Structural Prediction, Glucose-1-Phosphate Interaction and Influence of Broad Leaves Herbicides on Spinach Leaves α-glucan Phosphorylase: An *in Silico* Study

Ritu Jain, Neha Jain, Anil Kumar*

School of Biotechnology, Devi Ahilya University, India

Abstract

Starch phosphorylase is a α-glucan-phosphorylase that plays various important roles in the regulation of carbohydrate metabolism. It is mainly involved in the degradation of starch, however, it has also been reported to be involved in the synthesis of starch especially during abiotic stress conditions. Mutant studies revealed that disruption of this enzyme adversely affects the survival of plants. In the present study, we have established that starch phosphorylase is highly conserved among plant species in terms of protein sequence and 3-dimensional structure. Besides, we theorized degradation by starch phosphorylase as a common plausible mechanism through which herbicides may exert adverse effects on general plant population. Using in silico biology, we provided a clue that some of the commercially available herbicides have potential to bind with the active site of starch phosphorylase and these can competitively inhibit the activity of this vital enzyme.

Keywords Starch Phosphorylase, Docking, Glyphosate, Sethoxydin, Quizalofop, Glucose-1-phosphate

1. Introduction

Starch phosphorylase is an important enzyme of glucan catabolism in plants [1, 2]. However, Tickle et al. (2009) [3] proposed a model suggesting its importance in starch synthesis. Zeeman et al. (2004) [4] ascribed its involvement in the survival of the plants using various abiotic stresses by providing alternate route for starch degradation. Besides, it has also been speculated that starch phosphorylase also provides substrate(s) for the chloroplastic oxidative pentose phosphate pathway. It has been assumed that this pathway is involved to provide reducing agents during dark and these agents are necessary to tolerate stress [5]. In addition, it also helps to provide carbon compounds that replenish intermediates in Calvin cycle. The pathway is important under photo-respiratory conditions where such intermediates get depleted [6]. These are some highlights on the emerging importance of starch phosphorylase as our basic understanding is enhancing in this field. On the other hand, it is also important that starch phosphorylase enzyme from different plant species has been reported to be highly conserved in sequence as well as protein structural level [7]. Due to structural similarity, it is anticipated that the substrate binding affinity of this enzyme must be similar in all plant starch phosphorylases.

In India, agriculture is an important component of economy and majority of population is dependent on agricultural earnings. There is much loss of agricultural products due to presence of weeds in the farms. There is geographical and biological diversity of weeds which require much use of different types of herbicides [8, 9]. In the present study, we studied structural similarity and substrate affinity of starch phosphorylase using in silico approach. Besides, we also studied the influence of herbicides on this enzyme. To the best our knowledge, it is the first report on the influence of herbicides on starch phosphorylase.

2. Materials and Methods

2.1. Protein-BLAST (BLASTp)

Primary sequence of starch phosphorylase was downloaded from Protein database of NCBI with the GenBank ID: CAA59464.1. (http://www.ncbi.nlm.nih.gov/). The length of the sequence was 971 amino acids. The FASTA sequence of this protein is used further for orthologs identification and structure prediction. Protein BLAST was performed against non-redundant protein database of NCBI with the default settings of NCBI BLAST [10].

2.2. Structure Prediction of Starch Phosphorylase by using Homology Modelling
The tertiary structure of the sequence was predicted by using Modeller 9 v 14. [11]. First NCBI BLASTp (Altschul et al., 1990) was performed against Protein Data Bank (PDB) for template identification and then Modeller 9.14 was used for structure generation.

2.3. Structure Validation

Ramachandran plot was plotted for the modelled structure by using SAVES server’s Procheck algorithm. (http://nihserver.mbi.ucla.edu/SAVES/)

2.4. Structure Alignment with *Arabidopsis thaliana* Starch Phosphorylase

The 3D structure model generated was structurally aligned with the X-Ray crystallographic structure of starch phosphorylase (PDB ID – 4BQE) from A. thaliana by using PyMol Viewer.

