

Accepted Manuscript

Expression and Characterization of a new isoform of the 9 kDa allergenic Lipid Transfer Protein from Tomato (variety San Marzano)

Mariateresa Volpicella, Claudia Leoni, Immacolata Fanizza, Sara Rinalducci, Antonio Placido, Luigi R. Ceci



PII: S0981-9428(15)30065-6

DOI: [10.1016/j.plaphy.2015.07.019](https://doi.org/10.1016/j.plaphy.2015.07.019)

Reference: PLAPHY 4237

To appear in: *Plant Physiology and Biochemistry*

Received Date: 19 May 2015

Revised Date: 17 July 2015

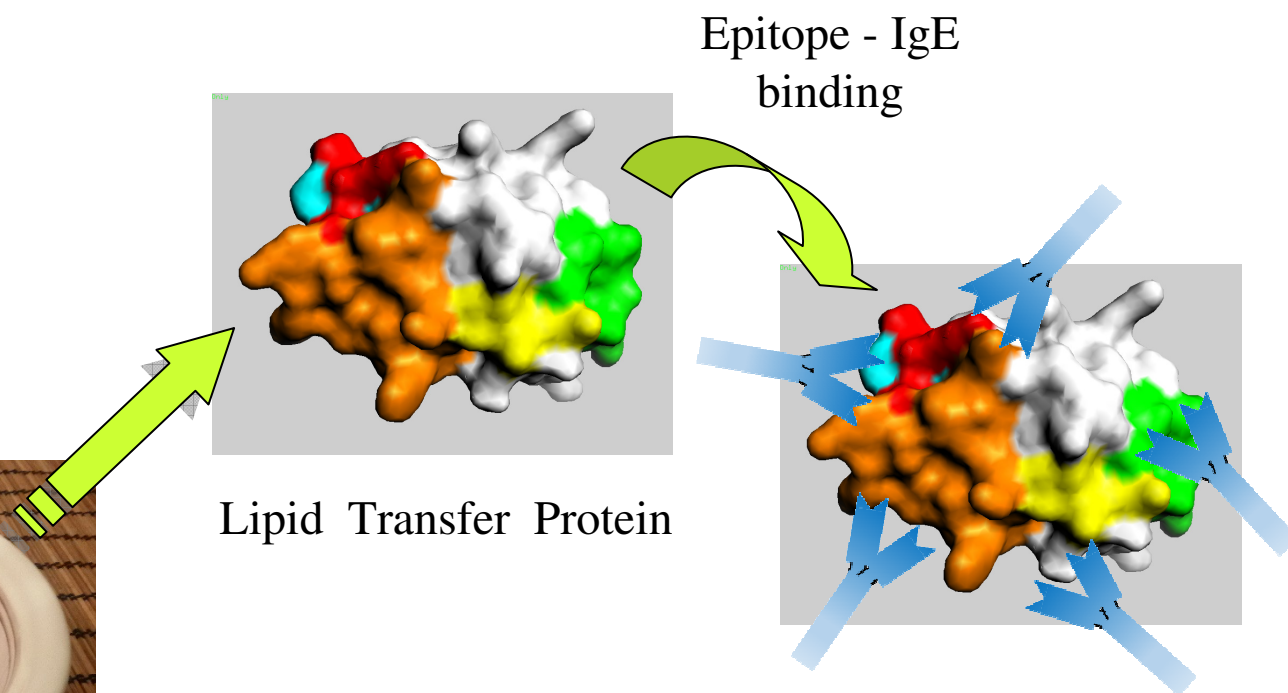
Accepted Date: 18 July 2015

Please cite this article as: M. Volpicella, C. Leoni, I. Fanizza, S. Rinalducci, A. Placido, L.R. Ceci, Expression and Characterization of a new isoform of the 9 kDa allergenic Lipid Transfer Protein from Tomato (variety San Marzano), *Plant Physiology et Biochemistry* (2015), doi: 10.1016/j.plaphy.2015.07.019.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



San Marzano tomatoes



Expression and Characterization of a new isoform of the 9 kDa allergenic Lipid Transfer Protein from Tomato (variety San Marzano)

Mariateresa Volpicella^{a,#}, Claudia Leoni^{b,#}, Immacolata Fanizza^a, Sara Rinalducci^c, Antonio Placido^b, Luigi R. Ceci^{b,*}

^aDepartment of Biosciences, Biotechnologies and Biopharmaceutics, University of Bari, Via Amendola 165/A, 70126 Bari, Italy.

^bInstitute of Biomembranes and Bioenergetics – National Research Council (CNR), Via Amendola 165/A, 70126 Bari, Italy.

^cDepartment of Ecological and Biological Sciences, University of Tuscia, Largo dell'Università' - 01100 Viterbo, Italy.

These authors contributed equally to this study.

* Corresponding author: L.R. Ceci, Phone: +39 080 5443311. Fax: +39 080 5443403. E-mail:

l.ceci@ibbe.cnr.it.

ABSTRACT

Lipid transfer proteins (LTPs) are food allergens found first in fruits of the *Rosaceae* family and later identified in other food plants. Their high structural stability causes them to behave as allergens in cooked and processed foods. Allergenic LTPs have been identified in tomato fruits as well, but studies of their thermal stability and structural characteristics are limited. In this article we report the identification of the coding region for a novel 9 kDa LTP isoform in the tomato variety San Marzano, together with the expression of the recombinant mature protein. The purified recombinant protein was further characterized for its thermal stability and was found to bind 1-palmitoyl-2-lysophosphatidylcholine (Lyso-C₁₆) after thermal treatments up to 105°C. Analysis of a modelling derived structure of the protein allowed the identification of possible epitope regions on the molecular surface.

KEYWORDS: Tomato; Lipid Transfer Protein; Food allergy

1. INTRODUCTION

Tomatoes are among the most important crops of the world with a wide number of cultivated varieties (Sardaro *et al.*, 2012). The tomato fruits are used as both raw and cooked ingredients in many food preparations. Tomatoes contain many constituent compounds with beneficial effects on human health (Bohn *et al.*, 2013), but are rich in allergens as well. Evaluations of the incidence of tomato allergies among populations are rare, however there is evidence of emerging allergy cases (Larramendi *et al.*, 2008; Westphal *et al.*, 2004). Currently five tomato food allergens are listed by the International Union of Immunological Society (IUIS, <http://www.allergen.org/>) (Radauer *et al.*, 2014)): Sola I 1 (Profilin), Sola I 2 (Beta-fructofuranoside), Sola I 3 (Non-specific lipid transfer protein 2, nsLTP), Sola I 4 (Pathogenesis-related protein, PR-10) and Sola I 6 (nsLTP 7S). Other possible allergens have been identified as well and are under investigation (Bencivenni *et al.*, 2012; Bencivenni *et al.*, 2014; Foetisch *et al.*, 2001) (see also the Allergome database at

27 <http://www.allergome.org/> (Mari *et al.*, 2009) and the AllergenPro database at
28 <http://nabic.rda.go.kr/allergen/> (Kim *et al.*, 2009)).

29 Among tomato fruit allergens, nsLTPs are of particular interest due to their high diffusion as
30 plant allergens and the ability to induce cross-reactivity based allergenic reactions in predisposed
31 subjects (Van Winkle and Chang, 2014). Lipid transfer proteins are ubiquitous in nature and are
32 involved in the generic role of lipid transfer between membranes. In plants however other functions
33 have been proposed, relating mostly to defence against bacteria and pathogenic fungi, as well as
34 abiotic stress response ((Kader, 1996; Wang *et al.*, 2012) and references therein). Non-specific
35 LTPs refer to a group of LTPs able to bind different classes of lipids (Kader, 1996). Structurally,
36 plant nsLTPs are separated into five distinct groups, with the two main groups (Type-I and Type-II)
37 common to all the species investigated so far. Due to specific deletions, Type-II LTPs (of about 7
38 kDa) are smaller than those of the Type-I (of about 9 kDa) (Ballmer-Weber, 2002; Wang *et al.*,
39 2012).

40 From an allergenic point of view plant nsLTPs belong to the large Prolamin protein
41 superfamily, which includes several allergen families (Egger *et al.*, 2010). Their most common
42 route of exposure is ingestion (see the AllFam Database of Allergene Families at
43 <http://www.meduniwien.ac.at/allergens/allfam/>). Allergens of the Prolamin superfamily share a
44 conserved scaffold consisting of bundles of four α -helices stabilized by four disulfide bridges
45 (Egger *et al.*, 2010). This compact structure is usually evoked to explain the persistence of
46 allergenicity for cooked food proteins (Mills *et al.*, 2009). Nonetheless, the allergenic potency of
47 some molecules is also due to the presence of linear epitopes, which maintain their allergenicity
48 even after food cooking or other transformation processes, as in the case of allergenic nsLTPs
49 identified in sterilized peach juice, polenta, and in fermented products such as wine and beer (Asero
50 *et al.*, 2001; Mills *et al.*, 2009; Pastorello *et al.*, 2003a; Pastorello *et al.*, 2003b). Most LTPs can be
51 considered as food allergens. Inhalation and skin contact have been identified as other possible
52 routes of exposure to allergenic LTPs (Egger *et al.*, 2010). Plant allergenic nsLTPs show a variable

degree of sequence identities (about 25-80%), even though proteins from evolutionarily close plant families have identities values higher than 65% (Radauer *et al.*, 2008). Under this complex scenario, the identification and analysis of new LTP molecules can provide additional data to unveil relationships among allergenicity and structural features, including stability to denaturing conditions. Furthermore, the availability of single allergens produced as recombinant proteins is considered of main value for allergology studies based on component-resolved diagnosis (Egger *et al.*, 2010; Hoffmann-Sommergruber and Mills, 2009; Luengo and Cardona, 2014; Valenta *et al.*, 1999; Van Winkle and Chang, 2014).

The two tomato LTP allergens listed in the IUIS database belong to the Type-I (Sola I 3) and Type-II (Sola I 6) groups. While the structural features and allergenicity of the Type-II Sola I 6 allergen have been recently described (Giangrieco *et al.*, 2015), studies aimed at better characterizing tomato LTP 9 kDa isoforms are limited and confusing. Indeed, the Sola I 3 allergen is reported as corresponding to a LTP protein (P93224) for which allergenicity has not been investigated (Plant *et al.*, 1991). Allergenicity of other tomato 9 kDa LTP isoforms from different tomato varieties, has been verified (Bencivenni *et al.*, 2014; Le *et al.*, 2006; Pravettoni *et al.*, 2009), but sequence and structural data are still missing.

