

ORIGINAL ARTICLE

Molecular Modeling of Pathogenesis-Related Proteins of Family 5

**Claudia E. Thompson,¹ Cláudia L. Fernandes,² Osmar N. de Souza,²
Francisco M. Salzano,¹ Sandro L. Bonatto,³ and Loreta B. Freitas^{1,*}**

¹Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil;

²Laboratório de Bioinformática, Modelagem e Simulação de Biosistemas, Faculdade de Informática, PPGCC, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil; and ³Centro de Biologia Genômica e Molecular, Faculdade de Biociências, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil

Abstract

The family of pathogenesis-related (PR) 5 proteins have diverse functions, and some of them are classified as thaumatin, osmotins, and inhibitors of α -amylase or trypsin. Although the specific function of many PR5 in plants is unknown, they are involved in the acquired systemic resistance and response to biotic stress, causing the inhibition of hyphal growth and reduction of spore germination, probably by a membrane permeabilization mechanism or by interaction with pathogen receptors. We have constructed three-dimensional models of four proteins belonging to the Rosaceae and Fagaceae botanical families by using the technique of comparative molecular modelling by homology. There are four main structural differences between all the PR5, corresponding to regions with replacements of amino acids. Folding and the secondary structures are very similar for all of them. However, the isoelectric point and charge distributions differ for each protein.

Index Entries: Pathogenesis-related proteins 5 (PR5); molecular modeling; protein structure.

INTRODUCTION

Plants have several active defence systems against pathogen attacks, including the accumulation of antimicrobial secondary metabolites, cell wall reinforcement, and expression of pathogenesis-related (PR) proteins. Sixteen PR protein families have been characterized from different species and classified according to their sequence similarity, serological relationship, enzymatic or biological activity, and structural similarities (1–3). In recent years, several PR5 proteins have been isolated from different species of plants, sequenced, and assayed for antifungal activity. They are an important group of human allergens (4–9), and their expression is regulated in plants in response to stressors such as drought, freezing temperatures, and infection by fungi, viruses, or bacteria (1,4,8,10). Because of their high sequence similarity

with thaumatin, the intensely sweet-tasting protein that accumulates in the fruit of *Thaumatococcus danielli* plants, these proteins are usually called “thaumatin-like proteins” (TLPs) (11). Several members of the PR5 group from tobacco and other plant species were shown to display in vitro significant activity inhibiting hyphal growth or spore germination of various fungi (7). Bioassay results indicate that there is specificity between structures of the PR5 proteins and those of target microbe organisms; therefore, their efficacy would depend on a combination of isoforms and pathogenic species (12–14). However, the mechanism of membrane permeabilization or of interaction with pathogen receptors has not yet been entirely identified (12). Although some PR5 proteins are inhibitors of phytopathogens, other members of this family are devoid of such activity, and the reason for the differences is still poorly understood (15). We have constructed three-dimensional (3D) models of four PR5 of the Rosaceae and Fagaceae botanical families, based on the tertiary structure of the antifungal protein zeamatin

* Author to whom all correspondence and reprint requests should be addressed. E-mail: loreta.freitas@ufrgs.br

(Protein Data Bank [PDB] code 1DU5) with the objective of better understanding the relationship between structure and function.

MATERIALS AND METHODS

Molecular Modeling

The sequences of all proteins used in the modeling process, including those of the possible templates and of the four PR5, were obtained from the National Center of Biotechnology Information (NCBI). Accession numbers are as follows: *Malus domestica* (Rosaceae: 1PR: Q9FSG7), *Castanea sativa* (Fagaceae: 2PR: Q9SMH2), *Prunus avium* (Rosaceae: 3PR: P50694), and *Pyrus pyrifolia* (Rosaceae: 4PR: O80327). Models were based on *Zea mays* (Fagaceae) antifungal protein zeamatin (PDB code 1DU5), found using Blastp (16,17), as its structure had been solved to a 2.5-Å resolution (18). The amino acid sequence alignments were performed with the ClustalW program (19), using the BLOSUM matrix for scoring, and manually adjusted to obtain optimal alignment. From the best alignment of template structures to target sequences, the program MODELLER6v2 (20) was used to build protein models according to the comparative protein modeling methodology (21), to satisfy spatial restraints, and to use standard and loop refinement protocols. The models were energy minimized with the AMBER 6.0 package (22). Finally, the best model for each protein was evaluated and selected on the basis of the results obtained by PROCHECK (23) and VERIFY-3D (24), which verify the stereochemistry and evaluate the environment of each amino acid residue in the model as a function of the environment encountered in the crystallographical structure that served as template. The cycle of realignment, modeling, and structure validation was repeated until no further improvement of the structure could be observed. The root mean square deviation (RMSD) between the template and each model was calculated, all figures drawn, and the molecular surface derived using the Swiss PDB Viewer (25).

