



Herbivory-inducible lipid-transfer proteins (LTPs) of *Cicer arietinum* as potential human allergens

Harshita Negi, Harshita Saxena, Indrakant K. Singh & Archana Singh

To cite this article: Harshita Negi, Harshita Saxena, Indrakant K. Singh & Archana Singh (2023): Herbivory-inducible lipid-transfer proteins (LTPs) of *Cicer arietinum* as potential human allergens, Journal of Biomolecular Structure and Dynamics, DOI: [10.1080/07391102.2023.2169353](https://doi.org/10.1080/07391102.2023.2169353)

To link to this article: <https://doi.org/10.1080/07391102.2023.2169353>



View supplementary material [↗](#)



Published online: 26 Jan 2023.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)



Herbivory-inducible lipid-transfer proteins (LTPs) of *Cicer arietinum* as potential human allergens

Harshita Negi^{a,b} , Harshita Saxena^{a,c} , Indrakant K. Singh^d  and Archana Singh^{a,e,f} 

^aDepartment of Botany, Hans Raj College, University of Delhi, Delhi, India; ^bDepartment of Biological Sciences, University of South Carolina, Columbia, SC, USA; ^cInstitute of Plant Breeding, Genetics and Genomics, University of Georgia, Griffin, GA, USA; ^dDepartment of Zoology, Deshbandhu College, University of Delhi, New Delhi, India; ^eJ C Bose Centre for Plant Genomics, Hansraj College, University of Delhi, Delhi, India; ^fDelhi School of Climate Change and Sustainability, Institution of Eminence, Maharishi Karnad Bhawan, University of Delhi, Delhi, India

Communicated by Ramaswamy H. Sarma

ABSTRACT

Lipid-transfer proteins (LTPs) are lipid-binding small proteins, ubiquitously distributed amongst plant kingdom. Apart from their involvement in plant defense, it has also been discovered that they induce allergic reactions in humans. A plethora of LTPs have been identified in vegetables, fruits, pollens, nuts, and latex, among which Pru p 3, a LTP allergen from peach fruit, is extensively studied and exhibits cross-reactivity with potential allergens from different species. In *Cicer arietinum*, a family of LTPs (CaLTPs) has been identified and their importance in plant defense during *Helicoverpa armigera*-infestation has been recognized. However, the determination of the allergenicity potential of CaLTPs has not been attempted. In this study, we aim to decipher the allergenicity potential of defense-related CaLTPs. The allergenicity potential prediction, and identification of B-cell epitope binding regions showed that the CaLTPs had conserved domains and B-cell epitopes in the same regions as Prup3 (a marker allergen for LTPs). Using molecular docking and simulations, we observed that the CaLTPs successfully interacted with the Immunoglobulin E (IgE) with docking energies ranging from -315.5 to -268.4 and the structures were stabilized within 10 ns of simulation. Through this study, we intend to embellish our present knowledge and understanding of the sensitization and allergenicity potential of CaLTPs.

ARTICLE HISTORY

Received 15 November 2022
Accepted 10 January 2023

KEYWORDS

Lipid-transfer proteins (LTPs); allergenicity potential; chickpea; *Helicoverpa armigera*; human allergens

Introduction

In plants, intracellular lipid transport is facilitated by lipid-transfer proteins (LTPs), which are a family of lipid-binding, low-molecular-weight proteins composed of basic polypeptides (about 91–95 amino acids long) with conserved three-dimensional structures and are abundantly present in the plant kingdom (Garino et al., 2016; Salcedo et al., 2004). LTPs play a significant role during plant-pathogen interactions and possess antibacterial and antifungal properties (Hoffmann-Sommergruber, 2000). They have been also classified under the family of pathogenesis-related (PR) proteins, as PR-14, due to their involvement in plant defense and differentially regulate during mechanical injury and herbivory (Kerchev et al., 2012; I. K. Singh et al., 2017). Concurrently, the expression of LTPs is also regulated by methyl jasmonate (MeJA), ethylene, and salicylic acid (SA), indicating their role in phytohormone-mediated signaling pathways operational during pest and pathogen attack (A. Singh et al., 2008; I. K. Singh & Singh, 2021; Yadav et al., 2022). Interestingly, LTPs are recognized as ubiquitous allergens and/or pan-allergens

in fruits, vegetables, nuts, seeds, latex, and pollens and pose several allergic reactions in humans. The implications due to LTP allergies are concentrated predominantly in the Mediterranean region (Egger et al., 2010; Giangrieco et al., 2015; Morales et al., 2014; Pastorello et al., 2001).

The first incidence of a LTP as an allergen was identified in peach, named Pru p 3 as per the IUIS Allergen Nomenclature Committee, by Lleonart et al. (1992). Pru p 3 is 10 kDa, low-molecular-weight non-specific LTP (nsLTP), localized in the outer epidermal layer or peel of the peach fruit (Hoffmann-Sommergruber, 2000; Lleonart et al., 1992; van Ree, 2002). After this discovery, several other LTPs were identified in the Rosaceae family with about 30% to 95% sequence similarity among them, linked with a phenomenon called birch pollinosis, which suggests cross-reactivity reactions of different LTPs specific to Immunoglobulin E (IgE) molecules (Salcedo et al., 2004, 2007; van Ree, 2002). The accumulation of allergens favorably takes place in the exposed plant surfaces like the outer epidermis; therefore, in Rosaceae fruits, the LTPs are concentrated in the fruit peels

CONTACTS Archana Singh  archanasingh@hrc.du.ac.in  Department of Botany, Hans Raj College, University of Delhi, Delhi, 110007, India; J C Bose Centre for Plant Genomics, Hansraj College, University of Delhi, Delhi, India; Delhi School of Climate Change and Sustainability, Institution of Eminence, Maharishi Karnad Bhawan, University of Delhi, Delhi, 110007, India; Indrakant K. Singh  iksingh@db.du.ac.in  Department of Zoology, Deshbandhu College, University of Delhi, New Delhi, India

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/07391102.2023.2169353>.

© 2023 Informa UK Limited, trading as Taylor & Francis Group

than in the pulps (Salcedo et al., 2004). LTPs are associated with a class of structurally related proteins or α -class proteins, having eight cysteine residues at conserved locations, and form a dense barrel-like structure with the help of four disulfide bridges. The conserved and compact structures of LTPs make them stable proteins and protect them from heat denaturation, pH changes, proteolysis, and other types of processing (Hoffmann-Sommergruber, 2000; van Ree, 2002). Along with a stable and common crystallographic structure, their basic isoelectric point, and high sequence similarity with allergenic proteins suggest presence of highly conserved regions, enabling them to cause severe allergic reactions and inducing IgE-mediated cross-reactivity (Pastorello et al., 2001).

