



Modeling and molecular dynamics simulation of PR-1 protein from *Solanum tuberosum*, an integral part of plant defense

Subarna THAKUR¹, Bothra K. ASIM², Sur SAUBASHYA¹ and Sen ARNAB^{1*}

¹NBU Bioinformatics Facility, Department of Botany, University of North Bengal, Siliguri 734013, India.

²Bioinformatics Cheminformatics Laboratory, Department of Chemistry, Raiganj College, Raiganj, India.

*Corresponding author. E-mail: senarnab_nbu@hotmail.com

ABSTRACT

Pathogenesis-related (PR) proteins are considered as major weapons in plant's defense tactics against pathogens. The PR-1 protein of *Solanum tuberosum* forms an integral part of the host defense system. We present here the 3D structure of PR-1 protein of *S. tuberosum* based on homology modeling technique. The model revealed a novel α - β - α sandwich arrangement of secondary structure. It contains four α helices (I-IV) and a β sheet arranged in a three stacked layer. The molecular dynamics simulation studies revealed the flexible nature of some residues (115-120) of fourth alpha helix. The flexibility of this region along with sequentially conserve nature provided clue in the identification of this region as functionally important. Amongst them, we identified His-117 as tentative active site based on position and previous experimental evidences. The structure based phylogenetic study showed; the protein is far apart from other non-plant members of PR-1 superfamily, although it shared some amount of sequence homology with them.

© 2010 International Formulae Group. All rights reserved.

Keywords: Homology modeling, template, molecular dynamics, GROMACS, Phylogenetic.

INTRODUCTION

The resistance response to a variety of pathogenic invasion in plants is often accompanied by production of several host proteins referred to as pathogenesis related proteins. These proteins are a major defense strategy of plants in pathological situations. These proteins are not incited by the type of pathogen rather they are determined by the type of reaction of the host plant (Chakraborty, 2008). PR proteins can be classified into 17 major groups on the basis of their sequence characteristics, immunological relationships and biological activities (Liu et al., 2005). Among the various PR proteins, PR-1 is the most abundant and reaching upto 1-2% of the total leaf proteins. Proteins belonging to PR-1 family were first identified from tobacco-mosaic virus infected tobacco plants (Van loon, 1976). Since then, many

proteins with similar properties have been isolated from many other dicot and monocot plant species. Many isoforms of PR-1 have been isolated, and many genes encoding them have been cloned and sequenced. Although the proteins have been shown to be associated with limited antifungal activities (Park et al., 2010), suggesting a potential role in plant defense, but their molecular mechanism and relationship to other protein are still relatively unknown. The physicochemical properties of these proteins enable them to survive in the harsh environmental conditions where they are found such as vacuolar compartment or intercellular spaces (Niderman et al., 1995). Although structures of some pathogenesis related proteins associated with tomato (Ghosh and Chakrabarti, 2005), (Fernandez et al., 1997) and tobacco (Koiwa et al., 1999) have been reportedly resolved by X-ray

© 2010 International Formulae Group. All rights reserved.

crystallography and NMR techniques, almost no such work has been reported on the 3-D structure of pathogenesis related proteins particularly PR-1 of *S. tuberosum* (potato).

In the present study, we have constructed the three dimensional model of PR-1 protein based on the NMR structure of pathogenesis related protein P14a (PDB code-1CFE) by homology modeling method. A suitable refined model of the protein was then subsequently used for molecular dynamics simulation to understand the motional and structural properties which can then be interpreted to understand the functionality and biological activity of these proteins. The constructed 3D model of the protein and its amino acid sequence was also utilized for the phylogenetic analysis to trace its relationship with other diverse proteins of the PR superfamily.

MATERIALS AND METHODS

Homology modeling

The amino acid sequence of the PR-1 protein of *S. tuberosum* bearing accession number CAB58263 was obtained from the NCBI (National Center for Biotechnology Information) database (www.ncbi.nlm.nih.gov). The first step of model building was to search for a number of related sequences to find a related protein that will be used as a template. Position specific iterative blast (PSI-BLAST) (Altschul et al., 1997) (<http://www.ncbi.nlm.nih.gov/blast/>) was carried out against database specification of PDB proteins to detect similarity. The program ClustalW 1.83 (Thompson et al., 1994) with default settings was used to carry out optimal alignment between the target sequence and template. Based on this alignment crude 3D model of the protein was created using the MODELLER9v4 program (Sali and Blundell, 1993). The crude models obtained are often erroneous and therefore refinement is indispensable. The constructed model was energetically minimized using the steepest descent (SD) and conjugate gradient (CG) methods. The computations were carried out *in vacuo* with the GROMOS96 43B1 factors using the Swiss-Pdb Viewer package (<http://expasy.org/spdv/program/spdv37sp5.zip>) (Kaplan and Littlejohn, 2001). Hydrogen bonds were ignored. The secondary structure of the protein was predicted with the help of HNN (Hierarchical Neural Network) methods.