Phylogenetic Analysis

For the comparative analysis, the amino acid sequences were also obtained from the NCBI. Accession numbers are as follows: *Avena sativa* (Poaceae: RST1AVESA: L39774; RST2AVESA: L39775; RST3AVESA: L39776; RST4AVESA: L39777), *Capsicum annuum* (Solanaceae: CAPANN_TLP: Q9ARG0), *Castanea sativa* (Fagaceae: Q9SMH2), *Hordeum vulgare* (Poaceae: HORVU1: CAB99485; HORVU3: P32938), *Lycopersicon esculentum* (Solanaceae: LYCES1: P12670; LYCES2: Q9M3X2; LYCES3: CAA50059), *Malus domestica* (Rosaceae: Q9FSG7), *Nicotiana tabacum* (Solanaceae:

OSMO_NICTAB1: P14170; OSMO_NICTAB2: P25871), *Prunus avium* (Rosaceae: 3PR: P50694), *Pyrus pyrifolia* (Rosaceae: 4PR: O80327), *Solanum commersonii* (Solanaceae: OSMOSOLCO1: P50701; OSMOSOLCO2: P50702), *Triticum aestivum* (Poaceae: TRIAES1: P27357), and *Zea mays* (Poaceae: ZEAMATIN: AAA92882). Alignments were performed using ClustalW. They were inspected, and manual changes made when necessary, using GeneDoc 2.6 (26). Phylogenetic trees, based on proportional distance among sequences, were obtained using the neighbor-joining (NJ) method (27), available in the MEGA program version 2.1 (28). In total, 1500 bootstrap replicates were performed to determine the reliability of each node of the tree.

RESULTS

Molecular Modeling

The amino acid sequences of the targets include a signal peptide of 24 residues that was not modeled, because a template does not exist for it. The degree of identity between the sequences of the selected template and of each of the four PR5 was approx 39%. The multiple alignment (Fig. 1) shows high similarity among the sequences. The degree of complete identity was 30%, with 15% showing strong similarity and 16% showing weak similarity.

All four target proteins have the signature of the thaumatin family, which has a consensual pattern corresponding to G-x-[GF]-x-C-x-T-[GA]-D-C-x-(1,2)-G-x-(2,3)-C (29). Ten models for each protein were initially created, and they were analyzed using the PROCHECK and VERIFY-3D programs, together with the values of the RMSDs. The best models resulting from this initial analysis were then submitted to an energy minimization for optimization of the side-chain geometry. No further improvement, however, was obtained (Fig. 2); moreover, in 4PR the energy algorithm has not converged. In the same way, the loop protocol that was used in an attempt to increase the quality of the models in the region 84-90 did not furnish the desirable effect (data not shown). The Ramachandran plots confirmed the excellent quality of the initial models, with the percentage of residues in most favored and additional allowed regions being no lower than 96.2%. The stereochemical parameters are shown in Table 1. The RMSDs between the backbones of the template and models 1PR, 2PR, 3PR, and 4PR were 0.36, 0.41, 0.39, and 0.41 Å, respectively, where the differences in the loop regions account for the largest differences. Together, these data suggest that the models were stereochemically valid.

The 3D models obtained are shown in Fig. 3. Three distinct domains are distinguished. The core domain (I)



Fig. 1. Multiple alignment between the PR5 targets (1PR–4PR) and the template (1DU5). Symbols: (*): complete identity (67/222; 30%); (:): strongly similar (33/222; 15%); (.): weakly similar (36/222; 16%); no symbol: different (87/222; 39%).

presents two bundles of antiparallel strands of β -sheets connected by loops to form a flattened β -sandwich, corresponding to the N-terminal part of the molecule; the second domain (II) has a main α -helix associated to shorter helical segments, whereas the third domain (III) consists of β -strands linked by a loop. Domains I and II together form a large central cleft, with stability of the structure being maintained by disulfide bonds. The total number of β -sheets varies between 14 and 18, 11 of which form the sandwich, representing the core of the protein.

The main structural differences between the four models are localized at positions 30–42, 40–54, 143–157, and 174–177, indicated in Fig. 3A–D, respectively. The first and third differences can be explained by the fact that no template is available for these regions. Furthermore, nonsynonymous substitutions have occurred there, causing structural changes. The second and the fourth differences are caused by amino acid

replacements. A β -strand structure begins at position 30 in 2PR and 4PR, but at position 36 in 1PR and 3PR; consequently, these proteins have an extended loop in that region (Fig. 4A). 2PR has three β -strands between positions 40 and 50, whereas the three others have one β -strand there (Fig. 4). All PRs, except 1PR, have two β -structures between positions 143 and 157 (Fig. 4B). Finally, only 2PR has a helix in position 174 (Fig. 4). The second and fourth differences are localized near the cleft in all PR5 and it is probably for this reason that they are the most important differences. The following substitutions in 2PR are responsible for these differences: Thr40, Gly42, and Lys48. And the helix in the position 174 of 2PR is formed by amino acid residues of Phe173, Asp174, Thr175, Ala176, and Arg177, which are absent in the other PR5 (Fig. 4).