Cross-reactivity studies, in which IgE antibody specific to one LTP binds with a different LTP (a potential allergen) of different origin, are conducted among LTPs and nsLTPs from different fruits, vegetables, and pollens to identify the homologous LTPs in the evolutionary unrelated species and determine their potential as allergens. The peach LTP, Pru p 3, was discovered to cross-react with proteins from other fruits and vegetables, such as corn, onion, peanut, hazelnut, and apple (Egger et al., 2010; Hauser et al., 2010; Morales et al., 2014; Wawrzęńczyk et al., 2020). LTPs exhibit sensitization, i.e. stimulating specific IgE molecules, and thus, occur as true food allergens (van Ree, 2002). The sensitization of LTPs is clinically essential due to their stability and conserved structures which may stipulate cross-reactivity among evolutionary unrelated plant species (Pastorello et al., 2001). The allergenicity of a protein is measured by its capability to induce immune responses with the synthesis of IgE and T helper 2 (Th2) cytokines specific to the allergens (Garino et al., 2016). Identification and modification (chemical or genetic) of IgE-binding epitopes can result in the reduction of the allergenicity of a protein. Several computational tools aid in the identification of epitopes of a potential allergen, which is based on different machine learning algorithms and sequence similarity searches (Arora et al., 2020; Sharma et al., 2013). The identification of IgE-binding epitopes is primarily mediated by the prediction of the B-cell epitope-binding residues that identifies the allergens. An understanding of the antigen-antibody interactions that take place between the B-cell epitopes and the allergens is essential for modulating the allergens for reduced effects. At the atomic level, molecular docking is one such computational technique that helps circumspect the interactions in the most stable form (Arora et al., 2020).

With this knowledge, this paper encompasses the prediction of the allergenicity potential of defense-related LTPs identified in *Cicer arietinum* (CaLTPs, from hereon) during its interaction with *H. armigera*. *In silico* allergenicity potential prediction tools, sequence similarity searches, identification of B-cell epitope binding regions, and their potential interaction with IgE were utilized to determine the allergenicity of these CaLTPs. Further, molecular docking and simulations were executed to investigate their stability as allergens. Prior knowledge of the binding sites in the B-cell epitopes, and confirmation of their interaction with IgE provided strong

evidence of the protein-protein interaction, which can serve as a basis for detecting allergenicity of a person and for discovering effective drugs in clinical trials to combat the allergic effects.

Materials and methods

Allergenicity potential prediction

Previously, 48 candidate LTP genes were identified using genome wide analysis of chickpea genome by *in silico* methods in chickpea and 14 of them showed their potential function in plant defense during *H. armigera*-infestation (unpublished data). Assessment of allergenicity of *H. armigera*-inducible 14 CaLTPs was conducted using AlgPred (<https://webs.iitd.edu.in/raghava/algpred/index.html>) with three different methods, support vector machine (SVM) method involving amino acid composition, SVM method depending on dipeptide composition, and BLAST (Saha & Raghava, 2006b). Allergenicity potential was also predicted using other web tools AllergenFP v.1.0 (<https://ddg-pharmfac.net/AllergenFP/>) (based on Tanimoto coefficient algorithms), and AllerTOP v. 2.0 (<https://www.ddg-pharmfac.net/AllerTOP/method.html>) (on the basis of auto cross-covariance (ACC) transformation, and classified by k-nearest neighbor algorithm) (Dimitrov et al., 2014a, 2014b). The CaLTPs that were predicted as potential allergens by at least two prediction tools were considered allergens for this study.

B-cell epitope prediction

Seven defense-related CaLTPs showing *H. armigera*-infestation responsiveness and identified as potential allergens were considered for B-cell epitope prediction. A sequence-based B-Cell Epitope Predictor, BepiPred-2.0 (<http://www.cbs.dtu.dk/services/BepiPred/cite.php>), was used to predict the B-cell epitopes in the selected CaLTP sequences (Jespersen et al., 2017). The sequences of currently known LTP allergens from Fabales (Ara h 9 (ABX56711), Ara h 16 (ASU04353), & Ara h 17 (ASU04352) of *Arachis hypogaea*, Len c 3 (AAX35807) of *Lens culinaris*, Lup an 3 (XP_019446786.1) of *Lupinus angustifolius*, and Pha v 3 (ADC80502) of *Phaseolus vulgaris*) and LTP marker allergen, Pru p 3 (P81402), from *Prunus persica* were retrieved from Allergen database (<http://www.allergen.org/>) and were used as reference LTPs to perform B-cell epitope prediction. BepiPred uses the propensity scale method in combination with the hidden Markov model to predict potential B-cell epitopes. To reduce the false positives, an epitope threshold of 0.5 i.e. any residues which scored above the threshold were considered potential epitopes.

The B-cell epitope prediction was also conducted using ABCpred (https://webs.iitd.edu.in/raghava/abcpred/ABC_submission.html), and BcePred (https://webs.iitd.edu.in/raghava/bcepred/bcepred_submission.html) (Saha & Raghava, 2004, 2006a, 2007). ABCpred uses Jordan network (a partial recurrent neural network) to obtain non-redundant B-cell epitopes and non-epitopes from BcePred and Swiss-Prot databases,

respectively. Here, a fixed overlapping filter with a window length of 16 residues and 0.5 thresholds was used for choosing probable B-cell epitopes. In the BcePred web tool, to increase the accuracy and specificity of the search, a combined threshold of 2 and physicochemical parameter antigenic propensity (Ap) was used. To procure a consensus sequence, epitopes from all three tools were compared and the residues common in all the three predicted sequences were chosen and their positions were also determined.

Sequence alignment of IgE binding domains

Multiple sequence alignment (MSA) was performed to access the conservation status of the B-cell epitopes in CaLTPs as well as other reference LTPs (Ara h 9, Ara h 16, Ara h 17, Len c 3, Lup an 3, Pha v 3, and Pru p 3), using the CLUSTAL W alignment algorithm in the software, Molecular Evolutionary Genetics Analysis (MEGA) 11 (Tamura et al., 2021). The web-based application for generating sequencing logos, WebLogo version 3.7.12 (<https://weblogo.berkeley.edu/logo.cgi>), was used to visualize the alignment in logo form (Crooks et al., 2004; Schneider & Stephens, 1990).