In order to test the internal quality and reliability, the model was subjected to some evaluations: ProSA (Wiederstein and Sippl, 2007) was performed to judge the accuracy of the modeled structures and to check the 3D models for potential errors. VERIFY3D (Eisenberg et al., 1997) was used to authenticate the refined structures. The Auto Deposition Input Tool (ADIT) (<http://deposit.pdb.org/validate>) was used to construct Ramachandran Plot to look into the possible conformations of ϕ and ψ angles for our modeled protein. (Ramachandran et al., 1963). SAVES (Structure analysis and verification server) (<http://nihserver.mbi.ucla.edu/SAVS/>) was used to verify the model quality. Presence of ligand binding pockets in the structures was predicted using CASTp server (Dundas et al., 2006).

Molecular dynamics simulation

The constructed refined model of the PR protein of *S. tuberosum* was used as starting structure for the Molecular dynamics simulation. The simulation was performed using the Groningen Machine for Chemical Simulation (GROMACS) program (Lindahl et al., 2001) running on linux platform. The starting structure was immersed in a cubic box with dimension of 10 Å³ consisting of SPC water molecules. The GROMOS96 force field parameters were used for the simulation. The Linear Constraint algorithm was used to constrain bond lengths, with a time step of 2 fs (Hess et al., 1997). The simulation was conducted at constant temperature (300 K), coupling each component separately to a temperature bath using the Berendsen coupling method (Berendsen et al., 1984). Non-bonded cutoffs of 1.0 nm were used for both Lennard-Jones and Coulomb potentials. The total MD simulation was performed for twelve nano seconds. Before running the simulation, the solvent was relaxed by energy minimization where 1000 steps of steepest descent (converging in 128 steps) was used followed by 1000 steps of conjugate gradient process (converging in 10 steps) and this was followed by 3.0 ns of simulation imposing positional restraints on the non-H atoms. The positional restraints were then released and 12 ns production run were obtained and analyzed. The trajectory files obtained during the whole simulation run was analyzed and RMSD (root mean square deviation), RMSF (root mean

square fluctuation), gyration radius, potential energy were calculated with the help of GROMACS package. The time evolution of secondary structure was evaluated with the help of DSSP package (Kabsch and Sander, 1983).

Sequence and structure based phylogenetic analysis

The DALI server (http://ekhidna.biocenter.helsinki.fi/dali_server/) was used to search for similarity between the 3D structures of the modeled protein with those of the proteins already present on the PDB. The three dimensional coordinates and amino acid sequences of the proteins showing structural similarities were obtained from RSCB Protein Data Bank. The structure-based and sequence-based relationships between different proteins were elucidated by constructing phylogenetic tree. Structure-based phylogenetic trees were constructed according to the Q values between the molecules after alignment with the help of the Multiseq tool (Roberts et al., 2006) where as the sequence-based phylogenetic trees were created according to ClustalW values. The Multiseq program has been incorporated in VMD version 1.8.6 as plugin tool. Multiple alignments of the selected proteins were performed using the Bioedit software (Hall, 1999) to identify the conserved residues.

RESULTS

Modeling

The 3D model of the protein was built taking PDB structure 1CFE (Fernández et al., 1997) as the template protein. The template was a 135 amino acid residue PR protein, P14a from *Lycopersicon esculentum*. Template selection was based upon its environmental and functional resemblance as well as phylogenetic similarity with the protein under consideration. The P14a protein shows a 90% identity with PR-1 protein from *Solanum tuberosum*. The pair wise alignment (Figure 1) revealed a high degree of conservation in the amino-acid residues amongst the two proteins. The major conserved regions lay between residues 25-51, 53-64, 66-82, 84-96 and 117-154. In the PR protein of *L. esculentum* (PDB code 1CFE) the residues Ser 3, Gln 5, His 48, Ser 49, His 903, Arg 100 and Asn 114 are strictly conserved regions of which His 48, Ser 49 and His 93 are functionally important sites. In our

modeled protein, the functionally important residues (i.e. active sites) corresponded to histidine in 72 position, serine in 73 position and histidine in 117 position.