2.5. Docking Studies

Lamarckian Genetic Algorithm (LGA) algorithm was used for ligand conformational searching which is a hybrid of genetic algorithm and local search algorithm. The modelled structure was first prepared for interaction analysis. An extended PDB format termed as PDBQT file was used for coordinate files which includes atomic partial charges. Kollman’s charges were given to the targets by using MGL AutoDockTools interface [12] so that the charges get uniformly attached with the protein surface. Then the target file was saved in PDBQT format which is a prerequisite for docking analysis. The enzyme interaction with glucose-1-phosphate was analyzed and thereafter its interaction studies with three herbicides namely glyphosate, sethoxydim and quizalofop were performed. All the ligand files were downloaded from PubChem database. (http://pubchem.ncbi.nlm.nih.gov/)

2.5.1. Starch Phosphorylase Interaction with Glucose-1-phosphate

Glucose-1-phosphate (G-1-P) is used as a substrate for the enzyme in vitro studies. Glucose-1-Phosphate 3D structure was downloaded from PubChem Database (compound ID: 65533) (https://pubchem.ncbi.nlm.nih.gov/) in SD file format and was converted into PDB format using PYMOL molecular viewer. G-1-P was given Gasteiger charges (opposite to that of the target) and was saved in PDBQT format using MGL AutoDockTools Interface (Version 1.5.6). The grid and Dock file was prepared by using Autodock Program (http://autodock.scripps.edu/) and the docking score values are predicted.

2.5.2. Starch Phosphorylase Interaction with Glyphosate, Sethoxydim and Quizalofop (Herbicides):

The 3D structures of glyphosate (compound ID: 3496), sethoxydim (compound ID: 52923) and quizalofop (compound ID: 53518) were downloaded from NCBI Pubchem database in SDF format and were converted into PDB format by using PYMOL Molecular Viewer. All the ligand files were converted into PDBQT format by using MGL AutoDockTools Interface after giving them the Gasteiger Charges. Grid file and dock file were prepared for each ligand (herbicide) by using the same procedure as that for G-1-P.

For validation, docking analysis was also performed using LeadIT (version 2.1.8) tool and the ligand binding affinity was assessed using Hyde tool.

3. Result and Discussion

### BLASTp Results

The protein sequence of spinach starch phosphorylase (glucan-phosphorylase) followed by BLAST showed more than 90 % sequence similarity and more than 90% query coverage with glucan phosphorylases from different sources of plant origin. The results revealed that Spinacia oleracea glucan phosphorylase sequence showed higher similarity with glucan phosphorylase from Citrus sinensis, Ricinus communis, Populus trichocama, Nicotiana sylvestrus, Solanum tuberosum, Lipmoea batatas, Jatropha curcas, Zea mays etc (Table 1). These data consistently revealed the conserved nature of starch phosphorylase among various plant species (Figure 1).

<table>
<thead>
<tr>
<th>Name</th>
<th>Query Coverage</th>
<th>E-Value</th>
<th>% Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-1,4 glucan phosphorylase [Citrus sinensis ]</td>
<td>97%</td>
<td>0.0</td>
<td>74%</td>
</tr>
<tr>
<td>Glycogen phosphorylase [Ricinus communis]</td>
<td>98%</td>
<td>0.0</td>
<td>72%</td>
</tr>
<tr>
<td>α-1 family protein [Populus trichocama]</td>
<td>97%</td>
<td>0.0</td>
<td>72%</td>
</tr>
<tr>
<td>α-1,4 glucan phosphorylase[Nicotiana sylvestrus]</td>
<td>99%</td>
<td>0.0</td>
<td>72%</td>
</tr>
<tr>
<td>α-1,4 glucan phosphorylase [Solanum tuberosum]</td>
<td>98%</td>
<td>0.0</td>
<td>73%</td>
</tr>
<tr>
<td>Starch phosphorylase [Lipmoea batatas]</td>
<td>94%</td>
<td>0.0</td>
<td>73%</td>
</tr>
<tr>
<td>JCGZ 19916 [Jatropha curcas]</td>
<td>98%</td>
<td>0.0</td>
<td>72%</td>
</tr>
<tr>
<td>α-1,4 glucan phosphorylase [Zea mays]</td>
<td>94%</td>
<td>0.0</td>
<td>71%</td>
</tr>
</tbody>
</table>
Figure 1. Putative conservatism obtained as a BLAST hit (BLAST hit Table snapshot) in response to the query sequence of spinach starch phosphorylase. All the results were of Red color depicting more than 200 bit score which is considered to be very good.

