Potato J (2017) 44 (1): 28-36

OSMOTIN AN ANTIFUNGAL CYTOTOXIC FROM *SOLANUM TUBEROSUM*: NEW INSIGHTS INTO STRUCTURAL AND ANTIFUNGAL PROPERTIES

Ritu Singh[1], Jagesh K Tiwari[1], Shashi Rawat[1] and SK Chakrabarti[1]

ABSTRACT: Pathogenesis related (PR) proteins are considered as the major sources in plants' defence response against pathogens as well as other biotic and abiotic stresses. We present here the 3-D structure of potato (*Solanum tuberosum* L.) osmotin-like protein (StOLP) based on the homology modelling using structure validation programs namely PROCHECK, PROSA, VERIFY3D and WHATIF. The model revealed that StOLP protein shared 16 conserved Cys, which formed eight disulfide bonds. The 3-D structure of StOLP showed three characteristic domains showing similarity fold with the thaumatin. Domain 1 consisted of a flattened 11-β strands (six antiparallel β strands present in the front end and other five antiparallel β strands are located at the back) which together constituted the β-sandwich. Domain II consisted of an α-helix and three shorter domains which formed loops exposed on the molecular surface of the protein. Together with domain I, domain II formed a large cleft region on the protein surface. Domain III consisted of a small loop of 32 residues which formed a long loop consisting of two β sheets. This modelling effort provided structural predictions of the catalytic domain of StOLP that could be verified via experimental techniques. Findings predicted the evidence of antifungal activity for StOLP against the oomycete *Phytophthora infestans* causes late blight disease in potato.

KEYWORDS: Antifungal alignment, homology modelling, osmotin-like proteins, procheck validation, 3D structure,

INTRODUCTION

Plants are sedentary in nature and unlike animals cannot physically escape from onslaught of biotic and abiotic stresses. Moreover, plants do not possess vascular defence system like warm blooded animals to resist pathogen attack. However, plants evolved intricate passive as well as active defence mechanisms to ward off or resist attacks from myriad of pathogens and pests. Passive defences are pre-existing in the particular plant species and usually create a barrier to alien invasion. On the contrary, active defences are triggered by recognition of invader at cell surface, signal transduction, and launching of defence related protein expression.

Extensive genetic and molecular work during the last two decades revealed that active responses can be broadly grouped into two major layers: i) those triggered by pathogen/microbe-associated molecular patterns that are conserved between species of a microbial group (PAMP/MAMP-triggered immunity, or PTI/MTI), and ii) those triggered by isolate-specific pathogen effectors (effector-triggered immunity, or ETI; (Akira *et al.*, 2006). MTI is based on PRRs that serve as receptors for MAMPs, whereas ETI invokes intracellular NB-LRR proteins that detect either the actions or structures of pathogen effectors (Apel and Hirt, 2004). Successful recognition of invading microorganisms by plants is followed by oxidative burst, protein kinase activity, ion channel gating and defence associated gene expression including pathogenesis related proteins (PRs). PRs have been defined as plant proteins that are induced in pathological response (van Loon *et al.*, 1994). Seventeen families of PR proteins have been recognized (van Loon *et al.*, 2006),

[1]ICAR-Central Potato Research Institute, Shimla, Himachal Pradesh - 171 001, India
Email: jageshtiwari@gmail.com

which are stress inducible proteins induced by abiotic stresses, UV-radiation, osmotic shock, low temperature, water deficit as well as biotic stresses from various infectious pathogens, insects, nematodes and herbivores (Antoniw *et al.*, 1980; van Loon *et al.*, 1994). The PR proteins are evolutionarily conserved in the plant kingdom. Antifungal activity of PR-5 proteins has been demonstrated against a variety of fungi, including *P. infestans*, *Candida albicans, Neurospora crassa, Trichoderma reesei, Trichoderma viridae, Fusarium oxysporum, Cercospora beticola, Alternaria solani, Verticillum alboatrum* and *V. dahliae* (Velazhahan *et al.*, 1999; Woloshuk *et al.*, 1991; Vigers *et al.*, 1992).

