Allergology International. 2013;62:447-462
DOI: 10.2332/allergolint.12-OA-0522

ORIGINAL ARTICLE

Characterization and Gene Cloning of an Acidic Thaumatin-Like Protein (TLP 1), an Allergen from Sapodilla Fruit (*Manilkara zapota*)

Hassan G Ashok Kumar[1], Venkatesh L Hegde[1,2], Santoshkumar M Shetty[1,3] and Yeldur P Venkatesh[1]

ABSTRACT

Background: Allergy to sapodilla (*Manilkara zapota*) fruit ingestion is rare. An independent study from our group has identified a basic thaumatin-like protein (TLP 2) as the major allergen. The present study was aimed at identifying and characterizing additional allergens from sapodilla.

Methods: Allergic subjects were identified by case history, skin prick test (SPT) and allergen-specific IgE. Sapodilla extract was fractionated using SP-Sepharose into 3 components (SP1, SP2 and SP3) which were analyzed by native/SDS-PAGE, IgE-immunoblot, isoelectric focusing (IEF) and N-terminal sequencing. The conserved regions of plant TLPs and the N-terminal sequence were used to design primers for PCR.

Results: SPT and ELISA confirmed a subject with oral allergy syndrome (OAS) to sapodilla and custard apple. Two proteins (26.9 and 24.5 kDa; reducing conditions) were detected as allergens, of which the latter in SP2 has already been identified as basic TLP (TLP 2). The 26.9 kDa protein present in SP1 was identified as an acidic TLP based on native PAGE, IEF and N-terminal sequencing. Presence of a basic β-1,3-glucanase in SP3 was inferred by zymography. Sequence analysis of the genomic clone of the acidic TLP gene revealed that it is intronless and non-glycosylated. Evolutionary relatedness to olive, grape and kiwi fruit allergenic TLPs were inferred by phylogenetic analysis.

Conclusions: An acidic TLP (TLP 1) was identified as a new allergen in sapodilla. TLP 1 is a single polypeptide (207 residues) belonging to the thaumatin family of the GH64-TLP-SF superfamily. Clinically, sapodilla should be considered in the list of fruits causing OAS.

KEY WORDS

acidic TLP, exotic fruit, food allergen, *Manilkara zapota*, oral allergy syndrome

INTRODUCTION

Sapodilla (*Manilkara zapota* syn. *Achras zapota*), also known as sapota, chikoo and by various other names, yields delicious fruits that are round, ovoid, ellipsoid or conical, flattened at the stem end, 5 to 10 cm wide.[1,2] The fruit is highly perishable and sensitive to cold storage; hence, the bulk of the produce is used for table purpose. Sapodilla is cultivated in the tropical parts of the world with India being the largest grower. Commonly available varieties in India are the round cricket ball variety and the oval Kalipatti variety. The fruits are rich in sugars (10.2%) and fiber (11.5%); however, the protein content is very low (~0.4-0.7%).

Fruits and vegetables form a major part of the food

[1]Department of Biochemistry and Nutrition, Central Food Technological Research Institute (CFTRI), Council of Scientific and Industrial Research, Mysore, Karnataka, India, [2]Present address: Department of Pathology, Microbiology and Immunology, University of South Carolina School of Medicine, Columbia, South Carolina, USA and [3]Present address: Centre for Research in Biotechnology for Agriculture (CEBAR), University Malaya, Kuala Lumpur, Malaysia.

Conflict of interest: No potential conflict of interest was disclosed.

Correspondence: Dr. Yeldur P Venkatesh, Department of Biochemistry and Nutrition, Central Food Technological Research Institute (CFTRI), Council of Scientific and Industrial Research, KRS Road, Mysore-570020, Karnataka, India.
Email: venkatyp@yahoo.com; ypv@cftri.res.in

Received 23 November 2012. Accepted for publication 10 May 2013.

pyramidal system. The importance of regular consumption of fruits and vegetables in the prevention of cancer and cardiovascular diseases has been widely acknowledged.[3] WHO recommends daily intake of approx. 400 g of fresh fruits and vegetables per person. Though uncommon, fruits and vegetables elicit allergic reactions in sensitized individuals; fruit allergy may increase in future due to increased consumption.[3-5] Although several common fruits (apple, banana, citrus, grape, and other Rosaceae fruits) have been studied with respect to their allergenicity, some exotic fruits, consumed commonly in a particular region have not been well studied in terms of oral allergy syndrome (OAS) and cross-reactive fruits.[6]

Generally, fruits cause allergic reactions confined to the oropharyngeal region, and is referred to as OAS or pollen-food allergy.[7,8] With import of several tropical fruits in many countries, sensitization to many uncommon fruits may occur more commonly. However, allergy to sapodilla is very rare, and only three cases of OAS have been reported; a 21 kDa allergen has been identified from sapodilla pulp.[9] Further, it has been shown that this allergen from sapodilla pulp is a basic thaumatin-like protein (TLP 2) based on its amino acid composition, isoelectric point and N-terminal sequence (Hegde VL *et al*., unpublished observations).

During the course of investigation of the basic TLP from sapodilla for further biochemical and immunological studies, we detected another protein with a similar molecular mass which eluted earlier to the basic TLP by SP-Sepharose cation-exchange chromatography of sapodilla extract. Therefore, it appeared interesting to purify this acidic protein from sapodilla fruit and characterize it in relation to fruit allergy with the objective of identifying additional allergens. The present study describes the characterization of the biochemically purified acidic TLP from sapodilla pulp and its genomic clone. Further, a comparative phylogenetic analysis has been carried out on acidic TLPs which have been identified as food allergens till date.

METHODS

MATERIALS

Sapodilla (cv. cricket ball) saplings were obtained from Indian Institute of Horticultural Research (Bangalore, India), and maintained at the nursery of this institute (CFTRI, Mysore, India). Round and oval sapodilla fruits were purchased from local market. InsTAclone™ PCR Cloning Kit was procured from Fermentas International (Burlington, ON, Canada). Primers were synthesized at Eurofins Genomics India Pvt. Ltd., Bangalore, India. Cetyltrimethylammonium bromide (CTAB), 3-(cyclohexylamino)-1-propanesulfonic acid (CAPS), laminarin (from *Laminaria digitata*), laminarinase (from *Trichoderma* sp.; 100-400 units/g solid), mouse monoclonal anti-human IgE-alkaline phosphatase (ALP) conjugate, Immobilon-P^{SQ} PVDF membranes, SP-Sepharose Fast Flow and JumpStart™ AccuTaq™ LA DNA Polymerase Mix were purchased from Sigma-Aldrich Chemical Co., St. Louis, MO, USA. 2,3,5-Triphenyltetrazolium chloride (TTC, A.R. grade) was a product of Reanal, Budapest, Hungary. Ultrafiltration membrane (PLGC, 10000 NMWL, 44.5 mm diameter) was purchased from Millipore Corporation, Bedford, MA, USA.

DIAGNOSTIC TESTS FOR SAPODILLA ALLERGY

This study was undertaken following approval by the Institutional Ethics Committee. Informed consent was obtained from normal and allergic subjects prior to collection of blood for *in vitro* tests. Diagnostic tests for confirming allergy to sapodilla included case history, skin prick test (SPT), and allergen-specific IgE by ELISA; these were carried out as described previously for diagnosis of allergy to the ingestion of sapodilla,[9] pomegranate, mango, a chewable drug containing excipient mannitol, and eggplant.[10-14] A total of 300 subjects in the age range 18-60 y were screened at random. A 22-year-old female always experienced allergy whenever she ate sapodilla fruit or custard apple (sweetsop; *Annona squamosa* of family Annonaceae). She first noticed itching, swelling and redness (mainly confined to the oropharyngeal area) to these fruits at the age of 10, within 15 min of ingestion of sapodilla or custard apple. Since then, she has avoided eating these fruits.

PURIFICATION OF AN ACIDIC THAUMATIN-LIKE PROTEIN

Pulp from round ripe sapodilla fruits (250 g) was homogenized in an equal volume (250 mL) of extraction buffer (20 mM KH_2PO_4, pH 7.4 containing 3 mM EDTA, 0.9% NaCl, and 0.02% NaN_3). The homogenate was clarified by cheese cloth filtration and centrifuged at 7000 × *g* at 4℃ for 10 min. The supernatant was dialyzed against the same buffer using 12-14 kDa cut-off dialysis tubing to remove low molecular weight substances. Proteins were precipitated by 75% ammonium sulfate saturation, and the precipitate was resuspended in a minimum volume of the starting buffer (20 mM sodium acetate buffer, pH 4.0) and dialyzed against the starting buffer using 12-14 kDa cut-off tubing. The dialyzed solution was loaded onto a cation-exchange column (SP-Sepharose FF, 1.4 × 6.5 cm; 10 mL bed volume) which had been pre-equilibrated with the starting buffer. After loading, the column was washed with the starting buffer to remove unbound proteins. The bound proteins were step-eluted as follows: (a) 20 mM sodium acetate buffer, pH 5.0, (b) buffer in step (a) + 0.1 M NaCl, (c) buffer in step (a) + 0.2 M NaCl, and (d) buffer in step (a) + 1 M NaCl. The eluted protein fractions were pooled, desalted and concentrated using a 10 kDa

mol. wt. cut-off membrane in an Amicon stirred-cell ultrafiltration system. Protein quantitation was performed by the dye-binding method using BSA as standard.[15]

NATIVE PAGE, SDS-PAGE AND ELECTROFOCUSING

The chromatographic components from SP-Sepharose chromatography of sapodilla extract were analyzed by 12% SDS-PAGE (dimensions: 14 cm width × 16 cm length for long gel, and 8 cm width × 6 cm length for mini gel) according to Laemmli[16] under both reducing and non-reducing conditions. PAGE at alkaline and acidic pH under native conditions were carried out using Tris-glycine buffer, pH 9.5 (pH 8.9 system or Davis system) and β-alanine-acetic acid buffer, pH 4.3 (pH 4.3 system or Reisfeld system), respectively.[17] Isoelectric point was evaluated by analytical electrofocusing in an immobilized pH gradient containing ampholyne covering the pH range 3.5-9.5 using a Multiphor II unit (Amersham Biosciences, Uppsala, Sweden). An isoelectric focusing (IEF) calibration kit (Amersham Pharmacia Biotech, Piscataway, NJ, USA) containing 11 proteins of known isoelectric points (pI) was used as reference. The proteins were allowed to focus initially at 500 V for 20 min, and then at 1500 V for 1 h.

ZYMOGRAPHY FOR β-1,3-GLUCANASE ACTIVITY

β-1,3-Glucanase activity on native polyacrylamide gels was carried out according to Stahmann *et al*..[18] After the native PAGE run, the gel was washed 3 times with distilled water and equilibrated in 0.5 M sodium acetate buffer, pH 4.5. The gel was then incubated in the same buffer containing laminarin (6.7 mg/mL) at 37℃ for 2 h. After washing 3 times in distilled water, the gel was immersed in a solution containing 0.15 g TTC in 100 mL of 1 M NaOH, and placed in a boiling water bath until red bands on a clear background appear. The gel was then immediately transferred to a solution of 7.5% acetic acid to conserve the clear background.