Structure prediction by Modeller

Upon performing NCBI BLAST against PDB database, the sequence of starch phosphorylase showed maximum similarity with PDB ID – 2C4M and with E-value 3 x 10^-111 and 81% query coverage, therefore it was taken as a template. Using this template, structure model was generated by using MODELLER9.14 `model-single.py` file. Five model structures were formed out of which the best one (Figure 2) was selected using DOPE score and MolPdf Score of Modeller tool.

Figure 2. Modeller structure in cartoon form visualized by using PYMOL [13]
Structure Validation

Ramachandran plot (Figure 3) for the modelled structure showed ~99% residues in the favourable region, thereby validating our predicted structure.

Starch Phosphorylase interaction with Glucose-1-phosphate: Glucose-1-phosphate was used as a ligand for two reasons, first because it is used as a substrate for the enzyme in in vitro studies having a good affinity with the enzyme and second to know the binding sites where the substrate can bind to the enzyme with lowest binding energy and maximum stability. Molecular in silico docking studies revealed that glucose-1-phosphate (G-1-P) formed three hydrogen bonds with starch phosphorylase enzyme. To form aforesaid H bonds, CYS180, VAL-178, LYS-185 residues were found to be involved (Figure 5).

Structural Alignment

The modelled structure was aligned with the A. thaliana starch phosphorylase to check the structure conservatism for the enzyme in plants. Both the structures were completely superimposed, except a very short region displayed in blue with 0.779 RMSD value.

Docking Analysis

From interactions study of above mentioned herbicides and starch phosphorylase using docking studies, it was found that glyphosate interacted with the same domain of starch phosphorylase where its original substrate G-1-P binds. Valine (VAL-178) was the key amino acid residue
found to be involved in the interaction of starch phosphorylase with its substrate as well as with glyphosate (Figure 6). These data are much important outcome of the study suggesting the blockade of starch phosphorylase activity with the binding of herbicide on the enzyme active site/cavity. It will be imperative to mention here that on one hand the structure of starch phosphorylase is highly conserved among various plant species and on the other hand the binding of herbicide with its active site may led to the agronomic losses in substantial proportions to the crops. Herbicides toxicities are well reported for animals, while also explored in microbes and plants [11, 12]. However, to the best of our knowledge, till date, no study has been conducted on the effects of herbicides on starch phosphorylase from plants.

Similarly, studies for other herbicide quizalofop also formed 3 hydrogen bonds with the enzyme. It was also bound almost near to the binding site of G-1-P and glyphosate with ARG-162, THR-191 and ARG38 amino acid residues involved in the hydrogen bonding. Another herbicide sethoxydim (CID 52923) binding pattern was also quite similar to the pattern of quizalofop, as this interaction also formed three hydrogen bonds with the involvement of the ARG38, TYR166 and ARG171 amino acid residues.

Comparative study of these three herbicides suggested that sethoxydim and quizalofop do not bind at exactly on the active site/ cavity, but still these are close to the binding site of either G-1-P or glyphosate (Figure 7 and 8). We speculated that these two above mentioned herbicides may be comparatively safer than glyphosate in terms of starch phosphorylase inhibition, but still cannot be concluded as completely safe. A summary of ligands, their structures and residues involved in hydrogen bond formation along with binding emery and ligand efficiency are given in Table 2.

Figure 6. Enzyme- glyphosate binding shown in surface view using PyMol Viewer.

Figure 7. Enzyme- quizalofop binding shown in surface view using PyMol Viewer.