Antifungal proteins and peptides have drawn the attention of a large number of researchers by virtue of their potential value in human medicine as well as for protecting economically important animals and crops from fungal attack. Osmotin is a 24 kDa protein belonging to the PR-5 family of pathogenesis related proteins and is often used as marker of the pathogen induced systemic acquired resistance (SAR) in plants (Singh *et al.*, 1987; Bol *et al.*, 1990). Osmotin show high sequence similarity to thaumatin, a sweet-tasting protein from the West African shrub *Thaumatococcus daniell. In vitro* antifungal activity of PR proteins has been demonstrated by bioassays against a variety of fungi, including *P. infestans*, *Candida albicans, Neurospora crassa* and *Trichoderma reesei* (Woloshuk *et al.*, 1991; Vigers *et al.*, 1992). The osmotin and osmotin-like proteins have *in vitro* antifungal activity toward *P. infestans* which causes late blight disease in potato and tomato (Woloshuk *et al.*, 1991; Liu *et al.*, 1994). Transgenic plants overexpressing PR-5 proteins showed resistance and delayed development of disease symptoms. *In vitro* assay showed that osmotin isolated from *Nicotiana tabacum* confers enhanced resistance to *P. infestans* in transgenic potato plants (Liu *et al.*, 1994). In tobacco and other plant species, antifungal activity against different pathogenic and non-pathogenic fungi by *in vitro* inhibiting hyphal growth, spore lysis, inhibition of spore germination reducing germling viability and dissipation of membrane potential (Roberts and Selitrennikoff, 1990; Vigers *et al.*, 1991). The identification of AP24 as osmotin, against *P. infestans* causes late blight in tomato prompted us to extend our studies to potato.

Potato is the third most important food crop in the world, after rice and wheat and is known to have antimicrobial, antioxidant and antifungal properties (Han *et al.*, 1996, Do *et al.*, 2004; Bontempo *et al.*, 2013). The exact mechanism of antifungal activity of osmotin protein is still unknown. The present study was performed with the aim to understand the 3D-structure of the potato osmotin by computational approach whose structure has not been reported yet. Further to check the reliability of the model and its validation, various tools and molecular dynamics method were used.

MATERIALS AND METHODS

Softwares

As a first step to guide selectivity studies we examined the molecular electrostatic potential and variations in the topology around the acidic cleft of PR-5 proteins. Then *Solanum tuberosum* osmotin-like protein (StOLP) was compared with other antifungal protein which can be further exploited for its antifungal activity. All computations and molecular modelling were carried out on a Intel® Xeon® Processor E5630 workstation with Windows 7 operating system using automated homology modelling program Modeller9.11 (http://salilab.org). Validations of 3D structure of modelled protein was further evaluated by PROCHECK (http://nihserver.mbi.ucla.edu/SAVES), PROSA (https://

prosa.services.came.sbg.ac.at/prosa. php) and Verify-3D (Eisenberg *et al.*, 1997). Interactive visualization and analysis of molecular structures was carried out in Pymol (http://pymol.sourceforge.net). Physical and chemical parameters for protein were calculated by Protoparam (http://www.expasy.ch/tools/protparam.html).

Sequence retrieval and analysis

The primary sequence of StOLP was obtained from the SWISS-PROT protein database (Accession No AAU95242.1, Q5XUH1). ProtParam was used to predict physiochemical properties. The parameters computed by ProtParam included the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathy (GRAVY).

3D modeling of StOLP

A NCBI-BLAST (http://ncbi.nlm.nih.gov/Blast) (Altschul *et al.*, 1990) and WU-Blast2 server {(Washington University Basic Local Alignment Search Tool) (http://www.ebi.ac. uk/blast2/)} were used to retrieve the corresponding template structure for modelling of StOLP against Protein Data Bank (PDB). The structurally conserved regions (SCR's) in template protein were identified and query was subsequently aligned to template structure in Modeller9v8. Subsequently, the alignment was manually checked for any mismatch. Further step consisted of performing multiple alignments of the selected templates and target sequence using the program ClustalW with the Blossum 62 substitution matrix with gap penalties of gap start 1 and gap extension 0.1. The model was constructed by using the program Modeller9v8. It uses an automated approach to comparative protein structure modelling. In brief, the modelling procedure begins with an alignment of the sequence to be modelled (target) with related known 3D structures (templates). The output is a 3D model for the target sequence containing all main chain and side chain non-hydrogen atoms. The program also employs probability density functions (PDFs) derived analytically using statistical mechanics and empirically using a database of known protein structures as the spatial restraints.

