**Computational molecular explanation of Soybean AHAS resistance from P197S mutation**

Gordon Chalmers1,\*, Rustem Ustun2, Daniel Tehrani1, Bulent Uzun2

1 Complex Carbohydrate Research Center, University of Georgia, Athens, GA, 30602

2 Department of Field Crops, Faculty of Agriculture, Akdeniz University, Antalya, Turkey

\*corresponding author: [gordoncs@uga.edu](mailto:gordoncs@uga.edu), ORCID: 0000-0003-0254-352X

**Abstract:**

Crop management through the use of herbicides is an integral part of modern high-yield farming. The identification of resistant mutated strains of normal plants amongst a background of weeds is important to the usage of herbicides, their development, and resistance. In this work a computational study of a point mutation of Proline-197-Serine of the Soybean AHAS enzyme explains the latter’s S197 resistance to the commonly used Chlorsulfuron. Just as a small molecular change to a therapeutic can alter drastically the binding and efficacy of a drug, a small substitution of one amino acid in a target can alter the binding and efficacy of an herbicide through conformation and binding changes of the bound inhibitor, e.g., eliminating a dominant bound herbicide conformation and de-localization generally of the binding. The computational approach here using large scale sampling and distributions from protein-ligand docking and AlphaFold generated protein structures is easily scaled to scan over the mutation possibilities of protein binding sites, similar to screening compounds for potential hits in therapeutic design.

Keywords: AHAS (ALS) enzyme, herbicides, protein-ligand interactions, small molecule inhibitors, crop management, protein engineering

**Introduction**

A molecular analysis of the binding of a popular herbicide Chlorsulfuron (CS) [1,2,3,4] is done on the Soybean AHAS [5,6] enzyme and a mutant to explain its immunity change from susceptible to resistant. Determining the source of resistance of plants to herbicides is important in the discovery and development of herbicides, and AHAS [7,8,9,10,11] is a common target in particular by the sulfonylureas. The X-ray structure of AHAS was first determined in *Saccharomyces cervisiae* (budding yeast) [12,1JSC], followed by five molecules from this class in complex with ScAHAS [13,CE 1N0H][14, TB 1T9A,CS 1T9B,SM 1T9C,MM 1T9D]. Due to problems with crystallizing *Arabidopsis Thaliana* the x-ray structures of complexes with the sulfunyureas were first [15, CE 1YBH, MM 1YHY, CS 1YHZ, SM 1YI0, 1YI1, IQ 1Z8N] (together with one inimidazolinone IQ) and then followed by the free AtAHAS [16, 5K6Q]. The history of inhibiting the AHAS’s function is that the herbicides were developed first, before the binding sites were discovered, and then it was found that there are 2 nearby sites: catalytic site having the function of the enzyme and an herbicide receptor site [10,14], the latter of which these small molecules bind to and are considered in this work. Developing pharmaceutical drugs traditionally has followed this pipeline and without much information about a binding site. There are now a plethora of x-ray structures of different AHAS’s complexed with small molecules.[[1]](#footnote-1) The complete cryo-em tetramer structure of AtAHAS in the presence and absence of valine was recently determined [24, 6VZ8,6U9H]. Due to advances in molecular modeling and force fields, the developments in docking software such as what is used in this [17], including an important use of large scale sampling with distributions [18,19], and the advent of AlphaFold for computationally generating protein structures [20,21,22], it is possible to find valuable information about the immunity of plant species to herbicides solely in computation and efficiently.

The regular use of well-known herbicides has led to the problem of plant resistant strains to herbicides caused by natural mutation throughout the generations [23]; for example, there are 2 plants involved in the same soil, plant and weeds. Responses to weed resistance at the molecular level include creating new small molecule herbicides or genetically altering wanted plants to be resistant to known or in development herbicides (such as the change P197S and Chlorsulfuron here). Weed, or other, control is an economical and well researched means of maximizing harvest. Testing large numbers of new possible herbicides against a plant, or for example, testing all possible point mutations iNj of the many amino acids N in a binding site to j (amino type) from i in a plant to find engineered resistance, can be done using computational binding studies as performed here. This is quicker than growing large numbers of tests in the field or making x-ray structures (of, for example, testing 100 point mutation combinations of the same AHAS enzyme).

The well-used herbicides include the classes sulfonylureas (chlorimuron-ethyl and CS), imidazolinones (imazaquin), pyrimidinyl-benzoates (bispyribac-sodium), triazolopyrimidines (florasulam) and sulfonylamino-carbonyl-triazolinones (thiencarbazone-methyl). A primary target of an herbicide is the acetohydroxyacid synthase (AHAS), also known as acetolactate synthase (ALS) for its production of an intermediate acetolactate.

The AHAS Inhibition in plants through the use of herbicides is regarded as one of the safest to human ways of plant control because the biological pathway of branched-chain amino acid production doesn’t exist in humans. However, mass use of an herbicide can promote resistance, or immunity, and generally will lead to lessening effects of use even with combined inhibitor treatments. Single-point amino acid mutations within the genome, i.e., simple ones, of the AHAS enzyme have been investigated in a number of plant species and in different amino acids with regards to resistance to these commonly used herbicides, e.g. in [24,25,26,27,28,29,30,31,32,33]. (22 amino acid substitutions (type) at 7 aminos in the sequence give resistance to weed herbicides known in 2010 [25], and 26 substitutions at 8 known in 2014 [26], although only a few tested for plant growth and health.) The fact that Proline 197 when altered can give resistance is not new in this work.

Although the small molecule herbicide binding effects and inhibition of AHAS catalytic activity is known [7,10], and x-ray structures exist of AHAS with and without these 5 classes of herbicides (primarily in *Arabidopsis Thaliana* and *Saccharomyces cerevisiae*), the computational molecular analysis of inhibition of the herbicide binding site is not complete, particularly in protein-ligand docking studies. In addition, x-ray structures are an entropically frozen state that does not guarantee modeling of thermal biologic conditions. In general the question as to whether a plant is resistant to an inhibitor is unaccompanied by any x-ray structure information at all. The use of computational molecular modeling techniques can aid in the search/selection of point mutations to improve crop performance via resistance factors, including quantification in dose. Molecular herbicide design and crop improvement can benefit substantially.

The point mutation effect of amino acid P197 to S197 and interaction with CS is investigated here in (Soybean and At) AHAS. This analysis uses x-ray structures as a check on related AlphaFold AHAS structures, but can be regarded as completely computational. The protein-ligand interactions are modeled computationally using CCDC GOLD [17] and the protein structures are generated using AlphaFold2 [18,19]. AlphaFold structures have been the subject of many tests of reliability and the calculations highly pass [20,21,22]; there is a confidence level given to regions of a generated structure and typically only flexible regions such as a tail or in the orientation of a multi-domain protein have an (expected) problem. AlphaFold uses a neural network trained on decades of x-ray and nmr data, such as from the NESG, and does calculate with this experimental input and homolog structural information from experimental structures. In principle any enzyme and small molecule herbicide can be examined using the methodology in this work, including mutants of enzymes pre-lab work. The computational analysis is very cost and time saving in practicality in narrowing down potential directions of planned field or lab studies of altered plants or the use of herbicides. Of course, there is no substitute to using experimentally determined structures (x-ray, nmr, cryo-em) and the approach in this work is a modeled approximation, however accurate in protein-ligand modeling and AlphaFold structure determination.

Visualization of molecules, figures, and some tools were used from the viewing packages Chimera, PyMol, and Nanome [34,35,36]. Molecule pdb and mol2 files were generated from SMILES using Corina Classic [37]. Matlab was used for scripting and analysis [38].