The molecular surface of these proteins is electrostatically polarized (Fig. 5). There is a concentration of acidic amino acids in the clefts in 1PR, 3PR, and 4PR

Table 1
Quality of Main-Chain and Side-Chain Parameters of the Modeled PR5 Proteins

Comparison values										
Stereochemical parameter	No. of data pts		Parameter value		Typical value		Band width		No. of band width from mean	
Main-chain	1PR	2PR	1PR	2PR	1PR	2PR	1PR	2PR	1PR	2PR
% residues in A, B, L	180	190	84.8	83.7	76.6	76.6	10.0	10.0	0.8	0.7
Omega angle SD	219	218	4.7	4.5	6.0	6.0	3.0	3.0	-0.4	-0.5
Bad contacts/100 residues	4	7	1.8	3.2	10.0	10.5	10.0	10.0	-0.9	-0.79
Zeta angle SD	202	205	1.6	1.4	3.1	3.1	1.6	1.6	-1.0	-1.1
H-bond energy SD	111	110	0.9	0.8	0.9	0.9	0.2	0.2	-0.3	0.4
Overall G-factor	222	221	-0.3	-0.3	-0.6	-0.6	0.3	0.3	1.0	1.1
Side-chain										
Chi-1 gauche minus SD	33	27	5.4	5.0	22.7	22.7	6.5	6.5	-2.7	-2.7
Chi-1 trans SD	55	57	10.5	8.7	22.7	22.7	5.3	5.3	-2.3	-2.6
Chi-1 gauche plus SD	71	81	6.4	9.9	21.3	21.3	4.9	4.9	-3.0	-2.3
Chi-1 pooled SD	159	165	7.9	8.8	22.0	22.0	4.8	4.8	-2.9	-2.7
Chi-2 trans SD	37	34	12.0	12.2	23.1	23.1	5.0	5.0	-2.2	-2.2
Main-chain	3PR	4PR	3PR	4PR	3PR	4PR	3PR	4PR	3PR	4PR
% residues in A, B, L	179	177	86.0	84.7	76.6	76.6	10.0	10.0	0.9	0.8
Omega angle SD	219	219	5.1	4.4	6.0	6.0	3.0	3.0	-0.3	-0.5
Bad contacts/100 residues	3	3	1.4	1.4	10.5	10.5	10.0	10.0	-0.9	-0.9
Zeta angle SD	201	194	1.5	1.8	3.1	3.1	1.6	1.6	-1.0	-0.8
H-bond energy SD	107	116	0.8	0.8	0.9	0.9	0.2	0.2	-0.6	-0.5
Overall G-factor	222	222	-0.3	-0.3	-0.6	-0.6	0.3	0.3	1.0	1.1
Side-chain										
Chi-1 gauche minus SD	38	29	10.9	7.7	22.7	22.7	6.5	6.5	-1.8	-2.3
Chi-1 trans SD	57	52	11.0	9.9	22.7	22.7	5.3	5.3	-2.2	-2.4
Chi-1 gauche plus SD	64	80	6.0	8.5	21.3	21.3	4.9	4.9	-3.1	-2.6
Chi-1 pooled SD	159	161	9.2	9.0	22.0	22.0	4.8	4.8	-2.6	-2.7
Chi-2 trans SD	31	27	11.4	11.3	23.1	23.1	5.0	5.0	-2.3	-2.4

SD, standard deviation; pts, points.

(Fig. 5). The isoelectric point (pI) calculated for these proteins were 4.98, 4.84, and 5.07, respectively. However, in 2PR there is a concentration of positive residues neighboring the cleft (Fig. 5), and the pI calculated was 8.57.

Relationships of the Modeled PR5 With Other TLPs

The amino acid phylogenetic tree comparing the four PR5 that have been modeled and other TLPs consists of three main branches (Fig. 6). The first branch comprises Solanaceae proteins, the second TLPs of the Poaceae, and the third the Rosaceae and Fagaceae proteins, which we have modeled. This pattern is similar to the TLP relationships reported previously (15).

DISCUSSION

The data presented here show that the proteins that have been modeled have a very high sequence identity and the same overall 3D fold characteristic of TLPs (Fig. 3). In spite of these structural similarities, the proteins differ by their surface properties (Fig. 5). PR5 can be categorized into three subclasses based on their isoelectric point: acidic, neutral, and very basic members with extracellular, extracellular, and vacuolar localization, respectively (30). Previous studies have revealed that the PRs modeled are apoplastic, because all of them presented a signal peptide and the absence of a C-terminal extension, as found in vacuolar forms (31–34). It is interesting to observe that unlike most basic PR proteins,