Model optimization, quality assessment and visualization

The stereochemistry of the models was further gauged by Ramachandran plot using PROCHECK program accessible at SAVES Server (http://nihserver.mbi.ucla.edu/SAVES). The PROCHECK program provides information about the stereo chemical quality of a given protein structure and was used to generate Ramachandran plot. All final models were inspected for the DOPE (discrete optimized potential energy) score of modeller output per residues of the model. DOPE score calculated by Modeller program is the distance dependent statistical potential based on probabilistic theory. The ProSA-Web-server (https://prosa. services.came.sbg.ac.at/prosa.php) was used to test the local and overall quality of the developed models from Modeller. The quality of the fold was inspected with PROSA that allowed all residues in negative energy regions very similar to the template protein, and indicates the possible correctness of the modelled structures. The overall quality of the structure was further determined by ERRAT, Verify 3D and WHAT_CHECK. Visualization of 3D structures, and superposition of query and template structure were performed in Pymol.

The electrostatic interaction is a crucial part of the non-covalent interaction energy between molecules. The electrostatic potential

on a molecular surface can be used to visually compare two molecules, guiding docking studies and identifying sites that interact with ligands because of opposite electrostatics. The following equation was used to calculate the MEP values.

$$EP(i) = \sum_{j=1}^{N} \frac{qi}{r\,ij}$$

Where EP (i) is the electrostatic potential at the surface point 'i' due to atom j having the partial charge 'qj' and separated by distance 'rij'. The electrostatic potential (EP) on the surface is generally colored according to the sign and magnitude of the potential. The color ramp for EP ranges from red (most positive) to blue (most negative). To understand more characteristic differences in active site of StOLP, MEP analysis was performed. Further the solvent accessibility of the amino acid residues in the modelled protein was determined by using ASA-view (http://gibk26.bse.kyutech.ac.jp/~shandar/netasa/asaview/) (Ahmad *et al.*, 2004).

RESULTS AND DISCUSSION

The present study focused on sequence, structural and functional analysis of StOLP as a model. ProtParam was used to analyze different physiochemical properties of the amino acid sequences. The 250 amino acids long StOLP was predicted to have a molecular weight of 27463.9 Daltons and an isoelectric point (pI) of 5.51 having more negatively charged protein and an instability index of 37.44, which classifies it as a stable protein. The negative GRAVY (Grand average of hydropathicity) index of -0.306 is indicative of a hydrophilic and soluble nature of the StOLP protein.

Homology modeling of StOLP

The first step toward constructing the homology model of StOLP was a template search that was carried out in NCBI-BLAST and WU-BLAST using Blossum 62 substitution matrix with e-value cut-off of 10. The program provided the most promising template structures for homology modelling. Two crystal structures with high percentage identity and resolution were selected as template for alignment, which included crystal structure of PR-5d *Nicotiana tabacum* (1AUN) and *Solanum lycopersicum* (2IOW) as templates. The selected templates (1AUN and 2IOW) were aligned pairwise (globally) to the query sequence StOLP using needle program in EMBOSS-GUI (http://bips.ustrasbg.fr/EMBOSS). The identity and similarity between the target and template protein 1AUN was 80.0% and 81.6%, respectively whereas template protein 2IOW showed 65.6% and 72.8%, respectively. A multiple sequence alignment view using multiple sequence algorithm ClustalW with the StOLP and the two selected templates is shown **Fig. 1**. The structurally conserved regions (SCR's) and structurally variable regions (SVR's) were defined and model was built using Modeller9v8. The program deduces the distance and angle constraints from a template structure and combines them with energy terms for adequate stereochemistry in an objective function that is later optimised in Cartesian space with conjugate gradients and molecular dynamic methods. Ten sets of models were generated. The fifth model has lowest DOPE score so this model was selected for further study.

