

Specific conformational epitope features of pathogenesis-related proteins mediating cross-reactivity between pollen and food allergens

Jose C. Jimenez-Lopez · Emma W. Gachomo ·
Oluwole A. Ariyo · Lamine Baba-Moussa ·
Simeon O. Kotchoni

Received: 24 December 2010 / Accepted: 23 April 2011
© Springer Science+Business Media B.V. 2011

Abstract Selected members of plant pathogenesis-related and seed storage proteins represent specific groups of proteins with potential characteristics of allergens. Efforts to understand the mechanism by which pathogenesis-related proteins mediate a broad cross-reactivity in pollen-plant food allergens are still limited. In this study, computational biology approach was used to reveal specific structural implications and conservation of different epitopes from members of Bet v 1 and nsLTP protein families mediating

cross-reactivity between pollen and food (fruits, vegetables, legumes, and nut/seeds) allergens. A commonly shared epitope conservation was found among all pollen and food Bet v 1 and nsLTP protein families, respectively. However, other allergenic epitopes were also specifically detected in each family. The implication of these conserved epitopes in a broad cross-reactivity for allergy clinical trials is here discussed.

Keywords Pathogenesis-related proteins · Allergens · Pollen · Plant food · Latex · Cross-reactivity · Conformational epitopes · Protein computational modeling

J. C. Jimenez-Lopez
Department of Biochemistry, Cell and Molecular Biology of
Plants, Estación Experimental del Zaidín, Consejo Superior de
Investigaciones Científicas (CSIC), Profesor Albareda 1,
E-18008 Granada, Spain

E. W. Gachomo
Department of Botany and Plant Pathology, Purdue University,
Lilly Hall, 915 West State Street, West Lafayette, IN 47907,
USA

O. A. Ariyo
Department of Biology, Division of Mathematics and Natural
Sciences, Allen University, Columbia, SC 29204, USA

L. Baba-Moussa
Laboratoire de Biologie et de typage Moléculaire en
Microbiologie, Département de Biochimie et Biologie
Cellulaire, FAST, Université D'Abomey-Calavi, 05BP1604
Cotonou, Benin

S. O. Kotchoni (✉)
Department of Agronomy, Purdue University, Lilly Hall, 915
West State Street, West Lafayette, IN 47907, USA
e-mail: skotchoni@gmail.com; skotchoni@yahoo.com

Present Address:

J. C. Jimenez-Lopez
Department of Biological Sciences, Purdue University, 201 S.
University Street, West Lafayette, IN 47907, USA

Introduction

Defense responses in higher plants are governed by a wide range of preformed or induced mechanisms to adequately respond to different pathogen attacks [1, 2]. These induced defense responses include the synthesis of defense-related proteins especially the pathogenesis-related proteins (PR-proteins). Several of these PR-proteins have also been found to be constitutively expressed in seeds and fruits. In addition to their specific conserved function across plant species, these seed and fruit specific PR-proteins have been reported to function against microbial pathogens and invertebrate pests [3, 4].

The PR-proteins represent a highly conserved protein superfamily composed of 17 families with potential allergen properties [5]. Members of the PR-2 (β -1,3-glucanase), PR-3 (Class I chitinases), PR-4 (chitinases), PR-5 (Thaumatin-like proteins), PR-8 (Class-III chitinases), PR-10 (Bet v 1-like proteins), PR-12 (defensin), PR-14 (lipid-transfer proteins), and PR-15 (oxalate oxidase) families are responsible for the wide cross-reactivity in patients with

pollen allergy, plant-derived food (fruits, vegetables, etc.,) allergy [6], and/or latex allergy [7]. These PR-proteins are mostly responsible for the wide cross-reactivity in pollen, food and latex allergic patients [8].

Members of the PR-10 family, related to the major birch pollen allergen Bet v 1, are ubiquitously distributed among dicotyledonous plants. They are also present in monocots [9] and gymnosperms [10]. PR-10-type allergens also known as panallergens, are most importantly present in many foods including fruits and vegetables. They are members of multifunctional protein family and subfamilies that share homology in tertiary structure despite a low sequence similarity found among them [11]. In addition, there are several isoforms of PR-10 proteins that are known to be tissue specific and/or developmentally regulated. These isoforms are highly expressed in pollen, seeds, fruits, or storage organs.

Another important family of the PR-proteins is the lipid-transfer protein, PR-14. This PR-14 is a small protein of 9 kDa and it widely occurs in the plant kingdom [12, 13]. PR-14 is known to be implicated in cuticle formation and defense mechanism of plants against pathogen attacks. PR-14 and its isoforms (especially, the LTP-lipid complexes) function by recognizing membrane receptors involved in the signaling transduction pathways of local defense responses. They are also responsible for many IgE cross-reactions even between unrelated pollen and plant food allergenic sources [14].

A broad cross-reactivity between allergen proteins in different plant species is generally due to 3D structural conservation of specific domains during evolution of the PR-proteins despite low amino acid sequence identity among them. We therefore hypothesized that members of PR-protein families from related or distant plant species will display various conserved epitopes mediating broad cross-reactivity among potential allergen proteins regardless their primary amino acid sequence divergence. To test our hypothesis, we use in this study a computational biology approach to depict conformational epitopes in different PR-protein families mediating cross-allergy. Computational biology (protein-modeling-based structural reconstruction) is a powerful tool in the post-genomic era to predict protein structures, functions and protein–protein or protein–ligand interactions. Domain organization of proteins is an intrinsic element of protein structure and functionality [15, 16]. In this study, we analyzed the structural relationship between members of Bet v 1-like protein family and nsLTP family responsible for a wide range of cross-reactivity between pollen, plant-derived food (fruits, vegetables, nuts/seeds, etc.,) allergen proteins to depict common and specific conformational epitopes of allergen proteins. In addition, the potential of the wide range of allergenic proteins to interact and/or bind to the MHC

Class-II (human HLA-DR) regions in the antigen sequence by considering the T-cell receptors mediating the allergy cross-reactivity is also discussed.