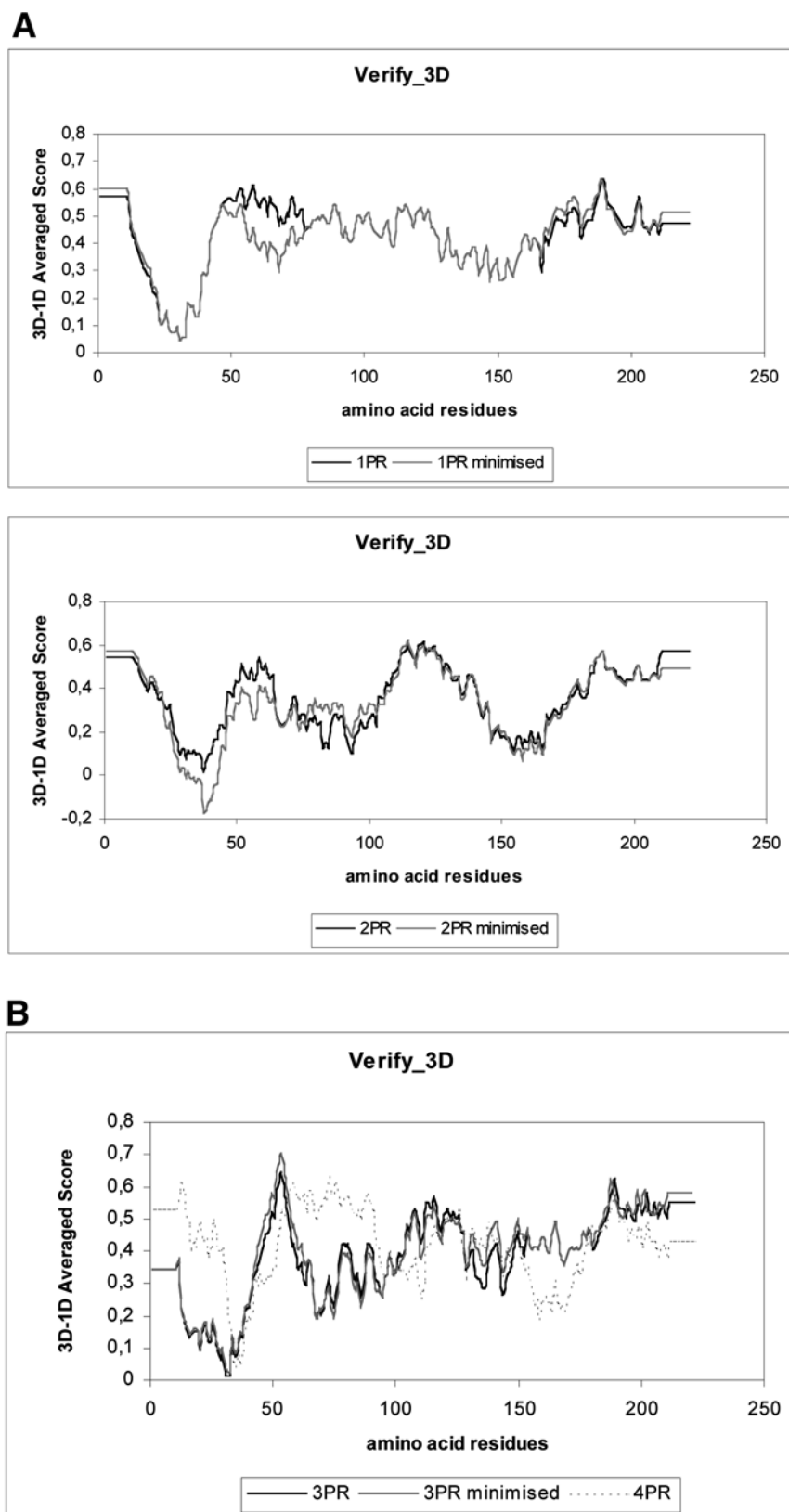


Fig. 2. VERIFY-3D evaluation of the initial and minimized models.

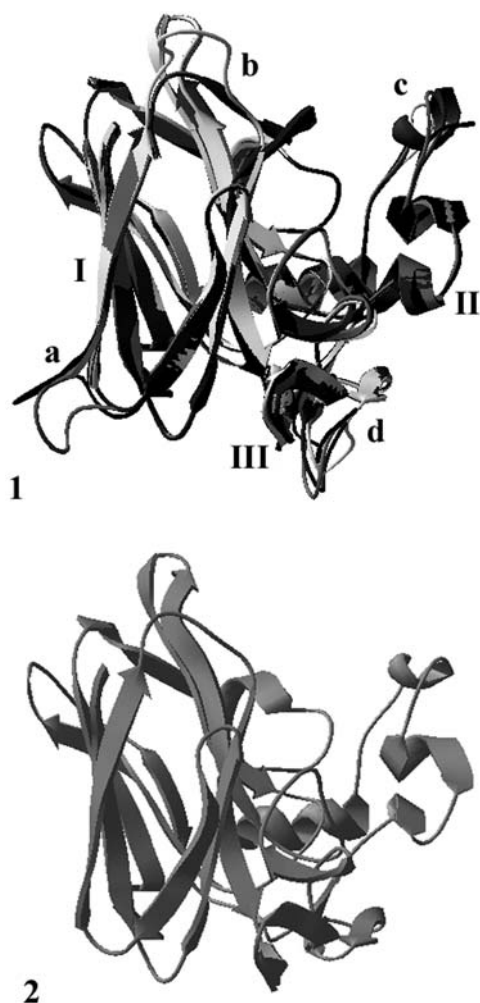


Fig. 3. (1) Ribbon representation of the superposition of the three-dimensional structures of 1PR (black), 2PR (dark gray), 3PR (gray), and 4PR (light gray) proteins, showing the domains I, II and III. The main differences are indicated as: a, b, c, and d. (2) View of the zeamatin 3D structure.

mature 2PR is localized in the extracellular matrix, as revealed by immunoelectron microscopy studies (32).

Three of the PR5 modeled exhibit function against fungal pathogens (31–33), except 3PR. This cherry TLP did not present antifungal activity in preliminary assays (34). Because the high level of expression of this protein has been observed in all varieties of cherries tested in both Europe and North America, it is unlikely that it arises as a pathogen reaction, and Fils-Lycaon et al. (34) suggest that it may have a role in the ripening process.