Energy minimization, quality assessment and visualization

The quality of the model was further assessed using Ramachandran plot in PROCHECK validation package. In this model, 89.3% of the amino acid residues were in the favourable regions of the plot, whereas 10.7% of the residues were in the

additionally allowed region and no residue in generously allowed region and disallowed region. Altogether 100% of the residues were in favoured and allowed regions. The phi and psi distributions of the Ramachandran plot of non-Glycine, non-Proline residues and main chain parameters plots for the model are summarized in Ramachandran plot summary (**Fig. 2**). Most of the bond angles, bond length and torsion angles were between values expected for naturally folded protein. Further refinement of the final StOLP model was done by energy minimization of the selected outlier residues using protein report tool. The final and refined 3-D structure of the StOLP superimposes well on the crystal structure of the templates taken for model building. The alpha-trace of StOLP generated was found to be similar to resolved crystal structure of the templates 1AUN and 2IOW in the same orientation (**Fig. 3**). Further, evaluated for recognition of error of the StOLP model was done with ProSA-Web-server. ProSA uses knowledge-based potentials of mean force to evaluate model accuracy. ProSA-Web-server analysis revealed that the modelled structure occupies a region of NMR predicted native

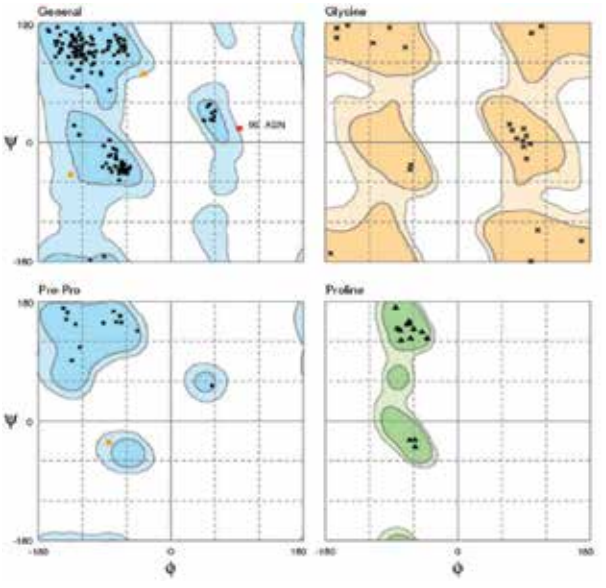

Fig. 2. Ramachandran plot analysis showing placement of residues in the deduced model (97.5% in favourable region). The structure orientation residues are separately considered for angle and torsions. The figure was generated using RAMPAGE web server.