β-1,3-GLUCANASE ASSAY

β-1,3-Glucanase activity for the SP-Sepharose chromatographic components was analyzed using laminarin as the substrate as per the protocol described by Sigma Chemical Co., St. Louis, MO, USA (L-5272 - Enzyme Assay). The assay mixture consisting of 100 mM sodium acetate buffer, pH 5 (0.6 mL), 2.5% laminarin in water (0.2 mL), and 5 µg test protein in water (0.2 mL) was incubated at 37℃ for 60 min. Release of reducing sugars was measured as glucose by dinitrosalicylic acid (DNS) method.[19] Briefly, 0.2 mL of DNS reagent and 0.2 mL of sample (test), or distilled water (blank), were taken in test tubes. The tubes were placed in a water bath at 100℃ for 10 min, and then allowed to cool to room temperature. After the addition of 0.8 mL distilled water to each tube and mixing, the absorbance of the solutions was measured at 540 nm. Appropriate controls were used (BSA as negative control and laminarinase as positive control), and a calibration curve was prepared using D-glucose as standard in the range 0-200 µg. The protein content of laminarinase was determined by Bradford assay.[15]

N-TERMINAL SEQUENCING

Protein components from SP-Sepharose chromatography were subjected to SDS-PAGE (12%, reducing). Multiple lanes of a sample were run; after electrophoresis, one half of the gel was stained using 0.3 M $CuCl_2$[20] for visualization of protein bands. The bands of interest were located on the gel based on molecular mass, and the corresponding regions in the other half of the gel were excised. The protein band of interest was transferred on to an Immobilon-P^{SQ} PVDF membrane by wet transfer in 10 mM CAPS buffer with 20% methanol overnight. After transfer, the membrane was transiently stained using 0.2% Ponceau S in 10% acetic acid for the visualization of the protein band; after excising the band, it was destained using water. The membrane was then subjected to N-terminal sequencing by Edman degradation on a PROCISE Protein Sequencer, Model 491 (Applied BioSystems, Foster City, CA, USA) at the Protein Sequencing Facility (Indian Institute of Science, Bangalore, and CSIR-CFTRI, Mysore, India).

CLONING OF A GENE ENCODING THE ACIDIC THAUMATIN-LIKE PROTEIN FROM SAPODILLA

Genomic DNA from the fresh leaves of sapodilla plant (cv. cricket ball) was prepared according to the method described by Saghai-Maroof *et al*..[21] Fresh leaves (1 g) were homogenized using an autoclaved mortar and pestle in 20 mL of extraction buffer (50 mM Tris-HCl buffer, pH 8, containing 0.7 M NaCl, 10 mM EDTA, 1% CTAB, 0.1% 2-mercaptoethanol) and incubated at 60℃ for 1 h with occasional mixing by gentle swirling. Ten milliliters of chloroform-isoamyl alcohol (24:1 v/v) were added and the solution was mixed by inversion to form an emulsion that was centrifuged at 6000 × *g* at 25℃ for 10 min. The aqueous phase was removed, and two-thirds volume of isopropanol was added and mixed 2-4 times by quick gentle inversions. The precipitated DNA was lifted out with a bent glass rod, transferred to a tube containing 20 mL of 76% ethanol/10 mM ammonium acetate, kept for 20 min, and then dissolved in 1.5 mL of 10 mM ammonium acetate/0.25 mM EDTA.

All the primers designed and used for cloning sapodilla acidic TLP gene are listed in Supplementary Table 1 available online. Two sets of nested degenerate primers were designed. N-Terminal sequence of the 26.9 kDa sapodilla acidic allergen (ATFDV

Table 1 Results of SPT and ELISA for diagnosis of sapodilla allergy

Sample	Wheal/flare diameter† (mm)	ELISA units‡ (A405 nm)	
		Normal serum (pooled)	Allergic serum
Negative control§	1/0	n.a.	n.a.
Positive control¶	5/25	n.a.	n.a.
Sapodilla extract	6/20	0.206	0.408
Custard apple extract	7/25‖	0.188	0.359

† SPT is considered positive if the wheal diameter is 3 mm or more.

‡ Mean of triplicate analysis; n.a., not applicable.

§ 50% glycerinated phosphate-buffered saline.

¶ 1.66 mg/mL histamine. 2HCl equivalent to 1 mg/mL histamine base.

‖ With pseudopods.

VNQCT FTVWA GASPG GGKQL-; identified as thaumatin-like protein) was used to design forward primer for set 1 (SAP2: 5' CAGTGTACCTTCACCGT CTGGGC 3'). Also, this sequence was used as a query to retrieve gene sequences and their respective protein translations from NCBI Database (Genbank ID: *Nicotiana tabacum* P07052.1, *Olea europaea* ACZ 57583.1, *Vitis vinifera* AAB61590.1, *Camellia sinensis* ABE01396.1, *Solanum tuberosum* AAU95244.1, *Sambucus nigra* AAK59278.1 and *Nepenthes rafflesiana* ACU31848.1). Multiple sequence alignment was performed using DIALIGN (http://bibiserv.techfak.uni-bielefeld.de/dialign/submission.html) for these gene and protein sequences. A highly conserved region was identified from the protein sequences' alignment (named conserved region A representing the sequence -CNNPCTV-) and a degenerate reverse primer was designed (SAP4: 5' TGTAACAACCCCTG TACCGTG 3'). Primers were analyzed using Integrated DNA Technologies (IDT) oligo analyzer software (http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/). PCR amplification was then carried out using JumpStart™ AccuTaq™ LA DNA Polymerase Mix according to the manufacturer's instructions.

PCR product formation was confirmed by 1% agarose gel electrophoresis and then cloned using InsTAclone™ PCR Cloning Kit and sequenced at Eurofins Genomics India Pvt. Ltd., Bangalore, India. The sequence of the product was used to design forward primer for set 2 (SAP5: 5' CGACATATCTCTGG TTGACGG 3'). Further, the sequence was also used as a query to retrieve sequences from NCBI Database (Genbank ID: *Camellia sinensis* ABE01396.1, *Vitis vinifera* AAB61590.1, *Olea europaea* ACZ57583.1, *Sambucus nigra* AAK59278.1, *Solanum tuberosum* AAU95246.1, *Sambucus nigra* AAK59276.1, *Coffea arabica* ABW76503.1 and *Actinidia deliciosa* ABQ 42566.1). Multiple sequence alignment was performed. A highly conserved region was identified at the C-terminus of plant TLPs and a reverse primer was designed (SAP9: 5' GGTTGTCTTCTGCCCTTGA AATTG 3'). Primer analysis and synthesis, PCR amplification and vector sequencing were carried out as described above.

NUCLEOTIDE SEQUENCE AND THE DEDUCED PROTEIN SEQUENCE ANALYSES

The two nucleotide sequences obtained were assembled using DIALIGN, and BLAST search was performed against NCBI database (http://blast.ncbi.nlm.nih.gov/). The cloned nucleotide sequence was translated into the protein sequence using ExPASy translate tool (http://web.expasy.org/translate/). The protein sequence was then analyzed for conserved domains using NCBI Conserved Domain Database (CDD) search (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). The physical and chemical characteristics of the deduced protein sequence were determined using ProtParam tool of ExPASy (http://web.expasy.org/protparam/). Phylogenetic relationship between sequences was analyzed using Phylogeny.fr one click mode (http://www.phylogeny.fr/). Putative *N*-glycosylation site prediction was analyzed using the NetNGlyc algorithm available online at http://www.cbs.dtu.dk/services.

RESULTS

IDENTIFICATION OF A SUBJECT HAVING ALLERGY TO SAPODILLA AND CUSTARD APPLE

The allergic subject showed a wheal/flare diameter (SPT response) of 6/20 mm for 50% sapodilla extract and 7/25 mm with pseudopods for 50% custard apple extract (Table 1); during the SPT, the allergic subject experienced irritation at the prick points. When allergen-specific IgE was carried out by ELISA using the allergic-subject's serum, there was a 2-fold increase in the absorbance for specific IgE compared to pooled normal sera for both these fruit extracts. These results indicate the confirmation of allergy to both sapodilla and custard apple fruits.

PURIFICATION OF THE ALLERGEN FROM SAPODILLA EXTRACT BY CHROMATOGRAPHY

An extract of sapodilla pulp followed by dialysis at neutral pH resulted in a protein yield of 95 mg from 250 g sapodilla pulp. After ammonium sulfate precipitation, reconstitution and dialysis at acidic pH, approx. 22 mg protein was recovered as a clear solution. The precipitated proteins at pH 4 amounting to ~77% of proteins in sapodilla extract could be resolubilized at neutral pH, and contained mainly high molecular weight proteins as analyzed by SDS-PAGE; further, the precipitated protein pool did not contain any low molecular weight proteins (<30 kDa) described hereinafter.

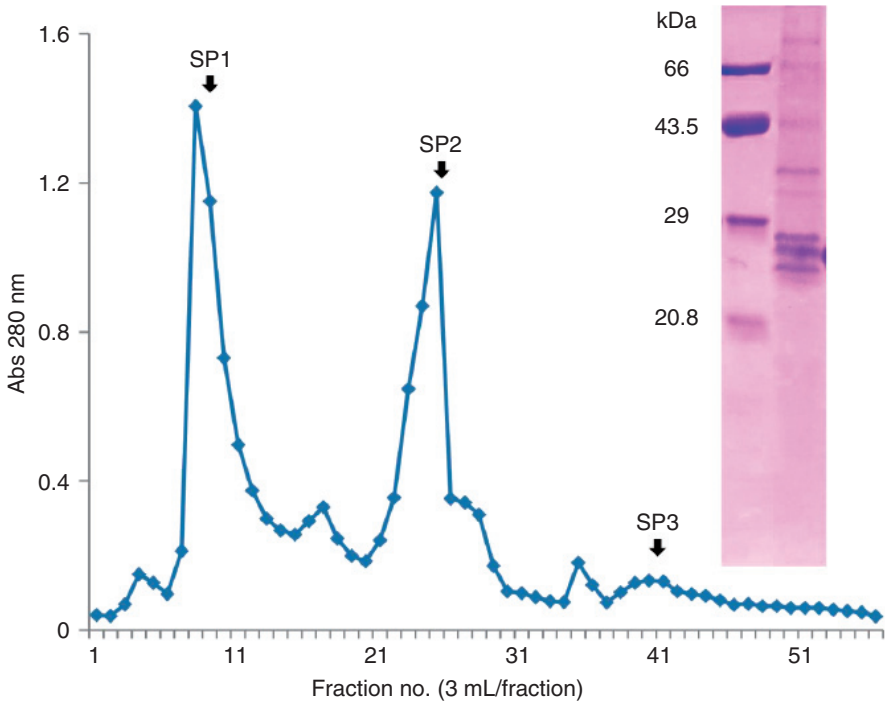


Fig. 1 Chromatographic profile of sapodilla extract on SP-Sepharose FF. Column size: 1.4 × 6.5 cm; protein load: 22 mg of 75% $(NH_4)_2SO_4$ precipitated sapodilla extract, reconstituted and dialyzed against 20 mM sodium acetate buffer, pH 4; flow rate: 20 mL/h. Component SP1 (1.7 mg): elution with 20 mM sodium acetate buffer, pH 5; component SP2 (1.4 mg): elution with 20 mM sodium acetate buffer, pH 5 + 0.1 M NaCl; component SP3 (1.9 mg): elution with 20 mM sodium acetate buffer, pH 5 + 0.2 M NaCl. Elution with 1 M NaCl in 20 mM sodium acetate buffer, pH 5 did not elute any protein (data not shown). Inset: 12% SDS-PAGE (reducing, long gel) pattern of sapodilla extract that was used for chromatography.