Materials and methods

The PR-sequence database search

The PR-protein sequences from Bet v 1-like and nsLTP families from different plant species were collected in Uniprot database [17]. Full-length amino acid sequences for PR-proteins were compiled and aligned using ClustalW [18]. This alignment was visualized by Bioedit V7.0.5.3 [19] to find consensual regions, and the percentages of identity between proteins were calculated. Protein sequence consensus were derived from this original alignment, and further analyzed for the presence of putative functional protein areas by using the PROSITE database [20, 21]. The biologically meaningful motif descriptors were derived from multiple alignments using ScanProsite program [22] and the Expert Protein Analysis System (ExPASy) proteomics server of the Swiss Institute of Bioinformatics [23].

PR-proteins: structural analysis and epitope characterization

To understand and compare the structural and molecular conformation between PR-proteins from different plant species, protein sequences of selected members were modeled using the top ten PDB closed template structures by SWISS-MODEL, a protein structure homology-modeling server, via the ExPASy web server [24–26]. Compared to other modeling software platforms, SWISS-MODEL is one of the best servers for protein structural prediction using the available PDB templates in the database. An initial structural model was generated for the different PR-protein members and subjected to energy minimization with GROMOS96 force field energy [27] implemented in DeepView/Swiss-PDBViewer v3.7 [28] to improve the van der Waals contacts and correct the stereochemistry of the model. For each sequence analyzed, the quality of the model was assessed by checking the protein stereology with PROCHECK [29] and the protein energy with ANOLEA [30]. The Ramachandran plot statistics for the models were calculated to show the number of protein residues in the favored regions. The 2D protein structural analysis, protein superimpositions and surface protein contours analysis were performed and visualized in PyMol 2006 software [31].

The Electrostatic Poisson-Boltzmann (PB) potentials for all the structures were obtained using APBS [32] molecular modeling software in PyMol 0.99 (DeLano Scientific LLC)

with AMBER99 [33] to assign the charges and radii to all of the atoms (including hydrogens), which were added and optimized with the Python software package PDB2PQR [34]. Fine grid spaces of 0.35 Å were used to solve the linearized PB equation in sequential focusing multigrid calculations in a mesh of 130 points per dimension at 310.00 K. The dielectric constants were two for the protein and 80.00 for water. The output mesh was processed in the scalar OpenDX format to render isocontours and maps onto the surfaces with PyMOL 0.99. Potential values are given in units of kT per unit charge (k Boltzmann's constant; T temperature).

The prediction of MHC Class-II binding regions in the antigen sequence, the T-cell receptors mediating the allergy cross-reactivity, was performed by using neuronal networks and quantitative matrices derived from published literature. The analysis was made by using the TEPITOPE software [35], with a threshold of 5% for the most common human HLA-DR alleles (DR1, DR2, DR3, DR4, DR5, DR7, and DR8) covering a large proportion of the peptides that bind with human HLA.

Results and discussion

Cross-reactive allergens in PR-10-type protein homologs: Bet v 1 homologous proteins

So far, 28 major groups of cross-reactive proteins have been identified. Six out of the 28 major groups represent the PR-proteins. PR-proteins, including the Bet v 1-related

proteins can be found in the pollen of Fagales trees and in foods from various botanical families [36]. They are responsible for adverse reactions upon interaction with both pollen and food allergen sources. Thus, this recognition (IgE dependent) defines the family of Bet v 1-related proteins as a cross-reactivity cluster. Therefore, this molecular definition cannot solely rely on sequence homology, but will require also structural characterization and experimental studies. This molecular classification offers the possibility to explain cross-reactivity to multiple pollen and pollen-food allergens [37]. For instance, the Bet v 1 superfamily is one of the allergenic proteins responsible for pollen-food allergy syndrome.

The first 3D structure of a protein from the Bet v 1 superfamily (PFAM accession number PF00407) ever reported was Bet v 1.57. The structural comparison studies have demonstrated that Bet v 1 folding structure is also found in diverse protein families outside the plant kingdom [38, 39]. This structural folding relationship could never have been detected by means of primary sequence comparison, due the low sequence similarity between the proteins. As shown in Fig. 1a,b it is easy to notice that various homologous allergens coming from different plant sources (pollen, fruits, legumes, nuts/seeds) display similar 3D structural features within each family despite the highly variable low range of primary amino acid sequence identity between each other (Fig. 2a). The reason why these proteins can easily produce cross-reactivity between themselves and with other allergenic proteins of close-related or unrelated plant species is the similarity in the 3D folding