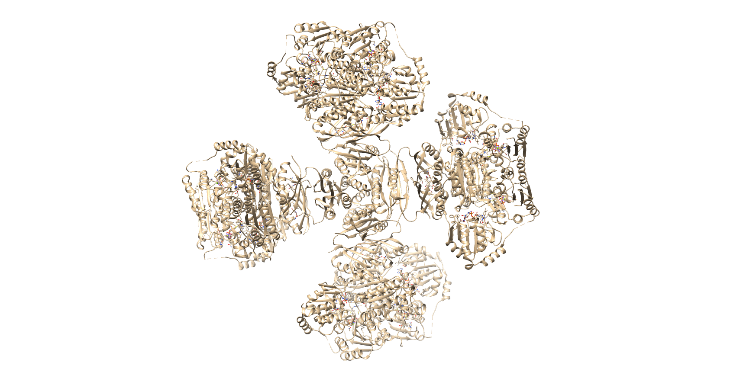
**Background**

The AHAS enzyme’s primary function is to start the process of biosynthesizing branched-chain amino acids [6,7], amino acids in which the aliphatic side-chain has a branch: valine, leucine, and isoleucine. These 3 are essential amino acids which are not synthesized in the human body but commonly found in plants, in particular Soybean. The action in the catalytic site for valine and leucine takes 2 pyruvate molecules CC(=O)C(=O)[O-] and converts these to acetolactate, CC(C(=O)O)OC(=O)C, through 2 CH3COCO2− → −O2CC(OH)(CH3)COCH3 + CO2 . The process for isoleucine uses 1 pyruvate and 1 2ketobutyrate, not 2 pyruvates. In addition to the use of cavity amino acids located in the catalytic site there are 3 cofactors to activate the AHAS enzyme: thiamine diphosphate (ThDP), ions, and flavine-adenine dinucleotide (FAD). These resulting V,L,I produced aminos are eventually used by the plant for different purposes including reproduction via DNA synthesis; clearly, blocking an organism of its natural production of amino acids is going to cause survival problems. The AHAS active cavity amino acids used in the catalytic action are Glycine 511, Valine 485, Methionine 513, and Histidine 643, and others in the cavity placement alter binding, interaction, and transport. The numbering follows from the genomic sequence of *Arabidopsis Thaliana* (aka mouse-ear cress), Appendix 1.

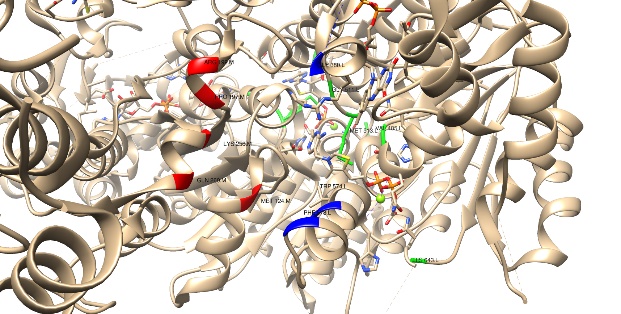
AHAS is a thiamine diphosphate dependent enzyme containing a catalytic and regulatory structural subunits (CSU,RSU). The CSU portion has three domains, alpha (85-269), beta (281-458), gamma (463-639), and a C-terminal tail (646-668). It occurs in yeast (ScAHAS) or bacteria as a dimer and in others such as *Arabidopsis Thaliana* as a ~700 kDa tetramer composed of a dimer of dimers. In a single dimer there are 2 alpha-domains and 2 gamma-domains creating the central region and a beta-domain on either side. The tetramer form is adopted in the presence of herbicides and with or without valine its cryo-em structure is presented in [24, PDB 6VZ8, 6U9H] and 6U9H is shown in Figure 1. The base of each bulb is the RSU portion, and each bulb is made of a dimer (of dimers). There are 2 tunnels in each dimer (of dimers), one on either side and in the cleft near the base attachment. The x-ray structure of the CSU monomer unit in the presence of CS is given as PDB ID: 1YBH [14] and without CS. It is in Figure 2.

**Figure 1:** Tetramer of *Arabidopsis Thaliana* AHAS. (a) PDB ID: 6U9H cryo-em structure. (b) Residue labeled herbicide binding site and the solvent accessible entry point to the catalytic site channel –- P197 is at the entry. Example chain M residues (monomer 1) are red and in chain L are blue (monomer 2). The catalytic amino acids G511, V485, M513, and H643 are green.

a.



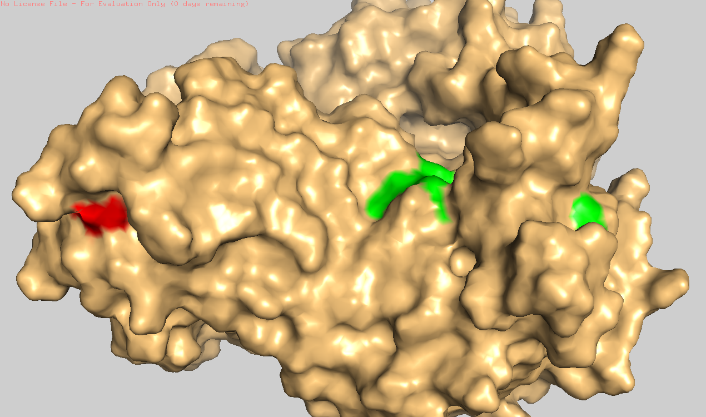
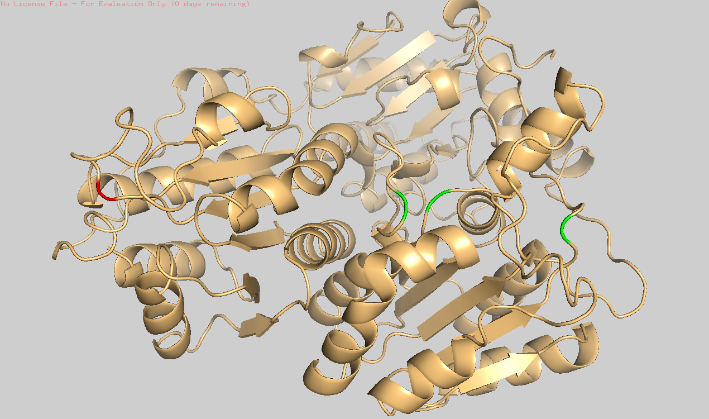
b.



The active catalytic site appears accessible being on the surface of AHAS. However, in the naturally occurring dimer or tetramer there are 2 active sites each of which are located at the interface of the CSUs; these 2 catalytic sites are not surface accessible by solvent and are led to the outside through a tunnel (or surface channel). At the entrance of the channel is the herbicide binding site, containing P197 and other amino acids connected to susceptibility or resistance, such as W574S and resistant. The channel is shown in Figure 1b; several residues are highlighted in red and blue coming from 2 monomers in the dimer and the 4 important amino acids in green which are responsible for the catalytic activity.

The types of AHAS inhibitors mentioned bind away from the catalytic site containing G511, V485, M513, and H643. This off-catalytic site binding of the 5 mentioned classes inhibitors is different than typical direct obstruction of an active site region on the surface or in the plant protein. Both types of inhibitors are observed also in therapeutics of human disease. The catalytic amino acids and the herbicide binding site are highlighted in green and red in the monomer depiction Figure 2. Pairs of these monomers dimerize and create the interface shown Figure 1. The entry point of the tunnel contains the herbicide binding site and allows transport of pyruvate molecules to the catalytic region; Glycine 511 is the first used in biosynthesis. This short tunnel can be used as a blockage point for the acetolactate production, and this is the high-binding region of AHAS to which these mentioned herbicides attach via intermolecular attraction.

**Figure 2.** The structure of Soybean AHAS is shown. This AHAS structure was generated by AlphaFold2 using the sequence in Appendix 1. The colored amino acids from the herbicide site (red) and the catalytic residues (green) are: Red is P197 and green is G511, M513, V485 and H643. The location from left to right of the green residues is the same as listed.