Here we report the identification of a novel 9 kDa LTP isoform in the tomato San Marzano variety. San Marzano tomatoes are typical of Southern Italy and used as components of many recipes. They are sold worldwide as a canned partially cooked product. The novel LTP isoform has been produced as a recombinant molecule and characterized for its lipid-binding activity at different temperatures. It has also been analysed by a modelling-based approach for its potential allergenicity. These data can contribute to a larger analysis of the diffusion and variability of LTP isoforms in tomato, useful in future investigations concerning their allergenic potency.

2. MATERIALS AND METHODS

2.1. DNA and RNA extraction

Seeds of tomato (variety San Marzano) were obtained from the Department of Agronomy (University of Naples) and sown in pots for plants production. Genomic DNA was extracted from young leaves, using the Gene Elute Plant genomic kit (Sigma Aldrich).

Total RNA was extracted from mature tomato fruits harvested at about 45-50 days after flowering, using a phenol method. 200 µg of starting material were homogenised in 1 ml of Trizol (Life Technologies) and incubated for 5 min at RT. After extraction with 200 µL of chloroform, RNA was precipitated with 500 µL of isopropanol. After centrifugation, the pellet was washed with 70% ethanol, air dried and then resuspended in 50 µL of DEPC-treated water. The resulting RNA was treated with Rnase DNase I (Invitrogen) according to the manufacturer's instructions.

All the DNA and RNA samples were analyzed by Nano-Drop ND1000 spectrophotometer (Thermo Scientific), and checked on 1-1.5% agarose gels.

2.2. PCR analysis

PCR amplifications of genomic DNA were carried out using 200 ng of total DNA, in the presence of 20 pmol of each primer (Table I), 0.2 mM of each dNTP, 2.5U of PCR Extender Polymerase Mix (5PRIME, Eppendorf) in 50 µL of the buffer solution indicated by the supplier. After a denaturation step at 94°C for 5 min, amplification reactions were run for 35 cycles for 1 min at 94°C, 1 min at the annealing temperature chosen for the two primers, and 1 min at 72°C. A final elongation step was run at 72°C for 7 min. The amplification products were cloned in pGEM-T vector (Promega) for sequencing analysis.

Table I. Primers used for PCR and RACE analysis

Name	Sequence (5'-3')
TomGe-LTP_F	GTGGTTGCACCCCATGCAGAG
TomGe-LTP_R	GGGCTGATCTTGTAAGGRATGTTGAC
TomSM-SM21_F	GCATGCAATTGCCTAAAATCA
TomSM-SM21_FA	AAAGCTGCTGGTCTCCCTAGTGTGTTG
TomSM-SM22_F	GACTACCGCGGATCGAAAGACA
TomSM-SM22_FA	GCTGCTAATGCTATAAAGGGAATTGATTT
Tom21_LTP_R	CTCATGAAACCAAATTCTGATG
Tom22_4_LTP_R	TSATCAACYTTACWRRACY
SM4_pET_F	AAAAACATATGGCACTAAGTTGCGGCCAGG
SM4_pET_R	TTTTTGATCCTTACAAAACCTTTCGAGCAGTCAGTAAAAGGGC
Anchor primer (AP)	GACCACGCGTATCGATGTC
Oligo-dT_AP	GACCACGCGTATCGATGTCT ₍₁₇₎

2.3. 3'-RACE (Rapid Amplification of cDNA ends analysis)

For 3' RACE, single strand cDNA was synthesized starting from 0.5 µg of total RNA and 1 µL of 50 µM anchored-oligodT, using the SuperScript First-Strand Synthesis System (Life Technologies) according to manufacturer's instructions.

The obtained cDNA was then used as substrate for a PCR reaction using a gene-specific primer (TomSM-22_F) and the anchor primer (AP), and then for a hemi-nested PCR using a gene-specific primer (TomSM-22_FA) and the AP primer (Table I).

2.4. Cloning, expression and purification of tomato LTP isoform

The nucleotide sequence corresponding to the mature LTP SM22 protein was PCR amplified from genomic DNA using a forward primer (SM4_pET_F) and a reverse primer (SM4_pET_R) containing the *NdeI* and *BamHI* restriction site sequences respectively (Table I). The amplified fragment was then inserted into the pET11b expression vector (Novagen) and the recombinant vectors were checked by sequencing analysis.

After transformation in *E. coli Origami* (DE3) cells (Invitrogen), the expression of LTP SM22 was induced by adding 1 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG). After overnight incubation at 28°C, cells were centrifuged and resuspended in 10 mL of 10 mM, pH 7.2

phosphate buffer (buffer A), 1% Triton. Cells were lysed using the Vibra-cell sonicator (Sonics) with 5 cycles of 15 sec at 30% amplitude. The lysate was centrifuged at 25.800xg for 30 min and the supernatant was subjected to ion exchange chromatography on a HiTrap SP Sepharose Fast Flow column (GE Healthcare). After loading cell lysate, the column was washed with five volumes of buffer A and proteins eluted with a step gradient (100, 200, 300, 500 and 1000 mM NaCl in buffer A). Eluted fractions were monitored at 280 nm and analyzed by SDS-PAGE.

2.5. Mass spectrometry analysis

Protein bands observed in SDS-PAGE were carefully excised and subjected to in-gel trypsin digestion according to Shevchenko and colleagues (Shevchenko *et al.*, 2006). Nano-LC MS/MS analyses were performed as already described (Pallotta *et al.*, 2013). Data were searched against the National Center for Biotechnology Information nonredundant database (NCBI nr, Version 20150127) using computer software (Mascot, in-house version 2.2, Matrix Science, London, UK) with the following constraints: taxonomy = Viridiplantae; enzyme = trypsin; missed cleavage = 1; peptide and fragment mass tolerance = ± 0.3 Da; fixed modifications = carbamidomethyl (Cys); variable modifications = Oxidation (Met); Gln->pyro-Glu (N-term Q), Glu->pyro-Glu (N-term E). For positive identification, the score of the result of $[-10 \times \text{Log}(P)]$ had to be over the significance threshold level ($p < 0.05$).

2.6. Fluorescence binding assays

Fluorescence of the recombinant LTP rSM-22 was measured at 25 °C with an excitation wavelength of 275 nm and recording the emission spectra in the 280-400 nm range, using the FP-8300 spectrofluorometer (JASCO). Before measurement, proteins in the 200 mM NaCl fractions were diluted to 10 μ M in buffer A, 25 mM NaCl.

Lipid binding assay for rSM-22 was performed using 1-palmitoil-2-lysophosphatidylcholine (Lyso-C₁₆, Sigma) as substrate (Gomar *et al.*, 1996). Emission spectra were recorded as for the purified protein.

To check the thermostability of rSM-22, aliquots of diluted protein were incubated at increasing temperatures (60, 75, 95, 105 and 120 °C) for 5 min and then put in ice-bath until measurements. Fluorescence was measured for both the purified protein and its Lyso C₁₆ -complex in the presence of four molar excess of substrate. All the experiments were done in technical triplicate.

3. RESULTS AND DISCUSSION

3.1. Identification of gene sequences

In tomato six mRNA sequences coding for 9 kDa LTPs have been identified (Table II), two of which (AM051295 and AM051296) were reported as coding for allergens (Le *et al.*, 2006). Multialignment of the six encoded polypeptides is shown in Figure 1.

Table II. Tomato Type-I LTP sequences

mRNA Acc. n.	Variety	Induction conditions/tissue	Protein Acc. n.	Ref.
U81996	Ailsa Craig	Drought and Abscissic acid/aerial vegetative tissues	P93224	(Plant <i>et al.</i> , 1991)
X56040	Rutgers-Marglobe	NaCl/stems	P27056	(Torres-Schumann <i>et al.</i> , 1992)
AM051295	Micro-Tom	--/fruit peel	Q4A1N1	(Le <i>et al.</i> , 2006)
AM051296	“ “	--/fruit	Q4A1N0	(Aoki <i>et al.</i> , 2010)
AK224707	“ “	“ “	K4D1U9	(Aoki <i>et al.</i> , 2010)
HM367634	cerasiforme	<i>Phytophthora infestans</i> /leaves	E2GLP6	Unpublished

In order to characterize Type-I LTPs of the tomato San Marzano variety, we started a systematic approach, described herein for the LTP SM-22 isoform. A gene portion containing the mature protein coding sequence (cds) (Figure 1) was identified according to the following strategy: i) PCR

172 amplification, using highly conserved sequences for primer design; ii) 3'-RACE analysis for the
173 identification of the *orf* 3'-terminal sequence using gene specific primers deduced from the first
174 amplification; iii) final PCR amplification of the mature protein cds using primers deduced from the
175 first two steps (see legend of Figure 1 for further details). It is worthwhile to mention that, due to
176 the presence in the tomato genome of a LTP gene family, all the amplification reactions above
177 indicated showed more than one amplification product.

178

179

180

181 **A**

182

183

```

184 Q4A1N1 MEMVNKIACFVLLCMVVVAPHAELTCGQVTSTLAPCLPYLMNRGPLRNCCDGVKGLLGQ
185 E2GLP6 -----ALTCGQVESGLAPCLPYLQKGKPLGGCCRGVKGLLGA
186 K4D1U9 MEMSSKIACFIVLCMIVVAPHGEALSCGQVESGLAPCLPYLQKGKPLGGCCRGVKGLLGA
187 Q4A1N0 MEMFGKIACFVFCMVVAPHAESLSCGEVTSGLAPCLPYLEGRGPLGGCCGGVKGLLGA
188 P93224 MEMFGKIACFVFCMVVAPHAESLSCGEVTSGLAPCLPYLEGRGPLGGCCGGVKGLLGA
189 P27056 MEMVSKIACFVLLCMVVVAPHAELTCGQVTAGLAPCLPYLQGRGPLGGCCGGVKNNLLGS
190          :*:***:*****:*** ** **
191
192 Q4A1N1 AKTTVDRQAACCTCLKSAASSFTGLNLGKAAALPNTCSVNIPIYKISPSTDCSKVQ---
193 E2GLP6 AETPADRKTAACCTCLKSAANAIGLNLGKAAGIPSACGVSIPIYKISPFTDCSKVQKTL
194 K4D1U9 AKTPADRKTAACCTCLKSAANAIGLNLGKAAGIPSACGVSIPIYKISPFTDCSKVQ---
195 Q4A1N0 AKTPEDRKTAACCTCLKSAANSIKGIDTGKAAGLPVCGVNIPIYKISPSTDCSTVQ---
196 P93224 AKTPEDRKTAACCTCLKSAANSIKGIDTGKAAGLPVCGVNIPIYKISPSTDCSTVQ---
197 P27056 AKTTADRKTAACCTCLKSAANAIGIDLNKAAGIPSVCKVNIPIYKISPSTDCSTVQ---
198          *: * **::*****:..*: ***: * * * ***** ***: **