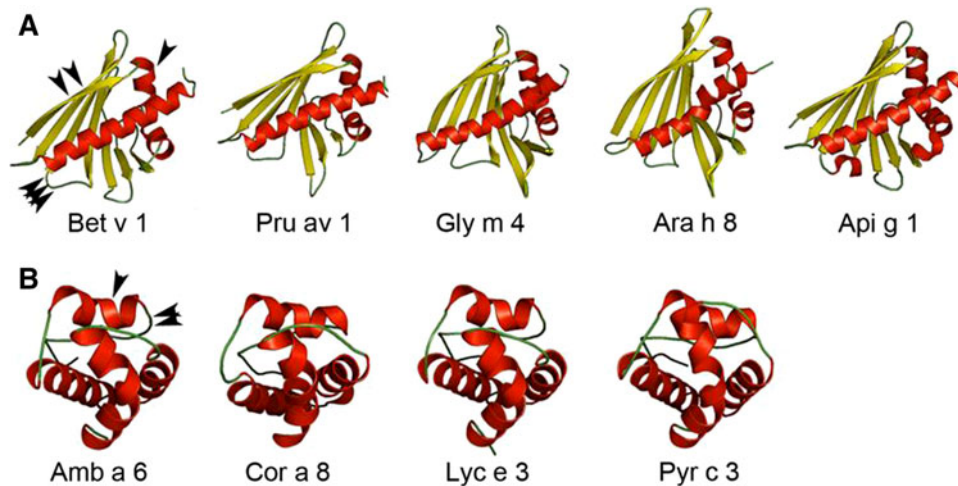


Fig. 1 Three-dimensional structures of selected members of Bet v 1 and nsLTPs superfamilies. **a** Structures show a typical alpha/beta fold responsible for IgE cross-reactivity among related and unrelated species. Secondary structure elements including α -helices (arrow head), β -sheets (double arrow head), and loops and turns (triple arrow head) are depicted. Members of Bet v 1 superfamily include representatives of pollen (Bet v 1, *Betula verrucosa*), fruits (Pru av 1, *Prunus avium*), legumes (Gly m 4, *Glycine max*), nuts/seeds (Ara h

8, *Arachis hypogea*), and vegetables (Api g 1, *Apium graveolens*). **b** nsLTPs structural features sharing a common fold composed of four α -helices (arrow head) and stabilized by four disulfide bonds (double arrow head) to form a central tunnel for ligand interaction. nsLTP proteins include representatives of pollen (Amb a 6, *Ambrosia artemisiifolia*), fruits (Pyr c 3, *Pyrus communis*), nuts/seeds (Cor a 8, *Corylus avellana*), and vegetables (Lyc e 3, *Lycopersicon esculentum*)

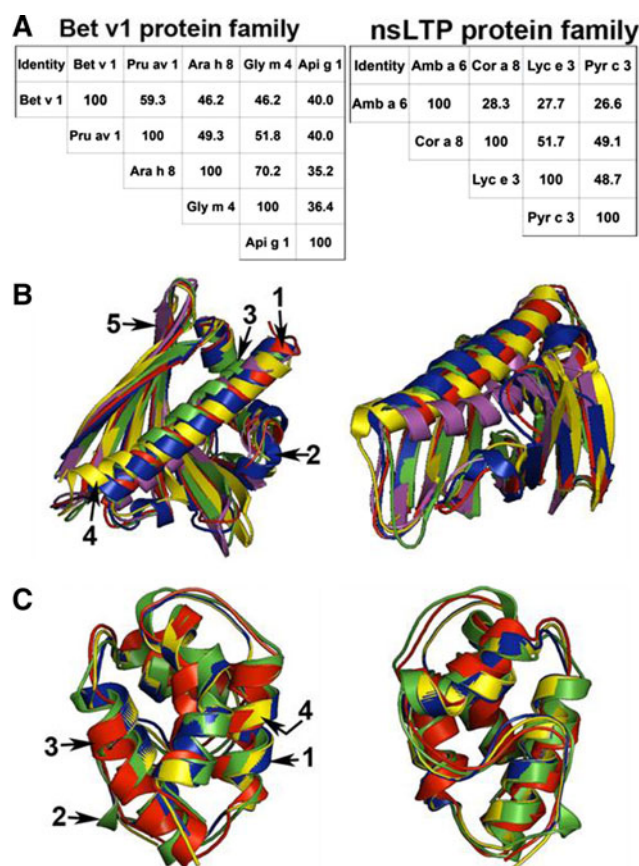


Fig. 2 Structural superimposition of selected members of Bet v 1 and nsLTPs superfamilies. **a** Amino acid sequence identity (in percentage) between members of Bet v 1 (Bet v 1, Pru av 1, Gly m 4, Ara h 8, and Api g 1) and nsLTPs (Amb a 6, Pyr c 3, Cor a 8, and Lyc e 3) superfamilies, respectively. **b** Protein structures superimposition between all selected members of the Bet v 1 superfamily. Each protein is specifically numbered: Bet v 1 (depicted by 1), Api g 1 (2), Pru av 1 (3), Gly m 4 (4), Ara h 8 (5). **c** Protein structures superimposition between all selected members of the nsLTP superfamily. Each protein is depicted by specific number: Amb a 6 (1), Pyr c 3 (2), Cor a 8 (3), and Lyc e 3 (4)

structural features (as demonstrated by structural superimposition of different proteins) of these proteins across diverse plant species as shown in Fig. 2b,c. This was also confirmed by the work of Sankian et al. [40] demonstrating that amino acid primary sequence homology is a poor predictive value for cross-reactivity between allergens. In addition, the work of Aalberse et al. [41] suggested that accurate knowledge of the full 3D structures of targeted protein(s) and related allergens would clearly be an advantage to better understand the molecular basis of cross-reactivity between allergens. The superimposition has indeed shown very similar 3D structures, with small spatial variations in some 2D structural elements, such as loops or turns in the proteins or the long N-terminal helix in particular allergenic members of Bet v 1 like Ara h 8 (magenta) (Fig. 2b).