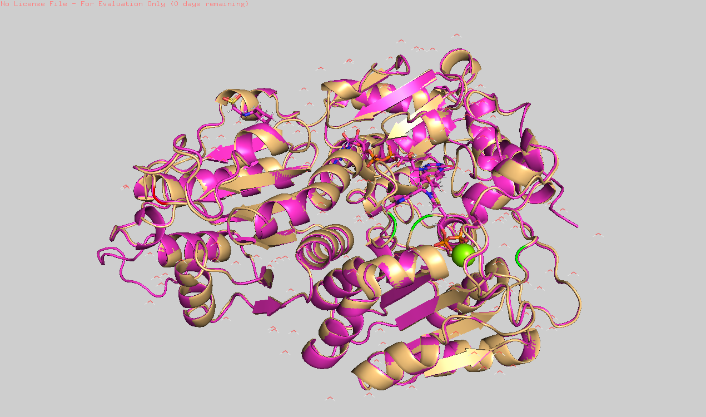


Due to the off catalytic binding site nature of the herbicidal small molecules with respect to the catalytic action it is interesting to understand what mutations do in regards to resistance. The simplest guess is that the entry door is closed by an herbicide and no pyruvates or 2ketobuyrate are allowed into the channel. In a typical inhibitor process the small molecule blocks the catalytic action by directly binding to the active amino acids, thus blocking natural function. A point mutation of an amino acid in this site could block the inhibitor from binding (non-covalently or covalently) or cause a conformational change in the region of the activity, both from a structural change of the catalytic site amino acids. With AHAS the catalytic site is unchanged, and a physical dynamical pathway to it is blocked by a small molecule if it bound strongly enough. Because the herbicide site is not a part of the AHAS function chemically in synthesis of amino acids, a simple explanation of resistance by point mutation is that the site is altered to not allow binding to the small molecule or to reduce it substantially. The wrong type of mutation, however, may *facilitate* the activity of the acetolactate production, by opening the entry door even wider; this would manifest through increased production of the 3 branched-chain amino acids [15].

The crystal structure of *Arabidopsis Thaliana* AtAHAS has been determined in the absence of any bound inhibitors, PDB ID: 5K6Q [16]. The structure of Soybean AHAS has not been crystallized and was instead determined by using AlphaFold2. AHAS structurally has strong similarities across species; most of its sequence across plant species is the same. The overlay of the x-ray AtAHAS and that of Soybean is overlayed in Figure 3. There are no noticeable differences in the secondary structure and the RMSD between the 2 is .299 A, well below the x-ray resolution of 2.9 A for 5K6Q. The global structural difference of .3 Angstroms is not significant in view of a conformation change.

**Figure 3.** The x-ray structure PDB ID: 5K6Q of *Arabidopsis Thaliana* AHAS (purplish) and Soybean AHAS from AlphaFold2 (tan). The RMSD between the 2 (5283 atoms) is .299 Angstroms. a. Overlay without the Connely surface. b. Overlay with the Connely surface. The amino acids in Figure 2 are highlighted in the same coloring. The crystal structure is obtained without herbicides but with the co-factors, as seen by examining the location near G511 (by zooming into the figures).

a. b.

A picture containing decorated

Description automatically generated

The comparison of these inhibited structures with the uninhibited AtAHAS structure concluded that there was no major structural change of significance, e.g., conformational change. The RMSD of the complete tetramer with and without the herbicide ligands was .23 Angstrom and that of the individual 3 domain subunits was .18 to .22 A [16]. In effect, for AtAHAS and the sulfonylureas or imidazolinone classes of small molecules, there is no noticeable conformational change in the x-ray structure upon the herbicide binding. Chlorsulfuron was a tested member of the sulfonylureas, and the absence of change supports the use of an uninhibited structure of both the P197 and S197 mutant form in making a binding study of CS. That is, it is reasonable to find the effects upon resistance to CS from an AHAS point mutation by using the free uninhibited different structures. The uninhibited Soybean AtAHAS is structurally almost the same as Soybean AHAS, having an RMSD of .3 A, and very similar in sequence, Figure 3.

Furthermore, in [16] it was pointed out that in the x-ray structures there are side-chain angular changes when bound versus unbound. This is very reasonable because side-chains, the tertiary structure, is flexible and a small molecule bound to a protein is going to disrupt it (energy minimum). Protein side-chains in GOLD were treated as flexible in this work, and the output files annotate any angular rotation from the input structure. Backbone phi, psi angles are fixed in the software GOLD, however, and these angles changing would potentially result in a conformational change. A short Amber [39] pre-minimization of the complex did not show any change.

To complete the picture of a general AHAS, not necessarily AtAHAS or Soybean AHAS, there is evidence in x-ray structures of a conformational change of AHAS in the *Saccharomyces cerevisiae* species. This is presumably due not to differences with AtAHAS in the catalytic site or the resistant site, but throughout the sequence and possibly the transport channel. This species of AHAS is not considered in the binding analysis.

**Computational modeling**

The alteration of one amino acid with a target binding site location to a small molecule does change the binding site, but how much and to what effect can only be understood by examining the protein’s interaction with the small molecule. Computational molecular modeling tools such as the CCDC GOLD protein-ligand docking software is used to study these interactions. Fortunately, and only available in the recent 2 years, is the availability of the AlphaFold protein structure modeling software; this computational tool relies on decades of x-ray and NMR data to train a neural network in successfully reproducing and predicting protein structures. While both of these tools are computational, they have passed successfully large numbers of tests in reliability, and in the case of AlphaFold, its structures have proven to be as confidently accurate as typical NMR structures. It is time consuming and costly to create x-ray structures of any plant’s enzyme, let alone mutating it and growing it to test. Both AlphaFold and molecular interaction analysis can provide a reasonable and much more efficient understanding of the effects of mutated enzymes.

The simplest effect of a single amino acid mutation in a protein is the modification of interactions: one amino acid has different interactions than another with an external ligand, the surface topography of the binding site may change. These interactions and their effects can be studied using computational docking. A different effect of a point mutation is a conformational change resulting from an amino change. The latter can be investigated with AlphaFold (without the ligand) and also to what extent the structural backbone in space changes, if at all. On top of these 2, there can be a protein conformational change brought about by the ligand attachment to the site and minimization techniques and long MD simulations sampling protein conformations can analyze that.

In the x-ray structure 1YHZ [15] the CS molecule is close to and appears to ring stack with Tryptophan 574. The structure is a monomer unit of the dimer however, and the adjacent monomer containing the rest of the herbicide and catalytic sites are not in it. The size of the entry point has to be modeled in, in this case by joining 2 AHAS x-ray determined monomers into a dimer and then adding the CS molecule and x-ray coordinates. The complete tetramer structure 6U9H (and 6VZ8), which does not have a Chlorsulfuron molecule in it, has a wide entry to the tunnel. P197 is not near the 4 catalytic amino acids of the opposite AHAS in the dimer in 6U9H using the 1YHZ CS coordinates; the distance from P197 to W574 is more than 15 A (nitrogen to nitrogen) in Figure 1b. The 5K6Q x-ray structure of AHAS [16] without CS is visually identical to the protein 1YHZ, which suggests a small dimer change to the corresponding one in structure 6U9H, if the illustrations (not x-ray structure) of the dimer interface [15,16] with CS, is that small. In the complete structure 6U9H, and in the vicinity of the interface, there are no hydrogen bond contacts between the 2 monomers in the dimer cavity region or at the point of entry.

The inhibition process investigated in this work is that of obstructing molecular transport of molecules such as pyruvate, 2ketobutyrate, or acetolactate/other components of V,L,I into or out of the site of catalytic activity. Proline 197 is the point of mutation and a fixed monomer AHAS backbone structure (with flexible amino acid side-chains) is used in the binding modeling. Rather than add in the completion of the dimer, only one monomer unit of the dimer is used to quantify the binding near P197 and in its change to S197. The entry point and region in the 6U9H tunnel is large enough to use only one monomer unit of the dimer, that containing P197, without sacrificing too much in the interaction with the opposite monomer. If investigating the mutation of W574 then the other monomer could be used and the binding examining near W574 and its mutant. In a more complete approximation of altering an amino acid in the herbicide site both monomers of the dimer would be used. In this case AlphaFold would be required to get the dimer orientation and interface correct, and this introduces a possible error in the modeling; AF is very robustly accurate with x-ray and nmr structures with monomers.

Several comments have to do with the location of the herbicide molecule in the x-ray complex of AHAS plus small molecule. Not presented in this work, a computational study was done in docking Chlorsulfuron with a docking radius of 4 and 10 A about the center point of the x-ray coordinates of CS in 1YHZ (with its CS stripped). The docking scores are generally smaller than the P197 region by about 3.5 PLP GOLD score, which is approximately .5 kcal and is a significant difference. (all docking calculations are available in the downloadable supplementary material). There could be several reasons why this site closer to the catalytic activity is found in an x-ray structure including the degree of flexibility the molecule has in binding, the amino acids it interacts with, and the lifetime of boundedness (x-ray conditions are not biological and mimic infinite relaxation time experiments).