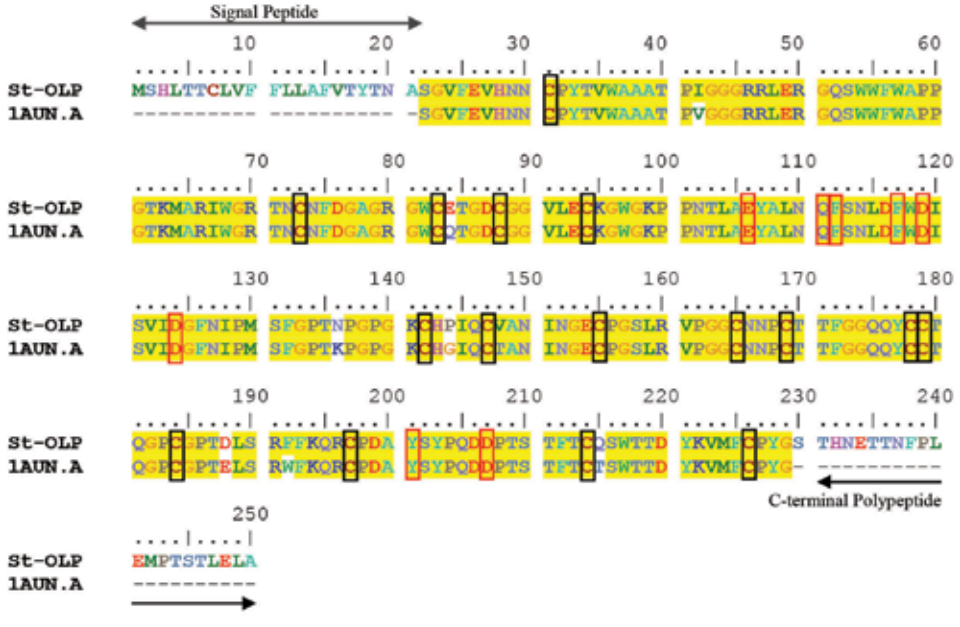

Fig. 1. Alignment of the amino acid sequences of StOLP with 1AUN generated using ClustalW. Conserved Cys are indicated with a black color boxes and conserved acidic amino acids in the acidic cleft are indicated with a box (red color) boxes.

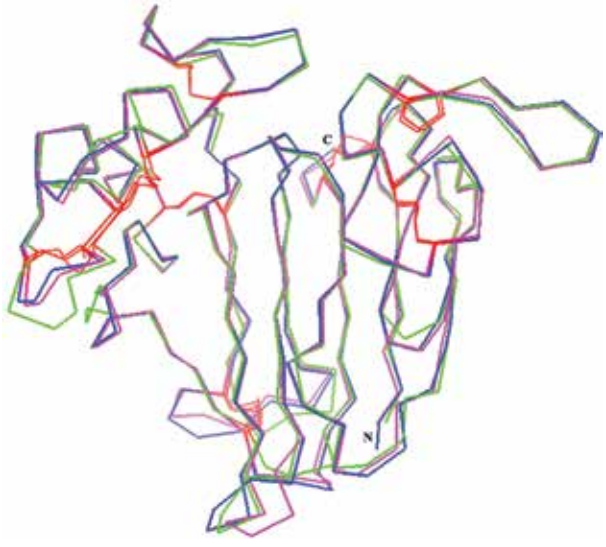

Fig. 3. Stereo diagram of the superposition of the Cα-traces of the StOLP (green), 1AUN (blue) and 2IOW (magenta) in the same orientation. The conserved cysteine residues are shown in red color. Figures were prepared with PyMol.

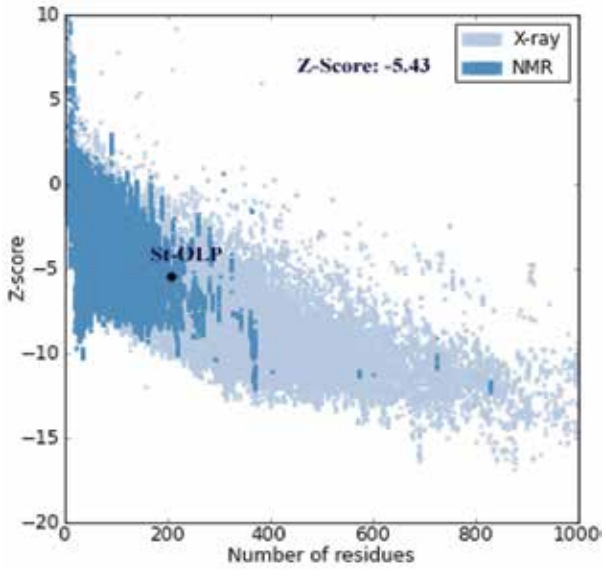

Fig. 4. Evaluation of the StOLP using ProSa web server (https://prosa.services.came. sbg.ac.at/prosa.php) displaying Z-Score value indicating nearness of the model structure with native structure.

protein structures of same size with Z score of -5.43. This indicates that the model was sufficiently accurate for the structure based analysis. ProSA-Web-server analysis revealed that the modelled structure occupies a region of NMR predicted native protein structures of same size with Z score of -5.43 (**Fig. 4**). This supports that the StOLP model is sufficiently accurate. The VERIFY 3D analysis revealed that 98.56% of the residues has an average 3D-1D score >0.2 showing that the model protein has good overall quality.