These similar 3D structures containing common conformational epitopes often leads to immediate-type (immune) reactions. These immune reactions are triggered by antigen-mediated cross-linking of IgE antibodies attached to the specific receptor on the surface of sensitized cells [42]. Without this cross-bridge formation between the IgE antibodies, allergic symptoms are not provoked even if an antigen is recognized by an IgE antibody on the membrane. The variable region of an IgE antibody recognizes the epitope of an antigen that generally exists on the exterior. Moreover, these structures would be conserved in the course of evolution regardless of the species. Generally, the extensive cross-reactivity is expected when such a conserved structure supplies epitopes for IgE antibodies [43].

Reasonably validated mathematical models for the prediction of core and binding properties of these peptides are few. These methods for core epitope prediction are generally based on genetic programming, discriminant analysis and matrix optimization techniques. The methods for the prediction of binders are based on motif, quantitative matrices and artificial neural networks. In this study, we have used a highly accurate prediction method based on virtual matrix. The matrix optimization technique has been used to detect the binding core of peptides to the MHC Class-II alleles. Subsequently support vector machine has been used for discrimination between binding and non-binding peptides. Unlike in the other mathematical approaches, this methodology is not limited by the polymorphism of the MHC Class-II alleles. In order to better understand the structural feature mediating cross-reactivity in the allergenic proteins, we have highlighted various specific epitopes mediating cross-reactivity between the allergenic proteins (Fig. 3) and refined the 3D conformational structures of the epitopes of Bet v 1 protein homologues (Fig. 4). Using computational biology, we have predicted and analyzed the binding regions for the most common MHC-II allelic receptors for T-cells, based on quantitative matrices and neuronal networks covering a large number of binding-peptides for human HLA-DR Class-II. In order to easily see and understand the specific epitope features, we depicted common epitopes of Bet v 1 homologues that are generally responsible for cross-reactivity in different colors (Figs. 3, 4). Bet v 1 provides common epitopes that are at the root of this extensive cross-reactivity to pollen and various vegetable foods (Figs. 3, 4). In addition, many of these PR-protein exhibit particular epitopes in black color (brown color for Bet v 1 allergen), which can potentially induce adverse symptoms in many atopic patients, or can be the cause of cross-reactivity with other allergens.

Previously, it has been confirmed that cross-reactivity of Bet v 1 with the major apple allergen Mal d 1 occurs not