```

199

200

201 **B**

202

203

```

204          *
204 GTGGTTGCACCCCATGCAGAGGCACTAAGTTGCGGCCAGGTTGAATCTGGCTTAGCTCCA 60
205 V V A P H A E A L S C G Q V E S G L A P
206 TGCTCCCTTATCTTCAGGAAAAGGCCCTCTAGGAGGGTGTGTCGTGGAGTTAAGGGT 120
207 C L P Y L Q G K G P L G G C C R G V K G
208 CTATTGGGCGCAGCCAAGACCCCGCGGATCGAAAGACAGCATGCACTTGCCTCAAATCA 180
209 L L G A A K T P A D R K T A C T C L K S
210 GCTGCTAATGCTATTAAAGGCCCTTAATCTAGGCAAAGCCGCTGGTATACCTAGTGCTTGT 240
211 A A N A I K G L N L G K A A G I P S A C
212 GCGTAAGCATTCTTACAAGATCAGCCCTTTTACTGACTGCTCGAAGtacgttatacat 300
213 G V S I P Y K I S P F T D C S K
214 ttcttttagtttaattcatatacgacgattaataaaattatgatttaacatgttaaaat 360
215 aatgatgtgtacttgcattgttgaatgttggttgaatgttatattgcagggttc 420
216 agtaaggttgatgaaagttcaattttccaaggcataattggggaaaaataagacggatat 480
217 atctatcgatcgatcgagtttaattccatcattttataatatgttgctctctttatattgt 540
218 atttttgtttgttgagttatattgttcgAGTTTGTAA 582
219          V L -
220

```

220

221 **Figure 1. Tomato 9 kDa LTP sequences.** A) Multialignment of LTPs. Alignment was obtained by
222 Clustal program (Sievers *et al.*, 2011). Conserved regions used for the construction of amplification

primers TomGe-LTP_F and TomGe-LTP_R (Table I) are boxed. An asterisk indicates the first amino acid of mature proteins. A vertical arrow indicates the amino acid whose codon is interrupted by an intron sequence. B) Tomato genomic sequence containing the mature LTP SM-22 cds. Part of the gene coding for the LTP SM-22 is shown. The sequence has been deposited at the European Nucleotide Archive (ENA) (<http://www.ebi.ac.uk/ena>) with the accession number LN810550. Exon sequences are reported in capital letters; intron sequence is reported in small letters. Bold nucleotides correspond to the region where the primers used for the first genomic amplification (TomGe-LTP_F and TomGe-LTP_R, see Table I) were designed (see also panel A). Sequences in light gray background correspond to the region where the primers used for the 3'-RACE analysis (TomSM-22_F and TomSM-22_FA) were designed. Nucleotides with gray background correspond to the degenerated primer TOM22_4_LTP_R used for final PCR on genomic DNA in couple with primer TomGe-LTP_F. Underlined nucleotides correspond to primers SM4_pET_F and SM4_pET_R (joining the two exons), respectively, used for amplifying the mature protein.

3.1.1. First genomic amplification. Four different sequences were obtained from the amplification of genomic DNA with primers TomGe-LTP_F and TomGe-LTP_R (see Figure S1 in Supplementary Material). In the comparison to known tomato LTP sequences one sequence (SM-16) is identical to Q4A1N1, the other three sequences correspond to previously unknown LTP genes. Sequences SM-17 and SM-21 show the highest identities with the Q4A1N1 sequence (98 and 87 % respectively); sequence SM-22 has the highest identity with P27056 (97%).

3.1.2. RACE analysis. On the basis of genomic sequences SM-21 and SM-22, we designed two pairs of forward primers (Table I) to be used for the PCR and hemi-nested PCR steps of the 3'-RACE. The analysis was carried out on single-strand cDNA synthesized on RNA extracted from San Marzano tomato fruits, starting with the oligo-dT primer. RNA was purified from complete fruit since the presence of allergenic LTPs in tomato pulp has also been reported (Pravettoni *et al.*, 2009). Double-stranded cDNA molecules were obtained by PCR using oligonucleotides TomSM-22_F (or TomSM-21_F) and the anchor primer (AP). PCR amplification products were then subjected to hemi-nested PCR, using the primers TomSM-22_FA (or TomSM-21_FA) as novel forward primers. Alignments of sequences obtained from the RACE experiments are reported in Supplementary Material (Figure S2).

3.1.3. Second genomic amplification. According to the sequences obtained by the 3'-RACE analysis, we designed new reverse primers to be coupled with primer TomGe-LTP_F (see Figure 1). Primer Tom22_4_LTP_R corresponds to the last ten nucleotides of the coding sequences,

plus the first nine nucleotides of the 3' untranslated region (UTR). Some degeneracies were introduced in order to take into account any sequence heterogeneity. Primer Tom21_LTP_R corresponds to the first nucleotides downstream of the stop codon.

Using the pair of primers TomGe-LTP_F and Tom22_4_LTP_R, a specific amplification product was obtained (Figure 1). PCR amplifications carried out using primers TomGe-LTP_F and Tom21_LTP_R always resulted in non-specific sequences (not shown).

3.2. Cloning and expression of LTP SM-22

Genes encoding LTPs listed in Table II are constituted by a first large exon of 335 nucleotides (coding for a 23 amino acid presequence, plus 88 amino acids of the mature protein) and a second small exon of only 10 nucleotides. To obtain the complete coding sequence for the mature form of LTP SM-22, the genomic sequence was used as substrate for a conclusive PCR using a forward primer (SM4_pET_F) containing the first 19 nucleotides of the SM-22 mature form coding sequence, and a reverse primer (SM4_pET_R) corresponding to the 10 nucleotides of the second exon joined to the last 22 nucleotides of the first exon. Both primers contain extra-sequences for specific restriction sites (*NdeI* and *BamHI*, respectively) for directional ligation into pET-11b vector. For the expression of the recombinant tomato LTP protein, whose peach homolog contains four disulfide bonds (Pasquato *et al.*, 2006), *Origami* (DE3) cells were used. Recombinant protein was expressed by IPTG induction and purified from lysed cells by ion-exchange chromatography through HiTrap Sepharose column using a stepwise NaCl gradient. Contents of the fractions was verified by SDS-PAGE (Figure 2). Fractions 5-9 were pooled, and protein concentration evaluated (0.33 ± 0.04 mg/mL), using the available tobacco LTP (Uniprot Q42952) molar extinction coefficient $\epsilon_{280} = 3480 \text{ M}^{-1} \text{ cm}^{-1}$. The identity of the protein in the major band at about 9 kDa in the 200 mM NaCl fractions was obtained by mass-spectrometry analysis of the corresponding tryptic digest. The analysis confirmed the presence of the recombinant LTP SM-22 protein (Supplementary material, Table SI).

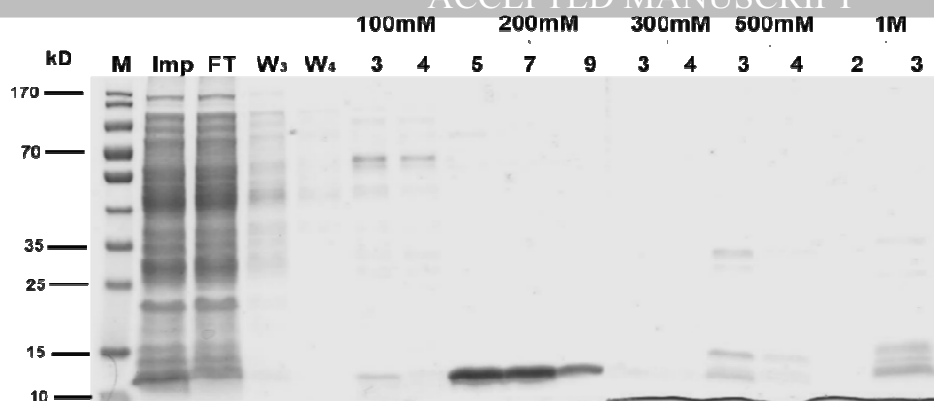


Figure 2. SDS-PAGE analysis of rLTP SM-22 expression and purification. From left to right: Molecular weight marker (M); Total soluble proteins from transformed cells (Imp); proteins not retained in the chromatographic column (FT); washing fractions (W₃, W₄); proteins from elution steps at different NaCl concentrations (numbers refer to respective fractions).

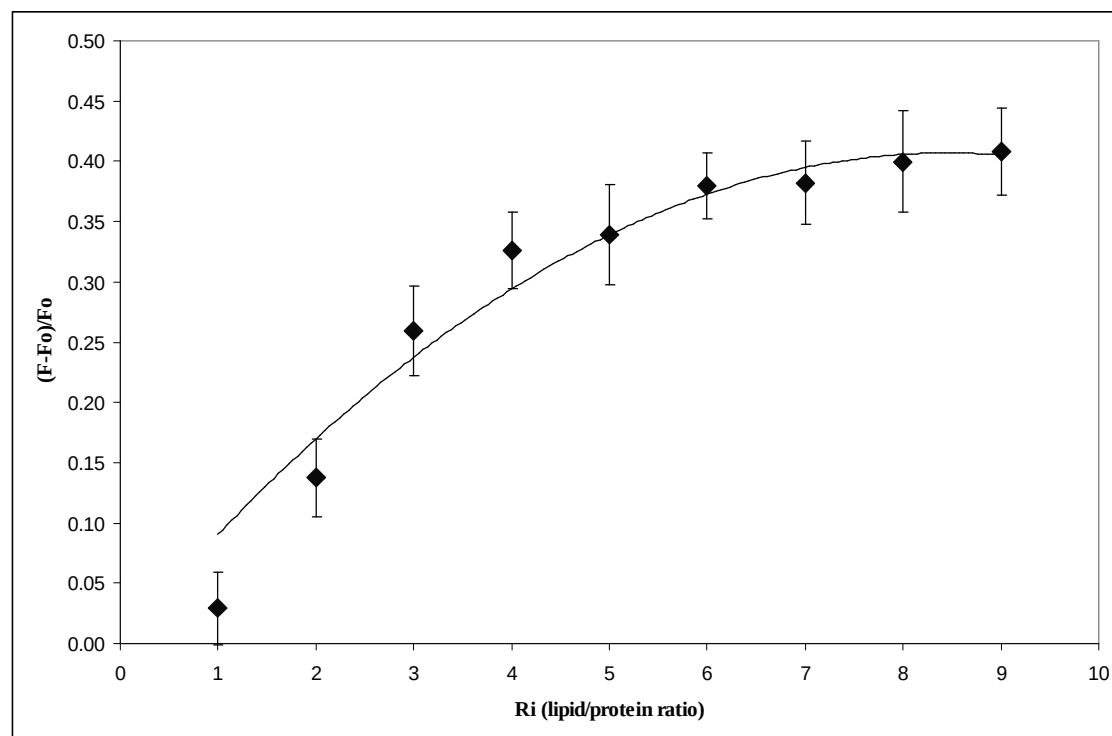
3.3. Functionality assays of recombinant LTP SM-22