Protein structure prediction and validation

The three-dimensional protein models of the candidate CaLTPs, as well as the seven reference LTPs, were predicted using two web-based servers, Iterative Threading ASSEmbly Refinement (I-TASSER, <https://zhanggroup.org/I-TASSER/>) (Yang et al., 2015; Yang & Zhang, 2015a; Zheng et al., 2021) and Protein Homology/analogY Recognition Engine V 2.0 (Phyre², <http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) (Kelley et al., 2015). The predicted structures were visualized and graphically represented using the PyMOL 2.5 software (downloaded from <https://pymol.org/2/>) (Schrödinger & DeLano, 2020). Structure validation was conducted to analyze the structural attributes and stability of the predicted proteins. The Structure Analysis and Verification Server (SAVES) v6.0 (<https://servicesn.mbi.ucla.edu/SAVES/>), an online platform, was used for assessing the stereochemical properties using the PROCHECK tool, and diatomic resolution coordinates were applied for three-dimensional structural analysis with VERIFY-3D tool, and finally, the ERRAT tool was used to determine the overall quality factor of non-bonded atoms in the predicted model (Bowie et al., 1991; Colovos & Yeates, 1993; Laskowski et al., 1993, 1996; Lüthy et al., 1992). In addition to this, the quality scores represented as 'Z-scores' were determined for the 3-D models on the Qualitative Model Energy ANALysis (QMEAN) server (<https://swissmodel.expasy.org/qmean/>) (Benkert et al., 2011).

Molecular docking and statistical analysis of allergen-antibody interaction

The structure of the IgE antibody was retrieved from the PDB database (www.rcsb.org/) (PDB Id.-2WQR). The online web tool, ClusPro2.0 (<https://cluspro.bu.edu/home.php>), was utilized to perform docking in antibody mode (Brenke et al., 2012; Desta et al., 2020; Kozakov et al., 2013, 2017; Vajda

et al., 2017). The IgE antibody was used as the receptor and the selected CaLTPs along with the 7 reference LTPs were used as ligands. All the generated models were evaluated as per their bound and unbound state, Root Mean Square Deviation (RMSD) values using structural alignment in PyMOL, as well as the weighted energy scores provided by ClusPro (Schrödinger & DeLano, 2020). The model with the highest negative interaction energy and least variance in RMSD value was selected as the most stable docking configuration. To summarize the dispersion and skewness of the dataset, residue lengths and the docking energies were represented as box plots and density maps, generated with the RStudio using the ggplot2 package (RStudio Team, 2020). Furthermore, to determine the correlation between the predicted B-cell epitope residues and the docking energy, a scatter plot was also generated using the ggplot2 and ggpubr libraries available in RStudio (Wickham, 2016).

MD simulation

Molecular dynamics (MD) simulation was conducted in real-time to get mechanistic insights into the docked CaLTP structures using the software, GROningen MAchine for Chemical Simulations (GROMACS) version 2021.3, under the force field OPLS-AA/L all-atom, which takes into account all the atoms of a protein (Bauer et al., 2022; Jorgensen et al., 1996; Kaminski et al., 2001; Arora et al., 2020). The cell volume and overall system size were defined by a cubic water box with a distance of 1 nm to the edges of the box, followed by solvation using SPC216 water model (Mark & Nilsson, 2001). The physiological salt concentration was maintained with the addition of sodium (Na⁺) and chloride (Cl⁻) ions to impersonate the intracellular conditions. To avoid steric hindrance or improper geometry of the solvated and electroneutral system, energy minimization was conducted by compiling the topology, structure, and simulation parameters using the steepest descent method of 500000 steps with 2 fs initial-step size. Newton's equations were ascribed to the MD simulation using the leap-frog integrator. The parameters for bond lengths involved the LINCS (LINear Constraint Solver) algorithm (Hess et al., 1997), and the periodic-boundary conditions (PBC) were set to 'xyz', considering all the directions for the system. A cut-off of 1 nm was chosen for short-range interactions while for the long-range electrostatics, the Particle Mesh Ewald method was considered (Essmann et al., 1995). The temperature and pressure coupling were regulated by V-rescale and Parrinello-Rehman types, respectively, with standard values of 300 K (temperature) and 1 bar (pressure). Before commencing with real dynamics, the solvent and ions surrounding the proteins require equilibration. The equilibration step was conducted in two phases: (i) 'isothermal-isochoric' phase or NVT ensemble (constant Number of particles, Volume, and Temperature), and a 100-ps NVT equilibration was used in this analysis; (ii) 'isothermal-isobaric' or NPT ensemble (constant Number of particles, Pressure, and Temperature). Following the equilibration, the well-defined system with desired temperature and pressure was then subjected to the

Table 1. Consensus sequences of the B-cell epitopes predicted in the 7 CalTPs and 7 reference LTPs using web tools BepiPred, ABCpred, and BcePred. The consensus sequences are highlighted in red and the residue numbers of the epitopes (Epi1-7) in each protein is mentioned separately.