Table 2 Analysis of SP-Sepharose components of sapodilla extract for allergen-specific IgE by ELISA

Sample	ELISA units† (A405 nm)	
	Normal serum (pooled)‡	Allergic serum‡
Sapodilla extract	0.197	0.419
SP-1 component	0.187	0.297
SP-2 component	0.183	0.258
SP-3 component	0.176	0.175

†Mean of triplicate analysis.
‡Serum dilution, 1 : 3.

The chromatographic profile of sapodilla protein fractionation is shown in Figure 1. The protein components eluted using 20 mM sodium acetate buffer, pH 5 is labeled as SP1 (1.7 mg); similarly, the protein components eluted using 20 mM sodium acetate buffer, pH 5 with 0.1 and 0.2 M NaCl are labeled as SP2 (1.4 mg) and SP3 (1.9 mg), respectively. When 1 M NaCl in 20 mM sodium acetate buffer, pH 5 was used, no protein was detected in the eluted pool.

The SP-Sepharose components were analyzed by ELISA for detection of the allergens using sapodilla-allergic subject's serum. The results are shown in Table 2. In comparison to the values obtained for normal serum, the sapodilla-allergic serum showed increased ELISA values for both SP1 and SP2 indicating the presence of allergenic proteins.

ANALYSES BY GEL ELECTROPHORESIS AND ELECTROFOCUSING

Upon analysis by SDS-PAGE under reducing conditions (Fig. 2, panel A), SP1 showed a major band of 26.9 kDa (band 'a'). SP2 showed two distinct bands - one of 24.5 kDa, (band 'b') and another of 28.2 kDa (band 'c'). Only a single protein band of 37.4 kDa (band 'd') is evident in SP3. Under non-reducing conditions (Fig. 2, panel A), the mobility of band 'a' shifted to 18.6 kDa, and that of band 'b' shifted to 19.9 kDa. However, the mobilities of the other two protein bands ('c' and 'd') did not alter much under non-reducing conditions (27.3 and 38.2 kDa, respectively). It should be noted here that in order to increase the resolution of the protein bands, a long

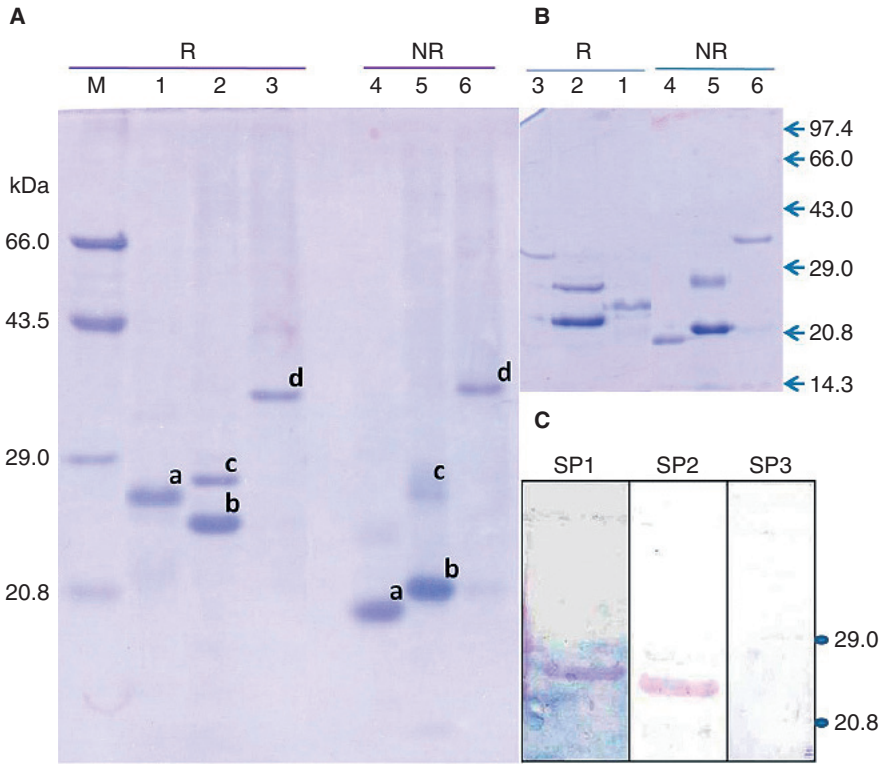


Fig. 2 Panel (**A**): SDS-PAGE (12% long gel) pattern of SP-Sepharose chromatographic components under reducing (R) and non-reducing (NR) conditions. Lane M, molecular weight markers; lanes 1 and 4: SP1 component; lanes 2 and 5: SP2 component; lanes 3 and 6: SP3 component. Protein bands labeled as 'a, b, c and d' in different components are for easy referral throughout the text. Protein load: 5 μg in all cases. Stain: Coomassie blue. Panel (**B**): 12% SDS-PAGE (mini gel, Coomassie stain) showing chromatographic pools from Figure 1. Lane numbers and other details are as shown in panel (**A**). Panel (**C**): IgE-immunoblot of SP-Sepharose pools with sapodilla-allergic serum (serum dilution, 1: 3; secondary antibody conjugate dilution, 1 : 2000).

SDS-PAGE gel was used; the protein bands 'a' and 'b' normally appear at 20-22 kDa position in the regular mini gel (Fig. 2, panel B). Only the protein bands 'a' in SP1 and 'b' in SP2 were detected in IgE-immunoblotting using sapodilla-allergic serum (Fig. 2, panel C); no bands were detected in SP3 pool.

Upon examining the protein components on native PAGE under basic system (Fig. 3, panel A), only band 'a' moved through the gel indicating its acidic nature, whereas bands 'b', 'c', and 'd' did not enter the gel, indicating their basic nature. On the other hand, when these components were subjected to native PAGE under acidic system (Fig. 3, panel B), band 'a' did not enter the gel, whereas bands 'b' and 'c' resolved as two closely spaced protein bands and band 'd' resolved as a single band indicating that all of them are basic proteins.

Electrofocusing analysis (Fig. 3, panel C) revealed that SP1 component showed 4 proteins of pI values 3.9, 4.1, 4.3 and 5.5 out of which the first three protein bands appear to represent variants of band 'a' shown in Figure 3 (panel A), since band 'a' runs as a single intense band in SDS-PAGE gels. The protein band with a pI of 5.5 is likely to be a different co-eluting protein in SP1. On the other hand, the protein bands in SP2 and SP3 components do not resolve in the electrofocusing gel and appear as a single protein band having a pI >9.5 (Fig. 3, panel C).

N-TERMINAL SEQUENCING AND IDENTIFICATION OF THE ALLERGENS AS THAUMATIN-LIKE PROTEINS

SP1 component was blotted onto PVDF membrane; upon N-terminal sequencing of the major band 'a' by Edman degradation, the sequence was found to be ATFDVVNQCTFTVWAGASPGGGKQL- (UniProt Accession No. B3EWX8; TLP 1b). BLAST search against protein sequence database picked up thaumatin-like protein sequences (Supplementary Table 2). Also, on another occasion, a separate purification was carried out using commercial sapodilla fruits and the purified SP1 component was sequenced. The

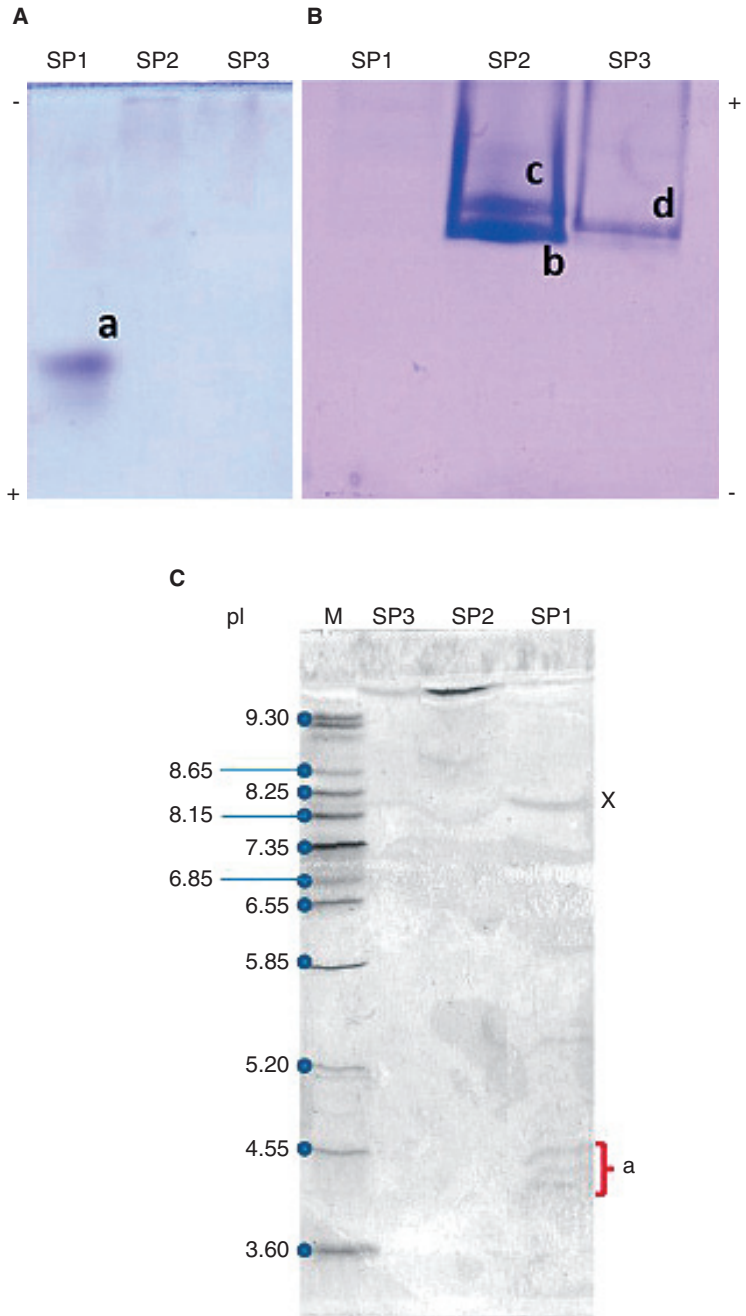


Fig. 3 Native (non-denaturing) PAGE (12%) of SP-Sepharose chromatographic components. Panel (**A**): native PAGE using basic system (Tris-glycine buffer system which runs at pH 9.5). Panel (**B**): native PAGE using acidic system (β-alanine-acetic acid buffer system which runs at pH 4.3). Protein bands labeled as 'a, b, c and d' correspond to the identical bands shown in Figure 2, panel **A**. Protein load: 5 μg in all cases; stain: Coomassie blue. Panel (**C**): isoelectric focusing of SP-Sepharose chromatographic components of sapodilla extract. The pI markers are indicated on the left. The 3 closely spaced bands grouped as 'a' have pI values of 3.9, 4.1 and 4.3. The band marked 'x' represents the position where sample strips are placed for IEF analysis.