The docked ligands in the modeled {P,S}197 AHAS enzymes fit in the 6U9H structure inside the tunnel close to the 1st monomer (containing amino acid 197). The binding in the approximation of one monomer instead of a total modeled dimer interface has only small corrections to the protein-ligand interaction if the interface is not thin. As in a box-like room, the modeled and docked CS ligands in the next sections are upon one wall and some distance from the other walls; these effects are distance dependent and the interactions are smaller to the 2nd monomer in the dimer.

The presence of a molecule non-covalently bound in the tunnel could potentially disrupt the transport of any of the previously mentioned molecules. Its presence generically would upset or block the biosynthetic process involving these transported molecules.

**Materials and Methods**

Development of EMS-mutated soybean populations and determination of resistance

Ten thousand soybean seeds of Arısoy variety (Glycine max L.) were applied a 0.1% EMS (Sigma Aldrich, M0880-25, Taufkirchen Germany) solution. All the treated seeds along with the wild-type as a control (no treatment) were sown in the experimental field of Akdeniz University (36°53’N, 38°30′E and altitude 33 m) in Antalya, Turkey in 2018. A commercial herbicide named Hammer10 WP (Doğal Kimya Company, Antalya, Turkey) with 10% chlorsulfuron active ingredient was used in this study. Three hundred thousand M2 soybean seeds were planted in experimental field and sprayer with chlorsulfuron active ingredient was applied with a concentration of 10 g da-1 when the plants were at V2-V3 stages. As a result of field trials for two years, one mutant plant in 300.000 M2 population was determined as completely resistant

Then, offspring were obtained from this resistant mutant and, the AHAS gene region sequence analysis of resistant and sensitive plants was performed. 197 amino acids found different because of point base mutation (paper is under review) in mutant and wild type sequence analysis. After the resistant was determined, herbicide was applied at doses of -10-20-30 to both sensitive wild type and resistant plants. To confirm the resistance, 10 seeds from the herbicide-resistant mutant named as AntSoy (sourced from the M2 population under field conditions) and 10 wild-type ArıtSoy seeds were planted in 5 L pots and grown in the greenhouse. When the plants reached the stage of V2-V3, the lowest dose to achieve full efficacy (10 g da-1) chlorsulfuron was applied in the greenhouse. The resistant progenies of AntSoy were also grown in a greenhouse to determine the resistance level at higher doses. For this purpose, 20 and 30 g da-1 (2X and 3X) herbicides were applied to the mutant progeny along with the wild-type.

*AlphaFold2 model of AHAS*

The soybean AHAS enzyme has not been crystallized. Although strongly similar to AHAS in other species, in particular AtAHAS, we used the AlphaFold computational tool to determine a 3D structure from the known sequence.

Three models of the resistant and susceptible plant were generated with progressively less confident models using Alphafold2’s neural network to model the structures against the known homologues from other AHAS species. The most confident model (first model) of each plant with either P197 or S197 was used for the modeling work in this paper. Alphafold2 builds structural models of a given input FASTA sequence using a two-track system: modeling the sequence based on its neural network and known homologues in the Protein Data Bank (PDB). Alphafold2 was run on the Sapelo2 supercomputer cluster at the Georgia Advanced Computing Resource Center (GARCC) using 1 GPU, 10 CPUs and about 6 hours per structure. All other parameters were defaulted. Documentation about using AlphaFold2 and the different parameters is available at https://wiki.gacrc.uga.edu/wiki/AlphaFold-Sapelo2.

**Chlorsulfuron and Soybean AHAS Binding**

The ligand binding to the AHAS active cavity is examined computationally using the docking software GOLD and the herbicide Chlorsulfuron, shown in Figure 4. The cavity was checked by: prior literature on AHAS enzymes in *Arabidopsis thaliana* and *Saccharomyces cerevisiasae*, x-ray structures of AHAS in complex with Chlorsulfuron (and others), and confirmed docking of Chlorsulfuron on the computational structure of soybean AHAS. The set of docking parameters used throughout this work are popsiz = 100, select\_pressure = 1.1, n\_islands = 8, maxops = 150000, niche\_siz = 2, for the population iterations, and pt\_crosswt = 95, allele\_mutatewt = 95, migratewt = 10 for the GA. The docking sphere is centered on the N of amino acid 197 (atom 1416) with radius 4, 7, and 15 Angstroms, and the increase in radius results in a larger region to scan for docked ligands and more active amino acids of the enzyme. The sphere region does not have a hard boundary and any amino acid touching it is included in the docking region. The active amino acids used in a radius 4 Angstrom sphere are, for P197 AHAS,

GLY23 ALA24 SER25 GLU27 SER70 GLY71 THR95 GLY96 GLN97 VAL98 PRO99 ARG100 ARG101 MET102 ILE103 GLY104 THR105 ALA107 PHE108 GLN109 GLU110 THR111 LEU126 ILE127 LEU128 ASP155 ILE156 PRO157 LYS158 ASP159 VAL160 GLN161 GLN162

and are offset by 98, i.e. PRO99 is PRO197 in usual nomenclature. There are 33 for radius 4A.

The standard GOLD PLP molecular scoring function is used. This parameter set with 8 islands and 150000 operations, and select pressure 1.1, gives a more thorough calculational search for the protein-ligand energy minimum than the default setting. All amino acids that contribute to the docking score are listed in the GOLD output files, in addition are the score components, hydrogen bonding contributions, and any rotations of protein target side-chains, available for download.

**Figure 4.** The Chlorsulfuron herbicide molecule. It is a very commonly used acetolactate synthase inhibitor. It has molecular mass 358 Daltons, 1 stereoisomer, 4 rotatable bonds, 2 H-bond donors, 7 H-bond acceptors, and is insoluble in water. In real application, however, it is mixed in water and shaken until dissolved.



Of the 3 AlphaFold structures generated for each 197 P and 197 S sequences, the most reliable one is used and this is ranked by AF in a confidence estimate. In order to determine any variability in the local binding of Chlorsulfuron using the docking software GOLD, 3 spheres with different radii were used to find any dependence on the search size in Angstroms. These 3 spheres explore the surface terrain of the protein to ensure the binding site is well-defined, and it is as was found due to no change in the binding distribution from altering the search radius too much. (A very large sphere permits the docked molecule to find spots far away that may have local energy minima but no effect on protein function, i.e. multiple energy minima in binding.) There is typically a small set of major conformational modes of the ligand in a binding site, and the 3 docking limited spheres of radii 4,7,15 contain the same set and binding location, as expected if the site is well-defined. There are 500 independent docking runs used to sample the computational epitopes, each randomly initialized with different dihedral angles, and there are 6 combinations of P197/S197 with sphere radii 4,7, and 15 A.

*Score distributions*

Docking score distributions from large sets of independent runs provide much information about the bound ligands [17]. Breaking the distribution into a set of peaks shows the conformational modes and the width, height allows for an interpretation of conformational flexibility and population of individual conformers. The distributions in score can be approximately translated into protein-ligand binding energy through substituting Interpreting the distribution as a sampling of the energy states gives a numerically calculated energy density of states, which with statistical mechanics can be used to calculate ensemble observables such as expected energy or entropy of binding, and at different temperatures.

The score distributions for each of the 6 combinations are shown in Figure 5. In each of the 6 combinations it is clear that there are 2 major conformations coming from 2 separated and large peaks. The P197 and S197 AHAS structures are different, even if only one amino acid is changed. In particular, Proline doesn’t have an amine or a side-chain, and this is very clear in the distributions; the major conformations have a center in score that are slightly shifted as a result. In P197 (susceptible, left column) the 2 conformations are centered at scores 34.5 and 42.3. In S197 (resistant, right column) the conformations don’t appear as well-defined and have center scores of 34.4 and 40.3; the less shaped peak is explained in the following. The score shifts and the population fraction in each does quantify the change of the interaction with the herbicide CS, in addition to peak shape (narrow, wide) and relative height.

