Sequence analysis and homology modeling of peroxidase from *Medicago sativa*

Vinita Hooda[1]*, Prasada babu Gundala[2] & Paramageetham Chinthala[3]

[1]Department of Botany, Maharshi Dayanand University, Rohtak 124001, India; [2]Department of Botany, Sri Venkateswara University,Tirupati-517502, India; [3]Department of Microbiology, Sri Venkateshwara University, Tirupati-517502, India; Vinita Hooda – Email: vinitahooda@yahoo.co.in; Phone: +91 9896795000; *Corresponding author

Abstract:
Plant peroxidases are one of the most extensively studied group of enzymes which find applications in the environment, health, pharmaceutical, chemical and biotechnological processes. Class III secretary peroxidase from alfalfa (*Medicago sativa*) has been characterized using bioinformatics approach Physiochemical properties and topology of alfalfa peroxidase were compared with that of soybean and horseradish peroxidase, two most popular commercially available peroxidase preparations. Lower value of instability index as predicted by ProtParam and presence of extra disulphide linkages as predicted by Cys_REC suggested alfalfa peroxidase to be more stable than either of the commercial preparations. Multiple Sequence Alignment (MSA) with other functionally similar proteins revealed the presence of highly conserved catalytic residues. Three dimensional model of alfalfa peroxidase was constructed based on the crystal structure of soybean peroxidase (PDB Id: 1FHF A) by homology modelling approach. The model was checked for stereo chemical quality by PROCHECH, VERIFY 3D, WHAT IF, ERRAT, 3D MATCH AND ProSA servers. The best model was selected, energy minimized and used to analyze structure function relationship with substrate hydrogen peroxide by Autodock 4.0. The enzyme substrate complex was viewed with Swiss PDB viewer and one residue ASP43 was found to stabilize the interaction by hydrogen bonds. The results of the study may be a guiding point for further investigations on alfalfa peroxidase.

Keywords: Homology modeling, Peroxidase, alfalfa, Docking, Hydrogen peroxide

Background:
Peroxidases (PODs, EC 1.11.1.7) are well known haem containing proteins that use hydrogen peroxide (H_2O_2) as the electron acceptor to catalyse a number of oxidative reactions [1-3]. PODs from bacteria, fungi, plants and animals constitute a superfamily consisting of three major classes [4]. Class I PODs include intracellular enzymes in plant, bacteria and yeast. Class II PODs consist of secretary fungal enzymes and Class III are typical secretary plant PODs, which have multiple tissue specific functions: e.g., removal of H_2O_2 from chloroplasts and cytosol, oxidation of toxic compounds, biosynthesis of cell wall, defence responses towards wounding, indole 3 acetic acid catabolism, ethylene biosynthesis etc. [5]. Class III PODs are glycoprotein's which belong to a large multigene family

containing four conserved disulphide bridges and two calcium ions [6]. Besides being vital for plants, the enzyme also has serious commercial applications in medicine as a component of diagnostic tools and in the bioremediation and bio bleaching industries [7, 8]. Horseradish POD and soybean POD are two important commercially exploited sources of the enzyme because of their higher stability [9]. A couple of preliminary experiments have already established the usefulness of alfalfa POD for biotechnological applications. The enzyme immobilized in self assembled monolayers has been successfully used for determination of thiodicarb [10] and the plant has also been studied for analysis of POD activity under abiotic stress [11]. Detailed study of alfalfa POD still needs to be done in order to raise the potential of its use for various

BIOINFORMATION

academic and industrial applications. Hence, the present paper enlists some of the physiochemical and functional properties of alfalfa POD and provides insight into its three dimensional structure. Docking studies were performed to predict the ligand binding site within the protein.

Methodology:
The study was conducted using Intel(R) Core(TM) i5-2400 CPU @ 3.10 GHz, 3101 MHz, 4 Core(s) processor and 64-bit Operating System.

Retrieval of target sequence
The 353 amino acid long sequence of POD from alfalfa (*Medicago sativa*) was obtained from the Protein sequence database of NCBI (GenBank Id: AAB41811) and blasted [12] against Protein Data Bank (PDB) entries to find similar sequences. Parameter values for BLAST 2.2.26 search were set as default.