Structural features responsible for the antifungal activity

At sequence level, comparison with StOLP showed a putative N-terminal elongation (1–21 aa length), a signal peptide sequence used for secretary transport and C-terminal elongated polypeptide (21 aa in length), which is possibly involved for vacuolar targeting (Neuhaus and Rogers 1998). All OLP proteins share 16 conserved Cys that are required for eight disulfide bonds and that play important roles in maintaining structure and activity (**Fig. 2**) (Koiwa *et al.*, 1999). The homology model of StOLP protein was compared with templates indicated structural similarity shared in these proteins. The three dimensional structure of StOLP (**Fig. 5**) showed three characteristic domain showing similarity fold with the thaumatin,

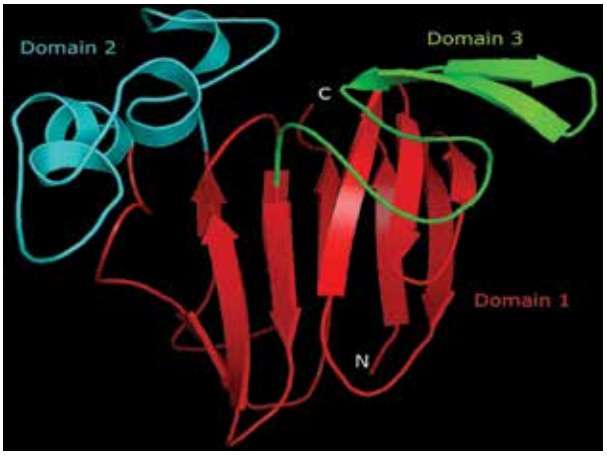

Fig. 5. Theoretical structure of StOLP constructed by homology modelling. The three domains are shown in red (domain1), (blue), cyan (domain 2) and green (domain3). The cleft region lies at the interface of domain 1 and 2.

TLP (PR-5d) from *Nicotiana tabacum* and NP24 from *Solanum lycopersicum*. Domain 1 consists of a flattened 11- β strands which form a β-sandwich. Six anti-parallel β strands (β2, β3, β5, β6, β7 and β10) are present in the front end and other five anti-parallel β strands (β1, β4, β8, β9 and β11) are located at the back, which together constitute the β-sandwich. Out of the eight disulfide bonds in the structure, three are present in this domain. Domain II consists of an α-helix (α4) and three shorter domains (α1, α2, α3) which are stabilized by three disulfide bonds (formed by six cysteine residues) and forms rich loops exposed on the molecular surface of the protein. Together with domain I, domain II formed a large cleft region on the protein surface. Domain III consisted of a small loop of 32 residues which formed a long loop consisting of two β sheets. This domain was stabilized by two disulfide bonds. All OLP proteins share 16 conserved Cys that are required for eight disulfide bonds (Cys10–Cys205, Cys52–Cys62, Cys67–Cys73, Cys134–Cys144, Cys148–Cys157, Cys158–Cys163, Cys126–Cys176, Cys193–Cys121), which play an important role in maintaining structure and activity (Koiwa *et al.*, 1999). Three-dimensional analysis of OLP proteins showed that the acidic cleft is important for β-1,3-glucan-binding activity (Koiwa *et al.*, 1999). The 3D model of StOLP indicated that the acidic amino acids Glu (E85), Asp (D98), Asp (D103) and Asp (D186) were conserved in the appropriate positions of the acidic cleft (**Fig. 6**). These hydrophilic residues in the acidic cleft region were important in determining its antifungal activity. This confirmed that OLP proteins are strongly conserved throughout the evolutionary process. Electronegative character of the acidic cleft of OLP's regions may be crucial in conferring antifungal activity (Grenier *et al.*, 1999; Leone *et al.*, 2006). Analysis of the electrostatic surface potential showed that the acidic cleft region formed by domain 1 and 2 comprised mainly of hydrophilic residues but was surrounded by aromatic residues Tyr86, Phe91, Phe96 and Tyr180, which contribute to the hydrophobicity were also conserved in the present model structure. Woloshuk *et al.* (1991) showed that the hydrophobic nature of the protein contributes interaction with