	Consensus sequence	Epi1	Epi2	Epi3	Epi4	Epi5	Epi6	Epi7
CalTPs								
CalTP5	MAHLVSFVTLIIMCPISSSCDHYVWDFVTCRLRYLAGYESDPTTRCCKTWKELNRDAKL HAATCQCIENANGIDINFDITRIEDLPNKCHESKSPISMS NCSKAH	18-26	47	NA	NA	NA	NA	NA
CalTP8	MAYVKVITCMAMMIMMILISSPMVNAISCEVIGEIDPCLEYLNGRYSAFSPRECEG VKNLSHIAHTTSARRATECECLSAAYSISDLNRHHDALARRCGVGLHYRVSTSIDCNRIRG	22-25	28-32	38	42-45	100-115	NA	NA
CalTP12	MGSSMLAKVTCMAMICFWMGSSIMKANGIEPCGOVLVASCVGLRSPGSPVPEP CCNGIRTVNSQSKTTADROQVNCCLKSTALSPGLNPQALALPANCNGVNLPIKISTSIDCNTIKY	29-34	47-48	51-53	105-111	NA	NA	NA
CalTP30	MGSKSVITNPISLFTLFLVLMFGLAYSVDNODKAETDKLILANCLPFTVGOAKSPNIDCC TGKDVNNRSKRLCLIKHDHDPNLGLKNITLALQLPNDCKTPTNTKCVGILHLPPKSPKAFIF	38-40	52-54	58-61	83-85	90-97	112-125	NA
CalTP36	EDIQNSMEKNSSTVPPASNTGKSTGRGQEKNGGWRKRWLMAEVLQWISASLISVSFVIF MESYRKVMIVGIMMLAIANAMLLVNGQSYCHMTKSGLNSCKPYVSGDNSVDQIPSNACCSAISK DLQCLRYMDSGLLSEFYGVDPKQAMELPVKCKLMDSFHC	38-39	42-46	74-76	79-85	91-100	NA	NA
CalTP40	MAHRKMMGLVFVIMAMLCAGAAAOSSCTNVLITLSPCLNITNGNSTPSSGCCSQLSSVWRSTPQC LCQVLGGGSSSLGININOTQALALPGACKVQTPPTSOCKNSSPANSAASPANSPTGTEAESPN VPSGGSKSTPTAGDGSNGNSIKLSIPLFLILAATYASVI	25-28	42	47	72	96-100	154-156	NA
CalTP45	MASKTCTSLALFFTLNLLFSLVSASGSCSNPNPNPKPSPTPKPNPNPKPNPSPKPNPKPTP SSGSCPSDTLKLGVCANVINGLLNIKLGQPPVAPCCSLLEGLTDLQAAACLTALKANILGINLNLPI LSLLLNVCSSKRVPRDFQCA	78	103	133-134	147-149	NA	NA	NA
Known allergens of Fabales and Pru p 3								
Pru p 3	ITCGQVSSSLAPCIPYVRGGGAVPPACCNIRNVNINLARTTPDROAACNCLKQLSASVPGVNPNA AALPGKGVSPYKISTNCAIVK	17-19	74-81	NA	NA	NA	NA	NA
Ara h 9	MASLKFAFVMLVCMAMVGMVNAISCGQVNSALAPCIPELTGKGAPPACCSGVRGLLGLARTTAD RQAAACNCLKAAAGSLRGLNOGNAALPGRCGVSPYKISTNCAIVK	39-43	98-105	NA	NA	NA	NA	NA
Ara h 16	VTCTPTELSPCLGALTGSSPPSCCKLRAQKPCLCNLIKPNALRTYVNSPGARRVASCGVPLPSC AIQCSFVTKSIAPCFGLYSGGTVSGPCCSGIQNINGTAKTTSDRQAVCNCLKSAAASLGSQINPNNA SLPGKGVSPYKISTINCSIK	13 18-19	19-21 23-24	30 76-83	32 NA	39-41 NA	NA NA	NA NA
Len c 3	MARGMKLACVVLVICMVVIAPIAEMEGASCGAVTSDLSPLCTLYLTGGPSPQCCGGVKKLAAANT TPDROAACNCLKSAAGSITKLTNNAAALPGKGVNIPYKISTTNCNTVKE	40-45	101-105	NA	NA	NA	NA	NA
Lup an 3	MAGIVKLACAVLIGMVVVSAPLTKAITCGQVNTANLAQCLNLYRSGGAVPAPCCNGIKNINLAKTTP DRRTACNCLKAAAANTPGLNPSNAGSLPGKGVNIPYKISTNCSIK	40-44	98-104	NA	NA	NA	NA	NA
Pha v 3	MASVKFACVVLICMVVVGAAHTAQMTCTGQVSNLPCVTFIQNGFVPAGCCNGVRNIMNSARS TADRRGICNCLKTAAGAVRGLNPNNAQALPGKGVNIPYKISTNCSIN	39-43	69	93-75	98-103	NA	NA	NA

MD simulation run under PBC. The MD simulation trajectories were assessed by analyzing the RMSD and radius of gyration (RG) plots generated using the 2D plotting tool, Grace 5.1.25 (Turner, 2015). The densities, potential energies, pressure, and temperature of each LTP model were also extrapolated using Grace.

Results

Identification of putative allergens from the LTP family of proteins of chickpea

A total of 48 LTPs named as CaLTPs are identified in chickpea, and out of them, 14 are differentially regulated during chickpea-*H. armigera* interaction (unpublished data). These herbivory-regulated and defense-related CaLTPs were assessed for their potential as allergens, using three prediction tools available online. According to the results obtained using AlgPred, the prediction done by the support vector machine (SVM) method, AllerTOP, and AllergenFP, yielded seven of the CaLTPs as allergens/potential allergens. These seven putative allergens were selected for further analysis, such that at least two prediction tools considered them as an allergen or a potential allergen (Supplementary Table 1).

Prediction and alignment of B-cell epitopes of allergenic CaLTPs

B-cell epitope prediction was performed for all the 7 LTPs that were predicted as probable allergens and differentially

regulated. To conduct a comparative analysis, epitopes were also predicted for Pru p 3 and 6 currently known allergen LTPs from Fabales. Three different web tools, BepiPred, ABCpred, and BcePred, were used to prevent discrepancies between these results (Supplementary Table S2). The amino acid residues identified as epitopes by all three web tools were taken as the final consensus sequence for each of the CaLTPs (Table 1). The major B-cell epitope binding regions in the consensus sequences were identified (as highlighted by red in Table 1) and labeled as epitope 1 (Epi1), epitope 2 (Epi2), and so on until epitope 7. The number of residues in epitopes ranged from 2–6.

Multiple sequence alignment (MSA) was performed using MEGA 11 software to see the conservation status of the CaLTPs' B-cell epitopes when compared to Pru p 3 (a well-studied LTP allergen from peach) (Figure 1a). Many researchers have shown that Pru p 3 has either of the three major B-cell binding sequence regions (Borges et al., 2008; García-Casado et al., 2003; Tordesillas et al., 2009) or two major regions ((Pacios et al., 2008) and Table 1). It was interesting to note that all the CaLTPs and 6 LTPs from Fabales showed a similar conservation pattern as that of Pru p 3 (as highlighted by red boxes in Figure 1a). A majority of the B-cell epitope binding regions of the CaLTPs and LTPs from Fabales were found in the same region as Pru p 3, suggesting similar roles in the hypersensitivity reaction. The web logo thus generated also indicated a strong conservation pattern for cysteine residues (Figure 1b). As scientists have previously shown that cysteine might play a role in the effective binding of the allergen, the presence of the conserved cysteine residues in CaLTPs,