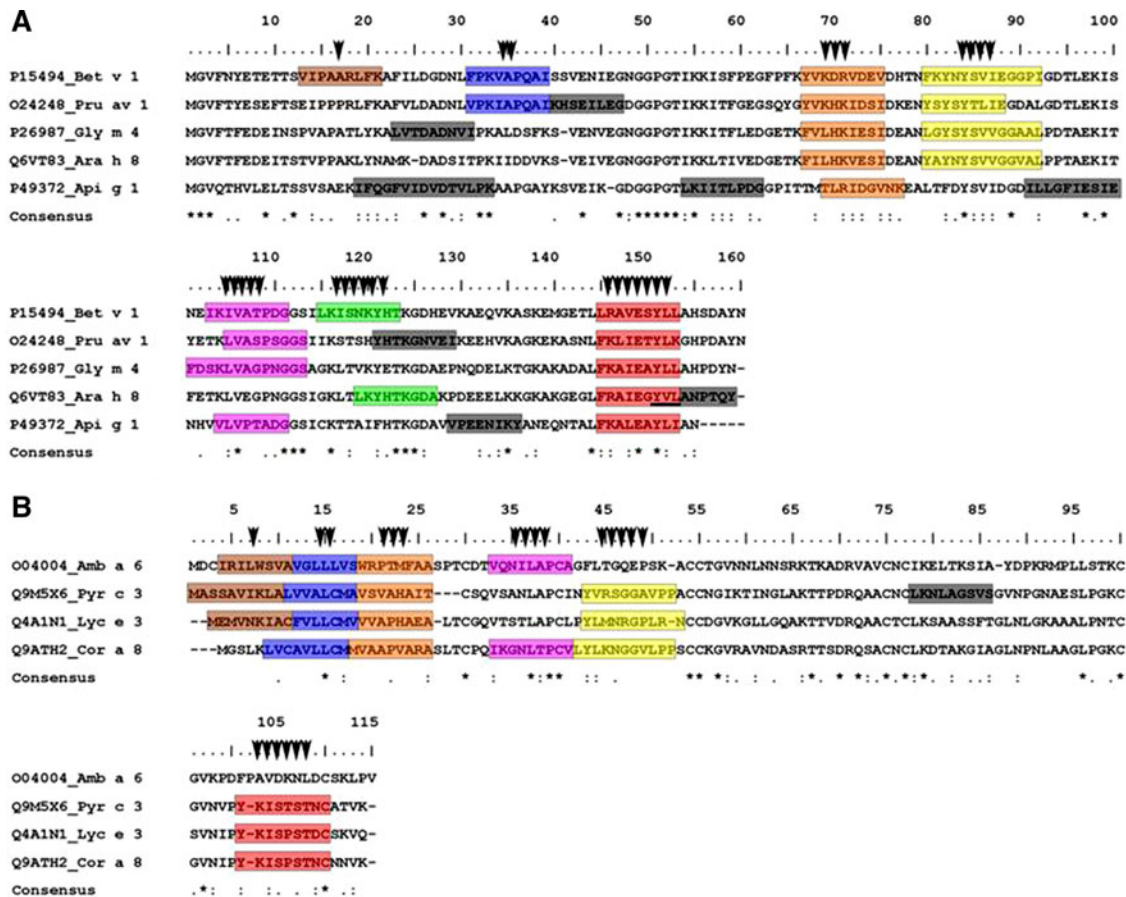


Fig. 3 Sequence alignment of selected members of the Bet v 1 and nsLTP superfamilies. **a** Alignment of pollen Bet v 1 (*Betula verrucosa*) protein allergen with Pru av 1 (*Prunus avium*), Gly m 4 (*Glycine max*), Ara h 8 (*Arachis hypogea*), and Api g 1 (*Apium graveolens*) proteins. An asterisk indicates amino acid identity, double point indicates conservative position (amino acid with the same properties), and a point indicates a semi-conservative position, whereas a *blank space* indicates a gap introduced to maximize the alignment. Different binding regions for the most common MHC-II allelic receptors for T-cells are highlighted by specific arrow head(s).

A specific epitope for Bet v 1 protein (arrow head), and in *black* color for the rest of the protein allergens. Common epitopes are subsequently depicted by indicated arrow heads (double, triple, quadruple, quintuple, etc.). **b** Alignment of pollen Amb a 6 (*Ambrosia artemisiifolia*) protein allergen with Pyr c 3 (*Pyrus communis*), Cor a 8 (*Corylus avellana*) and Lyc e 3 (*Lycopersicon esculentum*) proteins. Different binding regions for the most common MHC-II allelic receptors for T-cells are highlighted by the same arrow head(s) parameters. A specific epitope is colored in *black* for Pyr c 3 protein, and common epitopes are depicted by indicated arrow heads

only at the B-cell but also at the T-cell level [44]. Bet v 1 also contains the major T-cell reactivating region of Api g 1, confirming that Bet v 1 is responsible for initializing an allergic response to the major celery allergen [45]. In addition, it appears that in legumes and nuts/seeds, the epitopes of the hazelnut allergen Cor a 1.04 are less related to the hazel pollen allergen Cor a 1 than they are to Bet v 1 from birch pollen [46]. In addition, Bet v 1 homologous allergen Gly m 4 of soybean was found to be responsible for inducing an oral allergy syndrome of extraordinary severity and severe systemic reactions in individuals with birch pollen allergy [47].

Others Bet v 1 homologous proteins capable of binding anti-Bet v 1 IgE were described in unrelated plants like Rosaceous fruits as Pru av 1 from sweet cherry [48], Pru ar

1 from apricot (EMBL/GenBank AF020784 and AF134731), and Pyr c 1 from pear (EMBL/GenBank AF057030) and the *Apiaceae* vegetables (e.g., celery Api g 1, carrot Dau c 1) [49, 50]. Moreover, it worth mentioning that the IgE cross-reactivity between evolutionary related plants are characterized by homologous molecules of Bet v 1 found in pollen of Fagales trees such as alder Aln g 1, hornbeam Car b 1, chestnut Cas s 1, hazel Cor a 1, beech Fag s 1, and oak Que a 1 [51].

Allergens homologous to PR-14 type proteins: nsLTP proteins

Allergenic nsLTPs have been identified in pollen of trees and weeds, in plant food and latex, but they have not been

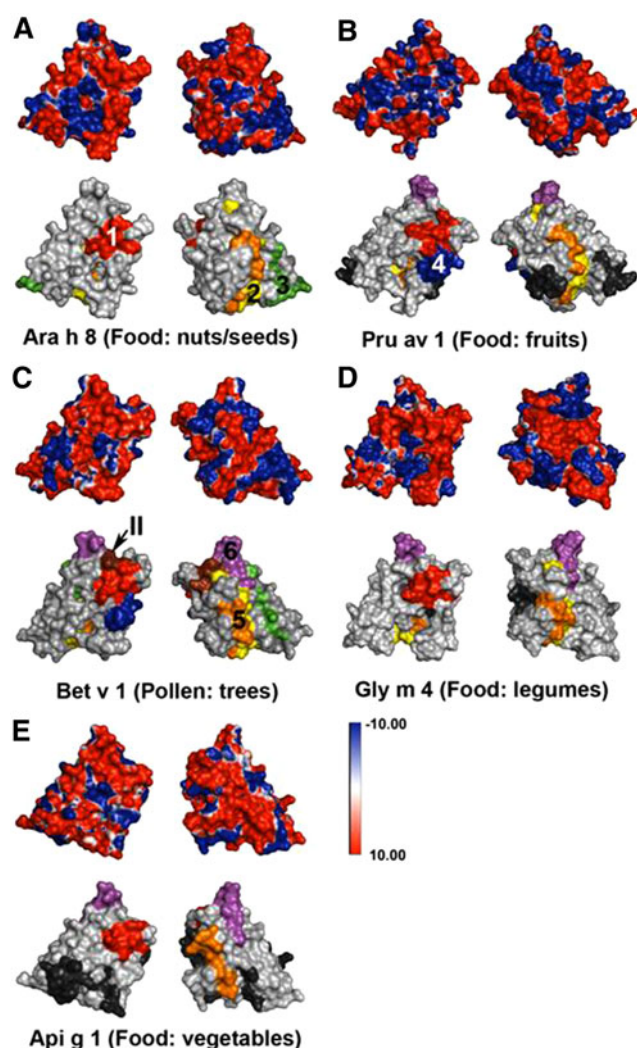


Fig. 4 Surface protein characterization and potential antigenic region prediction of selected members of Bet v 1 superfamily. The electrostatic Poisson-Boltzmann (PB) potential distribution of the surface of Ara h 8, Pru av 1, Bet v 1, Gly m 4, and Api g 1 structures are depicted (**a–e upper panels**). The surface electrostatic charges are clamped at -10 and $+10$ values, where -10 means negatively charged residues and $+10$ means positively charged residues. Both structural representations are showing 180° vertical rotation views for each protein. The surface potential binding MHC Class-II (human HLA-DR) regions are highlight in numbers (1–6) for common epitopes, but in roman number (II) for Bet v 1, and in *black* for specific epitopes (**a–e lower panels**)