The core structure of PR5 proteins is conserved in all published 3D crystal structures and models, and the presence of a long and deep cleft that traverses the surface of the proteins is an important feature of these pro-

teins (35). The proteins modeled in this work are composed of three domains and have a similar fold to thaumatin (36) and other PR5 proteins that have a known structure, such as zeamatin (18), tobacco PR-5d protein (14), and osmotin (30). The amino acid residues that are localized at the cleft are part of the β -barrel and loops of region II, and they are negatively charged. Because this charge is conserved in both acidic and basic proteins, the presence of a conserved region in the cleft of PR5 suggests that this particular area may be involved with function (35). This acidic cleft could interact with positively charged membrane proteins involved in water or ion transport across the fungal plasma membrane, causing disturbance of the water or ion flux (11). Besides, the cleft is characteristic of the polysaccharide-binding site of many glycoside hydrolases, so it could be a candidate region for (1,3)- β -D-glucan binding, a type commonly found in fungal walls (35). However, the role of this acidic cleft in the antifungal properties of PR5 proteins remains controversial because there are some proteins without antifungal activity that also possess such a cleft, like 3PR (TLP from cherry) (34) and Ban-TLP from banana (15). Therefore, it is possible that the structural differences not localized in the cleft encountered among the models developed here (Fig. 3) can also be important for the activity of these proteins.

The acidic residues involved in the formation of the acidic cleft of the mature tobacco PR-5d are Arg45, Glu85, Asp98, Asp103, and Asp186, and these residues are also found in 1PR (Arg50, Glu92, Asp105, Asp110, and Asp203), 2PR (Arg50, Glu91, Asp104, Asp109, and Asp202), 3PR (Arg50, Glu92, Asp105, Asp110, and Asp203), and 4PR (Arg50, Glu92, Asp105, Asp110, and Asp203). The most important structural differences between the PR5 modeled here are localized in the neighborhood of the acidic cleft (amino acids 40–54 and 174–177).

The extended extruding loops identified as possible epitope regions on the surface of thaumatin (37), which contain a sequence similar to the peptide sweeteners DAA (Asp-Ala-Ala) and DAG (Asp-Ala-Gly), methyl esters important to determine the sweet taste are not present in the PR5 modeled. Therefore, a site for the peptide sweeteners cannot be formed on the surface of those proteins and hence they are devoid of any sweet taste.

The biological significance of the relationship between PR5 isoforms and target microbes needs more investigation. Many pathogen recognition genes, such as the R-genes, undergo rapid adaptive evolution, providing evidence that these genes may play a critical role in plant–pathogen coevolution. Nevertheless, according to our previous studies, there is no evidence of the occurrence of positive selection acting in the PR5 (38).