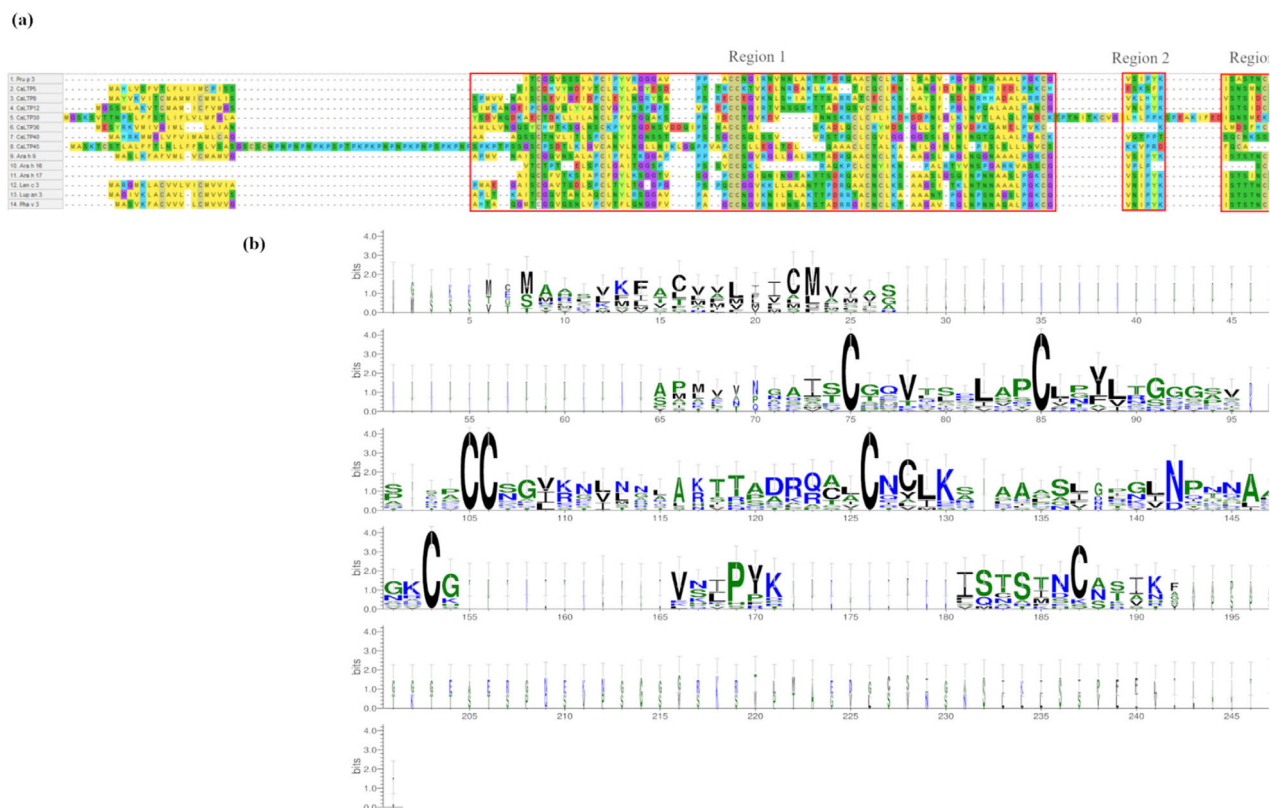


Figure 1. (a) Multiple sequence alignment of 7 CaLTPs and 7 reference LTPs performed with CLUSTAL W algorithm in the MEGA 11 software. The consensus regions are confined in boxes labeled regions 1–3. (b) Web logo of the 7 LTP sequences of chickpea with 7 reference LTPs showing the conserved Cysteine residues is all the LTPs.

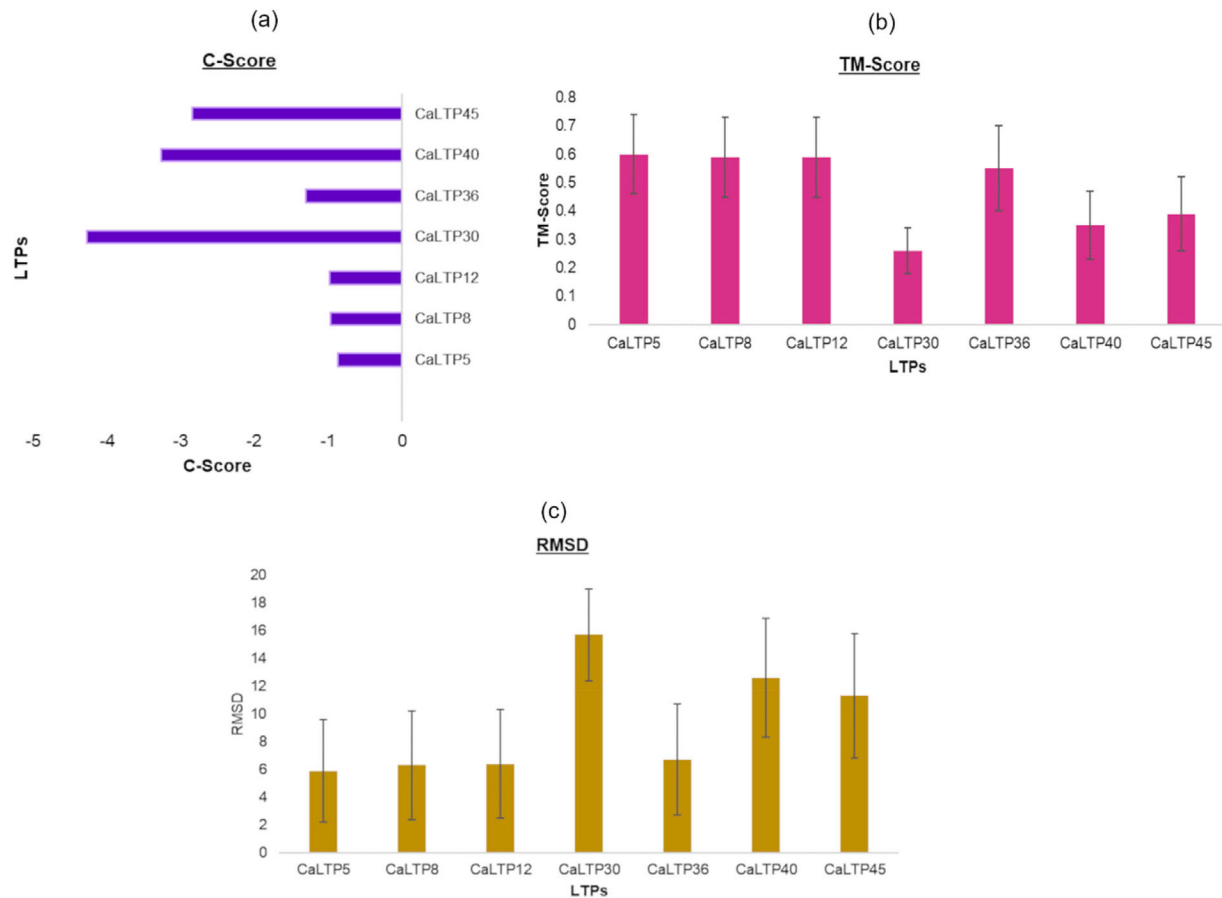


Figure 2. Graphical representation of the three-dimensional protein structure prediction scores for the 13 CaLTPs: (a) C-Score, (b) TM-Score, and (c) RMSD, using I-TASSER. The C-Score of all the LTPs is in the range $[-5, 2]$ which indicates the good quality of the models. The TM-Score of only a few LTPs is above the threshold of 0.5 suggesting confidence in the correct topology predictions of these models, however, the TM-Score of other LTPs is not below 0.17 which means the quality of the model predicted is average and not by random similarity. RMSD values of some of the LTPs are in the range of 2–5 Å suggesting medium resolution, while values greater than 5 Å are suggestive of the low resolution of the models predicted.

suggests that they are involved in allergenicity reactions (Banerjee et al., 2002; de Weerd et al., 2003).