The proper folding into a functionally active form of rSM-22 was verified by fluorescence assays carried out using 1-palmitoyl-sn-glicero-3-phosphocoline (Lyso-C₁₆) as substrate. The intrinsic fluorescence of LTP molecules, due to two highly conserved tyrosine residues, is enhanced after lipid binding, also allowing to measure the complex stability (Gomar *et al.*, 1996). The intrinsic fluorescence of rSM-22 at 10 μ M concentration, was recorded at room temperature using a 275 nm wavelength for excitation. It results in an emission curve (registered in the 280-400 nm range) with a maximum at 305 nm (Supplementary Material, Figure 3S), similar to those observed for other LTPs (Ramazzina *et al.*, 2012). In Figure 3 the fluorescence intensity of rSM-22 in the presence of increasing amounts of Lyso-C₁₆, is reported. The dissociation constant for the rSM-22/Lyso-C₁₆ complex was calculated by fitting the relative emission curves using GraphPad Prism software. A K_d of $(0.56 \pm 0.04)10^{-4}$ was obtained which is in good agreement with values reported for other plant LTPs (Gomar *et al.*, 1996).

The thermostability of the protein was evaluated by following its capacity to bind Lyso-C₁₆ after incubation at different temperatures. Even though increasing fluorescence intensities due to complex formation is quite evident up to 105 °C (Figure 3), a general decrease of the difference between emission values at 305 nm of lipid-associated and free protein can be observed, indicating

307 a slightly lowered complex stability. After incubation at 120 °C the protein shows a completely
308 different fluorescence capacity, in the presence of substrate or not. The thermostability of rSM-22
309 up to 105 °C is a good indication of the correct folding of the protein, including the possible
310 establishment of disulfide bridges among the pairs of cysteines 4-50, 14-27, 28-73 and 48-87, as
311 identified in the peach LTP (Pasquato *et al.*, 2006). The recombinant molecule however is somehow
312 unfolded after 120 °C heating and can no longer efficiently bind the lipid.

313



B

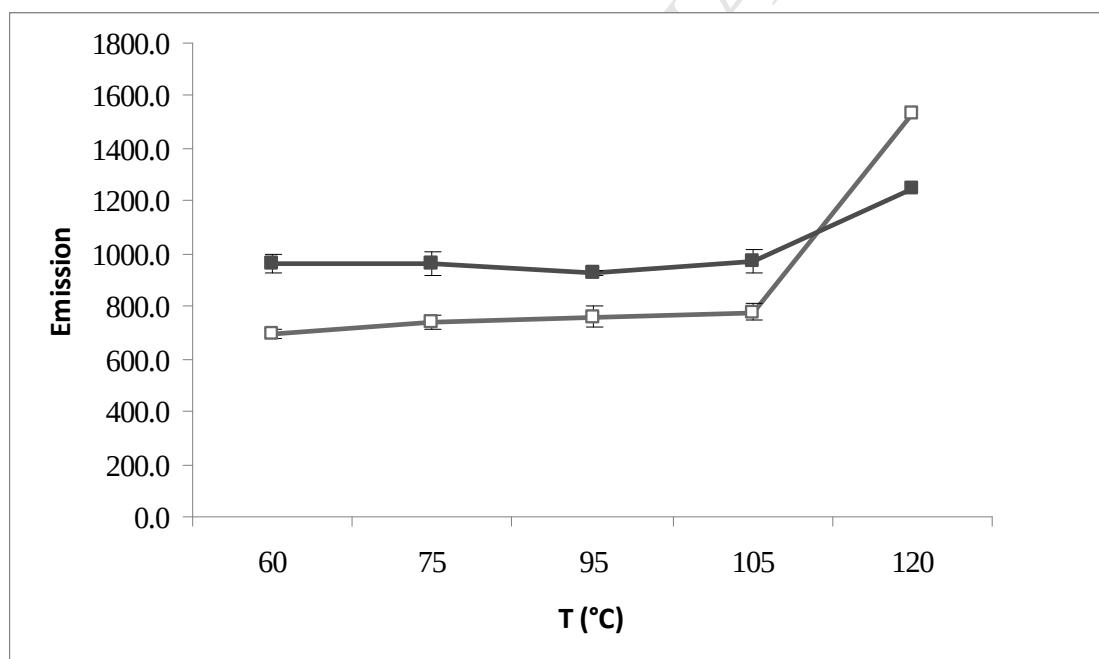


Figure 3. rSM-22 fluorescence assays. A) Relative fluorescence intensity of 10 μ M rSM-22 LTP at increasing amounts of Lyso-C₁₆. B) rSM-22 fluorescence intensity after heat-treatment. After cooling, fluorescence was measured at room temperature, using four-fold molar excess of Lyso-C₁₆. Open and full squares refer to measurements without and with substrate, respectively. In both experiments fluorescence was measured at 305 nm, after excitation at 275 nm. Points represent mean values from three independent experiments.

3.4. Structural analysis

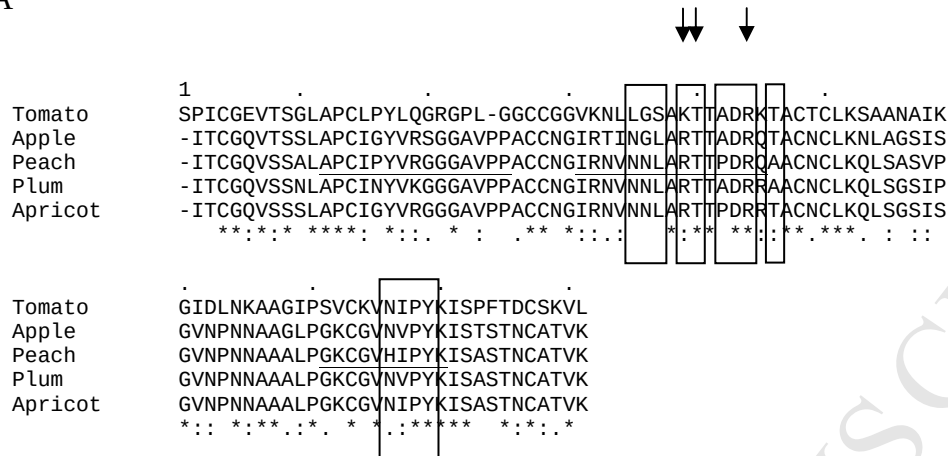
Allergenic plant LTP proteins are currently considered as pan-allergens. Attempts to identify possible epitopes have been carried out by immunoreactivity analysis of site directed mutants of the peach LTP allergen (Garcia-Casado *et al.*, 2003), synthetic peptides derived from LTP allergens of apple, apricots, peach and plum (Borges *et al.*, 2008; Garcia-Casado *et al.*, 2003), and IgE-reacting phage-displayed random dodecapeptides (Pacios *et al.*, 2008). According to these analyses a map of linear and conformational epitopes can be established (Figure 4). With the exception of a three nucleotide region (corresponding to the first boxed sequences), tomato LTP SM22 shows high degrees of sequence similarity with the allergenic LTPs identified in *Rosaceae*. For example, the three amino acids identified as strictly required for allergenicity in peach LTP (R39, T40 and R44) (Garcia-Casado *et al.*, 2003) are highly conserved in SM-22, which contains the amino acids K39, T40 and R44. Thanks to the epitope mapping studies and taking advantage from the available three-dimensional structure of peach LTP (Pasquato *et al.*, 2006) mapping of conserved sequences corresponding to possible epitopes on the three-dimensional structure of rSM-22 has been obtained by the One-to-One threatening tool of the Phyre2 (Protein Homology/analogy Recognition Engine) server (Kelley and Sternberg, 2009) using the structure of the peach LTP allergen as reference. It is interesting to note that they are all exposed on the surface of the protein, strengthening their possible involvement in IgE recognition.

Conclusion

This is the first report of a tomato 9 kDa LTP characterized in its recombinant form for activity, resistance to thermal denaturation and potential allergenic structural characteristics. Recombinant allergens are of fundamental interest for advanced research in allergology. They constitute a pure form of allergen to be used in diagnosis and are the basis of studies for establishing relationships between structural characteristics and allergenicity. Moreover in the cases of allergens which are highly diffused in nature, the availability of recombinant allergens can considerably accelerate the comparison of the allergenicity grade of the different isoforms. More LTP isoforms should be

characterized in the available tomato varieties, possibly also studying their capacity to react with IgEs of allergic patients. These studies will provide a comprehensive analysis of tomato LTPs allergenicity, useful for both clinical and plant breeding applications.

