones, sulfonylureas, triazolopyrimidines
3	Root-Growth Inhibitors	Microtubule Inhibitors	Benzamide, benzoic acid (DCPA), dinitroaniline, phosphoramidate, pyridine
4	Plant-Growth Inhibitors	Site of Action Unknown	Benzoic acid, phenoxycarboxylic acid, pyridine carboxylic acid, and quinoline carboxylic acid
5	Photosynthesis Inhibitors	Photosystem II Inhibitors	Triazine, triazinone, phenylcarbamates, pyridazinones, and uracils.
6	Photosynthesis Inhibitors	Photosystem II Inhibitors	Nitriles, benzothiadiazinones, and phenylpyridazines
7	Photosynthesis Inhibitors	Photosystem II Inhibitors	Phenyl, urea, and amides
8	Shoot-Growth Inhibitors	Lipid-Synthesis Inhibitors	Phosphorodithioates and thiocarbamates
9	Amino-Acid Synthesis Inhibitors	EPSP Synthase Inhibitors	Not designated by any specific chemical family
10	Nitrogen-Metabolism Inhibitors	Glutamine-Synthesis Inhibitors	Not designated by any specific chemical family
12	Pigment-Synthesis Inhibitors	HPPD Inhibitors	Amides, anilidex, furanones, phenoxybutan-amides, pyridiazinones, and pyridines
13	Pigment-Synthesis Inhibitors	Diterpene-Synthesis Inhibitors	Isoxazolidinone
14	Cell-Membrane Disrupters	PPO Inhibitors	Diphenylether, aryl triazolinone, N-phenylphthalimides, oxadiazoles, oxazolidinediones, phenylpyrazoles,

ISSN 2689-6389 (Print)
*ISSN*2687-7939 (Online)

			pyrimidindiones, and thiadiazoles.
15	Shoot-Growth Inhibitors	Very-Long-Chain Fatty Acid (VLCFA) Inhibitors	Chloroacetamide, acetamide, oxyacetamide, and tetrazolinone.
22	Cell-Membrane Disrupters	PSI Inhibitor	Bipyridilium
27	Pigment-Synthesis Inhibitors	HPPD Inhibitors	Isoxazole

Supplementary Table 2. List of selected organisms and accession number of homologous sequences with ALS protein sequences of *Arabidopsis thaliana* obtained by blastp.

Sl. No	Organisms Name	Accession Number	Sl. No	Organisms name	Accession Number
1	*Arabidopsis thaliana*	AAK68759.1	2	*Arabidopsis Lyrate subsp. lyrata*	XP_002877617.1
3.	*Camelina sativa*	XP_010503496.1	4.	*Camelina microcarpa*	AAR06607.1
5.	*Eutrema salsugineum*	XP_006404237.1	6.	*Descurainiasophia*	ACN62348.1
7.	*Brassica napus*	XP_013648961.1	8.	*Brassica carinata*	AJF23179.1
9.	*Brassica oleracea var. oleracea*	XP_013603602.1	10.	*Brassica rapa*	XP_009151430.1
11.	*Brassica juncea*	AJF23157.1	12.	*Capsella rubella*	XP_006293107.1
13.	*Carica papaya*	XP_021905626.1	14.	*Ricinus communis*	XP_002511176.1
15.	*Actinidia chinensis var. chinensis*	PSS07666.1	16.	*Cicer arietinum*	XP_004485753.1
17.	*Ziziphus*	XP_015882186.1	18.	*Tremaorientale*	PON72129.1

ISSN 2689-6389 (Print)
*ISSN*2687-7939 (Online)

	jujuba				
19	*Populus trichocarpa*	XP_002322262.1	20	*Prunus avium*	XP_021812945.1
21	*Jatropha curcas*	XP_012090695.1	22	*Prunus yedoensis var. nudiflora*	PQP93616.1
23	*Gossypiumarboreum*	XP_017649232.1	24	*Raphanus sativus*	XP_018471079.1
25	Raphanusraphanistrum	CAC86692.1	26	Theobroma cacao	EOY04840.1
27	Gossypium hirsutum	XP_016699120.1	28	*Camellia sinensis*	XP_028123767.1
29	*Citrus clementina*	XP_006436842.1	30	*Malus domestica*	RXH85640.1
31	*Gossypiumhirsutum*	XP_016699120.1	32	*Solanum tuberosum*	XP_006361740.1
33	*Cucurbita pepo subsp. pepo*	XP_023516040.1	34	*Pyrus x bretschneideri*	XP_009360722.1
35	*Eucalyptus grandis*	XP_010026500.1	36	*Gossypiumraimondii*	XP_012455043.1
37	*Euphorbia maculata*	AMB51356.1	38	*Corchorus capsularis*	OMO68011.1
39	*Ipomoea nil*	XP_019154060.1	40	*Xanthium sp*	AAA74913.1
41	*Duriozibethinus*	XP_022734870.1	42	*Chenopodium quinoa*	XP_021737451.1
43	*Punicagranatum*	OWM65493.1	44	*Cucurbita maxima*	XP_022987362.1
45	*Helianthus annuus*	XP_021984153.1	46	*Nicotiana tabacum*	XP_016478370.1
47	*Solanum pennellii*	XP_015081631.1	48	*Nelumbo nucifera*	XP_010274326.1
49	*Erythrantheguttata*	XP_012830574.1	50	*Parasponiaandersonii*	PON77293.1