Figure 2 shows the 3D structure of the modeled protein. The modeled protein was named as 1SOP. The 3D structure of the protein shows a novel α - β - α arrangement of α -helices and β -strands quite similar to that of 1CFE. The protein has four alpha helices and a beta sheet arranged between them. The HNN analysis of the secondary structure of the protein illustrated that the alpha helix portion consisted of 39 (24.53%), extended strand 28 (17.61%) and random coil 92 (57.86%) residues.

Figure 3 shows the superimposed figure of template and modeled protein.

The quality of the refined model assessed by different programs like ProSa, Verify 3D and ERRAT Score suggested that the predicted protein model was quite reliable and consistent. The overall quality score determined by ProSA for our structure is demonstrated in a plot (Figure 4) that shows the scores of experimentally determined protein chains available in the Protein Data Bank (PDB). The PR protein of *Solanum tuberosum* had a z score of -4.61. This result pointed out that the z-score of our model was within the range of scores normally found for proteins of comparable size. VERIFY 3D revealed that 80.63% of the residues had an averaged 3D-1D>0.2. The plot of average 3D-1D profile score of residues of our predicted protein model signifies that our model is reliable. CASTp revealed the presence of 16 pockets for binding regions with varying area and volume. These pockets play an important role in the protein functionality. The largest pocket comprises of the residues Thr11, leu13, Ala 17, Val15, Ser 83, Gly 84, Gly116, His 117 and Arg 135. The Ramachandran plot (Figure 5) representing the backbone conformations for the modeled protein reveals that 82.7% of the non-glycine and non-proline residues of the modeled protein reside in the most favored regions.

Dynamics study

Several parameters such as RMSD, Radius of gyration (Rg), RMSF, etc, were considered to assess the structural stability of the protein during the course of simulation. The RMSD was used to evaluate the deviation

of the structure from the original starting structure over the period of the simulation. Figure 6 shows a plot of RMSD over time. It is quite apparent from the figure that RMSD gradually increased up to 3000 ps, then suddenly decreased around 4000 ps and started increasing again. From 7000 ps onwards it is observed that there is a leveling off in deviation from the reference structure. The time evolution of Radius of gyration (Rg) is presented in the Figure 7. The Radius of gyration gives us an idea about the compactness of the structure. The figure clearly shows that there is quite a variation in the Radius of gyration during the simulation particularly during 3000 to 5000 ps time frame. This further proves that a conformational change is taking place during the simulation period of 3000 to 5000 ps.

The root mean square fluctuation is a measure of flexibility of each residue over the trajectory. Figure 8 demonstrates the RMSF of Ca atoms plotted against the residue number over the time period 7000-12000 ps. From the plot it is obvious that the first few residues of the proteins fluctuate considerably. Importantly a major fluctuation is noticed in the residue stretches 5-20, 110-120 and 130-140 during the simulation reflecting that this region of the protein shows most mobility. Apart from the first few residues, a profound fluctuation is observed in residues 110-120. Some of these residues are part of the alpha helix IV secondary structure of the protein. During the simulation the alpha helices particularly the fourth helix showed major conformational changes. Figure 9 shows the secondary structure evolution of the protein during the time frames 7000-12000 ps. The

secondary structure analysis was done using the DSSP program. The simulation time frame of 7000 ps to 12000 ps was selected to examine the changes in the secondary structure because the protein was structurally quite stable in this period as revealed by RMSD and Radius of gyration values. The figure reveals that three alpha helices were relatively quite stable during the simulation and did not undergo any structural changes but the alpha helix comprising of the residues 115-120 residues was significantly broken during this period and was never recovered. The β sheet present in the structure was also very much stable during the simulation. During the time course of the simulation, the total potential energy of the protein did not exceed -9.6×10^5 KJ mol⁻¹, indicating that the protein was energetically quite stable.

Phylogenetic analysis

The phylogenetic analysis based on amino acid sequences revealed that other than 1CFE (the template); 1SOP showed close relationship with the protein 1QNX. This protein is an insect allergen protein from the venom of *Vespula vulgaris* (Henriksen et al., 2001). These three proteins are present in the same clade of the phylogenetic tree. The protein 1SOP is also relatively close to the pathogenesis related protein from nematode parasite *Necator americanus* (1U53) (Asojo et al., 2005) in the phylogenetic tree. The tree in Figure 10a clearly exhibit that the protein 1SOP is sequentially quite distant from the other studied proteins of the PR superfamily, maintains quite a distance from them in the phylogenetic tree.

```

CLUSTAL W (1.83) multiple sequence alignment

1CFE      -----QNSPQDYLAVHNDARAQVGVGPMMSWDANLASRAQNY
1SOP      MGLFNISLLLTCLMVLAI FHSCDAQNSPQDYLAVHNDARAQVGVGPMMSWDAGLASRAQNY
          *****

1CFE      ANSRAGDCNLIHSGAGENLAKGGDFTGRAAVQLWVSEKPSYNYATNQCVGGKCRHYTQ
1SOP      ANSRTGDCNLIHSGAGENLAKSGDFTGRAAVQLWVGEKPNYNYGTNQCASGQVCGHYTQ
          ****:*****

1CFE      VVWRNSVRLGCGRARCNGWVFISCNYPVGNWIGQRPY
1SOP      VVWRNSVRLGCGRARCNGWVFISCNYPVGNWVQRPY
          *****
    
```