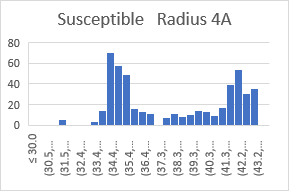
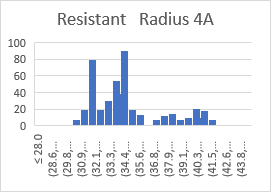
In all combinations, there is a local minimum at score 37.7 that separates the distribution into the 2 major conformational modes. The percentage of docking runs per the total in each combination is given in Table 1. In each set of different radii of computational docking, S197 (resistant) and P197 (susceptible), the percentages in the 2 conformations remains about the same, but the percentages are different across the 2 types P,S of sequences. In the susceptible strain P197, CS binds equally as much in both conformational modes, and in the resistant S197 strain, it binds significantly more to the 2nd less energetic mode by 3:1. The cavity has changed both in surface geometry and amino acid content/interactions and this is reflected in the difference from S197 to P197.

**Table 1.** Binding occupancy of the 2 major conformational modes in each combination of P197/S197 (susceptible/resistant) and R=4,7,15. All combinations used 500 runs and there were a couple of non-docked ligands.

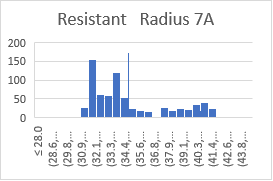
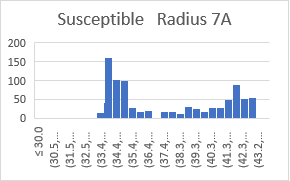
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | S197 Conf1 | Conf2 |  | P197 Conf1 | Conf2 |
| 4 A | 384/497=.77 | 113/497=.23 |  | 250/499=.5 | 249/499=.5 |
| 7 A | 728/955=.76 | 227/955=.24 |  | 513/950=.54 | 437/950=.46 |
| 15 A | 383/497=.77 | 114/497=.23 |  | 274/500=.55 | 226/500=.45 |

**Figure 5.** GOLD score distribution of Chlorsulfuron in a sphere of radius 4 Angstroms centered in the AHAS cavity. (a) The susceptible form of AHAS with P197. (b) The resistant form of AHAS with a S197. 500 extensively parameterized GOLD docking runs each.

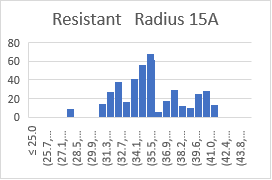
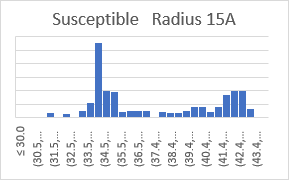
a. 4 A, P197 (susceptible) b. 4 A, S197 (resistant)

c. 7 A, P197 d. 7 A, S197



e. 15 A, P197 f. 15 A, S197



A more refined view of the set of 500 radius 4 scores, the distribution, is in Figure 6. For example, the relative population in the highest binding mode substantially decreased 3-fold (Table 1), and a new low scoring binding mode with splitting opens at score 32. The conformational make-up of bound CS is visibly seen to change much from P197 to S197.

**Figure 6.** This view of the distribution points to structural differences in the AHAS cavity and the influence it has on conformational differences, from the P197 AHAS type to the S197, as seen from many CS ligand docking calculations. Note the y-axis scale count change from a max of 60 to 35.

a. P197 AHAS b. S197 AHAS

Chart, histogram

Description automatically generated Chart, histogram

Description automatically generated

The distribution of the Chlorsulfuron non-covalently bound states is found from the random sampling of 500 docking runs. It is clear from the distributions in Figures 5,6 that the highest scoring binding mode in the resistant S197 AHAS has been almost eliminated from that in the natural P197 strain; the population fractions in Table 1 quantify the overall extent. This can be seen at radii 4,7, and 15 when the y-axis scale is taken into account. A more detailed quantitative analysis is given next.

*Expected binding energies at room temperature*

A Boltzmann statistical calculation of the expected interaction energy from the distribution of populated binding states gives [17],

,

for protein-ligand interaction and entropic energy contributions to the free energy,

.

The estimate of 6.5 PLP score change, which is approximately correct in numerous examples and ranges from IC50 data [17], to 1 kcal is used to translate. The interaction energy is approximately .4 kcal different between the 2 strains of AHAS and CS. This is =2.5 in bound affinity as affinity scales exponentially in base 10 with 6.5. The minimum dose was given to the Soybean strains to eliminate the susceptible one. A bound affinity of x2.5 would presumably require at least 2.5 more to have the same effect. In other words, out of the 500 docking runs, 50% were in the low scoring binding mode and 50% in the high one; to achieve the same number from a 77%, 23% mix in the resistant case would require a dose x2.2, close to the quantifiable x2.5. However, these calculations do not consider the conformational shape and location in or about the cavity containing 197.

Upon closer examination of the location and shape of the CS epitopes, there is more that can be said about the differences between P197 and S197 AHAS. The maximum of both distributions is 44.2 and 45.0. This naively indicates that CS appears to bind higher to the resistant strain. However the highest 5 scoring docked CS molecules in the S197 resistant calculation are epitopes outside the cavity and ‘down a hill’ from the interior region. These are far from the active site and should not be considered in the comparison. There are no outlier states at high score in the susceptible strain. Redoing the Boltzmann statistical calculation gives,

.

The interaction energy difference is now .5 kcal, and the entropic energy didn’t change within single digit precision. The binding affinity difference from interaction is approximately . x3.2 is the beyond the maximal dose given to these Soybean plants of x3.0.

*De-localization of bound epitopes, due to weakness of binding*

In further analysis, in order to justify an action of a ligand as an inhibitor it is not only necessary to examine its physical likelihood via binding attraction but also that the ligand is in fact bound in the relevant region, i.e., the cavity the protein uses to perform a function and if in a catalytic site that it does so to block the functional role of particular amino acids. Docking scores provide binding quantity estimates. Specific conformations/localized modes of interaction can be found by large sampling and in mimicking thermal conditions (not a frozen out isolated x-ray epitope).

There can be higher scoring, and more attracted to the protein, epitopes but not located specifically in the cavity to result in inhibition; examples are the outliers at high score in the resistant S197 strain. In the section the cluster of epitopes is examined from the full set of docking runs, and it is clear that with one amino acid change from a Proline to Serine that the binding goes from selective and confined in a small portion of the cavity to a much larger region extending to the boundaries and slightly beyond the cavity. RMSD calculations about the tight bunch of closely high scoring epitopes quantify the delocalization.

Figures 7 and 8 illustrate how the CS bound specificity to the herbicide site decreases in the highest binding conformational mode (right-hand peak in Figure 6a,b). The coloring of the surface goes from blue-white-red following levels of hydrophilicity to hydrophobicity. Figures 7a (P197) and 8a (S197) each shows a set of 20 bound ligands sampling the mode with score >37.3 to the maximum score found. The grouping in P197 (susceptible plant) are all grouped together highly and indicate that the tight Chlorsulfuron binding potential. In the S197 set, Figure 7a, the grouping is much more broad and there is much less visibility of tightness – the bound states have become delocalized in the P197S mutation. This is more pronounced if only the 20 scoring ligand bound states are used and not a thorough sampling over the entire 1st binding mode (in Figure 6a,b).

Figures 7b and 8b don’t show bound states near the maximum score in each P197 and S197, but rather a sampling of all binding states in the higher conformational mode. The docking solutions have been ranked and in the ranking a complete spread of docked Chlorsulfuron is selected to span the conformational mode. The score ranges are listed. These sets of ligand bindings show a much more diffuse binding in location and broader in the total region, as in the cluster of states near the maximum score. The ‘volume’ occupied by the hair (ligands superimposed on each other) has increased substantially.

**Figure 6.** Primary binding mode of the susceptible P197 AHAS strain with 20 scores at 38.49 to 43.4. These are a set of ranked docked CS runs from 4\*n+20 from n=0,…,19 and span most of the scores above 37.7 (113/497) and all of the highest binding mode in the Figure 1(a) distribution. P197 is colored in purple. (b) Random collection of 100 bound ligands of any score.

a. b.