Characterization of target sequence
Transmembrane helices in the query sequence were predicted by TMHMM 2.0, ABTMpro and TargetP 1.1 server. Subcellular localization of protein using amino acid composition was achieved by MultiLoc [13]. NetNGlyc 1.0 Server predicted N-Glycosylation sites using artificial neural networks. Physiochemical properties such as molecular weight, theoretical pI, total number of negatively (Asp+Glu) and positively (Arg+Lys) charged residues, extinction coefficients [14], instability index [15], aliphatic index [16] and grand average of hydropathicity (GRAVY) [17] of the mature protein were computed using Expasy's ProtParam Proteomics server.

Protein Topology prediction
Secondary structures within alfalfa POD were calculated with SOPMA (Self Optimized Prediction Method with Alignment) [18] and disulfide bonds were predicted by the Cys_REC tool from softberry.

Multiple Sequence Alignment
In order to know the key residues responsible for catalytic activity of the enzyme, amino acid sequence of alfalfa POD was compared with amino acid sequences from other PODs of known crystallographic structures. Pairwise sequence alignment server Clustal-W [19] was employed to compare POD form alfalfa, soybean (PDB Id: 1FHF), peanut (PDB Id: 1SCH) and horseradish (PDB Id: 1ATJ).

Homology modeling
High resolution (2.80A°) crystal structure of soybean POD (PDB Id: 1FHF) was selected as template to create the three dimensional model for POD from alfalfa. The homology modeling program, Modeller 9v9 [20] was used to generate a total of 10 models of target protein.

Model refinement and evaluation
The models constructed were solvated and subjected to energy minimization using the steepest descent and conjugate gradient technique to eliminate bad contacts between protein atoms and structural water molecules. Computations were carried out *in vacuo* with the GROMOS 96 43B1 parameters set, implementation of Swiss-PDB Viewer. The stereo chemical quality and accuracy of the predicted models was evaluated

with PROCHECK [21] by Ramachandran plot analysis [22]. The best model was selected on the basis of overall G-factor, number of residues in core, allowed, generously allowed and disallowed regions. The selected model was further put to analysis by VERIFY 3D [23], WHAT IF [24], ERRAT [25] and 3D Match program from softberry. ProSA [26] was used for the display of Z-scores and energy plots. Finally the protein was visualized with Swiss-PDB viewer [27].

Docking Studies
The structure of substrate H_2O_2 was obtained from NCBI PubChem in psf format and converted to pdb format using OpenBabel [28]. To understand the molecular interactions between H_2O_2 and the protein receptor, flexible small molecule-rigid protein docking experiments were performed using Autodock 4.0 [29]. In the protein non-polar hydrogen atoms were merged and total gasteiger charge of -3.9953 was added to the protein. It was made sure that there are no non-bonded atoms in the protein. Gasteiger partial charges were also assigned to the ligand and all torsions were allowed to rotate during docking. The grid box was centred at the modeled structure and affinity maps were calculated by AutoGrid. Fifty Lamarckian Genetic Algorithm runs with default parameter settings were performed and the docked confirmations were played ranked by energy. The lowest energy complex of ligand and protein was further viewed using Swiss PDB viewer.
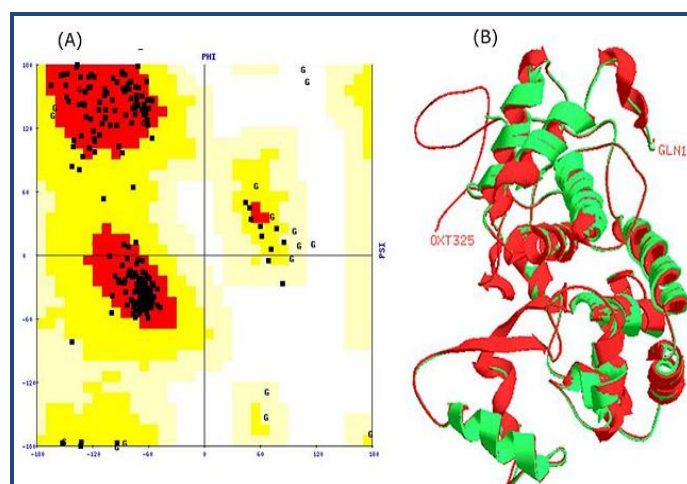