5th residue was found to be Ile rather than Val. Based on these results, it can be inferred that there are at least two acidic TLP genes encoding the two isoforms of acidic TLP (1a and 1b).

The 24.5 kDa protein band 'b' present in SP2 component was also sequenced and the N-terminal sequence was found to be ATFDIQNNC- (UniProt Accession No. B3EWE5; TLP 2). The isolation and characterization of the basic TLP (TLP 2) are described in a separate study (Hegde VL *et al*., unpublished observations).

N-Terminal sequencing of band 'c' present in SP2 component resulted in the sequence DVSXFA- which did not yield any significant matches upon BLAST search. In the case of sequencing of band 'd' present in SP3 component, no amino acids were released right from the first cycle indicating the presence of a blocked N-terminus.

Table 3 β-1,3-Glucanase activity of SP-Sepharose chromatographic components of sapodilla extract

Protein sample	Enzymic activity (Units/mg protein)†
SP1	1.6
SP2	1.6
SP3	10.8
Laminarinase	26.0
BSA	0.0

†One unit liberates 1 mg of reducing sugar (measured as glucose) from laminarin per min at pH 5.0 at 37 °C.

ANALYSIS OF β-1,3-GLUCANASE ACTIVITY

Table 3 shows the results of the β-1,3-glucanase assay performed on the SP-Sepharose chromatographic components of sapodilla extract. Since both SP1 and SP2 components showed significantly higher enzymic activity than the negative control (BSA), it was inferred that these components exhibit β-1,3-glucanase activity. On the other hand, SP3 shows β-1,3-glucanase activity which is significantly higher than the activity displayed by either SP1 or SP2. Laminarinase, which hydrolyzes both β-1,3 and β-1,4 glycosidic bonds, was used as a positive control in this assay. Since laminarin contains β-1,3 and β-1,6 linkages in the ratio 3:1, it appears that the protein present in SP3 component is likely to be β-1,3-glucanase.

Analysis of SP1 component for β-1,3-glucanase activity by zymography on native basic PAGE demonstrated detectable enzymic activity by acidic TLP shown as band 'a' (Fig. 4, panel A). In addition to the β-1,3-glucanase activity seen for acidic TLP, one can also observe a marked glucanase activity exhibited by a less acidic or a neutral protein (not detectable by Coomassie staining in Fig. 4, panel A) which appears to have co-purified with acidic TLP in SP1 component; this protein is likely to be an acidic or neutral β-1,3-glucanase in sapodilla pulp. Similarly, SP2 component shows detectable β-1,3-glucanase activity by zy-

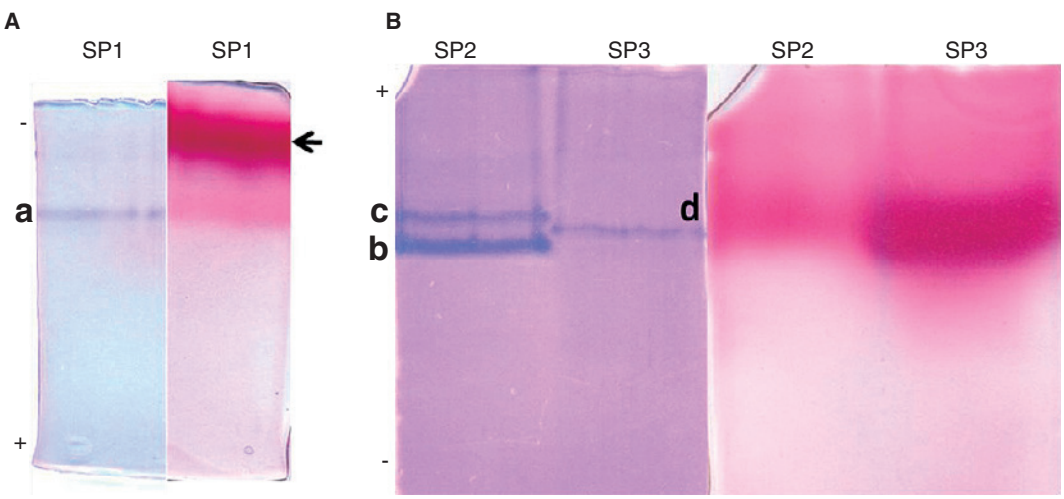


Fig. 4 Zymogram of SP-Sepharose chromatographic components of sapodilla extract. β-1,3-Glucanase activity staining of SP1, SP2 and SP3 components was carried out using laminarin as substrate. In each panel, the left half represents Coomassie-stained gel and the right half represents the zymogram. Protein load: 5 μg in all cases. Panel (**A**): native PAGE (12%), basic system. The band labeled 'a' shows acidic TLP having weak β-1,3-glucanase activity and the solid arrow shows an acidic or neutral protein with strong β-1,3-glucanase activity. Panel (**B**): native PAGE (12%), acidic system. SP2 component has 2 closely spaced bands ('b' and 'c') which are identical to the corresponding bands in Figure 2, 3 based on Coomassie staining intensity. SP3 component shows a single protein (band labeled 'd') which displays strong β-1,3-glucanase activity.

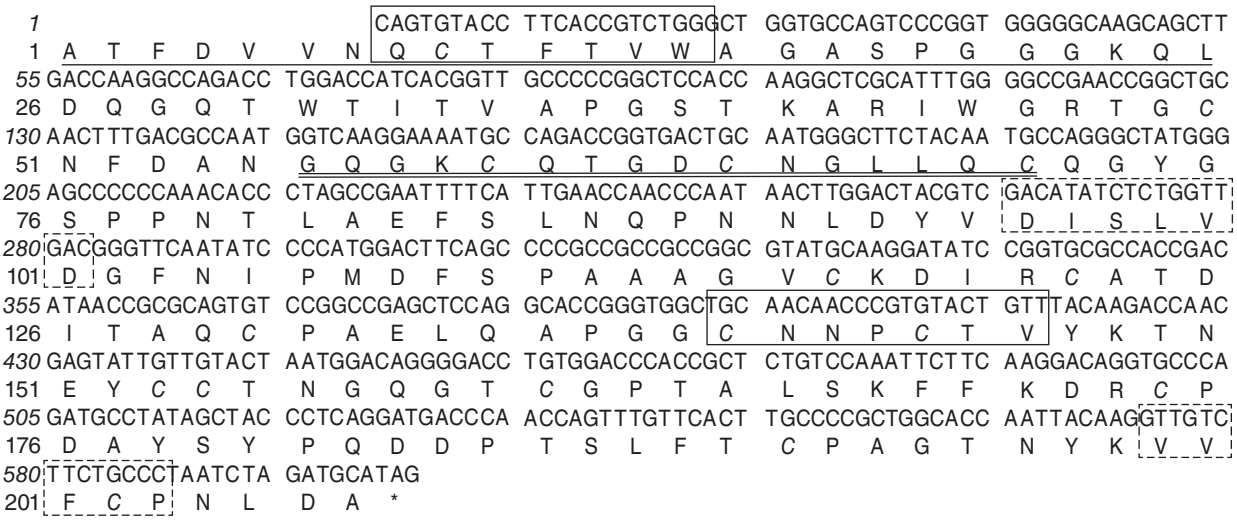


Fig. 5 The complete amino acid sequence of sapodilla acidic TLP (TLP 1). Residues 8 to 207 represent the deduced sequence of sapodilla acidic TLP gene cloned from genomic DNA (JN624813.1). The top line shows the nucleotide sequence (italicized numbers) and the bottom line represents the deduced amino acid sequence (non-italicized numbers). The stop codon (TAG) is indicated by an asterisk. The sequence of residues 1 to 7 taken from the N-terminal sequence of acidic TLP (SP1 component, band 'a') is underlined up to 25th residue (UniProt Accn. B3EWX8; sapodilla acidic TLP 1b). The 5th position is an Ile in another N-terminal sequence analysis up to 12 residues (UniProt Accn. B3EWS0; sapodilla acidic TLP 1a). Residues 56-71 shown in double underline represent the thaumatin family signature (PROSITE PS00316). The boxed regions of the sequence were used for designing the primers in 2 sets: boxes with solid line (set 1); boxes with broken line (set 2). The acidic TLP sequence does not contain an *N*-glycosylation sequon. Cysteine residues are shown in italics. Based on analogy with thaumatin (thaumatin-1, UniProt ID P02883; thaumatin-2, P02884), the disulfide bond pairing are as follows: 9-202, 50-60, 65-71, 117-191, 122-174, 130-140, 144-153, 154-161.

mography on native acidic PAGE (Fig. 4, panel B) indicating that the basic TLP (the more intense band 'b' in Coomassie staining) has weak β-1,3-glucanase activity. SP3 (protein band 'd') displays significant β-1,3-glucanase activity (Fig. 4, panel B) indicating that it is most likely a basic β-1,3-glucanase in sapodilla pulp; this seems to parallel the enzymic activity seen in the test tube assay (Table 3).

SEQUENCE ANALYSIS OF THE DEDUCED SAPODILLA ACIDIC TLP

The partial nucleotide sequence of the acidic TLP gene has 603 bp including the stop codon, and has been deposited at NCBI GenBank sequence repository (JN624813.1). BLAST search identified the deduced protein sequence (AEP84104.1) as belonging to TLP. CDD search revealed that the sapodilla acidic allergen has domains which fit into the superfamily glycoside hydrolase family 64 (GH64) and thaumatin-like proteins (TLP). The deduced partial protein sequence (200 residues) lacks the N-terminal 7 residues as well as the signal peptide. The composite sequence of sapodilla acidic TLP based on the N-terminal sequence (residues 1-25) and the deduced sequence (residues 8-207) is shown in Figure 5. The calculated molecular mass is 21922.4 with a theoretical pI of 4.44. Notably, the protein lacks histidine and contains only one methionine at position 107; the 3 highly conserved tryptophans are present at positions 14, 31 and 45. Sapodilla acidic TLP is rich in glycine (10.6%), threonine (9.7%), alanine (8.7%), cysteine/proline (7.7% each) and aspartic acid (7.3%), and lacks a putative N-glycosylation site (-N-X-S- or -N-X-T-).