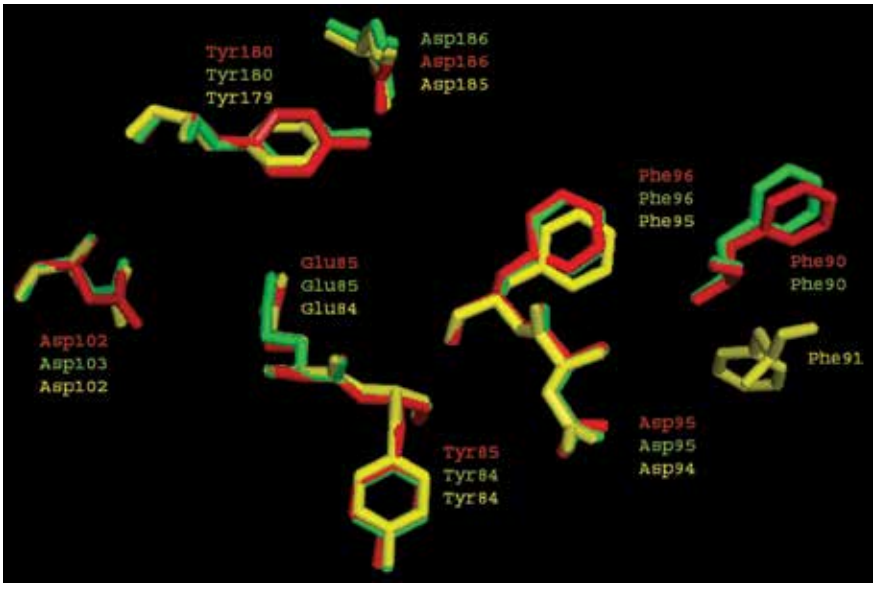

Fig. 6. Close-up view of conserved residues at and around the active site of StOLP (red), 1AUN (green) and 2IOW (yellow).

the fungal cell wall components. However, a similar electronegative central cleft was previously identified in the crystal structure of zeamatin from *Zea mays*, PR-5d and osmotin from tobacco. The negative charge in the cleft regions may be crucial in conferring antifungal activity. The back side of the clefts of StOLP as compared to PR-5d *Nicotiana Tabacum* (1AUN) and *Solanum lycopersicum* (2IOW) was basic in nature. Although all known antifungal PR-5 proteins are similar in structure and have an acidic cleft but they showed different specificity to their target cell. Therefore, it is assumed that acidic cleft of St-OLP interacts with its receptor in the plasma membrane of fungi. The electrostatic potential of tobacco PR 5d (1AUN) and *Solanum lycopersicum* (2IOW) varies from -48.075 to + 48.075 and -46.194 to +46.194 respectively which has less electrostatic potential than StOLP ranging from -54.694 to + 54.694. Osmotin causes membrane leakage and dissipation of pH gradient across the plasma membrane. Such small differences in the surface electrostatic potential around the acidic cleft of StOLPs have to be considered for explaining specificities of the OLPs towards target fungal cells. Based on crystal structure of zeamatin, it was suggested that all osmotin-like proteins have an electrostatically polarized surface that may be critical for antifungal activity of this group of proteins.

CONCLUSION

Structural features of PR-5 proteins have been well-characterized by three dimensional modelling of osmotin from *Solanum tuberosum*. Three-dimensional structure analysis allowed us to make a detailed comparison of antifungal and non-antifungal PR-5 proteins in order to determine their biological functions. Thus modelling effort provided structural predictions about the catalytic domain of StOLP that are being verified via experimental techniques. However this antifungal protein can be tested for use as pharmaceutical agents for the treatment of various fungal diseases. The antifungal protein gene characterized in this study has great potential to be used in agribusiness to create transgenic plants with increased fungal resistance.