ISSN **2689-6389 (Print)**
*ISSN***2687-7939 (Online)**

51	*Arabis alpina*	KFK34163.1	52	*Tarenayahassleriana*	XP_010550735.1
53	*Gossypium barbadense*	PPE02752.1			

Supplementary Table 3. A. Jalview analysis of conserved domain (1) TPP _Enzyme _N by using Clustal Omega to obtain amino acid conserve site.

Conserved Site	Amino acid Present	Physiochemical Property	Mutation Type	Effect on Phylogeny
141	i. Alanine	i. Neutral, Nonpolar, Hydrophobic	Mutation	NO
	ii. Serine	ii. Neutral, Polar, Hydrophilic		
143	i. Isoleucine	Both are Neutral, Nonpolar, Hydrophobic	Silent Mutation	-
	ii. Valine			
144	i. Leucine	Both are Neutral, Nonpolar, Hydrophobic	Silent Mutation	–
	ii. Isoleucine			
151	i. Glutamine	i. Neutral, Polar, Hydrophilic	Mutation	Yes
	ii. Glutamic acid	ii. Acidic, Polar, Hydrophilic		
	iii. Leucine	iii. Neutral, Nonpolar, Hydrophobic		
165	i. Methionine	All are Neutral, Nonpolar, Hydrophobic	Silent Mutation	-
	ii. Leucine			
	iii. Valine			

ISSN 2689-6389 (Print)
*ISSN*2687-7939 (Online)

177	i. Isoleucine	Both are Neutral, Polar, Hydrophobic	Silent Mutation	-
	ii. Valine			
178	i. Arginine	i. Basic, Polar, Hydrophilic	Mutation	Yes
	ii. Tyrosine	ii. Neutral, Polar, Hydrophobic		
	iii. Lysine	iii. Basic, Polar, Hydrophobic		
180	i. Valine	i. Neutral, Nonpolar, Hydrophobic	Mutation	Yes
	ii. Cysteine	ii. Neutral, Polar, Hydrophilic		
198	i. Serine	i. Neutral, Polar, Hydrophilic	Mutation	No
	ii. Alanine	ii. Neutral, Nonpolar, Hydrophobic		
199	i. Serine	Both are Neutral, Polar, Hydrophobic	Silent Mutation	-
	ii. Threonine			
202	i. Proline	i. Neutral, Nonpolar, Hydrophobic	Mutation only on *Gossypium hirsutum* (serine)	No
	ii. Valine	ii. Neutral, Nonpolar, Hydrophobic		
	iii. Serine	iii. Neutral, Polar, Hydrophilic		

***ISSN* 2689-6389 (Print)**
***ISSN*2687-7939 (Online)**

204	i. Valine ii. Isoleucine	Both are Neutral, Nonpolar, Hydrophobic	Silent Mutation	-
206	i. Isoleucine ii. Valine	Both are Neutral, Nonpolar, Hydrophobic	Silent Mutation	-
221	i. Alanine ii. Serine	i. Neutral, Nonpolar, Hydrophobic ii. Neutral, Polar, Hydrophilic	Mutation	No
224	i. Leucine ii. Methionine iii. Serine	i. Neutral, Nonpolar, Hydrophobic ii. Neutral, Nonpolar, Hydrophobic iii. Neutral, Polar, Hydrophilic	Mutation	Yes
228	i. Valine ii. Isoleucine iii. Cysteine	i. Neutral, Nonpolar, Hydrophobic ii. Neutral, Nonpolar, Hydrophobic iii. Neutral, Polar, Hydrophilic	Mutation	Yes
229	i. Proline ii. Alanine	Both are Neutral, Nonpolar, Hydrophobic	Silent Mutation	-
231	i. Valine	Both are Neutral,	Silent	-