A



B

R44 T40 K39

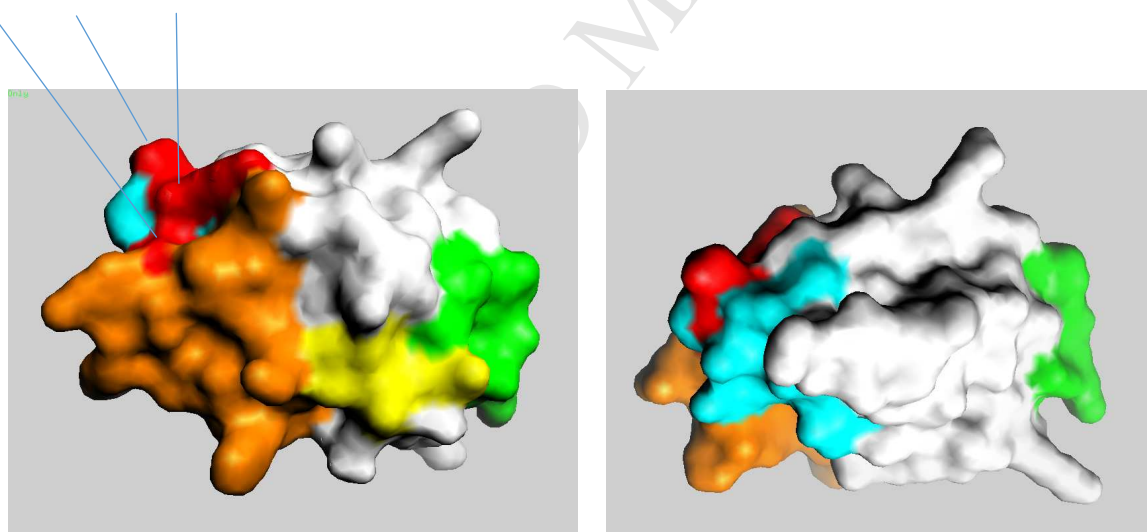


Figure 4. SM-22 structure analysis. A) Multialignment of LTPs and epitopes. Sequences refer to tomato LTP rSM-22, apple (EBI accession number Q9M5X7), peach (Pastorello *et al.*, 1999), plum (Pastorello *et al.*, 2001), apricot (Pastorello *et al.*, 2001). Multialignment was obtained by Clustal Omega (Sievers *et al.*, 2011). Sequences with a light gray background correspond to linear epitopes detected by using 15 amino acid long synthetic peptides (Borges *et al.*, 2008). Underlined sequences correspond to linear epitopes detected by using 10 amino acid long synthetic peptides (Garcia-Casado *et al.*, 2003). Vertical arrows indicate amino acids which have been identified as strictly required for allergenicity in peach fruits (Garcia-Casado *et al.*, 2003). Boxes indicate regions corresponding to possible conformational epitopes (Pacios *et al.*, 2008). Tomato sequences with a dark gray background correspond to conserved epitope regions. B) Localization of possible epitopes in the deduced structure of rLTP SM-22. The three-dimensional structure of rLTP SM-22,

as deduced by modelling using the peach LTP structure (PDB: 2ALG, chain A) as reference, is shown. The figures in the two panels are interconvertible by 180° rotation around a horizontal axis passing through the center of the structure. In green, amino acids 17-24 (see Figure 4A); in cyan, amino acids 37-51; in yellow, amino acids 64-69; in orange, amino acids 76-91. The three amino acids identified as necessary for allergenicity in peach LTP are in red and indicated.

CONTRIBUTIONS

MV, CL and LRC: design and interpretation of all experiments.

MV, CL, IF and AP: carried out molecular biology experiments and characterization of rSM-22.

SR: designed and carried out mass spectrometry experiments.

MV, CL and LRC: wrote the manuscript.

ACKNOWLEDGEMENTS

Authors wish to thank Prof. Rosa Rao (University of Naples, Italy) for seeds of the San Marzano tomato variety and Prof. Lello Zolla (University della Tuscia, Viterbo, Italy) for his continuing support. The technical assistance of Mr. Bruno Gattulli (IBBE, Bari, Italy) is also acknowledged. Dr Anthony B. Comeau (Brown University, Providence – Rhode Island) is warmly acknowledged for his critical reading of the manuscript.

This work was also supported by a grant from Fondazione Cassa di Risparmio di Puglia for the study of food allergies.

- Aoki K, Yano K, Suzuki A, Kawamura S, Sakurai N, Suda K, *et al.* (2010) Large-scale analysis of full-length cDNAs from the tomato (*Solanum lycopersicum*) cultivar Micro-Tom, a reference system for the Solanaceae genomics. *BMC Genomics* **11**: 210.
- Asero R, Mistrello G, Roncarolo D, Amato S and van Ree R (2001) A case of allergy to beer showing cross-reactivity between lipid transfer proteins. *Ann Allergy Asthma Immunol* **87**: 65-67.
- Ballmer-Weber BK (2002) Lipid transfer protein as a potential panallergen? *Allergy* **57**: 873-875.
- Bencivenni M, Faccini A, Bottesini C, Rao R, Detoraki A, Ridolo E, *et al.* (2012) Assessing allergenicity of different tomato ecotypes by using pooled sera of allergic subjects: identification of the main allergens. *Eur Food Res Technol* **234**: 405-414.
- Bencivenni M, Faccini A, Zecchi R, Boscaro F, Moneti G, Dossena A and Sforza S (2014) Electrospray MS and MALDI imaging show that non-specific lipid-transfer proteins (LTPs) in tomato are present as several isoforms and are concentrated in seeds. *J Mass Spectrom* **49**: 1264-1271.
- Bohn T, Blackwood M, Francis D, Tian Q, Schwartz SJ and Clinton SK (2013) Bioavailability of phytochemical constituents from a novel soy fortified lycopene rich tomato juice developed for targeted cancer prevention trials. *Nutr Cancer* **65**: 919-929.
- Borges JP, Barre A, Culerrier R, Granier C, Didier A and Rouge P (2008) Lipid transfer proteins from Rosaceae fruits share consensus epitopes responsible for their IgE-binding cross-reactivity. *Biochem Biophys Res Commun* **365**: 685-690.
- Egger M, Hauser M, Mari A, Ferreira F and Gadermaier G (2010) The role of lipid transfer proteins in allergic diseases. *Curr Allergy Asthma Rep* **10**: 326-335.
- Foetisch K, Son DY, Altmann F, Aulepp H, Conti A, Haustein D and Vieths S (2001) Tomato (*Lycopersicon esculentum*) allergens in pollen-allergic patients. *Eur Food Res Technol* **213**: 259-266.
- Garcia-Casado G, Pacios LF, Diaz-Perales A, Sanchez-Monge R, Lombardero M, Garcia-Selles FJ, Polo F, Barber D and Salcedo G (2003) Identification of IgE-binding epitopes of the major peach allergen Pru p 3. *J Allergy Clin Immunol* **112**: 599-605.
- Giangrieco I, Alessandri C, Rapaia C, Santoro M, Zuzzi S, Tuppo L, Tamburrini M, D'Avino R, Ciardiello MA and Mari A (2015) Structural features, IgE binding and preliminary clinical findings of the 7kDa Lipid Transfer Protein from tomato seeds. *Mol Immunol* **66**: 154-163.
- Gomar J, Petit MC, Sodano P, Sy D, Marion D, Kader JC, Vovelle F and Ptak M (1996) Solution structure and lipid binding of a nonspecific lipid transfer protein extracted from maize seeds. *Protein Sci* **5**: 565-577.
- Hoffmann-Sommergruber K and Mills EN (2009) Food allergen protein families and their structural characteristics and application in component-resolved diagnosis: new data from the EuroPrevall project. *Anal Bioanal Chem* **395**: 25-35.
- Kader JC (1996) Lipid-Transfer Proteins in Plants. *Annu Rev Plant Physiol Plant Mol Biol* **47**: 627-654.
- Kelley LA and Sternberg MJ (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4**: 363-371.
- Kim C, Kwon S, Lee G, Lee H, Choi J, Kim Y and Hahn J (2009) A database for allergenic proteins and tools for allergenicity prediction. *Bioinformatics* **3**: 344-345.
- Larramendi CH, Ferrer A, Huertas AJ, Garcia-Abujeta JL, Andreu C, Tella R, *et al.* (2008) Sensitization to tomato peel and pulp extracts in the Mediterranean Coast of Spain: prevalence and co-sensitization with aeroallergens. *Clin Exp Allergy* **38**: 169-177.
- Le LQ, Lorenz Y, Scheurer S, Fotisch K, Enrique E, Bartra J, Biemelt S, Vieths S and Sonnewald U (2006) Design of tomato fruits with reduced allergenicity by dsRNAi-mediated inhibition of ns-LTP (Lyc e 3) expression. *Plant Biotechnol J* **4**: 231-242.

- Luengo O and Cardona V (2014) Component resolved diagnosis: when should it be used? *Clin Transl Allergy* **4**: 28.
- Mari A, Rasi C, Palazzo P and Scala E (2009) Allergen databases: current status and perspectives. *Curr Allergy Asthma Rep* **9**: 376-383.
- Mills EN, Sancho AI, Rigby NM, Jenkins JA and Mackie AR (2009) Impact of food processing on the structural and allergenic properties of food allergens. *Mol Nutr Food Res* **53**: 963-969.
- Pacios LF, Tordesillas L, Cuesta-Herranz J, Compes E, Sanchez-Monge R, Palacin A, Salcedo G and Diaz-Perales A (2008) Mimotope mapping as a complementary strategy to define allergen IgE-epitopes: peach Pru p 3 allergen as a model. *Mol Immunol* **45**: 2269-2276.
- Pallotta V, D'Alessandro A, Rinalducci S and Zolla L (2013) Native protein complexes in the cytoplasm of red blood cells. *J Proteome Res* **12**: 3529-3546.
- Pasquato N, Berni R, Folli C, Folloni S, Cianci M, Pantano S, Helliwell JR and Zanotti G (2006) Crystal structure of peach Pru p 3, the prototypic member of the family of plant non-specific lipid transfer protein pan-allergens. *J Mol Biol* **356**: 684-694.
- Pastorello EA, Farioli L, Pravettoni V, Giuffrida MG, Ortolani C, Fortunato D, *et al.* (2001) Characterization of the major allergen of plum as a lipid transfer protein. *J Chromatogr B Biomed Sci Appl* **756**: 95-103.
- Pastorello EA, Farioli L, Pravettoni V, Ortolani C, Fortunato D, Giuffrida MG, Perono Garoffo L, Calamari AM, Brenna O and Conti A (2003a) Identification of grape and wine allergens as an endochitinase 4, a lipid-transfer protein, and a thaumatin. *J Allergy Clin Immunol* **111**: 350-359.
- Pastorello EA, Farioli L, Pravettoni V, Ortolani C, Ispano M, Monza M, *et al.* (1999) The major allergen of peach (*Prunus persica*) is a lipid transfer protein. *J Allergy Clin Immunol* **103**: 520-526.
- Pastorello EA, Pompei C, Pravettoni V, Farioli L, Calamari AM, Scibilia J, *et al.* (2003b) Lipid-transfer protein is the major maize allergen maintaining IgE-binding activity after cooking at 100 degrees C, as demonstrated in anaphylactic patients and patients with positive double-blind, placebo-controlled food challenge results. *J Allergy Clin Immunol* **112**: 775-783.
- Plant AL, Cohen A, Moses MS and Bray EA (1991) Nucleotide sequence and spatial expression pattern of a drought- and abscisic Acid-induced gene of tomato. *Plant Physiol* **97**: 900-906.
- Pravettoni V, Primavesi L, Farioli L, Brenna OV, Pompei C, Conti A, Scibilia J, Piantanida M, Mascheri A and Pastorello EA (2009) Tomato allergy: detection of IgE-binding lipid transfer proteins in tomato derivatives and in fresh tomato peel, pulp, and seeds. *J Agric Food Chem* **57**: 10749-10754.
- Radauer C, Bublin M, Wagner S, Mari A and Breiteneder H (2008) Allergens are distributed into few protein families and possess a restricted number of biochemical functions. *J Allergy Clin Immunol* **121**: 847-852 e847.
- Radauer C, Nandy A, Ferreira F, Goodman RE, Larsen JN, Lidholm J, *et al.* (2014) Update of the WHO/IUIS Allergen Nomenclature Database based on analysis of allergen sequences. *Allergy* **69**: 413-419.
- Ramazzina I, Amato S, Passera E, Sforza S, Mistrello G, Berni R and Folli C (2012) Isoform identification, recombinant production and characterization of the allergen lipid transfer protein 1 from pear (Pyr c 3). *Gene* **491**: 173-181.
- Sardaro ML, Marmiroli M, Maestri E and Marmiroli N (2012) Genetic characterization of Italian tomato varieties and their traceability in tomato food products. *Food Sci Nutr* **1**: 54-62.
- Shevchenko A, Tomas H, Havlis J, Olsen JV and Mann M (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* **1**: 2856-2860.
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**: 539.
- Torres-Schumann S, Godoy JA and Pintor-Toro JA (1992) A probable lipid transfer protein gene is induced by NaCl in stems of tomato plants. *Plant Mol Biol* **18**: 749-757.