The alignment of the complete sequence of sapodilla acidic TLP with the sequences of the top 10 matches from BLAST search and their percent identities are annotated in Figure 6. The residue numbers given henceforth refer to the residue numbers of sapodilla acidic TLP. The salient features of this annotation include the high degree of conservation of all the 16 cysteine residues and the 3 tryptophans. However, there is an additional Trp at position 74 in castor bean TLP, and two additional tryptophans at positions 34 and 74 in soybean and chickpea TLPs. The single Met at position 107 is highly conserved in all the sequences shown except in olive TLP. The thaumatin family signature[22] is clearly evident from the sequence stretch of -G-X-[G/F]-X-C-X-T-[G/A]-D-C-X(1,2)-[G/Q]-X(2,3)-C- which represents amino acid residues -GQGKCQTGDCNGLLQC- from 56-71 in the mature sapodilla acidic TLP.

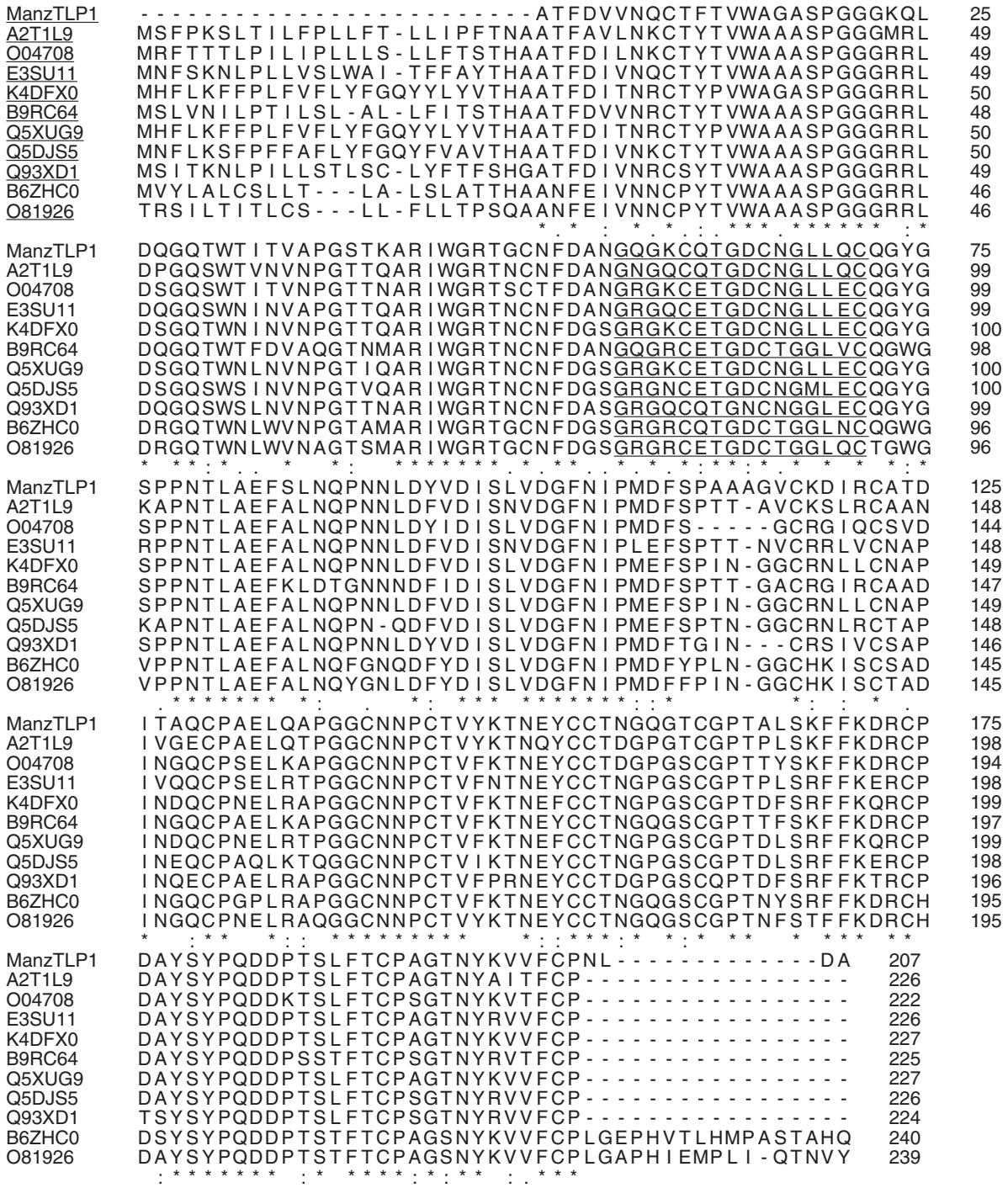


Fig. 6 Sequence alignment of sapodilla acidic TLP (Man z TLP1) with other plant TLPs. Alignments were created using T-Coffee multiple sequence alignment program. Symbols: (*), complete identity; (:), strongly similar; (.), weakly similar; no symbol, different. The region denoting thaumatin family signature is underlined. UniProt ID and the taxonomic name are shown below (the common name, percent identity with Man z TLP1 and theoretical pI of the mature sequence are shown in parentheses): A2T1L9 - *Camellia sinensis* (tea, 80%, 4.59); O04708 - *Vitis vinifera* (grape, 79%, 4.76); E3SU11 - *Olea europaea* (common olive, 78%, 4.76); K4DFX0 - *Solanum lycopersicum* (tomato, 78%, 4.76); B9RC64 - *Ricinus communis* (castor bean, 78%, 4.73); Q5XUG9 - *Solanum tuberosum* (potato, 76%, 4.76); Q5DJS5 - *Nicotiana tabacum* (common tobacco, 75%, 4.94); Q93XD1 - *Sambucus nigra* (European elder, 74%, 4.85); B6ZHC0 - *Glycine max* (soybean, 72%, 6.19); O81926 - *Cicer arietinum* (chickpea, 73%, 5.05). The TLPs with pI <6 is arbitrarily considered as acidic and shown as underlined in the first segment.

SEQUENCE COMPARISON AND PHYLOGENETIC ANALYSIS OF SAPODILLA ACIDIC TLP WITH KNOWN ALLERGENIC TLPs

The alignment of the complete sequence of sapodilla acidic TLP with those of allergenic (food and pollen) acidic TLPs along with their percent identities are shown in Figure 7. Met-107 is conserved in Cry j 3.4, Mal d 2, Pru av 2.0101, Pru p 2.0201 and Pru p 2.0301, whereas in all the other allergenic acidic TLPs it is replaced by Leu. Trp-45 is conserved in all the TLPs listed except Pru p 2.0201 which has Phe. The missing Trp-31 in some acidic TLPs (Cry j 3.5, Mal d 2, Pru av 2.0101, Pru p 2.0201, Pru p 2.0301 and Pyr py 2) appears to be compensated by Trp-40; these 6 TLPs also have a longer sequence which may provide them with additional structural or functional features.

The phylogenetic analysis of sapodilla acidic TLP with known allergenic acidic and basic TLPs from several food sources (olive, kiwi fruit, bell pepper, apple, pear, peach and sweet cherry) is shown in Supplementary Figure 1. It is readily seen that sapodilla acidic TLP is closely related to the TLPs from olive, grape, kiwi, bell pepper and tobacco; among these, tobacco TLP (osmotin) is an inhalant allergen. Further, it is apparent that sapodilla acidic TLP is evolutionarily related to the TLPs from olive and kiwi fruits, both of which belong to the same subclass as sapodilla (subclass Asterids). Interestingly, this order also includes families which produce several edible berries (blueberry, cranberry and bilberry) and fruits (persimmon, date plum, star apple and sapote).

DISCUSSION

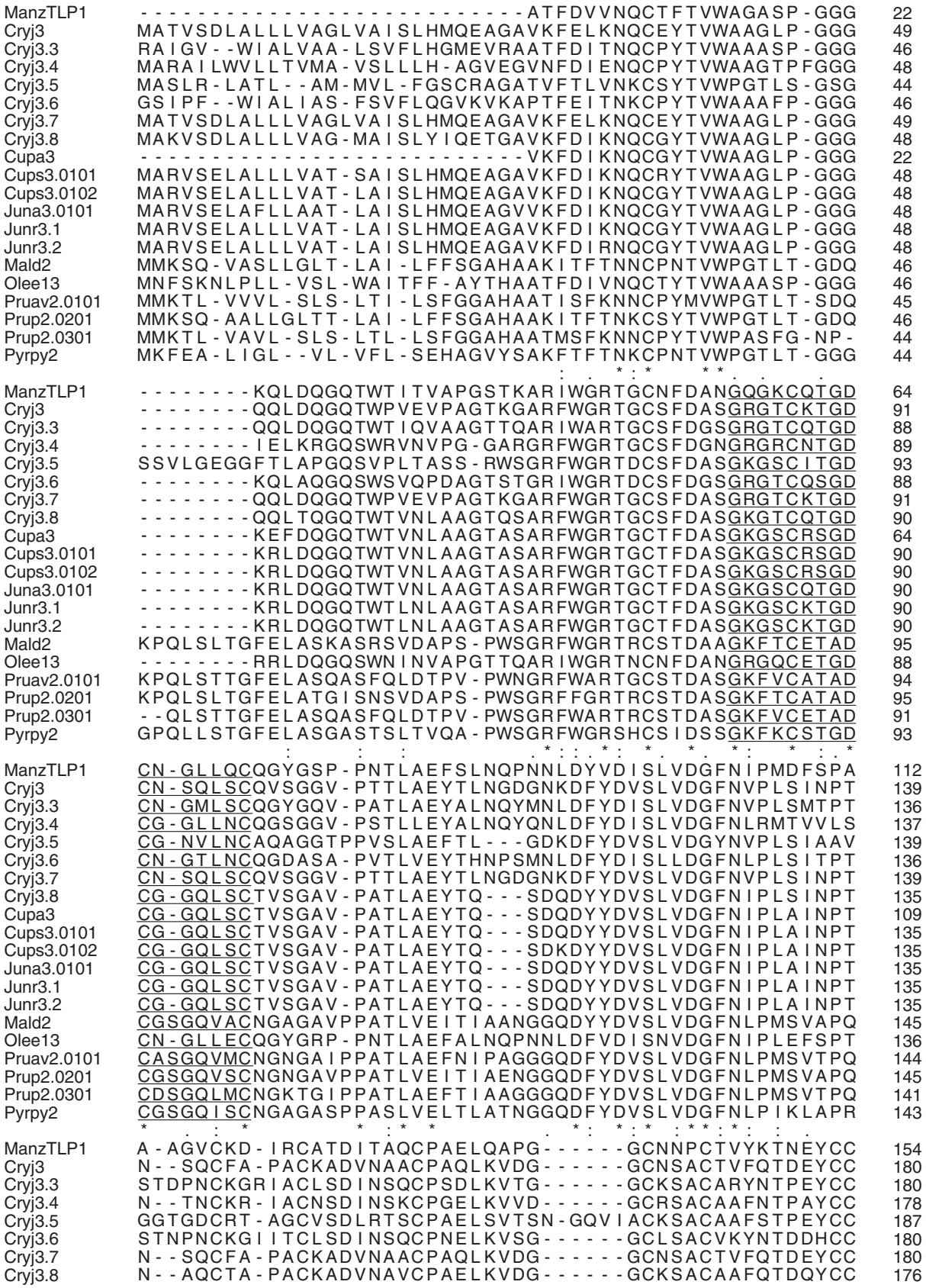
Allergic reactions to ingestion of sapodilla fruit are rare; only 3 cases of sapodilla allergy displaying oral allergy syndrome has been described earlier.[9] The major allergen in sapodilla causing OAS has been identified as a basic TLP in an extended study (Hegde VL *et al*., unpublished observations). An important caveat of that study is the non-detection of an acidic TLP in sapodilla extracts. In the present study, we have described an additional case of sapodilla allergy based on allergy diagnostic tests, and found that an acidic protein also reacted with the serum from the sapodilla-allergic subject. The allergic subject experienced symptoms of oral allergy to the ingestion of sapodilla as well as custard apple, another exotic fruit commonly consumed in India. Several reports of allergy to custard apple have appeared in the literature;[23-25] in one study, a 20-25 kDa protein was identified as the allergen by IgE-immunoblotting,[23] which may likely be TLP. An important constraint in the present study is the non-availability of sera from earlier sapodilla-allergic subjects to address the issue of prevalence of sapodilla allergy.