identified in grass pollen. The allergenic potential of nsLTPs is related to the cellular localization (cell wall fruit peels) as well as their thermal-stability and proteolysis resistance. This knowledge is especially important for sensitized allergic patients as they can reduce the risk of severe allergic reactions by avoiding certain kind of fruits or by consuming peeled-off fruits. This is also an important point concerning sensitization because nsLTPs can act as true food allergens with the capacity to induce severe

symptoms by surviving food processing and the harsh environment of the gastrointestinal tract [52], due to their high resistance to heat and proteolysis. The thermal stability of LTPs explains their behavior as allergens in processed drinks and foods such as peach juice, beer, wine or roasted hazelnuts.

Another reason for potential allergenic features of nsLTP is their 3D structural common characteristics, such as common fold of four α -helices stabilized by four disulfide bonds that form a central hydrophobic tunnel interacting with lipid molecules. This structural similarity is shown in different members of nsLTP (Amb a 6, Cor a 8, Lyc e 3 and Pyr c 3) as depicted in Fig. 1b. The superimposition of these four allergenic structures (Fig. 2c) displayed a very similar 3D structure, with small spacial variations occurring almost exclusively in loops or turns in the proteins (Fig. 2c). Surprisingly, the sequence identity between these proteins is very low (Fig. 2a).

Pollen or food nsLTPs could act as primary sensitizers by oral route. Cross-reactivity studies of nsLTPs have revealed that most *Rosaceae*-allergic and nsLTP mono-sensitized patients experience adverse reaction after ingestion of botanically unrelated plant-derived foods as well. In this regard, the most frequently reported causes of allergic symptoms were nuts (hazelnut, walnut, and peanut). Other sources of allergenic nsLTPs are from pollen (mugwort, Art v 3; ambrosia, Amb a 6; olive, Ole e 7; hazelnut, Cor a 8); fruits (grape, Vit v 1); vegetables (maize, Zea m 14; asparagus, Aspa o 1), and latex (*H. brasiliensis*, Hev b 12) [8]. A parallel study with Bet v 1 proteins superfamily was made to analyze the T-cell antigenic regions (conformational epitopes) that are responsible for the extensive cross-reactivity. The most common epitopes (depicted in brown, blue, and orange colors), as shown in Fig. 3 were, however, mostly absent in the 3D protein structures (Fig. 5). The absence of these epitopes mediating cross-reactivity in the 3D modeled proteins (Fig. 5) is due to the fact that the *N*-terminal regions of the best crystallographic template structures used to model the different LTPs protein allergens were not present in each individual PDB dataset. However, it is important to highlight that nsLTP homologues share other common epitopes as depicted in red, yellow, and magenta colors, or a specific epitope shown in black color for the Pyr c 3 allergen (Fig. 5). These epitopes are the primary cause of cross-reactivity between nsLTP protein allergens from different plants. They are also responsible for the pollen-food syndrome.

Conclusion and future perspectives in allergy therapy

A number of plant-derived allergens and PR-like proteins can be found in many plant species. These proteins are

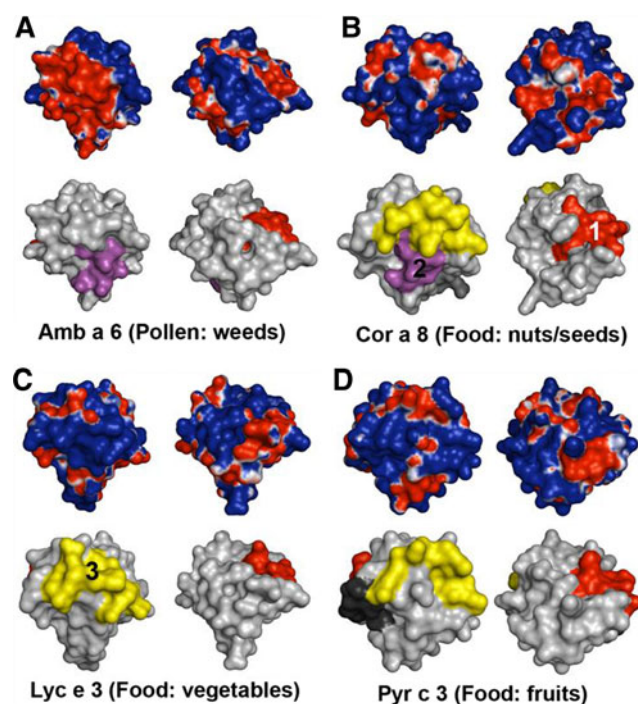


Fig. 5 Surface protein characterization and potential antigenic region prediction of selected members of nsLTP superfamily. The electrostatic Poisson-Boltzmann (PB) potential distribution of the surface of Amb a 6, Cor a 8, Lyc e 3, and Pyr c 3 structures are depicted (**a–d upper panels**). The surface electrostatic charges are clamped at -10 and $+10$ values, where -10 means negatively charged residues and $+10$ means positively charged residues. Both structural representations are showing 180° vertical rotation views for each protein. The surface potential binding MHC Class-II (human HLA-DR) regions are highlight in numbers (1–3) for common epitopes, and in black color for specific epitopes (**a–d lower panels**)

responsible for cross-reactions between different allergenic sources.