Figure 8. Enzyme- sethoxydim binding shown in surface view using PyMol Viewer.
Table 2. Residues forming hydrogen bonds with starch phosphorylase for all the 4 ligands.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Structure</th>
<th>Hydrogen Binding residues</th>
<th>Binding Energy</th>
<th>Ligand Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose-1-phosphate (G-1-P) [CID: 65533]</td>
<td><img src="image" alt="Glucose-1-phosphate" /></td>
<td>CYS180, VAL178, LYS185</td>
<td>-0.45</td>
<td>-0.03</td>
</tr>
<tr>
<td>Glyphosate (CID: 3496)</td>
<td><img src="image" alt="Glyphosate" /></td>
<td>VAL178</td>
<td>-1.05</td>
<td>-0.11</td>
</tr>
<tr>
<td>Sethoxydim (CID: 52923)</td>
<td><img src="image" alt="Sethoxydim" /></td>
<td>TYR166, ARG171, ARG38</td>
<td>-4.42</td>
<td>-0.21</td>
</tr>
<tr>
<td>Quizalofop (CID: 53518)</td>
<td><img src="image" alt="Quizalofop" /></td>
<td>ARG38, ARG162, TYR191</td>
<td>-5.09</td>
<td>-0.21</td>
</tr>
</tbody>
</table>

For validation, docking analysis was also carried out using LeadIT tool (Evaluation version) and thereafter, enzyme-ligand binding was analyzed using Hyde assessment tool [16]. In Hyde result analysis, Red colored atoms indicated an unfavorable positive score contribution, whereas green colored atoms indicated a favorable negative score contribution. Therefore, only favorable negative score contribution was considered. Only the poses giving negative (\(\Delta G\)) binding affinity and LE ligand efficiency > 15 were considered. During Hyde analysis, glucose-1-phosphate formed three hydrogen bonds with Phe181, Cys180 and Lys185, respectively having binding affinity -23 KJ/mol and Ligand efficiency of 0.36. The residues involved in this pose were similar to that of Autodock results except Phe181, however, the cavity was same and Hyde assessment further supported the ligand binding efficiency. Similarly glyphosate, sethoxydim and quizalofop showed the binding with the same residues as that of Autodock results. (Table 3)
Table 3. Docking poses as obtained using LeadIT software, binding affinity (ΔG) and ligand efficiency as obtained using Hyde.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Posed Structure</th>
<th>Binding affinity (ΔG)</th>
<th>Ligand efficiency (LE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose-1-phosphate (G-1-P) [CID: 65533]</td>
<td><img src="#" alt="Docking Pose" /></td>
<td>-23 KJ/mol</td>
<td>0.36</td>
</tr>
<tr>
<td>Glyphosate (CID: 3496)</td>
<td><img src="#" alt="Docking Pose" /></td>
<td>-9 KJ/Mol</td>
<td>0.22</td>
</tr>
<tr>
<td>Sethoxydim (CID: 52923)</td>
<td><img src="#" alt="Docking Pose" /></td>
<td>-14 KJ/Mol</td>
<td>0.30</td>
</tr>
</tbody>
</table>
4. Conclusions

Outcome of the present studies suggested that starch phosphorylase originated from different plant species are almost 90% similar in amino acid sequence and 3-D structures. For the first time, our study revealed binding of three different herbicides viz. glyphosate, sethoxydim and quizalofop on the active site/cavity or close to the active site/cavity of starch phosphorylase. It suggested inhibition of this enzyme and also provided a prototype model for the herbicide mediated potential toxic effects on plants through this novel mechanism. These three herbicides use starch phosphorylase as a common target without any differentiation among various plants whether they are commercially important ones or weeds. We speculated that unnecessary or excess use of herbicides may adversely affect the plants through this mechanism in generalized manner.

Acknowledgements

The authors acknowledge the facilities of the Department of Biotechnology, Ministry of Science and Technology, Government of India, New Delhi (DBT) under the M.Sc. Biotechnology program and Bioinformatics sub centre used in the present work. The authors also acknowledge UGC, New Delhi for partial financial support.

REFERENCES