Biochemical purification of the acidic allergen eluting at pH 5 (band 'a') from SP-Sepharose chromatography of sapodilla extract, and N-terminal sequencing up to 12 or 25 residues identified this protein as a TLP by BLAST search. Upon native PAGE using basic system, the TLP in SP1 moves as a single protein band indicating its acidic nature; conversely, this protein cannot be detected in native PAGE using acidic system. By N-terminal sequencing of the band 'a' on two occasions, it was found that the amino acid at position 5 can be either isoleucine or valine indicating that there are at least 2 variants of sapodilla acidic TLP (TLP 1a and 1b, respectively). Further, IEF analysis showed that TLP 1 resolves into 3 variants which differed in their pI values by ~0.2 units. Further, both TLP 1 and TLP 2 show a remarkable difference in electrophoretic mobility between reducing and non-reducing conditions of SDS-PAGE. A similar observation has been made for TLP from kiwi fruits.[26] This anomalous behavior of the protein indicates that it is highly rich in disulfide bonds which contribute to the stability of the protein. Purified acidic TLP has been obtained in a yield of 6.8 mg per kg sapodilla fruit.

Many TLPs have been identified as allergens in plant foods and pollens in the last decade.[27] Allergenic TLPs have been described from kiwi, apple, grape, sweet cherry, peach, bell pepper and almond.[26,28-33] Palomares *et al*.[34] have shown that occupational asthma can be caused by TLP from olive fruit. Many of these allergenic TLPs have been cloned and expressed as recombinant proteins for detailed studies on allergenicity in relation to protein structure and allergy diagnosis.

Since very little information is available on the proteins of sapodilla pulp, it appeared important to clone the gene coding for acidic TLP for allergological studies. A partial acidic TLP gene, found to be intronless, has been obtained by genomic cloning which codes for 200 amino acid residues. The deduced protein lacked the N-terminal heptapeptide as the primer was designed from the 8th residue in order to avoid the high degeneracy. Functional primers could not be designed for up to the 7th residue. Moreover, attempts with various combinations of primers (Supplementary Table 1) and temperatures were unsuccessful. Based on the position of the first cysteine occurring at residue 9 in TLPs, it was inferred that the deduced sequence of the partial acidic TLP represented residues 8-207 of the mature sapodilla acidic TLP. Residues 8 to 25 of the deduced sequence matched completely with the corresponding sequence obtained by N-terminal sequencing of biochemically purified acidic TLP from sapodilla pulp. As expected, the deduced protein sequence contains 16 highly conserved cysteines which are involved in disulfide bonding for providing structural stability in TLPs.[35]

TLPs occur in all organisms including fungi, animals and plants.[35] Although TLPs are known to be expressed by specific stress, non-specific general stress conditions as causative factors have also been consid-


ManzTLP1 ----------ATFDVVNQCTFTVWAGASP-GGG 22
Cryj3 MATVSDLALLLVAGLVAISLHMQEAGAVKFELKNQCEYTVWAAGLP-GGG 49
Cryj3.3 RAIGV--WIALVAA-LSVFLHGMEVRAATFDITNQCPYTVWAAASP-GGG 46
Cryj3.4 MARAILWVLLTVMA-VSLLLH-AGVEGVNFDIENQCPYTVWAAGTPFGGG 48
Cryj3.5 MASLR-LATL--AM-MVL-FGSCRAGATVFTLVNKCSYTVWPGTLS-GSG 44
Cryj3.6 GSIPF--WIALIAS-FSVFLQGVKVKAPTFEITNKCPYTVWAAAFP-GGG 46
Cryj3.7 MATVSDLALLLVAGLVAISLHMQEAGAVKFELKNQCEYTVWAAGLP-GGG 49
Cryj3.8 MAKVSDLALLLVAG-MAISLYIQETGAVKFDIKNQCGYTVWAAGLP-GGG 48
Cupa3 ----------VKFDIKNQCGYTVWAAGLP-GGG 22
Cups3.0101 MARVSELALLLVAT-SAISLHMQEAGAVKFDIKNQCRYTVWAAGLP-GGG 48
Cups3.0102 MARVSELALLLVAT-LAISLHMQEAGAVKFDIKNQCGYTVWAAGLP-GGG 48
Juna3.0101 MARVSELAFLLAAT-LAISLHMQEAGVVKFDIKNQCGYTVWAAGLP-GGG 48
Junr3.1 MARVSELALLLVAT-LAISLHMQEAGAVKFDIKNQCGYTVWAAGLP-GGG 48
Junr3.2 MARVSELALLLVAT-LAISLHMQEAGAVKFDIRNQCGYTVWAAGLP-GGG 48
Mald2 MMKSQ-VASLLGLT-LAI-LFFSGAHAAKITFTNNCPNTVWPGTLT-GDQ 46
Olee13 MNFSKNLPLL-VSL-WAITFF-AYTHAATFDIVNQCTYTVWAAASP-GGG 46
Pruav2.0101 MMKTL-VVVL-SLS-LTI-LSFGGAHAATISFKNNCPYMVWPGTLT-SDQ 45
Prup2.0201 MMKSQ-AALLGLTT-LAI-LFFSGAHAAKITFTNKCSYTVWPGTLT-GDQ 46
Prup2.0301 MMKTL-VAVL-SLS-LTL-LSFGGAHAATMSFKNNCPYTVWPASFG-NP- 44
Pyrpy2 MKFEA-LIGL--VL-VFL-SEHAGVYSAKFTFTNKCPNTVWPGTLT-GGG 44
ManzTLP1 --------KQLDQGQTWTITVAPGSTKARIWGRTGCNFDANGQGKCQTGD 64
Cryj3 --------QQLDQGQTWPVEVPAGTKGARFWGRTGCSFDASGRGTCKTGD 91
Cryj3.3 --------QQLDQGQTWTIQVAAGTTQARIWARTGCSFDGSGRGTCQTGD 88
Cryj3.4 --------IELKRGQSWRVNVPG-GARGRFWGRTGCSFDGNGRGRCNTGD 89
Cryj3.5 SSVLGEGGFTLAPGQSVPLTASS-RWSGRFWGRTDCSFDASGKGSCITGD 93
Cryj3.6 --------KQLAQGQSWSVQPDAGTSTGRIWGRTDCSFDGSGRGTCQSGD 88
Cryj3.7 --------QQLDQGQTWPVEVPAGTKGARFWGRTGCSFDASGRGTCKTGD 91
Cryj3.8 --------QQLTQGQTWTVNLAAGTQSARFWGRTGCSFDASGKGTCQTGD 90
Cupa3 --------KEFDQGQTWTVNLAAGTASARFWGRTGCTFDASGKGSCRSGD 64
Cups3.0101 --------KRLDQGQTWTVNLAAGTASARFWGRTGCTFDASGKGSCRSGD 90
Cups3.0102 --------KRLDQGQTWTVNLAAGTASARFWGRTGCTFDASGKGSCRSGD 90
Juna3.0101 --------KRLDQGQTWTVNLAAGTASARFWGRTGCTFDASGKGSCQTGD 90
Junr3.1 --------KRLDQGQTWTLNLAAGTASARFWGRTGCTFDASGKGSCKTGD 90
Junr3.2 --------KRLDQGQTWTLNLAAGTASARFWGRTGCTFDASGKGSCKTGD 90
Mald2 KPQLSLTGFELASKASRSVDAPS-PWSGRFWGRTRCSTDAAGKFTCETAD 95
Olee13 --------RRLDQGQSWNINVAPGTTQARIWGRTNCNFDANGRGQCETGD 88
Pruav2.0101 KPQLSTTGFELASQASFQLDTPV-PWNGRFWARTGCSTDASGKFVCATAD 94
Prup2.0201 KPQLSLTGFELATGISNSVDAPS-PWSGRFFGRTRCSTDASGKFTCATAD 95
Prup2.0301 --QLSTTGFELASQASFQLDTPV-PWSGRFWARTRCSTDASGKFVCETAD 91
Pyrpy2 GPQLLSTGFELASGASTSLTVQA-PWSGRFWGRSHCSIDSSGKFKCSTGD 93
ManzTLP1 CN-GLLQCQGYGSP-PNTLAEFSLNQPNNLDYVDISLVDGFNIPMDFSPA 112
Cryj3 CN-SQLSCQVSGGV-PTTLAEYTLNGDGNKDFYDVSLVDGFNVPLSINPT 139
Cryj3.3 CN-GMLSCQGYGQV-PATLAEYALNQYMNLDFYDISLVDGFNVPLSMTPT 136
Cryj3.4 CG-GLLNCQGSGGV-PSTLLEYALNQYQNLDFYDISLVDGFNLRMTVVLS 137
Cryj3.5 CG-NVLNCAQAGGTPPVSLAEFTL---GDKDFYDVSLVDGYNVPLSIAAV 139
Cryj3.6 CN-GTLNCQGDASA-PVTLVEYTHNPSMNLDFYDISLLDGFNLPLSITPT 136
Cryj3.7 CN-SQLSCQVSGGV-PTTLAEYTLNGDGNKDFYDVSLVDGFNVPLSINPT 139
Cryj3.8 CG-GQLSCTVSGAV-PATLAEYTQ---SDQDYYDVSLVDGFNIPLSINPT 135
Cupa3 CG-GQLSCTVSGAV-PATLAEYTQ---SDQDYYDVSLVDGFNIPLAINPT 109
Cups3.0101 CG-GQLSCTVSGAV-PATLAEYTQ---SDQDYYDVSLVDGFNIPLAINPT 135
Cups3.0102 CG-GQLSCTVSGAV-PATLAEYTQ---SDKDYYDVSLVDGFNIPLAINPT 135
Juna3.0101 CG-GQLSCTVSGAV-PATLAEYTQ---SDQDYYDVSLVDGFNIPLAINPT 135
Junr3.1 CG-GQLSCTVSGAV-PATLAEYTQ---SDQDYYDVSLVDGFNIPLAINPT 135
Junr3.2 CG-GQLSCTVSGAV-PATLAEYTQ---SDQDYYDVSLVDGFNIPLAINPT 135
Mald2 CGSGQVACNGAGAVPPATLVEITIAANGGQDYYDVSLVDGFNLPMSVAPQ 145
Olee13 CN-GLLECQGYGRP-PNTLAEFALNQPNNLDFVDISNVDGFNIPLEFSPT 136
Pruav2.0101 CASGQVMCNGNGAIPPATLAEFNIPAGGGQDFYDVSLVDGFNLPMSVTPQ 144
Prup2.0201 CGSGQVSCNGNGAVPPATLVEITIAENGGQDFYDVSLVDGFNLPMSVAPQ 145
Prup2.0301 CDSGQLMCNGKTGIPPATLAEFTIAAGGGQDFYDVSLVDGFNLPMSVTPQ 141
Pyrpy2 CGSGQISCNGAGASPPASLVELTLATNGGQDFYDVSLVDGFNLPIKLAPR 143
ManzTLP1 A-AGVCKD-IRCATDITAQCPAELQAPG------GCNNPCTVYKTNEYCC 154
Cryj3 N--SQCFA-PACKADVNAACPAQLKVDG------GCNSACTVFQTDEYCC 180
Cryj3.3 STDPNCKGRIACLSDINSQCPSDLKVTG------GCKSACARYNTPEYCC 180
Cryj3.4 N--TNCKR-IACNSDINSKCPGELKVVD------GCRSACAAFNTPAYCC 178
Cryj3.5 GGTGDCRT-AGCVSDLRTSCPAELSVTSN-GQVIACKSACAAFSTPEYCC 187
Cryj3.6 STNPNCKGIITCLSDINSQCPNELKVSG------GCLSACVKYNTDDHCC 180
Cryj3.7 N--SQCFA-PACKADVNAACPAQLKVDG------GCNSACTVFQTDEYCC 180
Cryj3.8 N--AQCTA-PACKADVNAVCPAELKVDG------GCKSACAAFQTDQYCC 176