Figure 1: (A) Ramachndran plot analysis for alfalfa POD. The plot statistics are: Total number of residues-325 with 92.1% residues in the core region (red); 6.6 % residues in allowed (yellow) and 1.4% in generously allowed (light yellow) region. Number of glycine residues (labelled as G) is 21 and Number of Pro residues is 16 **(B)** Superimposition of CA atoms of alfalfa POD (red) with soybean POD (green) as done with Deep View Swiss PDB viewer gave a RMS value of 0.26 A°, arrows within helices and sheets points towards the C-terminus. Gln1 is the first amino acid at N-terminus whereas Oxt 325 represents the terminal oxygen atom at C-terminus.

Results and Discussion:
The 353 amino acid long sequence of alfalfa POD retrieved from Protein: sequence database of NCBI (GenBank ID: AAB41811) had a putative signal peptide of 28 amino acids at N-terminus and a mature peptide of 325 amino acid residues. The sequence of mature protein was blasted against the PDB database for proteins with similar sequence and known three dimensional

structure using compositionally adjusted substitution matrices. A 304 amino acid long sequence of POD from soybean (PDB Id: 1FHF) had 69% residues identical with the target protein with 0% gaps. Structure of soybean POD has been solved at a resolution of 2.80 A°, hence it was chosen as template to build the 3D structure for target protein.

Characterization of POD
TargetP 1.1 designated alfalfa POD to be a protein of secretary pathway and TMHMM server 2.0 predicted the presence of one transmembrane domain within the signal peptide which may be required to direct the protein to secretary pathway. ABTMpro server also indicated that alfalfa POD is an alpha helical transmembrane protein. Multiloc predicted the protein to be extracellular in nature. Eight asparagines residues were found to be N-glycosylated by NetNGlyc server. Glycosylation of asparagines residues is required for correct folding of protein before being exported.

A comparison of physiochemical properties of POD from three different sources viz. horseradish, soybean and alfalfa computed using ProtParam tools is presented in **Table 1 (see supplementary material)**. Horseradish and soybean are two of the most popular sources for isolation of POD and hence were included in the study for comparison purpose. Data shown in **Table 1** suggested that alfalfa POD is superior over POD from both the sources. The computed isoelectric point (pI value) revealed alfalfa POD to be more acidic than horseradish POD but less acidic than soybean POD. The computed isoelectric point will be useful for separating the protein on a polyacrylamide gel by isoelectric focusing. The extinction coefficient of a protein as calculated by the program depends on the molar extinction coefficient of Tyr, Trp and Cys residues. The extinction coefficient can be used to calculate the concentration of a protein in solution. Stability of the three PODs was studied by analyzing the values for instability index, aliphatic index and Grand average of hydropathicity (GRAVY) index. Instability index relies upon the occurrence of certain dipeptides along the length of the protein to distinguish between the unstable and stable protein [16]. The value of instability index was least for alfalfa POD (42.94, 30.18 and 23.17 for horseradish, soybean and alfalfa POD respectively); hence it could be safely predicted as most stable of the three proteins. The aliphatic index refers to the relative volume of a protein that is occupied by aliphatic side chains and contributes to the increased thermo stability of protein. Aliphatic index of POD from all the three sources was comparable. GRAVY index indicates the solubility of proteins: a negative GRAVY value for POD from all the compared sources designates it to be hydrophilic in nature.

Topology of alfalfa POD
A comparison of secondary structure elements as calculated by SOPMA showed that random coil occupied the largest part of the protein followed by alpha helix, extended strand and beta turns for all three PODs **Table 2 (see supplementary material)**. Disulfide bridges play an important role in the folding and stability of proteins. CYS_REC version 2 revealed the presence of 10 disulfide linkages (at positions: 13, 14, 39, 72, 77, 119, 125, 204, 236 and 237) in alfalfa POD while the horseradish and soybean POD had eight disulphide linkages each (at positions: 12, 45, 50, 92, 98, 178, 210, 302 and 11, 44, 49, 91, 97, 176, 208, 299

respectively). The presence of extra disulphide bonds might contribute towards higher stability of POD from alfalfa.