ISSN 2689-6389 (Print)
*ISSN*2687-7939 (Online)

	ii. Isoleucine	Nonpolar, Hydrophobic	Mutation	
238	i. Proline	i. Neutral, Nonpolar, Hydrophobic	Mutation	No
	ii. Serine	ii. Neutral, Polar, Hydrophilic		
267	i. Leucine	Both are Neutral, Nonpolar, Hydrophobic	Silent Mutation	-
	ii. Methionine			
269	i. Valine	Both are Neutral, Nonpolar, Hydrophobic	Silent Mutation	-
	ii. Isoleucine			
274	i. Arginine	i. Basic, Polar, Hydrophilic	Mutation	No
	ii. Methionine	ii. Neutral, Nonpolar, Hydrophobic		
282	i. Lysine	Both are Neutral, Nonpolar, Hydrophobic	Silent Mutation	-
	ii. Isoleucine			
285	i. Serine	Both are Neutral, Polar, Hydrophilic	Silent Mutation	-
	ii. Threonine			
290	i. Proline	Both are neutral, polar, hydrophilic	Silent Mutation	-
	ii. Alanine			

ISSN 2689-6389 (Print)
*ISSN*2687-7939 (Online)

Supplementary Table 3. B.Jalview analysis of conserved domain (2) TPP _Enzyme _M by using Clustal Omega to obtain amino acid conserve site.

Conserved Site	Amino acid Present	Physiochemical Property	Mutation Type	Effect on Phylogeny
338	i. Valine ii. Isoleucine iii. Leucine	All are Neutral, Nonpolar, Hydrophobic	Silent Mutation	-
340	i. Leucine ii. Phenylalanine	Both are Neutral, Nonpolar, Hydrophobic	Silent Mutation	-
344	i. Serine ii. Alanine	i. Neutral, Polar, Hydrophilic ii. Neutral, Nonpolar, Hydrophobic	Mutation	Yes
345	i. Lysine ii. Arginine	Both are Basic, Polar, Hydrophilic	Silent Mutation	-
346	i. Lysine ii. Arginine iii. Methionine	i. Basic, Polar, Hydrophilic ii. Basic, Polar, Hydrophilic iii. Neutral, Nonpolar, Hydrophobic	Mutation	Yes
355	i. Cystine ii. Serine	Both are Neutral, Polar, Hydrophilic	Silent Mutation	-
370	i. Glycine ii. Leucine	i. Basic, Nonpolar, Hydrophilic ii. Neutral, Nonpolar, Hydrophobic	Mutation	Yes
405	i. Tyrosine ii. Phenylalanine	i. Neutral, Polar, Hydrophilic ii. Neutral, Nonpolar, Hydrophobic	Mutation	No
407	i. Valine	Both are Neutral,	Silent	-

ISSN 2689-6389 (Print)
*ISSN*2687-7939 (Online)

	ii. Isoleucine	Nonpolar, Hydrophobic	Mutation	

Supplementary Table 3. C.Jalview analysis of conserved domain (3) TPP _Enzyme _C by using Clustal Omega to obtain amino acid conserve site.

Conserved Site	Amino acid Present	Physiochemical Property	Mutation Type	Effect on Phylogeny
567	i. Isoleucine ii. Methionine ii. Valine	All are Neutral, Nonpolar, Hydrophobic	Silent Mutation	-
570	i. Alanine ii. Serine	i. Neutral, Nonpolar, Hydrophobic ii. Neutral, Polar, Hydrophilic	Mutation	Yes
572	i. Alanine ii. Glycine	i. Neutral, Nonpolar, Hydrophobic ii. Neutral, Nonpolar, Hydrophilic	Silent Mutation	-
578	i. Valine ii. Isoleucine	Both are Neutral, Nonpolar, Hydrophobic	Silent Mutation	-
591	i. Valine ii. Leucine	Both are Neutral, Nonpolar, Hydrophobic	Silent Mutation	-
608	i. Leucine ii. Isoleucine	Both are Neutral, Nonpolar, Hydrophobic	Silent Mutation	-
648	i. Leucine ii. Valine	Both are Neutral, Nonpolar, Hydrophobic	Silent Mutation	-
660	i. Arginine ii. Glutamine	Both are Neutral, Polar, Hydrophilic	Silent Mutation	-