Our findings indicate that the conserved PR-protein epitopes, rather than the primary amino sequence identity characteristics, are the major factors of cross-reactivity among allergens. A proper identification and characterization of these specific epitopes is therefore critical for efficient clinical trials, allergy diagnosis and immunotherapy. Our data demonstrate that computational biology mediating 3D structural analysis is a powerful and an important approach to predict and highlight hidden features of common and specific epitopes within related and diverged PR-proteins mediating cross-reactivity in allergy. In summary, our data suggest that computational biology is a post genomic approach to additionally improve allergy therapy and diagnosis. This modeling based methodology can assist in proper characterization of specific epitope features and patient-tailored immunotherapeutic prediction by engineer recombinant allergens for the ultimate goal of improving patient treatment. This approach will significantly alleviate the problems associated with current allergy cross-therapy.

References

- Gachomo WE, Shonukan OO, Kotchoni SO (2003) The molecular initiation and subsequence acquisition of disease resistance in plants. *Afr J Biotechnol* 2(2):26–32
- Gachomo WE, Seufferheld MJ, Kotchoni SO (2010) Melanization of appressoria is critical for the pathogenicity of *Diplocarpon rosae*. *Mol Biol Rep* 37(7):3583–3591
- Hoffmann-Sommergruber K (2000) Plant allergens and pathogenesis-related proteins: what do they have in common? *Int Arch Allergy Immunol* 122:155–166
- Kotchoni SO, Gachomo WE (2006) The reactive oxygen species network pathways: an essential prerequisite for perception of pathogen attack and disease resistance in plants. *J Biosci* 31(3): 389–404
- van Loon LC, Rep M, Pieterse CMJ (2006) Significance of inducible defense-related proteins in infected plants. *Annu Rev Phytopathol* 44:135–162
- Hofmann A, Burks AW (2008) Pollen food syndrome: update on the allergens. *Curr Allergy Asthma Rep* 8(5):413–417
- Blanco C (2003) Latex-fruit syndrome. *Curr Allergy Asthma Rep* 3(1):47–53
- Hoffmann-Sommergruber K (2002) Pathogenesis-related (PR)-proteins identified as allergens. *Biochem Soc Trans* 30(6):930–935
- Wang CS, Huang JC, Hu JH (1999) Characterization of two subclasses of PR-10 transcripts in lily anthers and induction of their genes through separate signal transduction pathways. *Plant Mol Biol* 40:807–814
- Yu XS, Ekramoddoullah AKM, Misra S (2000) Characterization of Pin m III cDNA in western white pine. *Tree Physiol* 20:663–671
- Pasternak O, Bujacz GD, Fujimoto Y, Hashimoto Y, Jelen F, Otlewski J, Sikorska MM, Jaskolski M (2006) Crystal structure of *Vigna radiata* cytokinin-specific binding protein in complex with zeatin. *Plant Cell* 18:2622–2634
- Pastorello EA, Pravettoni V, Farioli L, Spanio M, Fortunato D, Monza M, Giuffrida MG, Rivolta F, Scibola E, Ansaloni R, Incorvaia C, Conti A, Ortolani C (1999) Clinical role of a lipid transfer protein that acts as a new apple-specific allergen. *J Allergy Clin Immunol* 104:1099–1106
- Asero R, Mistrello G, Roncarolo D, de Vries SC, Gautier M-F, Ciurana CLF, Verbeek E, Mohammadi T, Knul-Brettlova V, Akkerdaas JH, Bulder I, Aalberse RC, van Ree R (2000) Lipid transfer protein: a pan-allergen in plant-derived foods that is highly resistant to pepsin digestion. *Int Arch Allergy Immunol* 122:20–32
- Sanchez-Monge R, Blanco C, Lopez-Torres G, Cumplido J, Recas M, Figueroa J, Carrillo T, Salcedo G (2006) Differential allergen sensitization patterns in chestnut allergy with or without associated latex-fruit syndrome. *J Allergy Clin Immunol* 118:705–710
- Kotchoni SO, Jimenez-Lopez JC, Gao D, Edwards V, Gachomo WE, Margam VM, Seufferheld MJ (2010) Modeling-dependent protein characterization of the rice aldehyde dehydrogenase (ALDH) superfamily reveals distinct functional and structural features. *PLoS One* 5(7):e11516
- Kotchoni SO, Jimenez-Lopez JC, Gachomo WE, Seufferheld MJ (2010) A new and unified nomenclature for male fertility restorer (RF) proteins in higher plants. *PLoS One*. doi:10.1371/journal.pone.0015906
- Apweiler R, Bairoch A, Wu CH (2004) Protein sequence databases. *Curr Opin Chem Biol* 8:76–80
- Thompson JD, Higgins DG, Gibson TJ (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* 22:4673–4680
19. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids* 41:95–98
 20. Sigrist CJA, Cerutti L, Hulo N, Gattiker A, Falquet L, Pagni M, Bairoch A, Bucher P (2002) PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief Bioinform* 3:265–274
 21. Sigrist CJA, Cerutti L, de Castro E, Langendijk-Genevaux PS, Bulliard V, Bairoch A, Hulo N (2010) PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res* 38:161–166
 22. de Castro E, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A, Hulo N (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res* 34:362–365