ClustalW analysis
The identification of catalytic residues is key to understanding the function of enzymes. The information from other functionally similar sequences with known crystallographic structures was exploited to identify the key catalytic residues. ClustalW alignment of alfalfa POD with soybean (PDB Id: 1FHF), peanut (PDB Id: 1SCH) and horseradish (PDB Id: 1ATJ) PODs as shown in **Table 3 (see supplementary material)** depicted a high degree of conservation among the sequences [30, 31]. Moreover POD in higher plants is a part of large multigene family where a number of isozymes are formed as a result of posttranscriptional and posttranslational changes. Hence, the compared sequences varied in length but essentially conserved the key catalytic residues which have been highlighted with an asterisk (*) symbol.

Model building, refinement and evaluation
Crystal structure of soybean POD was used as template to generate three dimensional coordinates for alfalfa POD. Ten models generated by Modeller 9v8 were viewed with Swiss PDB viewer and energy minimized. Ramachandran plot analysis of the ten models was obtained by PROCHECK server. The best model in terms of stereo chemical quality showed a overall G-factor value of -0.04 which indicates that geometry of the model corresponds to high probability confirmation with 92.1% residues in the core region of Ramachandran plot **(Figure 1A)**. The number of residues in allowed and generously allowed regions were 6.6% and 1.4% respectively. It is generally accepted that if 90% residues are in the allowed region, the quality of the model is evaluated as good and reliable. None of the residues was present in the disallowed region of the plot. Superimposition of the model with the template (used as reference layer) with Swiss PDB viewer as shown in **(Figure 1B)** showed a very low RMSD of 0.26 A°, suggesting high similarity between them. 3D Match program was also used for alignment of protein 3D structures. The RMSD score of 0.256A° with a Z score of 7.930 indicated that both template and target proteins have similar folds.

Verify 3D analysis revealed that 84.97 % of the residues had an average 3D-1D score of <0.2, predicting that the model is compatible with its sequence. The amino acid environment was evaluated using ERRAT plots, which assess the distribution of different types of atoms with respect to one another in the protein model and is used for making decision about its reliability. ERRAT showed an overall quality factor of 85.050, a result expected for crystallographic models with resolution >2.5A°. B-factor analysis is done with WHAT IF server reflected the mobility or flexibility of various parts of the molecule. Averaged B-factor deviation for protein backbone was 30.594 and averaged standard deviation was 38.491. Since average deviation value was less than standard deviation, so it reflected a good quality model.

Z score Analysis
RMS Z-scores for anomalous bond lengths and bond angles as determined by WHAT IF 0.895 and 1.209 respectively, which is very close to 1.0 suggesting high model quality. ProSA was used to check three dimensional model of alfalfa POD for

potential errors. The program displays two characteristics of the input structure: its Z-score and a plot of its residue energies. The Z-score indicates overall model quality and measures the deviation of the total energy of the structure with respect to an energy distribution derived from random conformations. As shown in **(Figure 2A)** the Z-score for alfalfa POD is -8.26 and for soybean POD is -7.04. The scores are well within the range of scores typically found for proteins of similar size indicating a