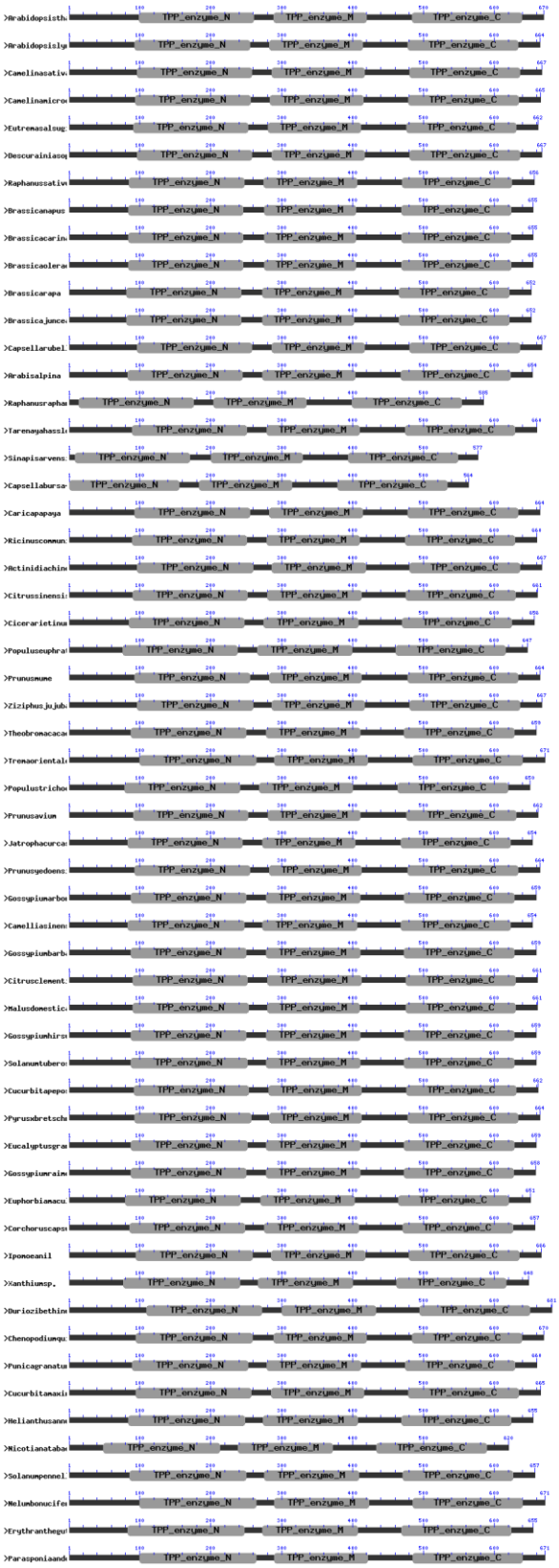
ISSN 2689-6389 (Print)
*ISSN*2687-7939 (Online)

678	i. Proline	i. Neutral, Nonpolar, Hydrophobic	Mutation	No
	ii. Glutamate	ii. Acidic, Polar, Hydrophilic		

Supplementary Table 4. Interacting amino acids of the receptor molecules responsible for hydrogen and hydrophobic bond and respective binding affinities (kcal/mol) between receptor and ligand.

Receptor	Ligand	Hydrogen bond	Hydrophobic bond	Binding Affinity
Acetolactate Synthase (ALS)	Lactyl -ThDP (TDL)	R279, R246,D397,P247,L183,S398	P281,Ile396,M280,G245, K220,G248	-7.0
ALS	Amidosulfuron	R279,R246	G245,D397,Y276,K220, L183	-6.3
ALS-Amidosulfuron	TDL	R279,S398,D397,R246,I396	P281,Y276,G245,P247, K220, L183	-6.1
ALS	Nicosulfuron	S454,T662	H646,G664,L648,D665,E663,K450, F451	-7.55
ALS-Nicosulfuron	TDL	T662,S454	H646,G664,L648,D665,E663,K450,F451,P452 V355,E645	-2.8
ALS	Cyclosulfamuron	R246,R279	L183,K220,G245,P281,M280,Y276,T219	-7.7
ALS-Cyclosulfamuron	TDL	H347,N497,D472,D342,K448	G477,E473,D476,K469,Q494,K499,F495,D341	-6.5

ISSN 2689-6389 (Print)
*ISSN*2687-7939 (Online)



SupplementaryFIG.1. Conserved domain structure obtained from NCBI Conserved Domain Search – NIH for all selected sequences homologous to ALS protein from *Arabidopsis thaliana*.

***ISSN* 2689-6389 (Print)**
***ISSN*2687-7939 (Online)**