Protein structure prediction and validation of allergenic CaLTPs

We conducted comparative modeling of the 7 CaLTPs and 7 reference LTPs using two prediction tools, amongst which the I-TASSER protein structure prediction yielded relatively better predicted models, in comparison to Phyre² (Supplementary Table S3). In I-TASSER, the secondary-structure Profile-Profile threading Alignment (PPA) method is used for predicting the three-dimensional structures of unknown proteins in comparison to representative proteins from the PDB structure library (Figure 3) (Zhang, 2008). The confidence of the models predicted was determined by three sets of scores that were standardized for the I-TASSER server: C-Score, TM-Score, and RMSD. The Confidence score (or C-Score) estimates the quality of the final structures and falls in the range of $[-5, 2]$, with the highest C-score value suggesting prediction with high confidence (Yang et al., 2015; Yang & Zhang, 2015a; Zheng et al., 2021). The C-score for the seven LTPs from chickpea fell in the range of $[-4.27, -0.87]$, with CaLTP30 having a C-score = -4.27 , a model with the least confidence, and CaLTP5 having a C-

score = -0.87 , a model with the highest confidence (Figure 2a, Supplementary Table S3). To measure the distance between two models (unknown and known) and their structural similarity to each other, TM-Score and RMSD values are benchmarked in I-TASSER. The correct topology of a model is marked by a TM-Score > 0.5 , and TM-score < 0.17 suggests random similarity (Yang et al., 2015; Yang & Zhang, 2015a; Zheng et al., 2021). The RMSD values falling in ranges 1–2 Å and 2–5 Å indicate high and medium-resolution models, respectively (Roy et al., 2010). From our prediction results, the TM-Score of the CaLTPs ranged from 0.26 ± 0.08 (least, CaLTP30) to 0.60 ± 0.14 (highest, CaLTP5), and RMSD values varied from 5.9 ± 3.7 (least, CaLTP5) to 15.7 ± 3.3 (highest, CaLTP30). The predicted structures have acquired a TM-Score > 0.17 , eradicating the chances of random similarity, however, only CaLTP 5, 8, 12, and 36 had TM-Score > 0.5 i.e. models have corrected topology (Figure 2b, Supplementary Table S3). Interestingly, all the CaLTPs have an RMSD > 5 Å, suggesting the low resolution of the predicted models (Figure 2c, Supplementary Table S3).

The validation of the predicted LTPs structures generated through both I-TASSER and Phyre² was conducted using SAVES and QMEAN web servers (Supplementary Table S3). The ERRAT scores for I-TASSER and Phyre² ranged between 51.2–89.3% and 34.2–89.3% respectively, while and the