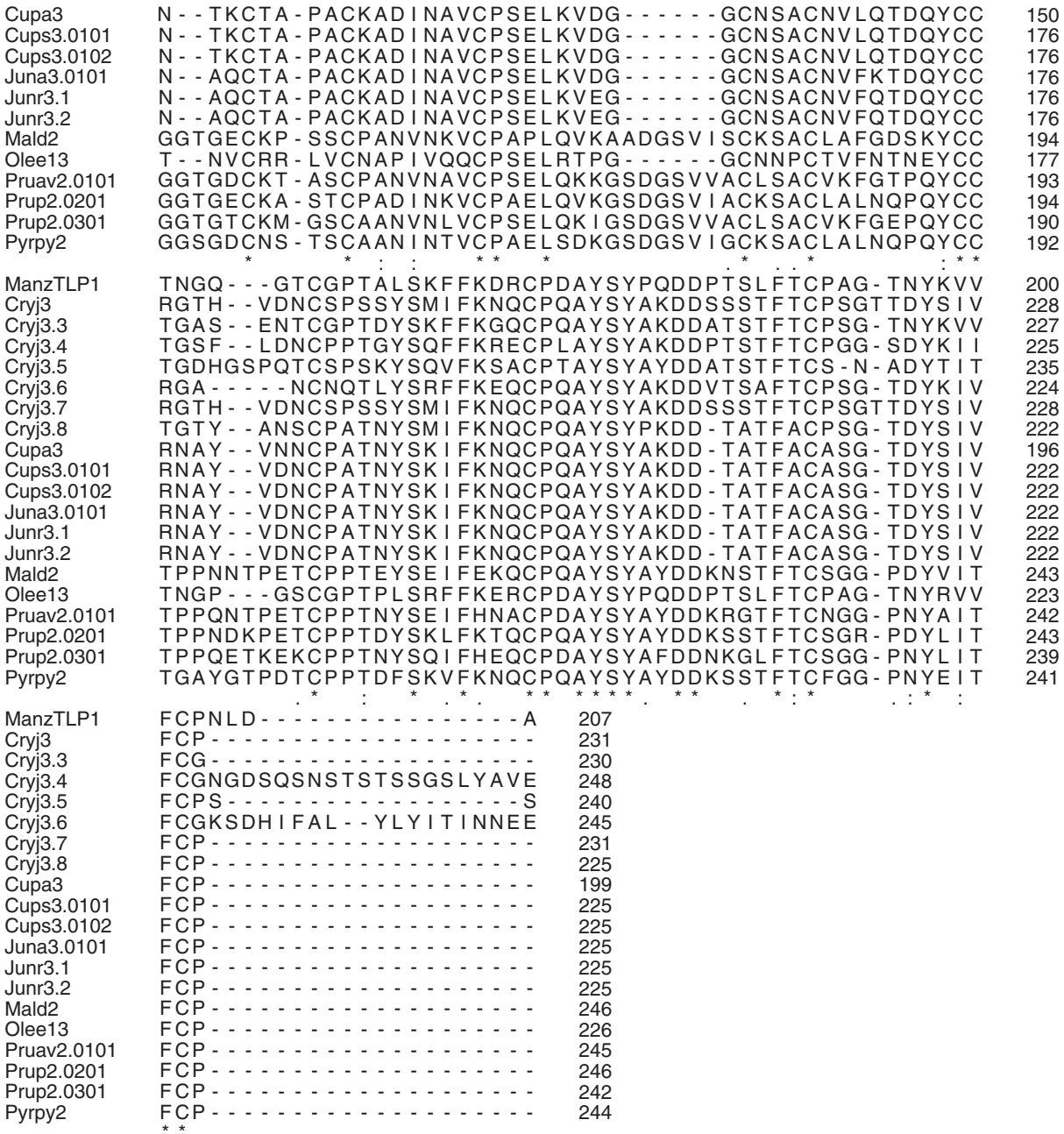


Fig. 7 Alignment of the complete sequence of sapodilla acidic TLP (Man z TLP1) with known acidic TLPs currently identified as pollen and food allergens. Alignments were created using T-Coffee multiple sequence alignment. Symbols: (*), complete identity; (:), strongly similar; (.), weakly similar; no symbol, different. The region denoting thaumatin family signature is underlined. Common name, allergome ID, percent identity with Man z TLP1 and theoretical pI of the mature sequence are shown in parentheses following the taxonomic name. Pollens: *Cryptomeria japonica* (Japanese cedar, Cry j 3, 54%, 4.62; Cry j 3.3, 65%, 4.63; Cry j 3.4, 53%, 5.47; Cry j 3.5, 42%, 4.54; Cry j 3.6, 53%, 4.80; Cry j 3.7, 54%, 4.62; Cry j 3.8, 53%, 4.43); *Cupressus sempervirens* (Mediterranean cypress, Cup s 3.0101, 53%, 5.32; Cup s 3.0102, 53%, 5.32); *Cupressus arizonica* (Arizona cypress, Cup a 3, 52%, 4.82); *Juniperus ashei* (mountain cedar, Jun a 3.0101, 55%, 4.79); *Juniperus rigidus* (Japanese juniper, Jun r 3.1, 54%, 4.82; Jun r 3.2, 54%, 4.82). Foods: *Malus domestica* (apple, Mal d 2, 39%, 4.72); *Olea europaea* (common olive, Ole e 13, 78%, 4.76); *Prunus avium* (cherry, Pru av 2.0101, 41%, 4.57); *Prunus persica* (nectarine/peach, Pru p 2.0201, 40%, 4.80; Pru p 2.0301, 42%, 4.58); *Pyrus pyrifolia* (apple pear, Pyr py 2, 43%, 5.01).

ered.[36] Because of their pathogen-induced expression, TLPs are classified under group 5 of the pathogenesis-related proteins (PR-5 family of PRPs),[37] which includes cysteine-rich proteins sharing a sequence similarity with thaumatin (THN), an intensely sweet-tasting protein from the fruit of a west African shrub *Thaumatococcus daniellii*, commonly known as miracle fruit or katemfe.[38] PR-5 proteins have a clearly recognizable structure and are annotated in protein databases as the THN domain which belongs to the thaumatin family (PROSITE PS00316).

Several TLPs have been shown to exhibit antifungal activity,[39] while others possess β-1,3-glucanase activity.[40,41] Some TLPs have also been shown to bind β-1,3-glucans.[42] Although our results are indicative of a weak β-1,3-glucanase activity by zymography for both acidic and basic TLPs, it cannot be ruled out that the weak enzymic activity is due to the presence of contaminating β-1,3-glucanases in our purified preparations of SP1 and SP2 components. Further work on obtaining highly purified acidic or basic TLPs as well as recombinant sapodilla TLPs should help to resolve the ambiguity of whether sapodilla TLPs possess inherent weak β-1,3-glucanase activity or not.

The major protein in SP3 component (~37 kDa) exhibits strong β-1,3-glucanase activity. Most basic β-1,3-glucanases are approx. 35 kDa in size and an examination of their N-terminus shows that Gln is present in most cases, and Glu in one case, viz., *Nicotiana plumbaginifolia* (Supplementary Fig. 2). Pyroglutamic acid formation occurs when the N-terminal residue of a protein is either glutamine or glutamic acid,[43] resulting in a blocked N-terminus. In view of the blocked N-terminus of SP3 component and strong enzymic activity, SP3 component has been suggested to be a basic β-1,3-glucanase.

A comparison of the complete sequence of sapodilla acidic TLP (207 residues) with other plant TLPs (BLAST search results) indicated 109 out of 207 residues are identical (52.7%). On the contrary, a similar comparison with food and pollen allergenic acidic TLPs revealed that only 47 out of 207 residues are identical (22.7%). This indicates that allergenic acidic TLPs seem to have more variation than plant TLPs that have not been identified as allergens till date. Phylogenetic analysis of the allergenic TLPs from foods indicated that sapodilla acidic TLP is closely related to the TLPs from olive, grape, kiwi, bell pepper and forms one cluster; this cluster is distinct from the other cluster containing allergenic acidic TLPs of the Rosaceae family. A recent study on the involvement of TLPs in plant food cross-reactivity using specific protein microarray indicated that TLPs are significant allergens in plant food allergy and should be considered when diagnosing and treating pollen-food allergy.[44]

In conclusion, an acidic form of TLP (TLP 1) has been isolated from sapodilla fruit which causes oral allergy syndrome in a sapodilla-allergic individual. Sapodilla TLP 1 is a single polypeptide occurring as at least two variants and exhibits weak β-1,3-glucanase activity. The intronless gene coding for sapodilla acidic TLP representing 97% of the mature sequence has been cloned. The complete sequence of sapodilla TLP 1 comprises of 207 residues stabilized by 8 disulfide bonds and lacks an *N*-glycosylation site. Phylogenetic analysis indicates close relationship with TLPs from olive, grape, kiwi fruit and bell pepper. The identification of an allergic subject experiencing oral allergy to sapodilla and custard apple may indicate cosensitization to TLP among fruits.