Figure 1: Pair-wise alignment of the template protein and PR1 protein of *Solanum tuberosum*. Residues marked as * are conserved.

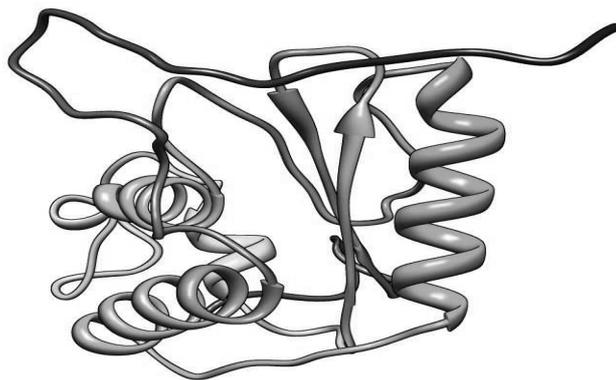


Figure 2: Three dimensional structure of PR-1 of *Solanum tuberosum*. The structure has four alpha helices and a beta sheet is interspersed between them.

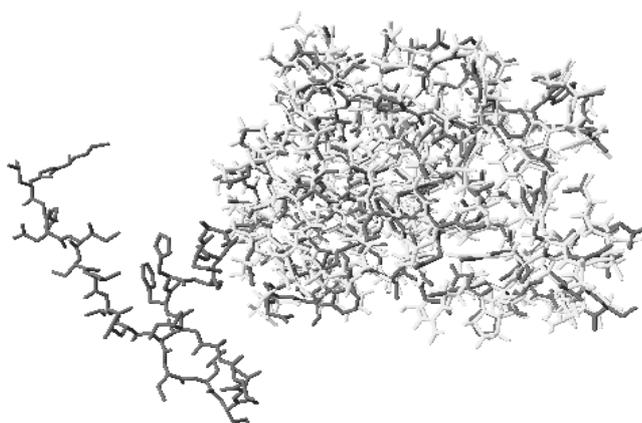


Figure 3: Superimposed figure of template (light grey) and modeled protein (dark grey).

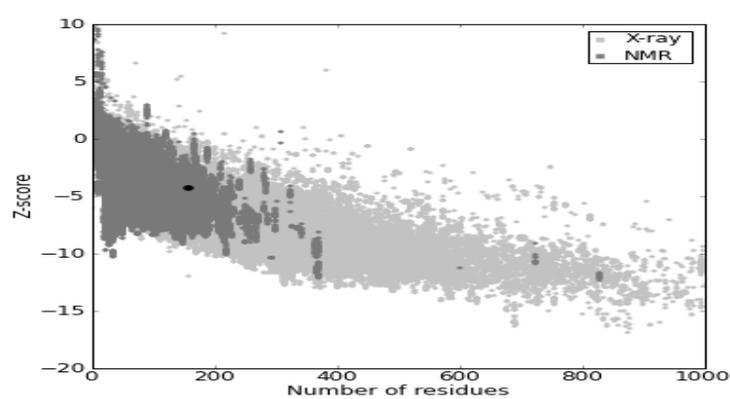


Figure 4: The quality of the protein model accessed by ProSA plot. The Z-score determined to 4.61.

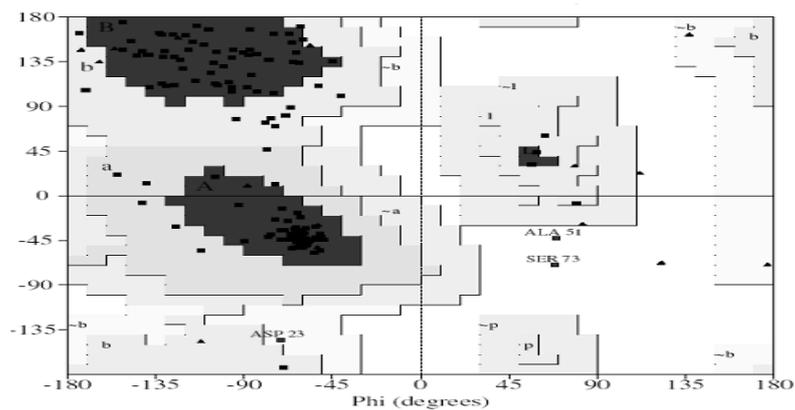


Figure 5: Ramachandran plot illustrating the backbone conformations of the modeled protein.