- Valenta R, Lidholm J, Niederberger V, Hayek B, Kraft D and Gronlund H (1999) The recombinant allergen-based concept of component-resolved diagnostics and immunotherapy (CRD and CRIT). *Clin Exp Allergy* **29**: 896-904.
- Van Winkle RC and Chang C (2014) The biochemical basis and clinical evidence of food allergy due to lipid transfer proteins: a comprehensive review. *Clin Rev Allergy Immunol* **46**: 211-224.
- Wang NJ, Lee CC, Cheng CS, Lo WC, Yang YF, Chen MN and Lyu PC (2012) Construction and analysis of a plant non-specific lipid transfer protein database (nsLTPDB). *BMC Genomics* **13 Suppl 1**: S9.
- Westphal S, Kempf W, Foetisch K, Retzek M, Vieths S and Scheurer S (2004) Tomato profilin Lyc e 1: IgE cross-reactivity and allergenic potency. *Allergy* **59**: 526-532.

Highlights

- A novel plant lipid protein gene has been identified in tomato, San Marzano variety;
- The recombinant protein (rSM-22) has been produced in *E. coli* cells;
- rSM-22 is active in lipid binding up to 105 °C;
- The protein shows potential epitome regions on its surface;
- Recombinant allergens can contribute to allergy diagnosis and treatment.

CONTRIBUTIONS

MV, CL and LRC: design and interpretation of all experiments.

MV, CL, IF and AP: carried out molecular biology experiments and characterization of rSM-22.

SR: designed and carried out mass spectrometry experiments.

MV, CL and LRC: wrote the manuscript.

Figure S1

```

SM-21      GTGGTTGCACCCCATGCAGAGGCAGCGATTTGTGGTCAAGTCACATCTGGTGTGGCTCCT 60
SM-22      GTGGTTGCACCCCATGCAGAGTCACCGATTTGCGGCGAGGTTACGTCTGGCTTGGCTCCT 60
SM-16      GTGGTTGCACCCCATGCAGAGGCTTTAACTTGTGGTCAAGTTACATCTACCTTGGCTCCT 60
SM-17      GTGGTTGCACCCCATGCAGAGGCATCAATTTGTGGTCAAGTTACTGCTGGCTTGGCTCCT 60
            ***** * * * * * * * * * *
SM-21      TGTCTGCCTTATCTTACCGGCCGCGGCCCTCGGAGGGTGCTGTGGCGGAATTAAGGGT 120
SM-22      TGCCCTCCCTTATCTTCAAGGCCGTGGCCCTCTAGGAGGCTGTTGTGGTGGTGTAAAGAAT 120
SM-16      TGTCTCCCTTATCTAATGAACCGCGGTCTCTCAGAACTGTTGTGATGGTGTAAAGGGT 120
SM-17      TGTCTCCCTTATCTAATGAACCGCGGTCTCTCAGAACTGTTGTGATGGTGTAAAGGGT 120
            ** * * * * * * * * * * * * * * * *
SM-21      CTTTTGGGTGCAGCCAGGACTCCAGCTGATCGAAAAATGGCATGCAATTGCCTAAAATCA 180
SM-22      CTGTTGGGTTTCAGCCAAGACTACCGGGATCGAAAGACAGCATGCACCTTGCCTGAAATCA 180
SM-16      CTTTTGGGTCAAGCCAAGACTACAGTAGACCGACAGGCCGCATGCACCTTGCCTAAAATCA 180
SM-17      CTTTTGGGTCAAGCCAAGACTACAGTAGACCGACAGGCCGCATGCACCTTGCCTAAAATCA 180
            ** * * * * * * * * * * * * * * * *
SM-21      GCAGCTACTGCTATTAAGGGCATTGATGCTGGTAAAGCTGCTGGTCTCCCTAGTGTGTTGT 240
SM-22      GCTGCTAATGCTATAAAGGGAATTGATTGAACAAAGCTGCTGGTATTCTAGTGTGTTGT 240
SM-16      GCTGCTTCTTCTTTTACAGGACTCAATTTGGGCAAAGCTGCTGCTCTCCCTAACACTTGT 240
SM-17      GCTGCTTCTTCTTTTACAGGACTCAATTTGGGCAAAGCTGCTGCTCTCCCTAACACTTGT 240
            ** * * * * * * * * * * * * * * * *
SM-21      GGCGTCAACATTCCCTTACAAGATCAGCCC 269
SM-22      AAAGTCAACATCCCTTACAAGATCAGCCC 269
SM-16      AGTGTCACATTCCCTTACAAGATCAGCCC 269
SM-17      AGTGTCACATCCCTTACAAGATCAGCCC 269
            *****

```

Figure S1. Clustal multialignment of genomic sequences obtained by PCR amplification using LTP primers TomGe-LTP_F and TomGe-LTP_R. Sequences highlighted in green correspond to the forward primers used for the PCR and hemi-nested PCR steps in 3'-RACE analysis.

SM21-3	AAAGCTGCTGGTCTCCCTAGTGTGTTGTAGTGCAACATCCCTTACAAGATCAGCCCCCTCT	60
SM21-6	AAAGCTGCTGGTCTCCCTAGTGTGTTGTAGTGCAACATCCCTTACAAGATCAGCCCCCTCT	60
SM21-4	AAAGCTGCTGGTCTCCCTAGTGTGTTGTAGTGCAACATCCCTTACAAGATCAGCCCCCTCT *****	60
SM21-3	ACTGACTGCTCTAAAGTTCAGTAAAGCTGATCATCAGAATTTGGTTTCATGAGGAGAATT	120
SM21-6	ACTGACTGCTCTAAAGTTCAGTAAAGCTGATCATCAGAATTTGGTTTCATGAGGAGAATT	120
SM21-4	ACTGACTGCTCTAAAGTTCAGTAAAGCTGATCATCAGAATTTGGTTTCATGAGGAGAATT *****	120
SM21-3	AAGAATAAGATAGATAGCATTGATCTTGCTTATGGATCCTTTCTTTCTATGTTGTATCAG	180
SM21-6	AAGAATAAGATAGATAGCATTGATCTTGCTTATGGATCCTTTCTTTCTATGTTGTATCAG	180
SM21-4	AAGAATAAGATAGATAGCATTGATCTTGCTTATGGATCCTTTCTTTCTATGTTGTATCAG *****	180
SM21-3	TTGTCACTTTCTCTTTTTTCTGTGTTTCGTTTAAAATCGTATGTACTCGAGTCTTGTTTC	240
SM21-6	TTGTCACTTTCTCTTTTTTCTGTGTTTCGTTTAAAATCGTATGTACTCGAGTCTTGTTTC	240
SM21-4	TTGTCACTTTCTCTTTTTTCTGTGTTTCGTTTAAAATCGTATGTACTCGAGTCTTGTTTC *****	240
SM21-3	GAAATTTGACGATTGATTATATTATATCAGTTGTTACTTTCTGTTTTCTGTGTTTCTTT	300
SM21-6	GAAATTTGACGATTGATTATATTATATCAGTTGTTACTTTCTGTTTTCTGTGTTTCTTT	300
SM21-4	GAAATTTGACGATTGATTATATTATATCAGTTGTTACTTTCTGTTTTCTGTGTTTCTTT *****	287
SM21-3	TAAATCGTGTGTAGTCGAGTCTTGTATCGAAATTTACGATTGGCTATGTTGTATTAAT	360
SM21-6	TAAATCGTGTGTAGTCGAGTCTTGTATCGAAATTTACGATTGGCTATGTTGTATTAAT	360
SM21-4	----- -----	
SM21-3	CTAATCTTTGATAATACACATCTATCTTATTTGGTATAAAAAAAAAAAAAAAAAA	413
SM21-6	CTAATCTTTGATAATACACATCTATCTTATTTGGTATAAAAAAAAAAAAAAAAAA	413
SM21-4	----- ----- *****	304

SM_22_15 GCTGCTAATGCTATAAAGGGAATTGATTTGAACAAAGCGCTGGTATTCCTAGTGTTTGT
SM_22_3 GCTGCTAATGCTATAAAGGGAATTGATTTGAACAAAGCGCTGGGCTCCCTGGAGTTTGT
SM_22_4 GCTGCTAATGCTATAAAGGGAATTGATTTGAACAAAGCGCTGGTATACCTAGTGCTTGT
***** * * * * *

SM_22_15 AAAGTTAACATTCTTACAAGATCAGCCCTTCCACTGACTGCTCCACGGTTCAGTAAAGT
SM_22_3 GGAGTCAACATTCTTACAAGATCAGCCCTTCCACTGATTGCTCAACGGTCCAGTAAAGT
SM_22_4 GCGCTAAGCATTCCTTACAAGATCAGCCCTTTACTGACTGCTCGAAAGTTTGTAAAGT
* * * * *

SM_22_15 TGATCATCAGAATATGGAACCATTGTGAGGAGAAGTTATAAGAATAAGA-----
SM_22_3 TGATGAAAGCTAAATCTGGCTTGTGGAGATGAAGAATAAG---ATGGATTATCTATC-
SM_22_4 TGATGAAAGTCAATTTTCCAAGGCATAATTGGGGAATAAAGACGGATATATCTATCG
* * * * *