ACKNOWLEDGEMENTS

The assistance rendered by the staff of the Health Centre, CSIR-CFTRI, Mysore, India, is gratefully acknowledged. We thank the Council of Scientific & Industrial Research (CSIR), New Delhi, for financial support (Y.P.V., institutional project; H.G.A. and V.L.H., research fellowship). S.M.S. thanks the Department of Biotechnology (DBT), New Delhi for the award of a research fellowship. We gratefully acknowledge the assistance received from Dr. Purnima Tiku (Protein Chemistry & Technology Department, CSIR-CFTRI, Mysore) for protein analysis by electrofocusing. Thanks are also due to Dr. Ram Rajasekaran (Director, CSIR-CFTRI) and Dr. Asha Martin (Food Safety & Analytical Quality Control Laboratory Department, CSIR-CFTRI) for protein sequencing of SP2 component (28 kDa) and SP3 component (37 kDa) in this study.

This work was presented at the 2012 WAO International Scientific Conference held on 6-9 December 2012 at Hyderabad, India (Abstract No. 2046).

NOTE ADDED IN PROOF

The sapodilla 28 kDa protein (band 'c' in SP2 component), upon analysis by LC-MS/MS of its tryptic digest, matched with class I chitinase from tall fescue grass (*Festuca arundinacea*) and basic chitinase from cocoa (*Theobroma cacao*), and the results are shown in Supplementary Fig. 3. Hence, the 28 kDa protein in sapodilla fruit is likely to be a basic chitinase. The study describing the basic TLP as an important allergen in sapodilla (Hegde VL *et al.*, unpublished observations) is now accepted for publication, and its full citation is as follows: Hegde VL, Ashok Kumar HG, Sreenath K, Hegde ML, Venkatesh YP. Identification and characterization of a basic thaumatin-like protein (TLP 2) as an allergen in sapodilla plum (*Manilkara zapota*). *Mol Nutr Food Res*. In press. DOI: 10.1002/mnfr.201300261.

SUPPLEMENTARY MATERIALS

Supplementary Table 1, 2 and Supplementary Figure

1, 2, 3 are available online.

REFERENCES

1. Paull RE. *Manilkara zapota*, chiku. In: Janick J, Paull RE (eds). *The Encyclopedia of Fruit & Nuts.* Wallingford, Oxfordshire, UK: CAB International, 2008;828-31.
2. Small E. Sapodilla. In: *Top 100 Exotic Food Plants.* Boca Raton, USA: CRC Press, 2011;515-8.
3. Marzban G, Mansfeld A, Hemmer W, Stoyanova E, Katinger H, da Câmara Machado ML. Fruit cross-reactive allergens: a theme of uprising interest for consumers' health. *Biofactors* 2005;**23**:235-41.
4. Ballmer-Weber BK, Hoffmann-Sommergruber K. Molecular diagnosis of fruit and vegetable allergy. *Curr Opin Allergy Clin Immunol* 2011;**11**:229-35.
5. Salcedo G, Díaz-Perales A, Sánchez-Monge R. Fruit allergy: plant defence proteins as novel potential panallergens. *Clin Exp Allergy* 1999;**29**:1158-60.
6. Knudsen I, Søborg I, Eriksen F, Pilegaard K, Pedersen J. Risk management and risk assessment of novel plant foods: concepts and principles. *Food Chem Toxicol* 2008; **46**:1681-705.
7. Webber CM, England RW. Oral allergy syndrome: a clinical, diagnostic, and therapeutic challenge. *Ann Allergy Asthma Immunol* 2010;**104**:101-8.
8. Katelaris CH. Food allergy and oral allergy or pollen-food syndrome. *Curr Opin Allergy Clin Immunol* 2010;**10**:246-51.
9. Hegde VL, Venkatesh YP. Oral allergy syndrome to sapodilla (*Achras zapota*). *J Allergy Clin Immunol* 2002;**110**: 533-4.
10. Hegde VL, Mahesh PA, Venkatesh YP. Anaphylaxis caused by mannitol in pomegranate (*Punica granatum*). *ACI Int* 2002;**14**:37-9.
11. Hegde VL, Venkatesh YP. Anaphylaxis following ingestion of mango fruit. *J Investig Allergol Clin Immunol* 2007; **17**:341-4.
12. Hegde VL, Venkatesh YP. Anaphylaxis to excipient mannitol: evidence for an immunoglobulin E-mediated mechanism. *Clin Exp Allergy* 2004;**34**:1602-9.
13. Harish Babu BN, Mahesh PA, Venkatesh YP. A cross-sectional study on the prevalence of food allergy to eggplant (*Solanum melongena* L.) reveals female predominance. *Clin Exp Allergy* 2008;**38**:1795-802.
14. Harish Babu BN, Venkatesh YP. Clinico-immunological analysis of eggplant (*Solanum melongena*) allergy indicates preponderance of allergens in the peel. *WAO Journal* 2009;**2**:192-200.
15. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 1976;**72**: 248-54.
16. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970; **227**:680-5.
17. Gabriel O. Analytical disc gel electrophoresis. *Methods Enzymol* 1971;**22**:565-78.
18. Stahmann KP, Schimz KL, Sahm H. Purification and characterization of four extracellular 1,3-β-glucanases of *Botrytis cinerea*. *J Gen Microbiol* 1993;**139**:2833-40.
19. Gonçalves C, Rodriguez-Jasso RM, Gomes N, Teixeira JA, Belo I. Adaptation of dinitrosalicylic acid method to microtiter plates. *Anal Methods* 2010;**2**:2046-8.
20. Sullivan ML, Hatfield RD, Thoma SL, Samac DA. Cloning and characterization of red clover polyphenol oxidase cDNAs and expression of active protein in *Escherichia coli* and transgenic alfalfa. *Plant Physiol* 2004;**136**:3234-44.
21. Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proc Natl Acad Sci U S A* 1984; **81**:8014-8.
22. Edens L, Heslinga L, Klok R *et al*. Cloning of cDNA encoding the sweet-tasting plant protein thaumatin and its expression in *Escherichia coli*. *Gene* 1982;**18**:1-12.
23. Gonzalo MA, Moneo I, Ventas P, Polo F, García JM. IgE-mediated hypersensitivity to custard-apple. *Allergy* 1997; **52**:597.
24. Sánchez-Guerrero IM, Escudero AI, Tortosa JA, Lombardero M. Anaphylaxis to cherimoya. *Allergy* 2000;**55**: 976-7.
25. Sánchez-Morillas L, Moneo I, Sedano E, Iglesias A, Caballero ML, Rodríguez M. Oral allergic syndrome after ingestion of custard apple. *Allergy* 2003;**58**:260-1.
26. Gavrović-Jankulović M, ćIrković T, Vucković O *et al*. Isolation and biochemical characterization of a thaumatin-like kiwi allergen. *J Allergy Clin Immunol* 2002;**110**:805-10.
27. Breiteneder H. Thaumatin-like proteins – a new family of pollen and fruit allergens. *Allergy* 2004;**59**:479-81.
28. Inschlag C, Hoffmann-Sommergruber K, O'Riordain G *et al*. Biochemical characterization of Pru a 2, a 23-kD thaumatin-like protein representing a potential major allergen in cherry (*Prunus avium*). *Int Arch Allergy Immunol* 1998;**116**:22-8.
29. Krebitz M, Wagner B, Ferreira F *et al*. 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* 2003;**329**: 721-30.
30. Pastorello EA, Farioli L, Pravettoni V *et al*. Identification of grape and wine allergens as an endochitinase 4, a lipid-transfer protein, and a thaumatin. *J Allergy Clin Immunol* 2003;**111**:350-9.
31. Fuchs HC, Bohle B, Dall'Antonia Y *et al*. Natural and recombinant molecules of the cherry allergen Pru av 2 show diverse structural and B cell characteristics but similar T cell reactivity. *Clin Exp Allergy* 2006;**36**:359-68.
32. Smole U, Bublin M, Radauer C, Ebner C, Breiteneder H. Mal d 2, the thaumatin-like allergen from apple, is highly resistant to gastrointestinal digestion and thermal processing. *Int Arch Allergy Immunol* 2008;**147**:289-98.
33. Palacín A, Tordesillas L, Gamboa P *et al*. Characterization of peach thaumatin-like proteins and their identification as major peach allergens. *Clin Exp Allergy* 2010;**40**:1422-30.
34. Palomares O, Alcántara M, Quiralte J, Villalba M, Garzón F, Rodríguez R. Airway disease and thaumatin-like protein in an olive-oil mill worker. *N Engl J Med* 2008;**358**: 1306-8.
35. Liu JJ, Sturrock R, Ekramoddoullah AK. The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. *Plant Cell Rep* 2010; **29**:419-36.
36. Narasimhan ML, Bressan RA, D'Urzo MP, Jenks MA, Mengiste T. Unexpected turns and twists in structure/function of PR-proteins that connect energy metabolism and immunity. *Adv Bot Res* 2009;**51**:439-89.
37. Hoffmann-Sommergruber K. Pathogenesis-related (PR)-proteins identified as allergens. *Biochem Soc Trans* 2002;

30 (Pt 6):930-5.
38. Temussi PA. Natural sweet macromolecules: how sweet proteins work. *Cell Mol Life Sci* 2006;**63**:1876-88.
39. Van Loon LC, Rep M, Pieterse CMJ. Significance of inducible defense-related proteins in infected plants. *Annu Rev Phytopathol* 2006;**44**:135-62.
40. Grenier J, Potvin C, Trudel J, Asselin A. Some thaumatin-like proteins hydrolyse polymeric β-1,3-glucans. *Plant J* 1999;**19**:473-80.
41. Menu-Bouaouiche L, Vriet C, Peumans WJ, Barre A, Van Damme EJM, Rougé P. A molecular basis for the endo-β-1,3-glucanase activity of the thaumatin-like proteins from edible fruits. *Biochimie* 2003;**85**:123-31.
42. Trudel J, Grenier J, Potvin C, Asselin A. Several thaumatin-like proteins bind to beta-1,3-glucans. *Plant Physiol* 1998;**118**:1431-8.
43. Liu YD, Goetze AM, Bass RB, Flynn GC. N-terminal glutamate to pyroglutamate conversion *in vivo* for human IgG2 antibodies. *J Biol Chem* 2011;**286**:11211-7.
44. Palacín A, Rivas LA, Gómez-Casado C *et al*. The involvement of thaumatin-like proteins in plant food cross-reactivity: A multicenter study using a specific protein microarray. *PLoS ONE* 2012;**7**:e44088.