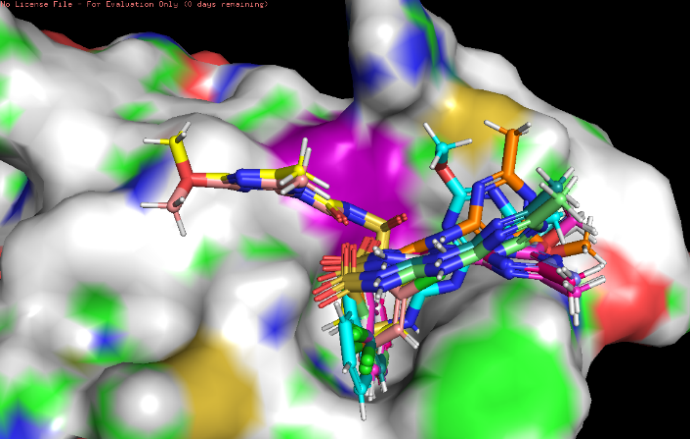
**A picture containing cake, decorated, colorful

Description automatically generated**A picture containing cake, decorated, birthday, chocolate

Description automatically generated

**Figure 7.** Primary binding mode of resistant S197 AHAS strain with 20 scores at 42.3 to 43.4. These are the docked CS from 6\*n+20 from n=0,…,19 and span half of the scores above 37.7. (b) Random collection of 100 bound ligands of any score.

a. b.

****A picture containing cake, decorated, birthday, colorful

Description automatically generated

The 500 docked CS molecules to the resistant S197 AHAS occupy a volume containing the herbicide site which is ~6.7 that of the P197 AHAS strain; the radius difference, seen visually and quantifiable by inspecting the bound states, is approximately twice. This factor can be calculated rigorously by an RMSD analysis, next. Just as the highest binding mode in P197 has been knocked out from that in Figure 6(a) to Figure 6(b) after the change P197S, most ligands can’t find a very specific repeatable pose in S197 in a random collection of docking runs. The non-specific distribution of bindings in the S197 cavity would require a concentration of x8 more to populate the P197 region the same amount. This is beyond the binding affinity difference of x3.2. This total of (6.7\*3.2) ~ x20 is a measure of the resistance of S197 AHAS in comparison to the minimal dose required for non-survival of P197 in the presence of CS.

There is a raw RMSD quantity that measures the difference between 2 docked ligands,

where d is the distance between the atom of 2 distinct ligands and N is the number of atoms. From a reference docked CS conformation all 500 RMSDs can be generated from the reference viewpoint. The reference of the highest scoring ligand can be used if the grouping is strong at high interaction score. A centroid over a set of docked ligands can be used also. But if there is a diffuse set of states in localization then more care is required. The 5th highest scoring epitope in P197 is used in this case, and the 10th highest is used the S197 case due to the much broadening. However, these also represent the potentially longest bound epitope in each case due to the higher scores of each in the modeled interaction. The distributions of these RMSDs for P197 and S197 are given in Figures 9a and 10a. Note the large difference between the 2 in the number at individual RMSD. The comparison of the 2 clearly show a large grouping near the highest binding epitope in the case of P197 as these are small RMSDs and a large number of RMSDs in the range [6,8] A. In the case of S197 most of the 500 docking runs are widely different than the highest scoring ligands; the RMSD is not directly connected to nor the same as binding interaction. Chlorsulfuron has become strongly delocalized in the P197S mutation with RMSDs in a range of [6,8] Angstroms.

Next, there is a running average RMSD that demonstrates a convergence of the RMSDs from high score to low score. First the 500 docking scores are ranked in descending order from the highest. Then a running average is calculated,

where is the number of scores, j. (j-4 or j-9 for P,S197)

From the 500 ranked ligands, a more interacting set of [1,50] can be compared to a more broadly interacting set of [1,100] or [1,200], i.e., the running average includes more scores less than the maximum. Both curves (~500 connected calculated points) share the same characteristics: possible convergence near 500 and 2 bumps at small ligand number. Figure 10c plots the running RMSD average with the P197 docked ligand 5 as a reference point; this gives some information about comparing a specific set of atomic coordinates and the docked CS molecules to S197; note the fluctuation at high rank and this is due to loss of delocalization and specificity at high score in S197 (ligand conformation jitter).

At ligand (score) number 100 or 250 it can be seen that the difference in the running average RMSD is approximately 2. The running average isn’t just a displacement of molecules by some distance but also includes the orientation and conformational change, in this case with respect to the reference molecule. There is clearly a much larger spread of docked ligands in the case of S197 than P197. The specificity of binding and localization has both been largely diminished in the P197S mutation. The change of x2 in an RMSD is similar to the change in occupied binding dimension of x2 (volume x2x2x2=x8) and the effect of delocalization is estimated roughly as x8 or approximate, as earlier in using binding volume. This estimate using volume is heuristically linear and the net result of delocalization and less binding interaction could be much more than x20. Having a more free and less interacting ligand is a result of an inappropriate to that region ligand which will slip away if the net protein-ligand interaction potential does not trap it and the CS molecule is not trapped in S197.

Note that the running average RMSD does depend on the reference ligand, or centroid if that is the case. If the scores are not highly grouped near the reference ligand then there can be large fluctuations in the nearby in score pose RMSDs and also a greater difference in the overall curve. This has been eliminated by starting the curves at ligand score 5 and 10 in P197 and S197. Some of the docked ligands in S197 from ranked 1 to 10 are quite different on average than the rest, and the first 4 poses on P197 are not locally close by nor in the cavity (previous section).

**Figure 9.** P197 RMSD calculations.

a. RMSD distribution from 500 runs P197.. b. Running average from ranked ligand #5. P197.

Chart, histogram

Description automatically generatedChart, line chart

Description automatically generated

**Figure 10.** S197 RMSD calculations.

a. RMSD distribution from 500 runs S197.. b. Running average from ranked ligand #10. S197.

Chart, histogram

Description automatically generatedChart, line chart

Description automatically generated

c. S197 running average from ranked ligand #1 and using the reference epitope of P197 ligand 5.

Chart, line chart

Description automatically generated

**Conclusions**

A detailed binding study of a commonly used herbicide Chlorsulfuron is performed on a naturally occurring plant (and weed) AHAS enzyme together with a mutated form. It is known that a multitude of point mutations in this enzyme can lead to resistance from many different herbicides, and the amino acid substitution P197S is considered in Soybean AHAS. The work shows that resistance to CS is obtained after this substitution, by examining the binding statistically of CS, and is quantified to be in the x20 range or more as far as dose amount to have an effect.

This enzyme has 2 sites of interest. The first is the catalytic site from which the enzyme initiates the production process of branched-chain amino acids (Val, Leu, Ile). The second is the herbicide binding site and this is connected by a transport channel to the former. The action of the herbicide is to block this channel from being used, stopping the transport of pyruvate or 2ketobyruvate to be used at the catalytic site. There is no evidence that the point mutation P197S or attachment of CS causes a conformational change to *Arabidopsis Thaliana* AHAS in comparing x-ray structures uninhibited and inhibited, nor also in computational energy minimization of Soybean S197 AHAS and in complex with CS; the 2 share much of their sequences and are visually indistinguishable in the backbone structure.

Large numbers of docking runs in a thorough study show that the major conformational mode of binding of CS to Soybean AHAS is eliminated upon altering P197S. The amino acid Proline is very different from Serine in that there is no amine or side-chain. The binding site changes slightly, together with the solvent accessible surface, and this alters quite much the binding modes of the ligand and the epitopes in shape. The point mutation also has an effect of delocalizing the attachment of bound CS to its herbicide site – it generally loses a specific target location and binds more diffusely to a larger surface area containing the sub-region that CS binds to in P197. Both the loss of most of the major binding mode and the delocalization block the CS molecule from binding highly and specifically in location. This mutation weakens the blockage of the entry to the transport channel to the catalytic site and thus the plant becomes resistant.

This study used molecular computational tools and methods to examine one amino acid point mutation in the Soybean AHAS enzyme. The approach is generally applicable to any other setup of mutation and protein. Although experimental x-ray structures were not used directly, known x-ray structures of AtAHAS were compared against. The computational aspect could allow for a faster analysis of enzymes involved in crop research and development genetically, and in the discovery and development of herbicides for combined benefit. There are a number of amino acids interacting in the herbicide site of AHAS - a potential computational study of mutating one at a time and analyzing the effect with a set of herbicides individually could elucidate an optimal choice of point mutation for individual or collective herbicidal resistance (not just CS used in this work). Although lab work is always necessary and in the plant context the growing of crops and testing, the computational contribution does not necessarily require either in order to find potentially interesting information about the small molecule herbicides binding to and resistance of plants with mutations.