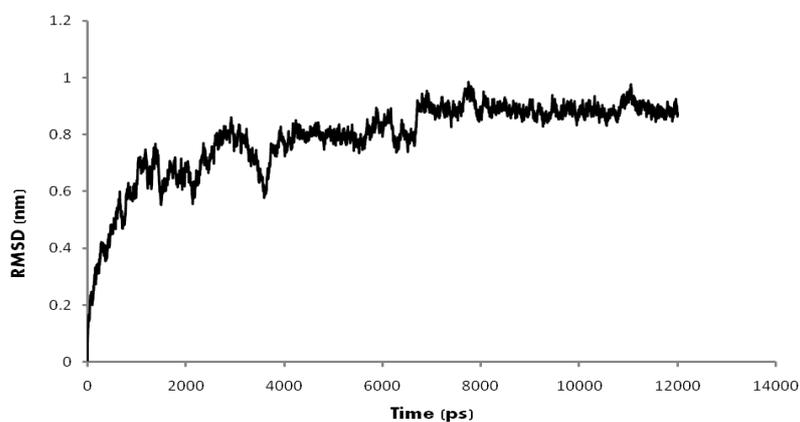


Figure 6: The time evolution of RMSD during the whole simulation.

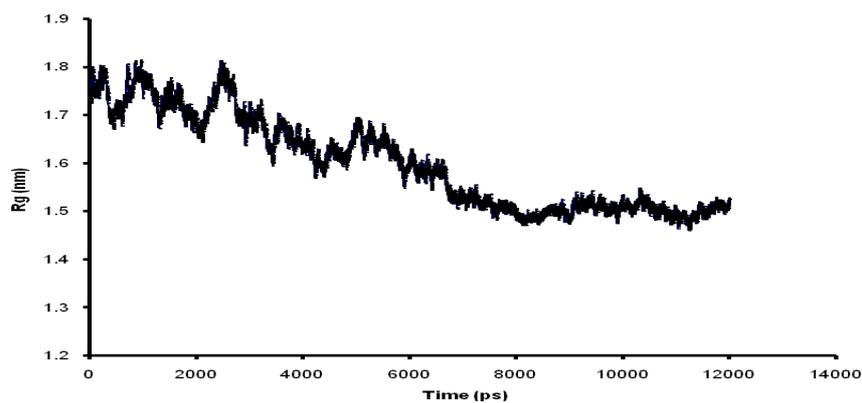


Figure 7: Time evolution of changes in radius of gyration of the protein in aqueous medium.

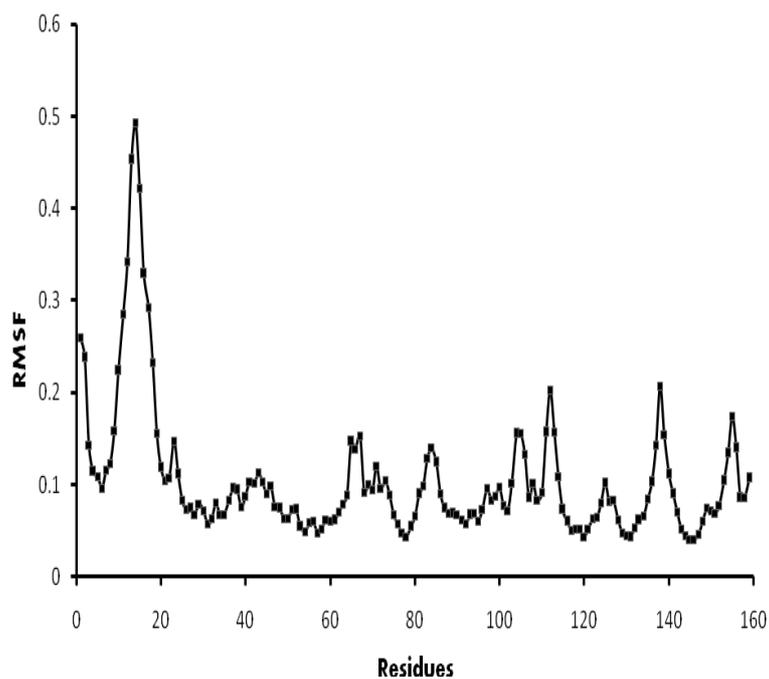


Figure 8: RMS fluctuation in the C α atoms of the residues of the protein in aqueous medium.