highly reliable structure. The energy plot shows the local model quality by plotting energies as a function of amino acid sequence position. In general, positive values correspond to problematic or erroneous part of a model. **Figure 2B** displays a comparable energy plot for both the target and template structures.
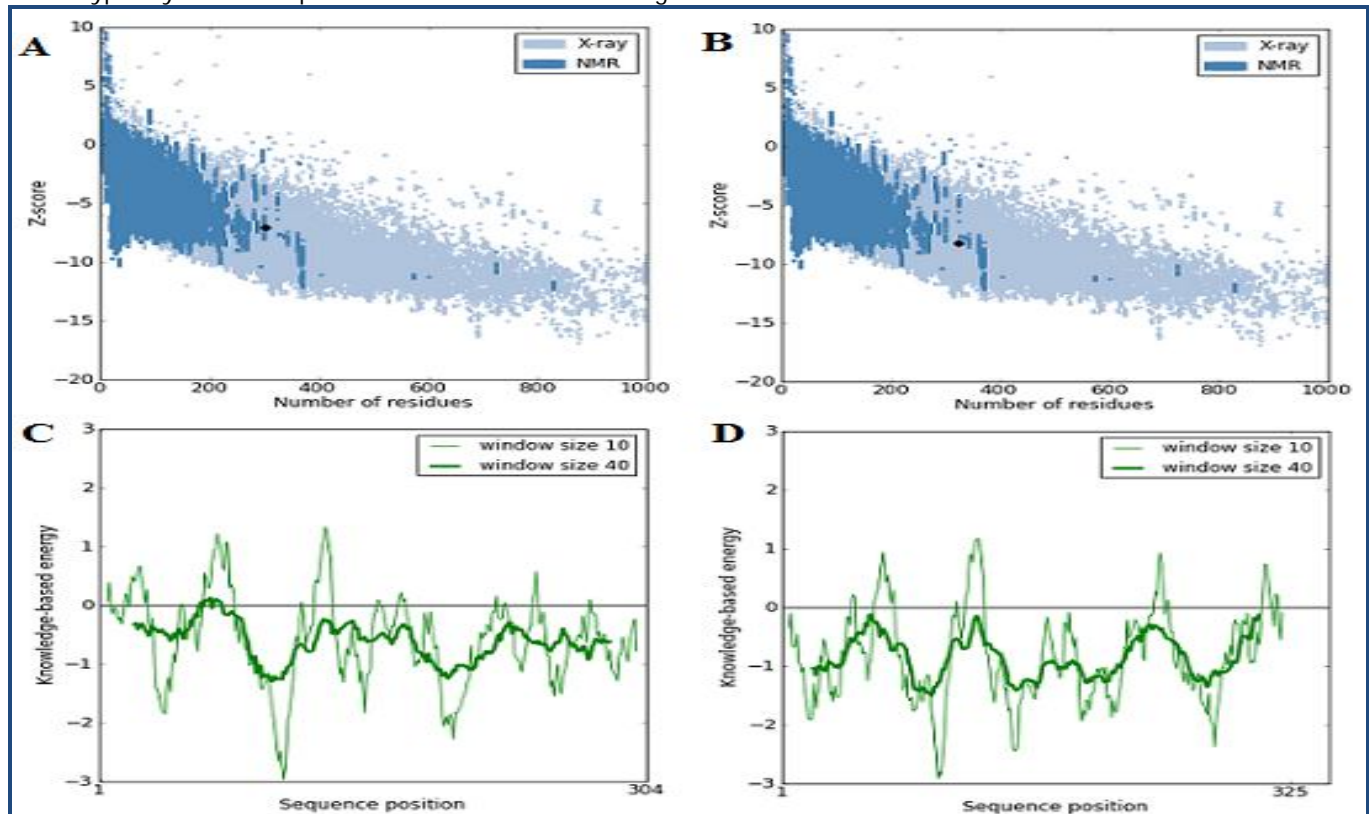

Figure 2: ProSA-web service analysis of alfalfa and soybean peroxidase. ProSA-web z-scores of all protein chains in PDB determined by X-ray crystallography (light blue) or NMR spectroscopy (dark blue) with respect to their length. The z-scores of alfalfa **(A)** and soybean **(B)** peroxidase are highlighted as large dots **(C)** Energy plot of alfalfa peroxidase **(D)** Energy plot of soybean peroxidase.