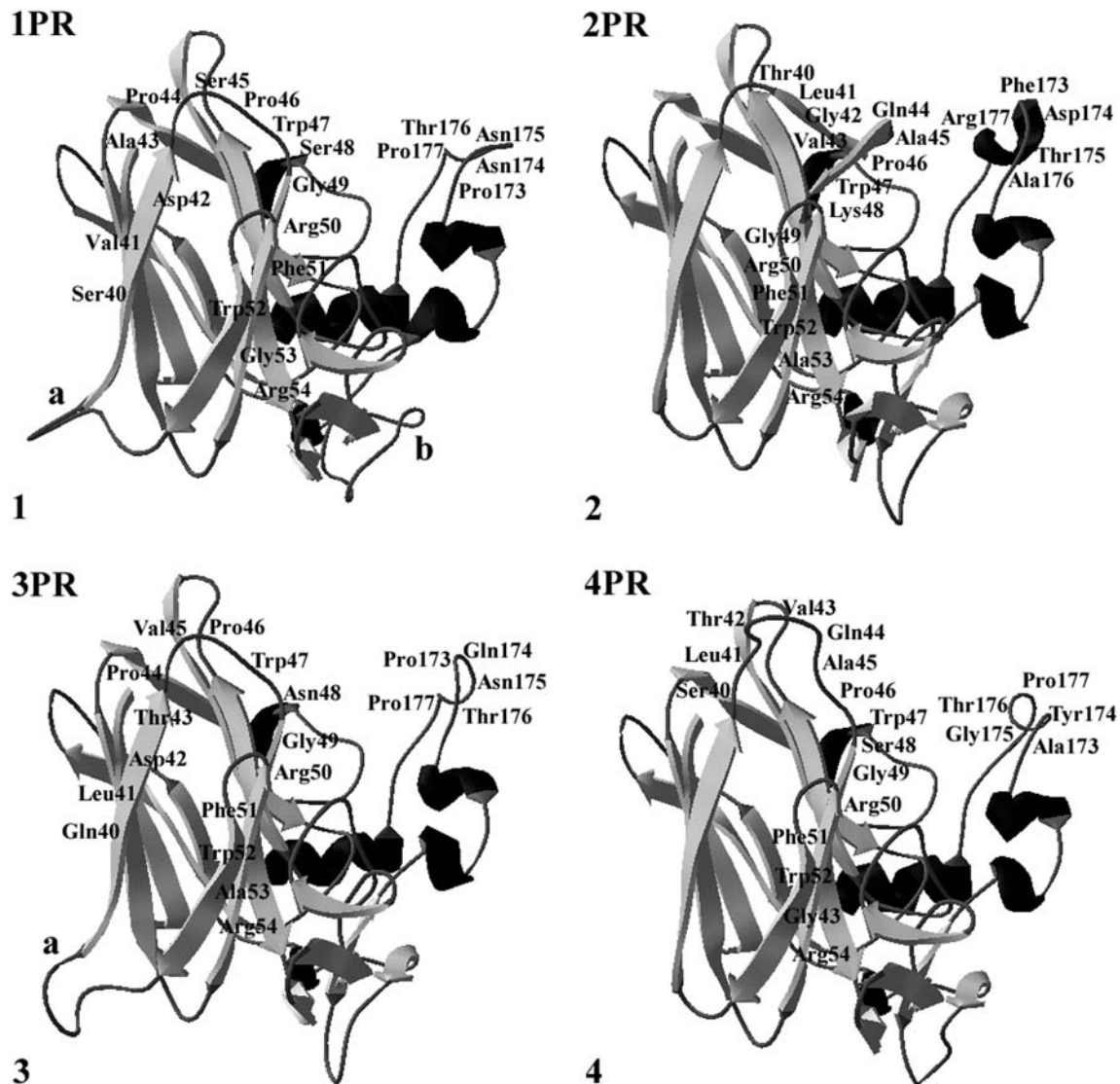


Fig. 4. Ribbon representation of 1PR (1), 2PR (2), 3PR (3), and 4PR (4) models in the same orientation shown in Fig. 3. The position of the amino acids 40-54 and 173-177 are labeled. Secondary structures are marked by different colors: light gray = β -sheets; black = α -helices, and dark gray = coil. In 'a' the presence of an extended loop in the 1PR and 3PR is indicated. Absence of a β strand in 1PR is indicated in 'b'.

ACKNOWLEDGMENTS

This work was supported by Programa de Apoio a Núcleos de Excelência (PRONEX), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Financiadora de Estudos e Projetos (FINEP), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), Pró-Reitoria de Pesquisa da Universidade Federal do Rio Grande do Sul (PROPESQ-UFRGS), and Pró-Reitoria de Pesquisa e Pós-Graduação da Pontifícia Universidade Católica do Rio Grande do Sul (PRPPG-PUCRS).

REFERENCES

1. Van Loon, L. C. and van Strien, E. A. (1999) The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiol. Mol. Plant Pathol.* **55**, 85-97.
2. Van Loon, L. C., Pierpoint, W. S., Boller, T., and Conjero, V. (1994) Recommendations for naming plant pathogenesis-related proteins. *Plant Mol. Biol. Rep.* **12**, 245-265.
3. Park, C. J., An, J. M., Shin, Y. C., Kim, K. J., Lee, B. J., and Paek, K. H. (2004) Molecular characterization of pepper germin-like protein as the novel PR-16 family of pathogenesis-related proteins isolated during the resistance