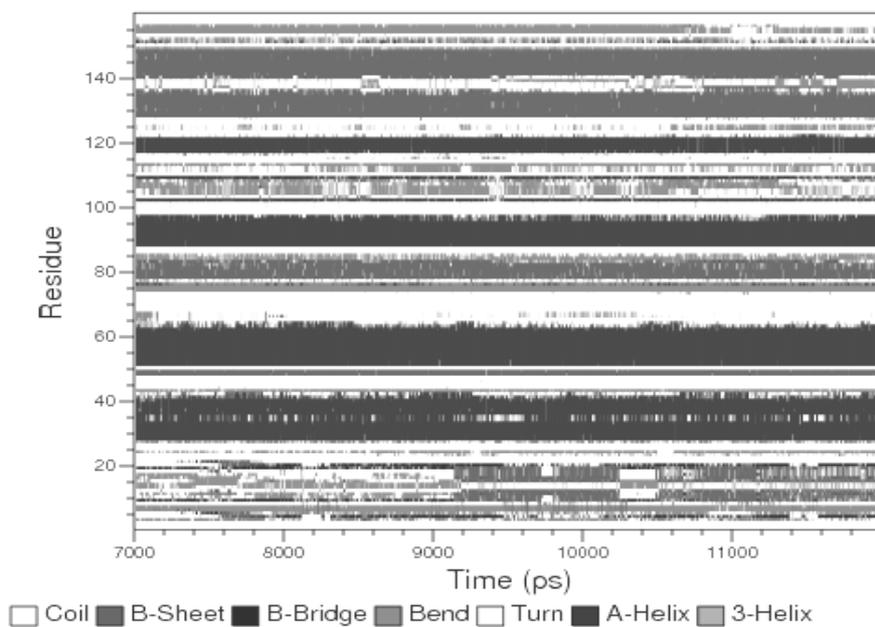


Figure 9: Secondary structure evolution of the protein during the time frame 7000-12000 ps. The helix between the residues 115-120 is clearly broken during this period.

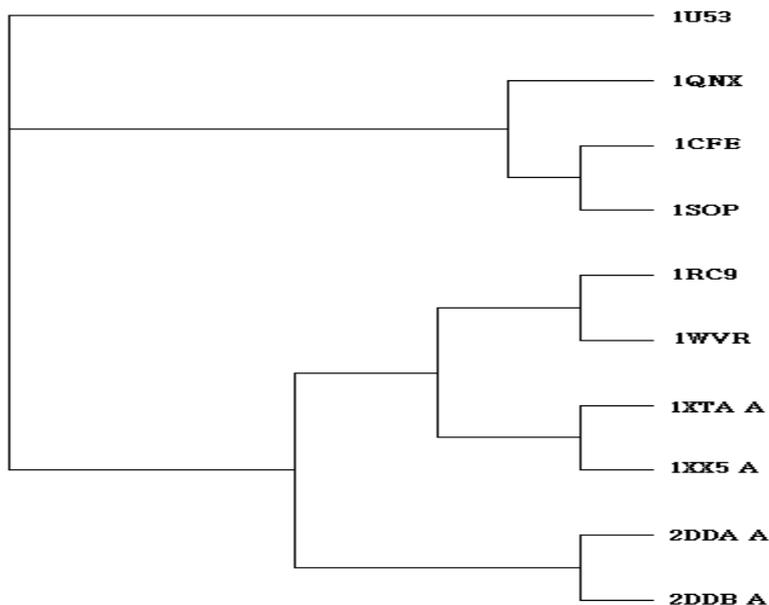


Figure 10a: Phylogenetic tree based on amino acid sequence.