Docking Studies

Autodock was used to generate fifty different conformations of the ligand H_2O_2. The confirmations were played ranked by their energy and the confirmation having lowest binding energy (-1.88 kcal/mol) was chosen for further analysis of protein-ligand complex. Docking site interaction was stabilized by hydrogen bonds between ASP43 OD2 and two hydrogen atoms of the ligand **(Figure 3A)**. Residues His40 and Cys49 were found within a distance of 3 A° from the ligand. When the distance was increased to 4A°, residues Leu101, Ala98, Asp50 and Cys44 were found to interact with the ligand. In the template protein both ASP43 and ASP50 are involved in binding of calcium ions, which in turn may help in binding the H_2O_2 molecule. The key ligand binding residues (Arg38, Phe41 and His42) as found in other functionally similar proteins (PDB Id: 1SCH & 1ATJ) were found within a distance of 5A° from H_2O_2, which might be due to rigid biomolecule docking. Complex of alfalfa POD with H_2O_2 as viewed by Deep view Swiss PDB viewer is shown in **(Figure 3B)**. Further analysis revealed the presence of fourteen alpha helices in the modelled protein. Surface view of the modeled structure **(Figure 3C)** showed that ligand binding site was buried inside a channel.
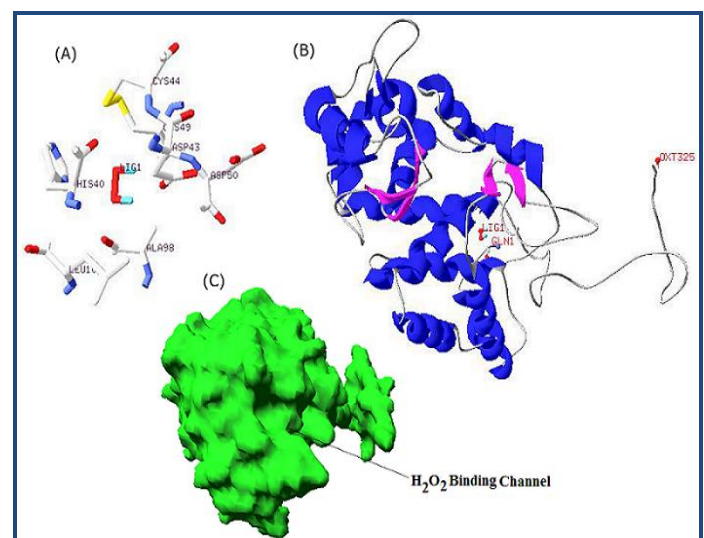

Figure 3: (A) Residues found within a distance of 4A° from the ligand. The bonding between H_2O_2 and protein was stabilized by hydrogen bond between the ligand and ASP43 **(B)** Predicted three dimensional structure of alfalfa POD. Secondary structure

BIOINFORMATION

elements of proteins are presented as ribbon. Sitting inside is a ball and stick model of ligand H_2O_2 labeled as Lig1. Gln1 is the first amino acid at N-terminus whereas Oxt 325 represents the terminal oxygen atom at C-terminus. **(C)** Molecular surfaces showing channel for substrate binding.

Conclusion:

The present study analyses the sequence of alfalfa POD *in silico* with respect to presence of transmembrane helices, glycosylation sites, subcellular localization, physiochemical properties, conserved regions and topology. Instability index value and presence of ten disulphide linkages suggest the superiority of alfalfa POD over other popular PODs from horseradish and soybean. Three dimensional structure of alfalfa POD was predicted using soybean POD as template (PDB Id: 1FHFA) with a RMSD value of 0.26 A°. Structure validation by PROCHECK, Verify 3D, WHAT IF, ERRAT, 3D Match and ProSA confirmed the reliability of the model. Docking of H_2O_2 in the vicinity of predicted key catalytic residues further suggested that model is of high quality. The predicted properties of alfalfa POD may help in better utilization of enzyme for various biotechnological applications.

Acknowledgement:

References:

[1] Li H & Poulos TL, *Structure.* 1994 **2**: 461 [PMID: 7922023]

[2] González-Rábade N *et al. Appl Biochem Biotechnol.* 2012 **166**: 1644 [PMID: 22350935]

[3] Bhatt I & Tripathi BN, *Biotechnol Adv.* 2011 **29**: 850 [PMID: 21777667]

[4] Dietz KJ, *Antioxid Redox Signal.* 2011 **15**: 1129 [PMID: 21194355]

[5] Poulos TL, *Arch Biochem Biophys.* 2010 **1:** 3 [PMID: 20206121]

[6] Lüthje S, *Phytochemistry.* 2011 **72**: 1124 [PMID: 21211808]

[7] Alvarado B & Torres E, *Recent Pat Biotechnol.* 2009 **3**: 88 [PMID: 19519565]

[8] Colonna S *et al. Trends Biotechnol.* 1999 **17**: 163 [PMID: 10203775]

[9] Bódalo A *et al. Bioprocess Biosyst Eng.* 2008 **31**: 587 [PMID: 18270748]