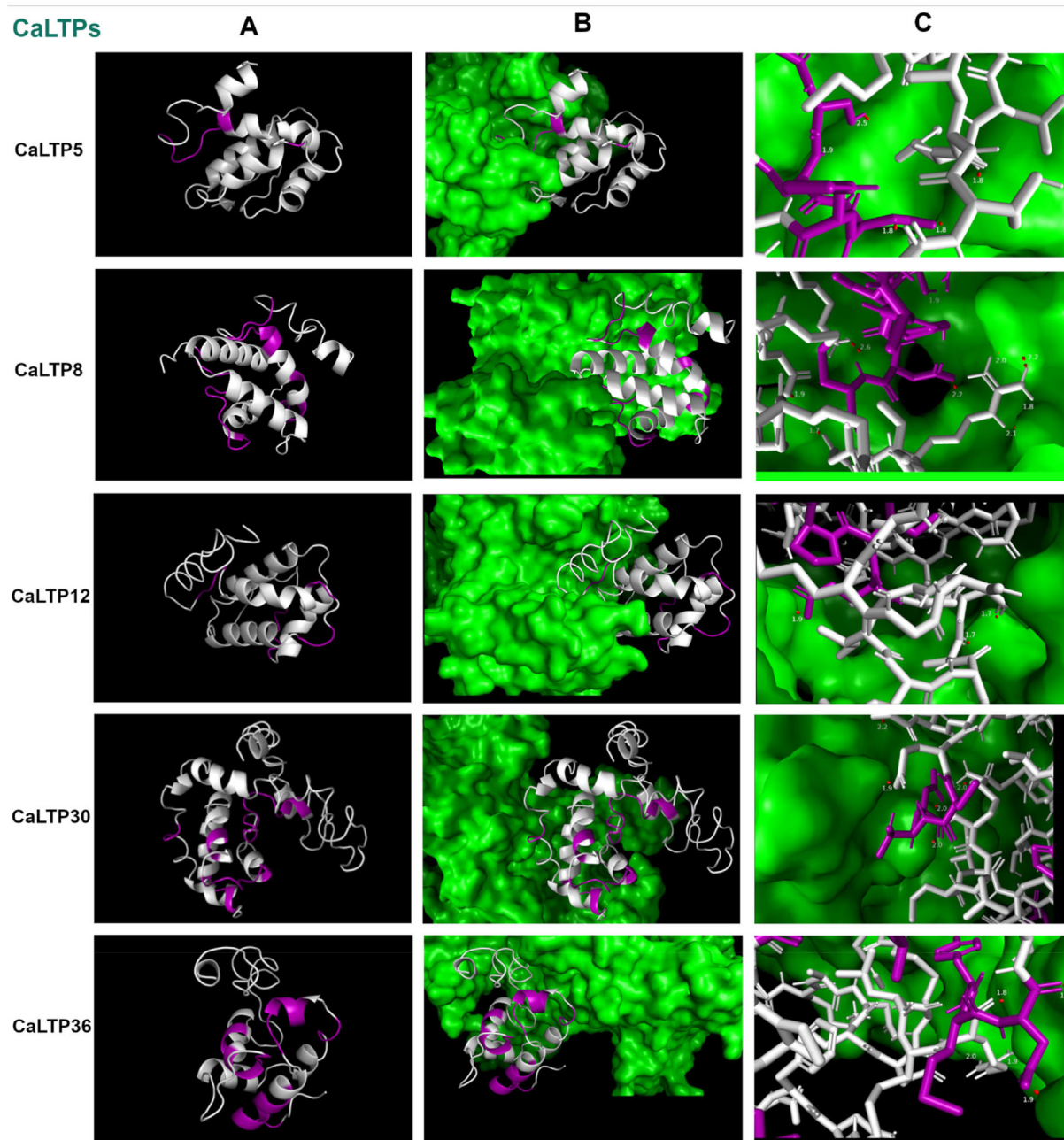


Figure 3. Protein structure prediction of the 7 candidate allergen LTPs in chickpea along with the reference LTPs (in column A); docking of IgE antibodies with the predicted structures (in column B), and interaction between the antigen (LTP)-antibody (IgE) at the atomic level. The ribbon diagram of the protein backbone is represented with white color, with B-cell epitope binding regions highlighted with purple color. The green globular structures are those of the IgE molecules. The conserved cysteine residues are marked with red color.

VERIFY-3D scores ranged between 65.61–93.02% and 36.08%–79.07% for I-TASSER and Phyre², respectively, which indicates a medium-to-high quality of the model. The Ramachandran plot statistics indicate an average quality of the models as the number of amino acids in the most favored regions ranged from 44.7–80.8%, while 10.8–40.8%, 3.9–13.8%, and 0–7.3% amino acids were present in the additionally allowed regions, generously allowed regions, and disallowed regions, respectively. The QMEAN Z-score of the structures generated via I-TASSER varied from –14.93 to –4.15 and Phyre² varied from –6.39 to –2.76. The overall quality of the structures generated by means of I-TASSER

was better than Phyre². Furthermore, in the CASP (Critical Assessment of Structure Prediction) experiments, I-TASSER has invariably ranked on top for Template based modelling. Therefore structures generated by I-TASSER were used for docking and simulation.

Molecular docking of CaLTPs with IgE

The epitope regions identified were highlighted in the generated 3-D protein structures using the color tools in PyMOL software. The ClusPro web server aided in the docking of the

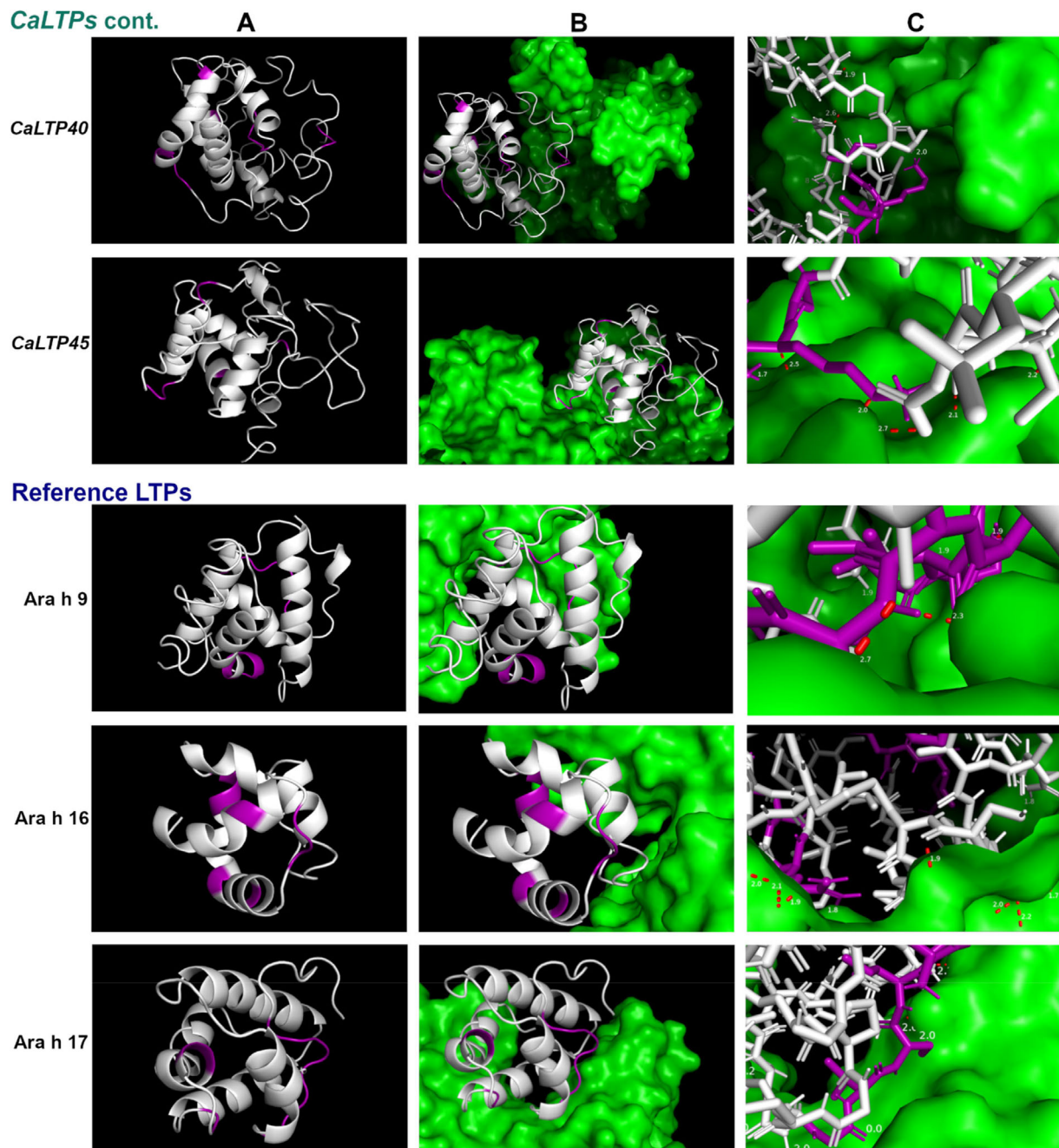


Figure 3. Continued.

potential allergenic CaLTPs with the IgE antibodies and the structures generated were visualized using PyMOL (Figure 3). The B-cell binding residues predicted were highlighted in purple and the IgE was highlighted in green color. The interaction of the B-cell epitope with the IgE model was also predicted using tools provided in PyMOL and was highlighted with red dotted lines. The threshold limit for the interaction length was 3 Å and anything above this was not considered.