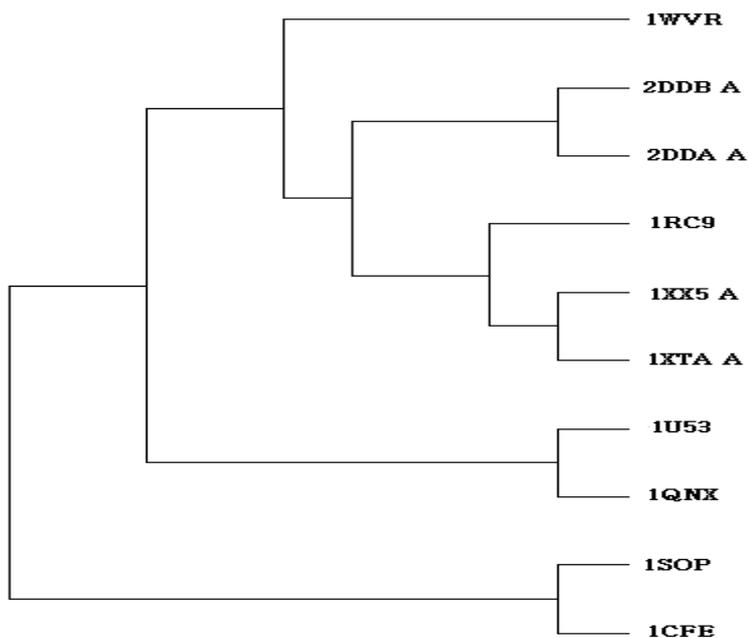


Figure 10b : Phylogenetic tree based on three dimensional structures.

The 3D structure of 1 SOP was further used for structure based phylogenetic analysis. The structure based phylogenetic tree (Figure 10b) discloses that the 1SOP protein remains close with 1CFE protein in a separate cluster

from all other member proteins of PR-superfamily. This clade seems to have diverged from the main stock quite early during the course of evolution. The two proteins, 1U53 and 1QNX, which were close

to 1SOP in sequence based phylogenetic tree are also close to them in this tree but present in an entirely separate clade along with the others.

DISCUSSION

The model of the PR-1 protein from *S. tuberosum* built on the template 1CFE offers reliable insight into its conformational properties and structure-function relationship. The α - β - α sandwich molecular architecture of the protein represents a rigid scaffold offering higher stability. The higher stability is pivotal for the working of the PR-1 in extreme extra cellular environment. The α - β - α core sandwich structure is quite common in some proteins belonging to pathogenesis-related protein superfamily characterized from plants, reptiles, insects, fungi, and even mammals like human. Higher degree of sequence conservation indicates that most PR-1 proteins assume a common three dimensional fold with hydrophilic residues occupying the surface and hydrophobic residues buried within the core. The 16-pockets identified by CASTP server are important functional site of the protein. The Ramachandran plot of the protein implies that the stereochemical properties and quality of the modeled structures are quite appropriate. The allotment of main chain torsion angles phi and psi evidently illustrated that bulk of the amino-acids are in a phi-psi distribution more or less reliable with right handed alpha helices.

The MD simulation data analysis revealed that the protein was overall structurally quite stable during the whole period of simulation. The RMSD and Radius of Gyration parameters illustrated that there is a variation in the structure of the protein during 3000-5000 ps. The RMSF analysis of C α atoms during 7000-12000 ps showed that the residues 115-120 belonging to IV alpha helix were rather flexible. The study of changes in secondary structure of the protein further substantiates the fact that these residues undergo certain conformational changes. The flexibility of these residues can be attributed to their role in the functionality

of the protein perhaps they constitute the binding sites of the protein. It also reemphasizes the fact that alpha IV helix plays a vital role in the α - β - α sandwich arrangement. The 115-120 residues show a great extent of conservation amongst various members of PR-superfamily as revealed by multiple alignments. Some of these conserved residues are solvent exposed and part of a large cavity in the 3D structure of the protein as revealed by Castp server results. Amongst these residues the Histidine residue at position 117 was conserved in all the studied proteins. This His 117 residue which corresponds to His 156 residue of 1QNX has already been indicated as a putative active site residue by Fernández et al. (1997).

The sequence based phylogenetic analysis revealed that PR-1 proteins from plant kingdom shared sequence similarities (30-35%) with various other non-plant proteins belonging to invertebrate and vertebrate, perhaps belonging to a large PR-superfamily. It also possibly shares some amount of biological function with them (Van Loon and Van Strien, 1999). But the structure based evolutionary tree discloses the fact that the plant PR-1 proteins separated at the root from their animal counterparts. For proteins sharing relatively low sequence identity as in this case, 3D structures are better than primary sequences for modeling of protein evolution (Balaji and Srinivasn, 2007). So it can be presumed that though there is sequence homology, plant PR-1 protein assumes some uncommon folds in relation to their functionality in plants which sets them apart.

Conclusion

The three dimensional structure of the PR-1 protein from *S. tuberosum* provided an insight into its structural and functional properties. The structure presented here is quite consistent with their biochemical and motional properties. The sequence and structure based phylogenetic studies offers a view about phylogenetic relationship of this protein. The plant PR-1 proteins have been shown to follow a different course during

evolution of its 3D structure. Molecular dynamics analysis revealed higher degree of stability but the residues of alpha IV helix are quite flexible during the simulation. The flexibility together with conservation of the residues 115-120 suggests their role as functionally important site. These results are vital for describing the mode of action of PR-1 of plants and can be further substantiated with experimental evidences.