LITERATURE CITED

Ahmad S, Gromiha M, Fawareh H, Sarai A (2004) ASAView: Database and tool for solvent accessibility representation in proteins. *BMC Bioinformatics* **5**: 51

Akira S, Uematsu S, Takeuchi O (2006) Pathogen recognition and innate immunity. *Cell* **124**: 783-801

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **3**: 403-410

Antoniw J F, Ritter C E, Pierpoint WS, van Loon LC (1980) Comparison of three pathogenesis-related proteins from plants of two cultivars of tobacco infected with TMV. *J Gen Virol* **47**: 79-87

Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu Rev Plant Biol **55***:* 373-399

Bol JF, Linthorst HJM, Cornelissen BJC (1990) Plant pathogenesis-related proteins induced by virus infection. *Annu Rev Phytopathol* **28**: 113-138

Bontempo P, Carafa V, Grassi R, Basile A, Carlo G et al. (2013) Antioxidant, antimicrobial and anti-proliferative activities of *Solanum tuberosum L. var. Vitelotte*. *Food Chem Toxicol* **55**: 304-312

Do JR, Kang SN, Kim KJ, Jo HJ, Lee SW (2004) Antimicrobial and antioxidant activities and phenolic contents in the water extract of medicinal plants. *Food Sci Biotech* **13**: 640–645.

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

Grenier J, Potvin C, Trudel J, Asselin A (1999) Some thaumatin-like proteins hydrolyse polymericb-1, 3-glucans. *Plant J* **19**: 473-480

Han J, Lee J D, Jiang Y, Li Z, Feng L, Ulevitch R J (1996) Characterization of the structure and function of a novel MAP kinase kinase (MKK6). *J Biol Chem* **271**: 2886–2891

Koiwa H, Kato H, Nakatsu T, Oda J, Yamada Y, Sato F (1999) Crystal Structure of tobacco PR-5d protein at 1.8 Å resolution reveals a conserved acidic cleft structure in antifungal thaumatin-like proteins. *J Mol Biol* **286**: 1137–1145

Leone P, Menu-Bouaouiche L, Peumans WJ, Payan F, Barre A, Roussel A, Van Damme EJM, Rouge P (2006) Resolution of the structure of the allergenic and antifungal banana fruit thaumatin like protein at 1.7-A. *Biochimie* **88**: 45-52

Liu D, Raghothama KG, Hasegawa PM, Bressan RA (1994) Osmotin overexpression in potato delays development of disease symptoms. *Proced Natl Acd Sci USA* **91**: 1888-1892

Neuhaus JM, Rogers JC (1998) Sorting of proteins to vacuoles in plant cells. *Plant Mol Biol* **38**: 127-144

Roberts WK, Selitrennikoff CP (1990) Zeamatin, an antifungal protein from maize with membrane-permeabilizing activity. *J Gen Microbiol* **136**: 1771–1778

Singh N, Bracker CA, Hasegawa PM, Handa AK, Buckel S, Hermodson MA, Pfankoch E, Regnier FE, Bressan RA (1987) Characterization of osmotin. *Plant Physiol* **85**: 529–536

van Loon LC, Pierpoint WS, Boller T, Conejero V (1994) Recommendations for naming plant pathogenesis-related proteins. *Plant Mol Bio Rep* **12**: 245-264

van Loon LC, Rep M, Pieterse CM (2006) Significance of inducible defense-related proteins in infected plants. *Annu Rev Phytopathol* **44**: 135-62

Velazhahan R, Datta SK, Muthukrishnan S (1999) The PR-5 family: thaumatin-like proteins in plants. In: Datta SK, Muthukrishnan S (eds). Pathogenesis-related proteins in plants. Boca Raton, FL, CRC press, p. 107-29.

Vigers AJ, Roberts WK, Selitrennikoff CP (1991) A new family of plant antifungal proteins. *Mol Plant-Microbe Inter* **4**: 315–323

Vigers AJ, Wiedemann S, Roberts WK, Legrand M, Selitrennikoff CP, Fritig B (1992) Thaumatin-like pathogenesis-related proteins are antifungal. *Plant Sci* **83**: 155-161

Woloshuk CP, Meulenhoff JS, Sela-Buurlage M, van den Elzen PJM, Cornelissen BJC (1991) Pathogen-induced proteins with inhibitory activity toward *Phytophthora infestans*. *Plant Cell* **3**: 619-628

MS received: 26 July 2016; Accepted: 20 April 2017