SM_22_15 -TGAATGGAGTGTGATCTGCTTATGGATCCATTAGTTATATGTTGTGTTGTGTCTTT
SM_22_3 -----GAGTTTGATCCATCATTATATATATATATATATATATATAT-----
SM_22_4 ATCAATCGAGTTTAATCCATCATTTTATAATA-----TGTTGTC-----
* * * * *

SM_22_15 TTGTTTCTTTTTTCGTTTTGTCTTTTGGGGGCTTGTATCGAAA-----CTTCAC
SM_22_3 ATGTTGTCTCTTTTCTTTTGAATTTGGAGTACTTATATATTGTCGAGTCTTGAATGAA
SM_22_4 TCTTTTATTTTGTATTTTGTGTTGTTGGAGTACTTATATGTTGTCGAGTATTGTAATGAA
* * * * *

SM_22_15 TATCGATGGATTGTATTACTTTAATGCAATCTTGATA--ATATATATCTATCTTTTTTA
SM_22_3 CATTGGTGGTTGCTTATGTTACAACCCAATCTTCATATAGTAATACATGTTAGT---TT
SM_22_4 CTTTGGTGGTTGGTTGTTATTACAACCCAATCTTCATAGTAATATATACATCATATTTT
* * * * *

SM_22_15 GTACCTTTAGAAGGAAAAA-----
SM_22_3 CTGCTTTTAAAAAAGTTCGACATCGATATACGGGGTC
SM_22_4 ACTTTAAGAAAAAAGTTCGACATCGATACGCGTGGTC

Figure S2. Clustal multialignment of cDNA sequences obtained by 3'-RACE analysis. Sequences derive from the hemi-nested PCR amplification, using the TomSM-22_FA (A) and

TomSM-21_FA (B) as forward primers. Forward primers are highlighted in green. Stop codons are highlighted in pink.

ACCEPTED MANUSCRIPT

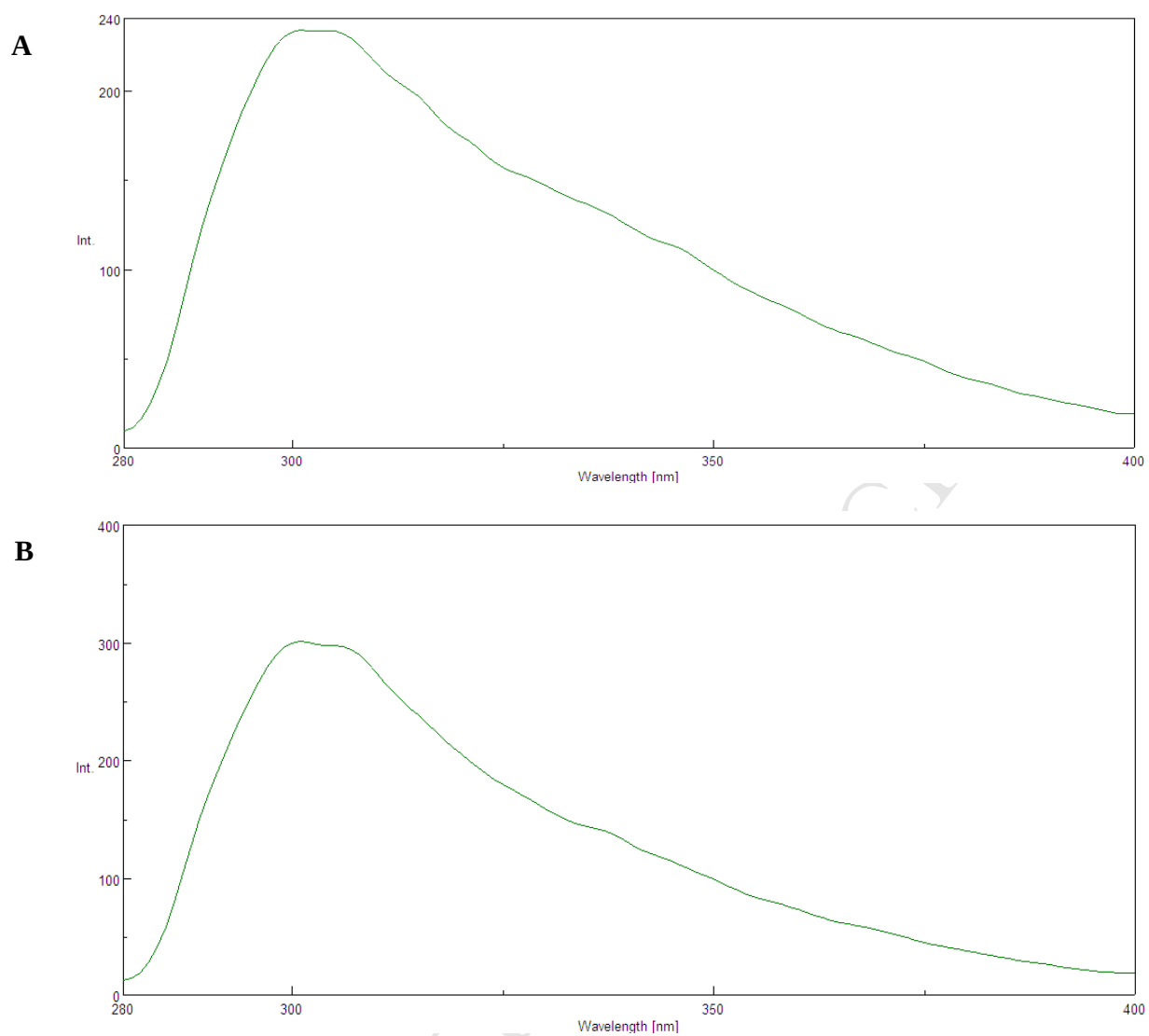


Figure S3. Fluorescence spectra of rSM-22. A) emission of a 10 μ M rSM-22 solution (excitation at 275 nm); B) as in A plus four molar excess of lyso-C₁₆.

Table SI. MS/MS protein identifications of SDS-PAGE bands (see Fig. 3)

Gel bands	Theor/Exp Mw (kDa)	Matches ^a	No. of unique peptides ^a	Seq. Cov.	Sequences	Mascot Score	NCBI Gi No.	Protein ID [Solanum lycopersicum]
Lines 5-9	12.07/9.0	76 (35)	9 (6)	58%	K.GLLGAAK.T K.GPLGGCCR.G R.GVKGLLGAAK.T K.ISPFTDCSK.V K.GLLGAAKTPADR.K K.SAANAIAKGLNLGK.A K.AAGIPSACGVSIPIYK.I K.TACTCLKSAANAIAK.G K.AAGIPSACGVSIPIYKISPFTDCSK.V	473	gi 460407212	PREDICTED: non-specific lipid-transfer protein 2-like

^a Values between brackets indicate statistically significant peptides ($p < 0.05$)

Supplementary Material

Figure S1

```

SM-21      GTGGTTGCACCCCATGCAGAGGCAGCGATTTGTGGTCAAGTCACATCTGGTGTGGCTCCT 60
SM-22      GTGGTTGCACCCCATGCAGAGTCACCGATTTGCGGCGAGGTTACGTCTGGCTTGGCTCCT 60
SM-16      GTGGTTGCACCCCATGCAGAGGCTTTAACTTGTGGTCAAGTTACATCTACCTTGGCTCCT 60
SM-17      GTGGTTGCACCCCATGCAGAGGCATCAATTTGTGGTCAAGTTACTGCTGGCTTGGCTCCT 60
            ***** * * * * * * * * * *
SM-21      TGTCTGCCTTATCTTACCGGCCGCGGCCCTCTCGGAGGGTGCTGTGGCGGAATTAAGGGT 120
SM-22      TGCTCCCTTATCTTCAAGGCCGTGGCCCTCTAGGAGGCTGTTGTGGTGGTGTAAAGAAT 120
SM-16      TGTCTCCCTTATCTAATGAACCGCGGTCTCTCAGAACTGTTGTGATGGTGTAAAGGGT 120
SM-17      TGTCTCCCTTATCTAATGAACCGCGGTCTCTCAGAACTGTTGTGATGGTGTAAAGGGT 120
            ** * * * * * * * * * * * * * * * *
SM-21      CTTTGGGTGCAGCCAGGACTCCAGCTGATCGAAAAATGSCATGCAATTGCCTAAAATCA 180
SM-22      CTGTTGGGTTTCAGCCAAEACTACCGCGGATCGAAAGACA GCATGCACCTTGCTTGAAATCA 180
SM-16      CTTTGGGTCAAGCCAAGACTACAGTAGACCGACAGGCCGCATGCACCTTGCTTGAAATCA 180
SM-17      CTTTGGGTCAAGCCAAGACTACAGTAGACCGACAGGCCGCATGCACCTTGCTTGAAATCA 180
            ** * * * * * * * * * * * * * * * *
SM-21      GCAGCTACTGCTATTAAGGGCATTGATGCTGGTAAAGCTGCTGGTCTCCCTAGTGTGTTG 240
SM-22      GCTGCTAATGCTATAAAGGGAATTGATTGAACAAAGCTGCTGGTATTCTAGTGTGTTG 240
SM-16      GCTGCTTCTTCTTTACAGGACTCAATTTGGGCAAAGCTGCTGCTCTCCCTAACACTTGT 240
SM-17      GCTGCTTCTTCTTTACAGGACTCAATTTGGGCAAAGCTGCTGCTCTCCCTAACACTTGT 240
            ** * * * * * * * * * * * * * * * *
SM-21      GGCGTCAACATTTCCTTACAAGATCAGCCC 269
SM-22      AAAGTCAACATCCCTTACAAGATCAGCCC 269
SM-16      AGTGTCACATTTCCTTACAAGATCAGCCC 269
SM-17      AGTGTCACATCCCTTACAAGATCAGCCC 269
            *****

```

Figure S1. Clustal multialignment of genomic sequences obtained by PCR amplification using LTP primers TomGe-LTP_F and TomGe-LTP_R. Sequences highlighted in green correspond to the forward primers used for the PCR and hemi-nested PCR steps in 3'-RACE analysis.