ACKNOWLEDGEMENTS

The authors are grateful to the Department of Biotechnology, Government of India, for providing financial help in setting up Bioinformatics Facility, in the Department of Botany, University of North Bengal. ST acknowledges the receipt of UGC Meritorious Fellowship.

REFERENCES

- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSIBLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**(17): 3389-3402.
- Asojo OA, Goud G, Dhar K, Loukas A, Zhan B, Deumic V, Liu S, Borgstahl GE, Hotez PJ. 2005. X-ray structure of Na-ASP-2, a pathogenesis-related-1 protein from the nematode parasite, *Necator americanus*, and a vaccine antigen for human hookworm infection. *J. Mol. Biol.*, **346**(3): 801-814.
- Balaji S, Srinivasan N. 2007. Comparison of sequence-based and structure-based phylogenetic trees of homologous proteins: inferences on protein evolution. *J. Biosci.*, **32**(1): 83-96.
- Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, **81**(8): 3684-3690.
- Chakraborty BN. 2008. Plant defense proteins. *NBU J. Pl. Sci.*, **2**(1): 1-12.
- Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J. 2006. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.*, **34**(Web Server issue): W116-W118.
- Eisenberg D, Luthy R, Bowie JU. 1997. VERIFY3D: Assessment of protein models with three-dimensional profiles. *Metho. Enzymol.*, **277**: 396-404.
- Fernández C, Szyperski T, Bruyere T, Ramage P, Mosinger E, Wuthrich K. 1997. NMR Solution Structure of the Pathogenesis-related Protein P14a. *J. Mol. Biol.*, **266**(3): 576-593.
- Ghosh R, Chakrabarti C. 2005. Crystallization and preliminary X-ray diffraction studies of NP24-I, an isoform of a thaumatin-like protein from ripe tomato fruits. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.*, **61**(8): 806-807.
- Hall TA. 1999. Bioedit: a user friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucl. Acids. Symp. Ser.* **41**: 95-98.
- Henriksen A, King TP, Mirza O, Monsalve RI, Meno K, Ipsen H, Larsen JN, Gajhede M, Spangfort MD. 2001. Major venom allergen of yellow jackets, Ves v 5: structural characterization of a pathogenesis-related protein superfamily. *Proteins.*, **45**(4): 438-448
- Hess B, Bekker H, Berendsen HJC and Fraaije JGEM. 1997. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.*, **18**(12): 1463-1472.
- Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**(12): 2577-2637.
- Kaplan W, Littlejohn TG. 2001. Swiss-PDB Viewer (Deep View). *Brief Bioinform.*, **2**(2): 195-197.
- 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**(4): 1137-1145.

- Lindahl E, Hess B, Vander Spoel D. 2001. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Model.*, **7**(8): 306-307.
- Liu Z, Du L, Wan B. 2005. Pathogenesis-related proteins in higher plants. *Natur. Prod. Res. Dev.*, **17**(1): 229-234.
- Niderman T, Genetet I, Bruyère T, Gees R, Stintzi A, Legrand M, Fritig B, Mosinger E. 1995. Isolation and Characterization of Three 14-Kilodalton Proteins of Tomato and of a Basic PR-1 of Tobacco with Inhibitory Activity against *Phytophthora infestans*. *Plant Physiol.* **108**(2): 17-27.
- Park SC, Lee JR, Kim JY, Hwang I, Nah JW, Cheong H, Park Y, Hahm KS. 2010. Pr-1, a novel antifungal protein from pumpkin rinds. *Biotechnol. Lett.*, **32**(1): 125-130.
- Ramachandran GN, Ramakrishnan C, Sasisekharan, V. 1963. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.*, **7**(1): 95-99.
- Roberts E, Eargle J, Wright D, Luthey-Schulten Z. 2006. MultiSeq: unifying sequence and structure data for evolutionary analysis. *BMC Bioinf.* **7**(1): 382.
- Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.*, **234**(3): 283-291.
- Thompson JD, Higgins DG, Gibson T. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**(22): 4673-4680.
- Van Loon LC, Van Strein EA. 1999. The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiol. Mol. Plant Pathol.*, **55**(2): 85-97.
- Van Loon LC. 1976. Specific soluble leaf proteins in virus-infected tobacco plants are not normal constituents. *J. Gen. Virol.* **30**(3): 375-379.
- Wiederstein M, Sippl MJ. 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.*, **35**: W407- W410.