[10] Moccelini SK *et al. Talanta.* 2010 **30**: 164 [PMID: 20685452]

[11] Wang WB *et al. Plant Physiol Biochem.* 2009 **47**: 570 [PMID: 19318268]

[12] Altschul SF *et al. J Mol Biol.* 1990 **215**: 403 [PMID: 2231712]

[13] Hoeglund A *et al. Bioinformatics.* 2006 **22**: 1158 [PMID: 16428265]

[14] Gill SC & Von Hippel PH, *Anal Biochem.* 1989 **182**: 319 [PMID: 2610349]

[15] Guruprasad K *et al. Protein Eng.* 1990 **4**: 155 [PMID: 2075190]

[16] Ikai A, *J Biochem.* 1980 **88**: 1895 [PMID: 7462208]

[17] Kyte J & Doolittle RF, *J Mol Biol.* 1982 **157**: 105 [PMID: 7108955]

[18] Geourjon C & Deléage G, *Comput Appl Biosci.* 1995 **11:** 681 [PMID: 8808585]

[19] Thompson JD *et al. Nucleic Acids Res.* 1994 **11:** 4673 [PMID: 7984417]

[20] Sali A & Blundell TL, *J Mol Biol.* 1993 **234**: 779 [PMID: 8254673]

[21] Laskowski RA *et al. J Biomol NMR.* 1996 **8:** 477 [PMID: 9008363]

[22] Ramachandran GN *et al. J Mol Biol.* 1963 **7:** 95 [PMID: 13990617]

[23] Eisenberg DR *et al. Methods Enzymol.* 1997 **277**: 396 [PMID: 9379925]

[24] Vriend GJ, *Mol Graph.* 1990 **8**: 52 [PMID: 2268628]

[25] Colovos C & Yeates TO, *Protein Sci.* 1993 **2**: 1511 [PMID: 8401235]

[26] Wiederstein M & Sippl MJ, *Nucleic Acids Res.* 2007 **35**: 407 [PMID: 17517781]

[27] Guex N & Peitsch MC, *Electrophoresis.* 1997 **18**: 2714 [PMID: 9504803]

[28] O'Boyle NM *et al. J Cheminform.* 2011 **3**: 33 [PMID: 21982300]

[29] Rosenfeld RJ, *Comput Aided Mol Des.* 2003 **17**: 525 [PMID: 14703123]

[30] Schuller DJ *et al. Structure.* 1996 **4:** 311 [PMID: 8805539]

[31] Gajhede M *et al. Nat Struct Biol.* 1997 **4:** 1032 [PMID: 9406554]

BIOINFORMATION

Supplementary material:

Table 1: Comparison of properties of POD from three different sources as predicted by ProtParam program

S.No.	Parameters	Peroxidase		
		Armoracia rusticana	*Glycine max*	*Medicago sativa*
1.	Source botanical name			
2.	Source common name	Horseradish	Soybean	Alfalfa
3.	GenBank Accession No.	CAA00083.1	13399943	AAB41811
4.	PDB Id	1HCH	1FHF	-
5.	Sequence length	309 a.a.	304 a.a.	325 a.a.
6.	Molecular weight (nonglycosylated forms)	34049.4	33117.1	35389.6
7.	Theoretical pI	6.30	4.61	5.76
8.	-R*	28	29	32
9.	+R*	27	18	28
10.	Extinction coefficients** ($M^{-1}cm^{-1}$ at 260 nm)	13450-12950	13450-12950	14940-14440
11.	Instability index	42.94	30.18	23.17
12.	Aliphatic index	83.98	87.57	84.62
13.	GRAVY	-0.177	-0.098	-0.226

* –R: total number of negative residues. +R: total number of positive residues.
** First value is based on the assumption both cysteine residues form cystines and the second assumes that both cysteine residues are reduced.