 23. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 31:3784–3788
 24. Peitsch MC (1995) Protein modelling by e-mail. *Nat Biotechnol* 13:658–660
 25. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modeling. *Bioinformatics* 22:195–201
 26. Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T (2009) The SWISS-MODEL repository and associated resources. *Nucleic Acids Res* 37:D387–D392
 27. Christen M, Hünenberger PH, Bakowies D, Baron R, Bürgi R, Geerke DP, Heinz TN, Kastenholz MA, Kräutler V, Oostenbrink C, Peter C, Trzesniak D, van Gunsteren WF (2005) The GROMOS software for biomolecular simulation: GROMOS05. *J Comput Chem* 26(16):1719–1751
 28. Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18(15):2714–2723
 29. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereo-chemical quality of protein structures. *J Appl Cryst* 26:283–291
 30. Melo F, Feytmans E (1998) Assessing protein structures with a non-local atomic interaction energy. *J Mol Biol* 277(5):1141–1152
 31. DeLano WL (2002) The PyMOL molecular graphics system. DeLano Scientific, San Carlos
 32. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98:10037–10041
 33. Wang J, Cieplak P, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 21:1049–1074
 34. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, Baker NA (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucl Acids Res* 35:522–525
 35. Zhao B, Mathura VS, Rajaseger G, Mochhala S, Kishore M (2003) A novel MHCp binding prediction model. *Hum Immunol* 64(12):1123–1143
 36. Breiteneder H, Ebner C (2000) Molecular and biochemical classification of plant-derived food allergens. *J Allergy Clin Immunol* 106:27–36
 37. Radauer C, Breiteneder H (2007) Evolutionary biology of plant food allergens. *J Allergy Clin Immunol* 120:518–525
 38. Neudecker P, Schweimer K, Nerkamp J, Scheurer S, Vieths S, Sticht H, Rösch P (2001) Allergic cross-reactivity made visible: solution structure of the major cherry allergen Pru av 1. *J Biol Chem* 276:22756–22763
 39. Aalberse RC (2006) Structural features of allergenic molecules. *Chem Immunol Allergy* 91:134–146
 40. Sankian M, Varasteh A, Pazouki N, Mahmoudi M (2005) Sequence homology: a poor predictive value for profilins cross-reactivity. *Clin Mol Allergy* 3:13
 41. Aalberse RC (2007) Assessment of allergen cross-reactivity. *Clin Mol Allergy* 5:2
 42. Huby RDJ, Dearman RJ, Kimber I (2000) Why are some proteins allergens? *Toxicol Sci* 55:235–246
 43. Sicherer SH (2001) Clinical implications of cross-reactive food allergens. *J Allergy Clin Immunol* 108:881–890
 44. Fritsch R, Bohle B, Vollmann U, Wiedermann U, Jahn-Schmid B, Krebitz M, Breiteneder H, Kraft D, Ebner C (1998) Bet v 1, the major birch pollen allergen, and Mal d 1, the major apple allergen, cross-react at the level of allergen-specific T helper cells. *J Allergy Clin Immunol* 102:679–686
 45. Bohle B, Radakovics A, Jahn-Schmid B, Hoffmann-Sommergruber K, Fischer GF, Ebner C (2003) Bet v 1, the major birch pollen allergen, initiates sensitization to Api g 1, the major allergen in celery: evidence at the T cell level. *Eur J Immunol* 33:3303–3310
 46. Lüttkopf D, Müller U, Skov PS, Ballmer-Weber BK, Wüthrich B, Skamstrup Hansen K, Poulsen LK, Kästner M, Hausteine D, Vieths S (2002) Comparison of four variants of a major allergen in hazelnut (*Corylus avellana*) Cor a 1.04 with the major hazel pollen allergen Cor a 1.01. *Mol Immunol* 38:515–525
 47. Kleine-Tebbe J, Vogel L, Crowell DN, Hausteine UF, Vieths S (2002) Severe oral allergy syndrome and anaphylactic reactions caused by a Bet v 1-related PR-10 protein in soybean, SAM22. *J Allergy Clin Immunol* 110:797–804
 48. Scheurer S, Metzner K, Hausteine D, Vieths S (1997) Molecular cloning, expression and characterization of Pru a 1, the major cherry allergen. *Mol Immunol* 34:619–629
 49. Breiteneder H, Hoffmann-Sommergruber K, O’Riordain G, Susani M, Ahorn H, Ebner C, Kraft D, Scheiner O (1995) Molecular characterization of Api g 1, the major allergen of celery (*Apium graveolens*) and its immunological and structural relationships to a group of 17 kDa tree pollen allergens. *Eur J Biochem* 233: 484–489
 50. Hoffmann-Sommergruber K, O’Riordain G, Ahorn H, Ebner C, Laimer Da Camara Machado M, Pühringer H, Scheiner O, Breiteneder H (1999) Molecular characterization of Dau c 1, the Bet v 1 homologous protein from carrot and its cross-reactivity with Bet v 1 and Api g 1. *Clin Exp Allergy* 29:840–847
 51. Hauser M, Egger M, Wallner M, Wopfner N, Schmidt G, Ferreira F (2008) Molecular properties of plant food allergens: a current classification into protein families. *Open Immunol J* 1:1–12
 52. Zuidmeer L, van Ree R (2007) Lipid transfer protein allergy: primary food allergy or pollen/food syndrome in some cases. *Curr Opin Allergy Clin Immunol* 7:269–273