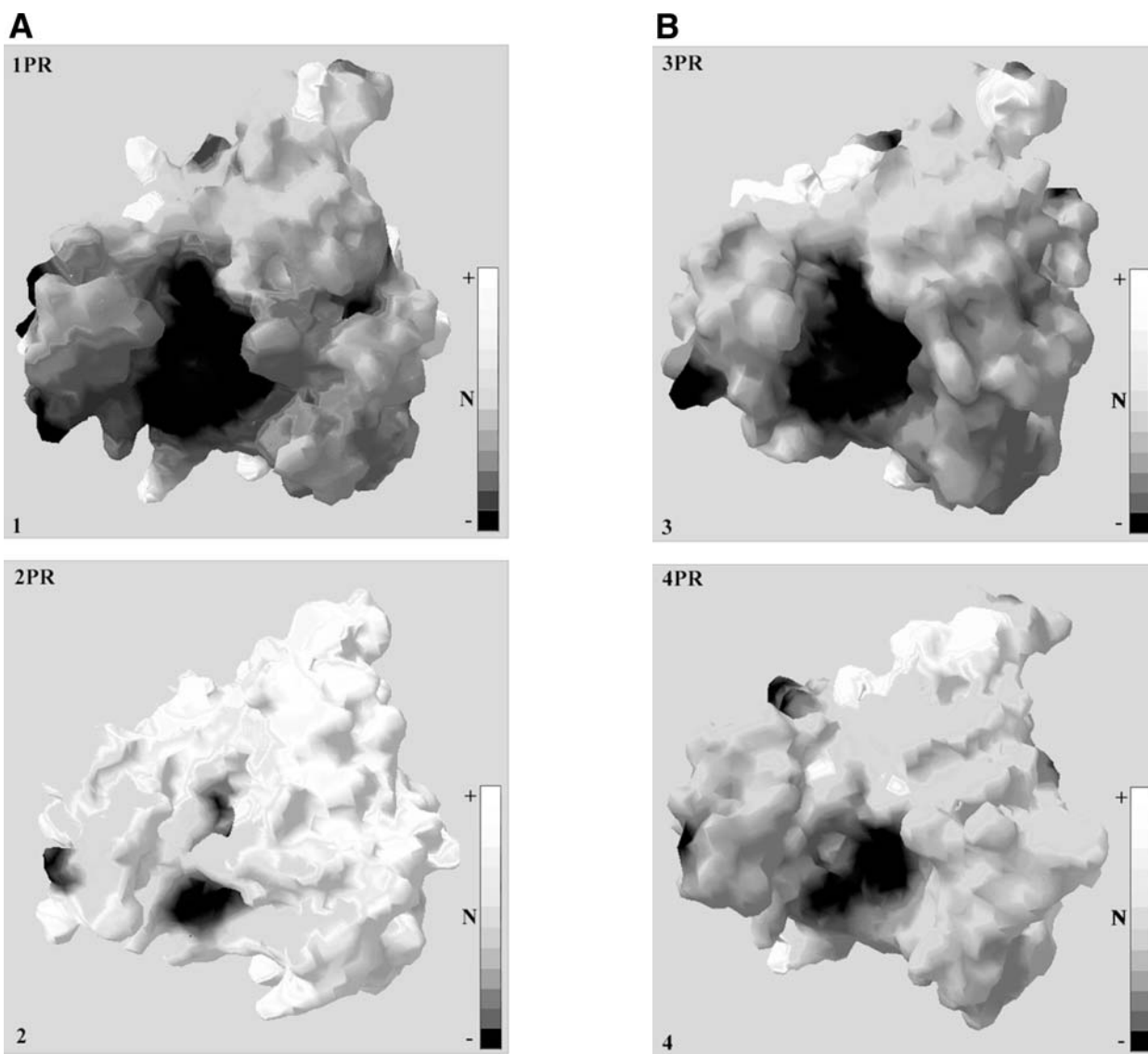


Fig. 5. View of the surface topology of the four PR5 models with the electrostatic potential represented as dark colors (most negative, -), N (neutral), and light colors (most positive, +).

- response to viral and bacterial infection. *Planta* **219**, 797–806.
4. Hoffmann-Sommergruber, K. (2002) Pathogenesis-related (PR)-proteins identified as allergens. *Biochem. Soc. Trans.* **30**, 930–935.
 5. Midoro-Horiuti, T., Goldblum, R. M., Kurosky, A., Wood, T. G., and Brooks, E. G. (2000) Variable expression of pathogenesis-related protein allergen in mountain cedar (*Juniperus ashei*) pollen. *J. Immunol.* **164**, 2188–2192.
 6. Soman, K. V., Midoro-Horiuti, T., Ferreon, J. C., et al. (2000) Homology modelling and characterization of IgE binding epitopes of mountain cedar allergen Jun a 3. *Biophys. J.* **79**, 1601–1609.
 7. Midoro-Horiuti, T., Brooks, E. G., and Goldblum, R. M. (2001) Pathogenesis-related proteins of plants as allergens. *Ann. Allergy Asthma Immunol.* **87**, 261–271.
 8. Breiteneder, H. and Radauer, C. (2004) A classification of plant food allergens. *J. Allergy Clin. Immunol.* **113**, 821–830.
 9. Cortegano, I., Civantos, E., Aceituno, E., et al. (2004) Cloning and expression of a major allergen from *Cupressus arizonica* pollen, Cup a 3, a PR-5 protein expressed under polluted environment. *Allergy* **59**, 485–490.
 10. Stintzi, A., Heitz, T., Prasad, V., et al. (1993) Plant 'pathogenesis-related' proteins and their role in defense against pathogens. *Biochimie* **75**, 687–706.
 11. Van Damme, E. J. M., Charels, D., Menu-Bouaouiche, L., et al. (2002) Biochemical, molecular and structural analysis of multiple thaumatin-like proteins from the elderberry tree (*Sambucus nigra* L.). *Planta* **214**, 853–862.
 12. Vigers, A. J., Wiedemann, S., Roberts, W. K., Legrand, M., Selitrennikoff, C. P. and Fritig, B. (1992) Thaumatin-like

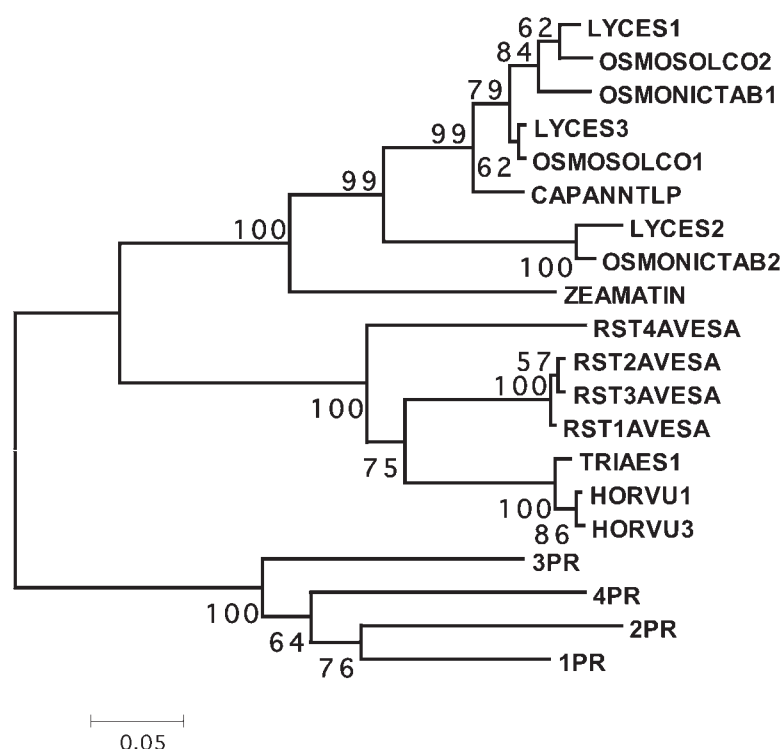


Fig. 6. Phylogenetic tree of the PR5 that were modelled with other TLPs. Numbers refer to bootstrap values.