**Acknowledgements**

G.C., R.U., D.T., and B.U. are grateful for the use of the Sapelo2 cluster at the Georgia Advanced Computing Resource Center and for the hospitality given at the Complex Carbohydrate Research Center. This work is partially supported by the Scientific Research Projects Coordination Unit of Akdeniz University (Grant No. FDK-2018-3775).

**Author Information**

Gordon Chalmers, *Complex Carbohydrate Research Center, University of Georgia, Athens, GA, 30602.* E-mail: [gordoncs@uga.edu](mailto:gordoncs@uga.edu)

Rustem Ustun, *Department of Field Crops, Faculty of Agriculture, Akdeniz University, Antalya, Turkey.* Email: [rustemustun@akdeniz.edu.tr](mailto:rustemustun@akdeniz.edu.tr)

Daniel Tehrani, *Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602.* E-mail: [daniel.tehrani@uga.edu](mailto:daniel.tehrani@uga.edu)

Bulent Uzun, *Department of Field Crops, Faculty of Agriculture, Akdeniz University, Antalya, Turkey.* E-mail: [bulentuzun@akdeniz.edu.tr](mailto:bulentuzun@akdeniz.edu.tr)

**Conflict of Interest Statement**

There are no conflicts of interest.

**Data Availability Statement**

All data generated or analyzed during this study are included in this published article.

**Supplementary information**

Calculations, docking results files, figures, and the paper are located at http:// … (2 Gigabytes)

**Appendix 1.**

*Protein structure*

The acetolactate synthase (ALS) enzyme (also known as acetohydroxy acid or acetohydroxyacid synthase, abbr. AHAS) is a protein found in plants and micro-organisms. ALS catalyzes the first step in the synthesis of the branched-chain amino acids. These are the sequences used to create the structure in AlphaFold2. The point mutation of Proline to Serine makes the plant resistant to a class of herbicides.

Chlorsulfuron is an ALS (acetolactate synthase) inhibitor herbicide and is a sulfonylurea compound. (Used herbicide).

*Resistance plant, red Serine, and susceptible plant, blue Proline. Soybean ALS enzyme.*

MAATASRTTRFSSSSSHPTFPKRITRSTLPLSHQTLTKPNHALKIKCSISKPPTAAPFTKEAPTTEPFVSRFASGEPRKGADILVEALERQGVTTVFAYPGGASMEIHQALTRSAAIRNVLPRHEQGGVFAAEGYARSSGLPGVCIATSGPGATNLVSGLADALMDSVPVVAITGQV{S,P}RRMIGTDAFQETPIVEVSRSITKHNYLILDVDDIPRVVAEAFFVATSGRPGPVLIDIPKDVQQQLAVPNWDEPVNLPGYLARLPRPPAEAQLEHIVRLIMEAQKPVLYVGGGSLNSSAELRRFVELTGIPVASTLMGLGTFPIGDEYSLQMLGMHGTVYANYAVDNSDLLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKQAHVSVCADLKLALKGINMILEEKGVEGKFDLGGWREEINVQKHKFPLGYKTFQDAISPQHAIEVLDELTNGDAIVSTGVGQHQMWAAQFYKYKRPRQWLTSGGLGAMGFGLPAAIGAAVANPGAVVVDIDGDGSFIMNVQELATIRVENLPVKILLLNNQHLGMVVQWEDRFYKSNRAHTYLGDPSSESEIFPNMLKFADACGIPAARVTKKEELRAAIQRMLDTPGPYLLDVIVPHQEHVLPMIPSNGSFKDVITEGDGRTRY

*For reference this sequence is also given, NP\_190425.1 Arabidopsis thaliana AHAS gene Region protein sequence*

MAAATTTTTTSSSISFSTKPSPSSSKSPLPISRFSLPFSLNPNKSSSSSRRRGIKSSSPSSISAVLNTTTNVTTTPSPTKPTKPETFISRFAPDQPRKGADILVEALERQGVETVFAYPGGASMEIHQALTRSSSIRNVLPRHEQGGVFAAEGYARSSGKPGICIATSGPGATNLVSGLADALLDSVPLVAITGQVPRRMIGTDAFQETPIVEVTRSITKHNYLVMDVEDIPRIIEEAFFLATSGRPGPVLVDVPKDIQQQLAIPNWEQAMRLPGYMSRMPKPPEDSHLEQIVRLISESKKPVLYVGGGCLNSSDELGRFVELTGIPVASTLMGLGSYPCDDELSLHMLGMHGTVYANYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKTPHVSVCGDVKLALQGMNKVLENRAEELKLDFGVWRNELNVQKQKFPLSFKTFGEAIPPQYAIKVLDELTDGKAIISTGVGQHQMWAAQFYNYKKPRQWLSSGGLGAMGFGLPAAIGASVANPDAIVVDIDGDGSFIMNVQELATIRVENLPVKVLLLNNQHLGMVMQWEDRFYKANRAHTFLGDPAQEDEIFPNMLLFAAACGIPAARVTKKADLREAIQTMLDTPGPYLLDVICPHQEHVLPMIPSGGTFNDVITEGDGRIKY

**References**

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

**2.** Ray TB. Site of action of chlorsulfuron-inhibition of valine and isolucine biosynthesis in plants, *Plant Physiol.* 75 **(1984)** 827–831.

**3.** Schloss JV, Ciskanik LM, Van Dyk DE. Origin of the herbicide binding site of acetolactate synthase, *Nature* 331 **(1988)** 360–362.

**4.** Chaleff RS, Mauvais CJ. Acetolactate synthase is the site of action of two sulfonylurea herbicides in higher plants, *Science* 224 **(1984)** 1443–1445.

**5.** A.K. Chang, R.G. Duggleby. Expression, purification and characterization of Arabidopsis thaliana acetohydroxyacid synthase, *Biochem. J.* 327 **(1977)** 161–169.

**6.** Duggleby RG, Pang SS. Acetohydroxyacid synthase, *J. Biochem. Mol. Biol.* 33 **(2000)** 1–36.

**7.** Gutteridge S, Thompson ME, Ort O, Shaner DL, Stidham M, Singh B, Tan S, Johnson TC, Mann RK, Schmitzer PR et al. **(2012)** Acetohydroxyacid synthase inhibitors (AHAS/ALS). In Modern Crop Protection Compounds (Kr€amer W, Schirmer U, Jeschke P & Witschel M, eds), pp. 29–162. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany.

**8.** Whitcomb CE. An introduction to ALS-inhibiting herbicides. *Toxic. Ind. Health* 15 **(1999)** 232–240.

**9.** Duggleby RG, McCourt JA & Guddat LW. Structure and mechanism of inhibition of plant acetohydroxyacid synthase. *Plant Physiol. Biochem.* 46 **(2008)** 309–324.

**10.** McCourt JA, Duggleby RG. Acetohydroxyacid synthase and its role in the biosynthetic pathway for branched-chain amino acids. Amino Acids 31 **(2006)** 2:173-210. PMID: 16699828 DOI: [10.1007/s00726-005-0297-3](https://doi.org/10.1007/s00726-005-0297-3)

**11.** Liu Y, Li Y, Wang X. Acetohydroxyacid synthases: evolution, structure, and function. *Applied Microbiology and Biotechnology* 100 **(2016)** 8633-8649*.* DOI: 10.1007/s00253-016-7809-9.

**12.** Pang SS, Duggleby RG, Guddat LW. Crystal structure of yeast acetohydroxyacid synthase: a target for herbicidal inhibitors. *J. Mol. Bio.* 317 **(2002)** 2:249-62. PMID: 11902841 DOI: 10.1006/jmbi.2001.5419

**13.** JA, Pang SS, Guddat LW & Duggleby RG. Crystal structure of yeast acetohydroxyacid synthase in complex with a sulfonylurea herbicide, chlorsulfuron. *Biochemistry* 44 **(2005)** 2330–2338.