Table 2: Secondary structure elements as predicted by SOPMA for PODs

Structural element	Horseradish POD	Soybean POD	Alfalfa POD
Alpha helix	35.60	36.84	33.23
Extended strand	15.21	16.78	17.85
Beta turn	6.80	5.59	7.08
Random coil	42.39	40.79	41.85

Table 3: CLUSTAL 2.1 multiple sequence alignment of the POD from alfalfa (*Medicago sativa*) with soybean (PDB Id: 1FHF), horseradish (PDB Id: 1ATJ) and peanut (PDB Id: 1SCH) PODs

```
Medicago    QLDNSFYRDTCPNVHSIVREVLRNVSKTDPRILASLMRVHFHDCFVQGCDASILLNTTST  60
1FHF│C      QLTPTFYRETCPNLFPIVFGVIFDASFTDPRIGASLMRLHFHDCFVQGCDGSVLLNNTDT  60
1ATJ│F      QLTPTFYDNSCPNVSNIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTTS  60
1SCH│A      XLSSNFYATKCPNALSTIKSAVNSAVAKEARMGASLLRLHFHDCFVQGCDASVLLDDTSN  60
            *   .** .***    :    .:.  .   .:.*: **::*:*********:***.*:**: * .

Medicago    ITSEQTAFGNNNSIRGLDVVNQIKTAVENACPNTVSCADILALAAEISSVLAHGPDWKVP  120
1FHF│C      IESEQDALPNINSIRGLDVVNDIKTAVENSCPDTVSCADILAIAAEIASVLGGGPGWPVP  120
1ATJ│F      FRTEKDAFGNANSARGFPVIDRMKAAVESACPRTVSCADLLTIAAQQSVTLAGGPSWRVP  120
1SCH│A      FTGEKTAGPNANSIRGFEVIDTIKSQVESLCPGVVSCADILAVAARDSVVALGGASWNVL  120
            :   *: *   * ** **: *:: :*: **. ** .*****:*::**. : .   *..* *

Medicago    LGRRDSLTANLTLANENLPSPAFNLSELKKNFDRQGLD-TTDLVALSGAHTIGRGQCRFF  179
1FHF│C      LGRRDSLTANRTLANQNLPAPFFNLTQLKASFAVQGLN-TLDLVTLSGGHTFGRARCSTF  179
1ATJ│F      LGRRDSLQAFLDLANANLPAPFFTLPQLKDSFRNVGLNRSSDLVALSGGHTFGKNQCRFI  180
1SCH│A      LGRRDSTTASLSSANSDLPAPFFNLSGLISAFSNKGFT-TKELVTLSGAHTIGQAQCTAF  179
            ******   *     ** :**:* *.*. *    *    *:   : :**:***.**:*: :* :

Medicago    VDRLYNFSNTGNPDPTLNTTYLQTLRTICPNGGPGSTLTDLDPTTPDTFDSAYYSNLRIQ  239
1FHF│C      INRLYNFSNTGNPDPTLNTTYLEVLRARCPQNATGDNLTNLDLSTPDQFDNRYYSNLLQL  239
1ATJ│F      MDRLYNFSNTGLPDPTLNTTYLQTLRGLCPLNGNLSALVDFDLRTPTIFDNKYYVNLEEQ  240
1SCH│A      RTRIYNESN-------IDPTYAKSLQANCPSVGGDTNLSPFDVTTPNKFDNAYYINLRNK  232
            *:** **        ::.** : *:  **   .    *   :*  **  **. ** ** **

Medicago    KGLFESDQVLASTS-GADTIAIVNSFNNNQTLFFEAFKASMIKMSKIKVLTGSQGEIRKQ  298
1FHF│C      NGLLQSDQELFSTP-GADTIPIVNSFSSNQNTFFSNFRVSMIKMGNIGVLTGDEGEIRLQ  298
1ATJ│F      KGLIQSDQELFSSPNATDTIPLVRSFANSTQTFFNAFVEAMDRMGNITPLTGTQGQIRLN  300
1SCH│A      KGLLHSDQQLFN---GVSTDSQVTAYSNNAATFNTDFGNAMIKMGNLSPLTGTSGQIRTN  289
            :**:.*** *  .  ...* . * :: .. .  *   *   :* :*.::   *** .*:** :

Medicago    CNFVNGNSGLATKVIRESSEDGIVSSY  325
1FHF│C      CNFVNG--------------------  304
1ATJ│F      CRVVNS--------------------  306
1SCH│A      CRKTN---------------------  294
            *.  .*
```