- pathogenesis-related proteins are antifungal. *Plant Sci.* **83**, 155–161.
13. Yun, D. J., Zhao, Y., Pardo, J. M., et al. (1997). Stress proteins on the yeast cell surface determine resistance to osmotin, a plant antifungal protein. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7082–7087.
 14. Koiwa, H., Kato, H., Nakatsu, T., Oda, J., Yamada, Y., and 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.
 15. Barre, A., Peumans, W. J., Menu-Bouaouiche, L., et al. (2000) Purification and structural analysis of an abundant thaumatin-like protein from ripe banana fruit. *Planta* **211**, 791–799.
 16. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. Z., Miller, W. and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402.
 17. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.
 18. Batalia, M. A., Monzingo, A. F., Ernst, S., Roberts, W. and Robertus, J. D. (1996). The crystal structure of the antifungal protein zeamatin, a member of the thaumatin-like, PR-5 protein family. *Nat. Struct. Biol.* **3**, 19–23.
 19. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weight-
 - ing, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.
 20. Sali, A. and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815.
 21. Martí-Renom, M. A., Stuart, A. C., Fiser, A., Sánchez, R., Melo, F., and Sali, A. (2000) Comparative protein structure modelling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 291–325.
 22. Pearlman, D. A., Case, D. A., Caldwell, J. W., et al. (1995) AMBER, a computer program for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to elucidate the structures and energies of molecules. *Comp. Phys. Commun.* **91**, 1–41.
 23. Laskowski, R. A., McArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
 24. Lüthy, R., Bowie, J. U., and Eisenberg, D. (1992) Assessment of protein models with three-dimensional profiles. *Nature* **356**, 83–85.
 25. Guex, N. and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modelling. *Electrophoresis* **18**, 2714–2733.
 26. Nicholas, K. B. and Nicholas, H. B., Jr (1997) GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the authors (www.psc.edu/bio-med/genedoc).

27. Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425.
28. Kumar, S., Tamura, K., Jacobsen, I. B. and Nei, M. (2001) MEGA: molecular evolutionary genetics analysis. Version 2.1. Distributed by the authors (www.megasoftware.net).
29. Gattiker A., Gasteiger E., and Bairoch A. (2002) ScanProsite: a reference implementation of a PROSITE scanning tool. *Appl. Bioinform.* **1**, 107–108.
30. Min, K., Ha, S. C., Hasegawa, P. M., Bressan, R. A., Yun, D. J., and Kim K. K. (2004) Crystal structure of osmotin, a plant antifungal protein. *Proteins* **54**, 170–178.
31. Krebitz, M., Wagner, B., Ferreira, F., et al. (2003) Plant-based heterologous expression of Mal d 2, a thaumatin-like protein and allergen of apple (*Malus domestica*), and its characterization as an antifungal protein. *J. Mol. Biol.* **329**, 721–730.
32. Garcia-Casado, G., Collada, C., Allona, I., et al. (2000) Characterization of an apoplastic basic thaumatin-like protein from recalcitrant chestnut seeds. *Physiol. Plant* **110**, 172–180.
33. Sassa, H. and Hirano, H. (1998) Style-specific and developmentally regulated accumulation of a glycosylated thaumatin/PR5-like protein in Japanese pear (*Pyrus serotina* Rehd.). *Planta* **205**, 514–521.
34. Fils-Lycaon, B. R., Wiersma, P. A., Eastwell, K. C., and Sautiere, P. (1996) A cherry protein and its gene, abundantly expressed in ripening fruit, have been identified as thaumatin-like. *Plant Physiol.* **111**, 269–273.
35. Osmond, R. I. W., Hrmova M., Fontaine, F., Imbert, A., and Fincher G. B. (2001) Binding interactions between barley thaumatin-like proteins and (1,3)- β -D-glucans. Kinetics, specificity, structural analysis and biological implications. *Eur. J. Biochem.* **268**, 4190–4199.
36. De Vos, A. M., Hatada, M., Van der Wel, H., Krabbendam H., Peerdeman, A. F., and Kim, S. H. (1985) Three-dimensional structure of thaumatin I, an intensely sweet protein. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1406–1409.
37. Sloodstra, J. W., De Geus, P., Haas, H., Verrips, C. T., and Melen, R. H. (1995) Possible active site of the sweet-tasting protein thaumatin. *Chem. Senses* **20**, 535–543.
38. Scherer, N. M., Thompson, C. E., Freitas, L. B., Bonatto, S. L., and Salzano, F. M. (2005) Patterns of molecular evolution in pathogenesis-related proteins. *Genet. Mol. Biol.* **28**, 645–653.