**14.** Pang SS, Guddat LW & Duggleby RG. Crystallization of Arabidopsis thaliana acetohydroxyacid synthase in complex with the sulfonylurea herbicide chlorimuron ethyl. *Acta Crystallogr D* 60 **(2004)** 153–155.

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

**16.** Garcia MD, Wang JG, Lonhienne T, Guddat LW. Crystal structure of plant acetohydroxyacid synthase, the target for several commercial herbicides. *FEBS Journal* 284 **(2017)** 13:2037-2051. PMID: 28485824 DOI: 10.1111/febs.14102

**17.** Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* 267 **(1997)** 3:727-748. DOI: 10.1006/jmbi.1996.0897 PMID: 9126849

Cambridge Crystallographic Data Centre. **(2021)** CCDC Discovery GOLD. Retrieved from GOLD Protein Ligand Docking Software: https://www.ccdc.cam.ac.uk/solutions/csd-discovery/Components/Gold/

**18.** Chalmers G. Dynamic docking in protein-ligand modeling, *to appear* *in the Journal of Computational Chemistry*, 10.26434/chemrxiv-2022-6m0q5

**19.** Chalmers G. Computational Study of Paxlovid in Ligand GA, *submitted to Journal of Medicinal Chemistry*. 10.26434/chemrxiv-2022-p2phq

**20.** Jumper, J., Evans, R., Pritzel, A. et al. Highly accurate protein structure prediction with AlphaFold. Nature 596 **(2021)** 583–589. <https://doi.org/10.1038/s41586-021-03819-217>.

**21.** Mihaly Varadi, Stephen Anyango, Mandar Deshpande, Sreenath Nair, Cindy Natassia, Galabina Yordanova, David Yuan, Oana Stroe, Gemma Wood, Agata Laydon, Augustin Žídek, Tim Green, Kathryn Tunyasuvunakool, Stig Petersen, John Jumper, Ellen Clancy, Richard Green, Ankur Vora, Mira Lutfi, Michael Figurnov, Andrew Cowie, Nicole Hobbs, Pushmeet Kohli, Gerard Kleywegt, Ewan Birney, Demis Hassabis, Sameer Velankar, AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models, Nucleic Acids Research, Volume 50, Issue D1, 7 January **2022**, Pages D439–D444, <https://doi.org/10.1093/nar/gkab1061>.

**22.** CASP-14. *Proteins: Structure, Function, and Bioinformatics.* 89 **(2021)** 12:1603-1997. <https://predictioncenter.org/casp14/index.cgi>

**23.** Tranel PJ, Wright TR. Resistance of weeds to ALS-inhibiting herbicides: what have we learned? *Weed Science.* 50 (2002) 6:700-712. DOI: <https://doi.org/10.1614>

**24.** Lonhienne T, Low YS, Garcia MD, Croll T, Gao Y, Wang Q, Brillault L, Williams CM, Fraser JA, McGeary RP, West NP, Landsberg MJ, Rao Z, Schenk G, Guddat LW. Structures of fungal and plant acetohydroxyacid synthases. *Nature* **(2002)** 586:317–321. DOI: 10.1038/s41586-020-2514-3

**25.** Qi Y, Han H, Vila-Aiub MM, Powles SB. AHAS herbicide resistance endowing mutations: effect on AHAS functionality and plant growth. J. Exp. Bot. 61 (2010) 14:3925-34. PMID: 20627897 PMCID: PMC2935867 DOI: 10.1093/jxb/erq205

**26.** Qi Y, Powles SB. Resistance to AHAS inhibitor herbicides: current understanding. Pest Management Science 70 (2014) 9:1340-1350.

**27.** Wang JG, Lee PK-M, Dong YH, Pang SS, Duggleby RG, Li ZM & Guddat LW. Crystal structures of two novel sulfonylurea herbicides in complex with Arabidopsis thaliana acetohydroxyacid synthase. *FEBS J.* 276 **(2009)** 1282.

**28.** Subramanian MV, Hung HY, Dias JM, Miner VW, Butler JH, Jachetta JJ. Properties of mutant acetolactate syntheses resistant to triazolopyrimidine sulfonanilide, *Plant Physiol.* 94 **(1990)** 239–244.

**29.** Chang AK & Duggleby RG. Herbicide-resistant forms of Arabidopsis thaliana acetohydroxyacid synthase: characterization of the catalytic properties and sensitivity to inhibitors of four defined mutants. *Biochem. J.* 333 **(1998)** 765–777.

**30.** Preston C, Stone LM, Rieger MA, Baker J. Multiple effects of a naturally occurring proline to threonine substitution within acetolactate synthase in two herbicide-resistant populations of Lactuca serriola, *Pestic. Biochem. Physiol.* 84 **(2006)** 227–235.

**31.** Mourad G, King J. Effect of four classes of herbicides on growth and acetolactate-synthase activity in several variants of Arabidopsis thaliana, *Planta* 188 **(1992)** 491–497.

**32.** Subramanian MV, Hung HY, Dias JM, Miner VW, Butler JH, Jachetta JJ. Properties of mutant acetolactate syntheses resistant to triazolopyrimidine sulfonanilide, *Plant Physiol.* 94 **(1990)** 239–244.

**33.** Preston C, Stone LM, Rieger MA, Baker J. Multiple effects of a naturally occurring proline to threonine substitution within acetolactate synthase in two herbicide-resistant populations of Lactuca serriola, *Pestic. Biochem. Physiol.* 84 **(2006)** 227–235.

**34.** Pettersen EF, Goddard GT, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera--a visualization system for exploratory research and analysis. J. Comput. Chem. 25 **(2004)** 13:1605. PMID: 15264254 DOI: 10.1002/jcc.20084

University of California at San Francisco (UCSF) - Resource for Biocomputing, V. a. (Current). UCSF Chimera, an Extensible Molecular Modeling System. Retrieved from UCSF Chimera: https://www.cgl.ucsf.edu/chimera/

**35.** The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC. Retrieved from PyMOL: http://www.pymol.org/pymol

**36.** Kingsley LJ, Brunet V, Lelais G, McCloskey S, Milliken K, Leija E, Fuhs SR, Wang K, Zhou E, Spraggon G. Development of a virtual reality platform for effective communication of structural data in drug discovery. *Journal of Molecular Graphics and Modeling* 89 **(2019)** 234-242. DOI:10.1016/j.jmgm.2019.03.010

Nanome, Inc. **(2021)** Nanome. Retrieved from Nanome: <https://nanome.ai/>

**37.** Sadowski J, Gasteiger J., Klebe G. Comparison of Automatic Three-Dimensional Model Builders Using 639 X-Ray Structures. *J. Chem. Inf. Comput. Sci.* 34 **(1994)** 1000-1008 DOI: 10.1021/ci00020a039

Schwab CH. Conformations and 3D pharmacophore searching*. Drug Discovery Today: Technologies* 2010, 7(4), Winter **2010**, e245-e253 DOI: 10.1016/j.ddtec.2010.10.003

Molecular Networks GmbH, Altamira, LLC. (2021). Corina. Retrieved from MN-AM Corina: <https://www.mn-am.com/products/corina>

Schwab CH. Molecular Structure Representation in Chemoinformatics Applications, Schwab CH, BigChem Autumn 2017 School, Modena, Italy. Received from http://bigchem.eu/sites/default/files/School3\_Schwab.pdf

**38.** MathWorks, Inc. MATLAB. (2022b). Retrieved from MathWorks: <https://www.mathworks.com/>

**39.** D.A. Case, H.M. Aktulga, K. Belfon, I.Y. Ben-Shalom, J.T. Berryman, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, G.A. Cisneros, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, K. Kasavajhala, M.C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, A. Shajan, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, J. Wang, H. Wei, R.M. Wolf, X. Wu, Y. Xiong, Y. Xue, D.M. York, S. Zhao, and P.A. Kollman. **(2022)**, Amber 2022, University of California, San Francisco.

1. X-ray structures give those of a single frozen mode in complex with the protein. These are effectively low entropic and low temperature conformations and although informative do not give a full picture of a thermally populated set of bound conformations in biological conditions. Large sampling of conformational states through docking runs provide a more complete picture [18,19]. [↑](#footnote-ref-1)