Figure S2

A

```

SM21-3      AAAGCTGCTGGTCTCCCTAGTGTTC TAGTGTCAACATCCCTTACAAGATCAGCCCTCT 60
SM21-6      AAAGCTGCTGGTCTCCCTAGTGTTC TAGTGTCAACATCCCTTACAAGATCAGCCCTCT 60
SM21-4      AAAGCTGCTGGTCTCCCTAGTGTTC TAGTGTCAACATCCCTTACAAGATCAGCCCTCT 60
            *****

SM21-3      ACTGACTGCTCTAAAGTTCAGTAAAGCTGATCATCAGAATTTGGTTTCATGAGGAGAATT 120
SM21-6      ACTGACTGCTCTAAAGTTCAGTAAAGCTGATCATCAGAATTTGGTTTCATGAGGAGAATT 120
SM21-4      ACTGACTGCTCTAAAGTTCAGTAAAGCTGATCATCAGAATTTGGTTTCATGAGGAGAATT 120
            *****

SM21-3      AAGAATAAGATAGATAGCATTGATCTTGCTTATGGATCCTTTCTTTCTATGTTGTATCAG 180
SM21-6      AAGAATAAGATAGATAGCATTGATCTTGCTTATGGATCCTTTCTTTCTATGTTGTATCAG 180
SM21-4      AAGAATAAGATAGATAGCATTGATCTTGCTTATGGATCCTTTCTTTCTATGTTGTATCAG 180
            *****

SM21-3      TTGTCACTTTCTCTTTTTCTGTGTTTCGTTTAAATCGTATGTAAGTCTGTTGTTTC 240
SM21-6      TTGTCACTTTCTCTTTTTCTGTGTTTCGTTTAAATCGTATGTAAGTCTGTTGTTTC 240
SM21-4      TTGTCACTTTCTCTTTTTCTGTGTTTCGTTTAAATCGTATGTAAGTCTGTTGTTTC 240
            *****

SM21-3      GAAATTTGACGATTGATTATATTATATCAGTTGTTACTTTCTGTTTCTGTTTCTTTT 300
SM21-6      GAAATTTGACGATTGATTATATTATATCAGTTGTTACTTTCTGTTTCTGTTTCTTTT 300
SM21-4      GAAATTTGACGATTGATTATATTATATCAGTTGTTACTTTCTGTTTCTTTT----- 287
            *****

SM21-3      TAAATCGTGTGTAGTCGAGTCTTGATCGAAATTTACGATTGGCTATGTTGTATTAAT 360
SM21-6      TAAATCGTGTGTAGTCGAGTCTTGATCGAAATTTACGATTGGCTATGTTGTATTAAT 360
SM21-4      -----

SM21-3      CTAATCTTTGATAATACACATCTATCTTATTTGGTATAAAAAAAAAAAAAAAAAA 413
SM21-6      CTAATCTTTGATAATACACATCTATCTTATTTGGTATAAAAAAAAAAAAAAAAAA 413
SM21-4      -----AAAAAAAAAAAAAAAAA- 304
            * *****

```

B

```

SM-22_15    GCTGCTAATGCTATAAAGGGAATTGATTTGAACAAAGCTGCTGGTATTCCTAGTGTTCG 60
SM_22-3     GCTGCTAATGCTATAAAGGGAATTGATTTGAACAAAGCTGCTGGTATTCCTAGTGTTCG 60
SM_22-4     GCTGCTAATGCTATAAAGGGAATTGATTTGAACAAAGCTGCTGGTATTCCTAGTGTTCG 60
            *****

SM-22_15    AAAGTTAACATTCCTTACAAGATCAGCCCTCCACTGACTGCTCCACGGTTCAGTAAAGT 60
SM_22-3     GGAGTCAACATTCCTTACAAGATCAGCCCTCCACTGATTGCTCAACGGTTCAGTAAAGT 60
SM_22-4     GCGTAAGCATTCTTACAAGATCAGCCCTTTACTGACTGCTCGAAAGTTTGTAAAGT 60
            ** * ***** * ***** * ** * ** *

SM-22_15    TGATCATCAGAATATGGAACCATTTGAGGAGAAGTTATAAGAATAAGA----- 60
SM_22-3     TGATGAAAGCTAAATCTGGCTTGTGGAGATGAAGAATAAG---ATGGATTTATCTATC- 60
SM_22-4     TGATGAAAGTTCAATTTTCCAAGGCATAATTGGGGAAAAATAAGACGGATATATCTATCG 60
            *** * * * *

SM-22_15    -TGAATGGAGTGTGATCCTGCTTATGGATCCATTAGTTATATGTTGTGTTTGTGCTTT 60
SM_22-3     -----GAGTTTGATCCATCATTATATATATATATATATATATAT----- 60
SM_22-4     ATCAATCGAGTTTAAATCCATCTTTATAATA-----TGTTGTC----- 60
            ***** * * * *

SM-22_15    TTGTTTCTTTTTCTGTTTGTCTTTTGGGGGCTTGTATCGAAA-----CTTCAC 60
SM_22-3     ATGTTGCTCTTTTCTTTTGTAAATTTGGAGTACTTATATATTGTCGAGTCTTGTAAATGAA 60
SM_22-4     TCTTTTATTTTGTATTTTGTGTTTGGAGTACTTATATGTTGTCGAGTATTGTAATGAA 60
            ** * * * * ***** * * * *

SM-22_15    TATCGATGGATTGTATTACTTTAATGCAATCTTGATA---ATATATATCTATCTTTTTTA 60
SM_22-3     CATTGGTGGTTGCTTATGTTACAACCAATCTTCATATAGTAATACATGTTAGT----TT 60
SM_22-4     CTTTGGTGGTTGTTGATTACAACCAATCTTCATAGTAATATATATACATCATATTTT 60
            * * * * * * * * * *

SM-22_15    GTACCTTAGAAGGAAAAA----- 60
SM_22-3     CTGCTTTAAAAAAAAAAAAAAAAAGTCGACATCGATATACGGGTC 60
SM_22-4     ACTTTAAGAAAAAAAAAAAAAAAAAGTCGACATCGATACGCGTGGTC 60
            ** *

```

Figure S2. Clustal multialignment of cDNA sequences obtained by 3'-RACE analysis. Sequences derive from the hemi-nested PCR amplification, using the TomSM-22_FA (A) and

TomSM-21_FA (B) as forward primers. Forward primers are highlighted in green. Stop codons are highlighted in pink.

Figure S3

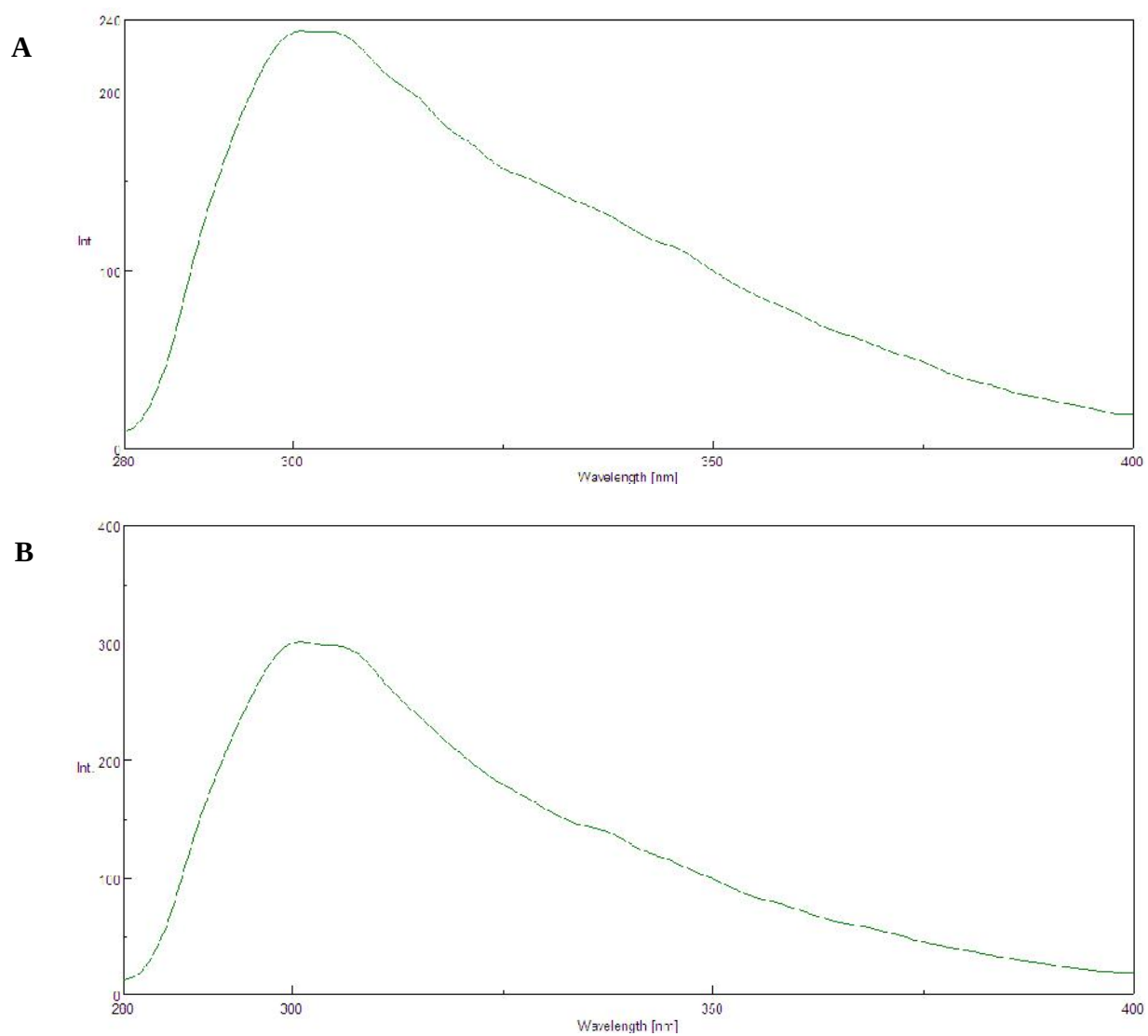


Figure S3. Fluorescence spectra of rSM-22. A) emission of a 10 μ M rSM-22 solution (excitation at 275 nm); B) as in A plus four molar excess of lyso-C₁₆.

Table SI. MS/MS protein identifications of SDS-PAGE bands (see Fig. 3)

Gel bands	Theor/Exp Mw (kDa)	Matches ^a	No. of unique peptides ^a	Seq. Cov.	Sequences	Mascot Score	NCBI Gi No.	Protein ID [Solanum lycopersicum]
Lines 5-9	12.07/9.0	76 (35)	9 (6)	58%	K.GLLGAAK.T K.GPLGGCCR.G R.GVKGLLGAAK.T K.ISPFTDCSK.V K.GLLGAAKTPADR.K K.SAANAIGLNLGK.A K.AAGIPSACGVSIPIYK.I K.TACTCLKSAANAIG.G K.AAGIPSACGVSIPIYKISPFTDCSK.V	473	gi 460407212	PREDICTED: non-specific lipid-transfer protein 2-like

^a Values between brackets indicate statistically significant peptides ($p < 0.05$)