The docked CaLTPs were also accessed for their docking energies. Models with the most negative docking energy were selected as the final models for each CaLTP and other reference allergens (Supplementary Table S4). It was interesting to note that CaLTP 45 had the highest negative docking energy of -315.5 and CaLTP 36 had the least

negative docking energy of -268.4 . To further visualize the distribution of the residue count (number of B-cell epitopes) and docking energies, boxplots and density graphs were produced (Figure 4). The median for the residue count and docking energies was estimated to be 12.5 and -287.4 , respectively. Both the boxplot and density graph revealed that more than half of our residue count data was skewed toward the lower number of residues with two outliers, CaLTP8 and CaLTP30 while the docking energy graph skewed towards the higher negative docking energies. Pearson's correlation test was also performed to see the correlation between the docking energies and the B-cell epitope residue count and the results were depicted in form of a scatter plot (Figure 4). The r-value for the B-cell

Reference LTPs cont.

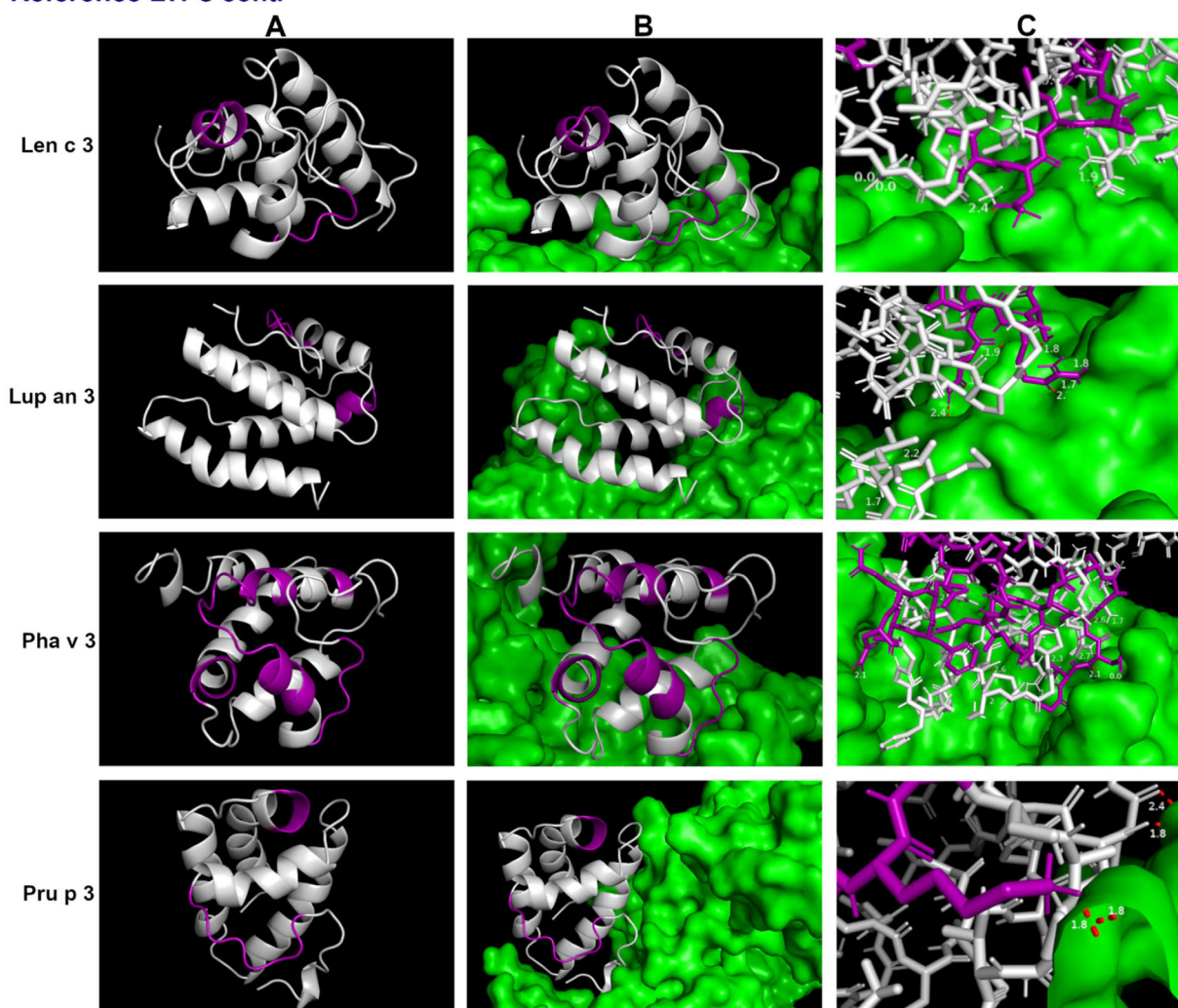


Figure 3. Continued.

epitope and the residue count was found to be 0.17 with a p-value of 0.56, this suggested that there was a very weak positive linear correlation.

Molecular dynamics simulation of CaLTPs with IgE

MD simulation for the docked complexes of the seven CaLTPs and seven reference LTPs was performed and evaluated using the trajectories. The different steps during simulation including energy minimization, the temperature during NVT equilibration, and pressure and density during NPT equilibration were assessed (Supplementary Figures S1–S4). The energy minimization step in MD simulation ensures the identification of a set of coordinates at which the docked structure acquires the minimum energy configuration (Gautam, 2020). The potential energy curves of the 7 CaLTPs (Supplementary Figure S1a) and 7 reference LTPs (Supplementary Figure S1b) show a steep descent in energy and the minimum energy is attained around -2.5×10^6 kJ/mol for all of them. The fluctuations in temperature during the system equilibration reach the target goal of 300 K over

a period of 100 ps for all the LTPs (Supplementary Figure S2) and stay stable during the remaining equilibration process. The pressure and density during the second phase of equilibration for the CaLTPs and reference LTPs were also assessed with a graph (Supplementary Figures S3 and S4), and we observed that during a period of 100 ps, the pressure fluctuations are quite extensive but eventually stabilizes near the target value of 1 bar (Supplementary Figure S3). For density, the fluctuations are not as extensive as the pressure but reach a value falling in the range of 1020–1025 kg/m³ over a course of 100 ps, near the expected value of ~ 1000 kg/m³ (Supplementary Figure S4). The final analysis of the simulated systems of the CaLTPs and reference LTPs involved extrapolation of RMSD and radius of gyration graphs. The RMSD fluctuations of the candidate complexes stabilize during 1 ns to a value falling in the range of 0.175–0.275 nm (~ 1.75 – 2.75 Å), suggesting majorly a high resolution and stability of the models (Figures 5 and 6). The radius of gyration (Rg) values help in defining the compactness of the protein complex after the MD simulation and are inversely proportional to each other. The decrease in Rg values indicates an increase

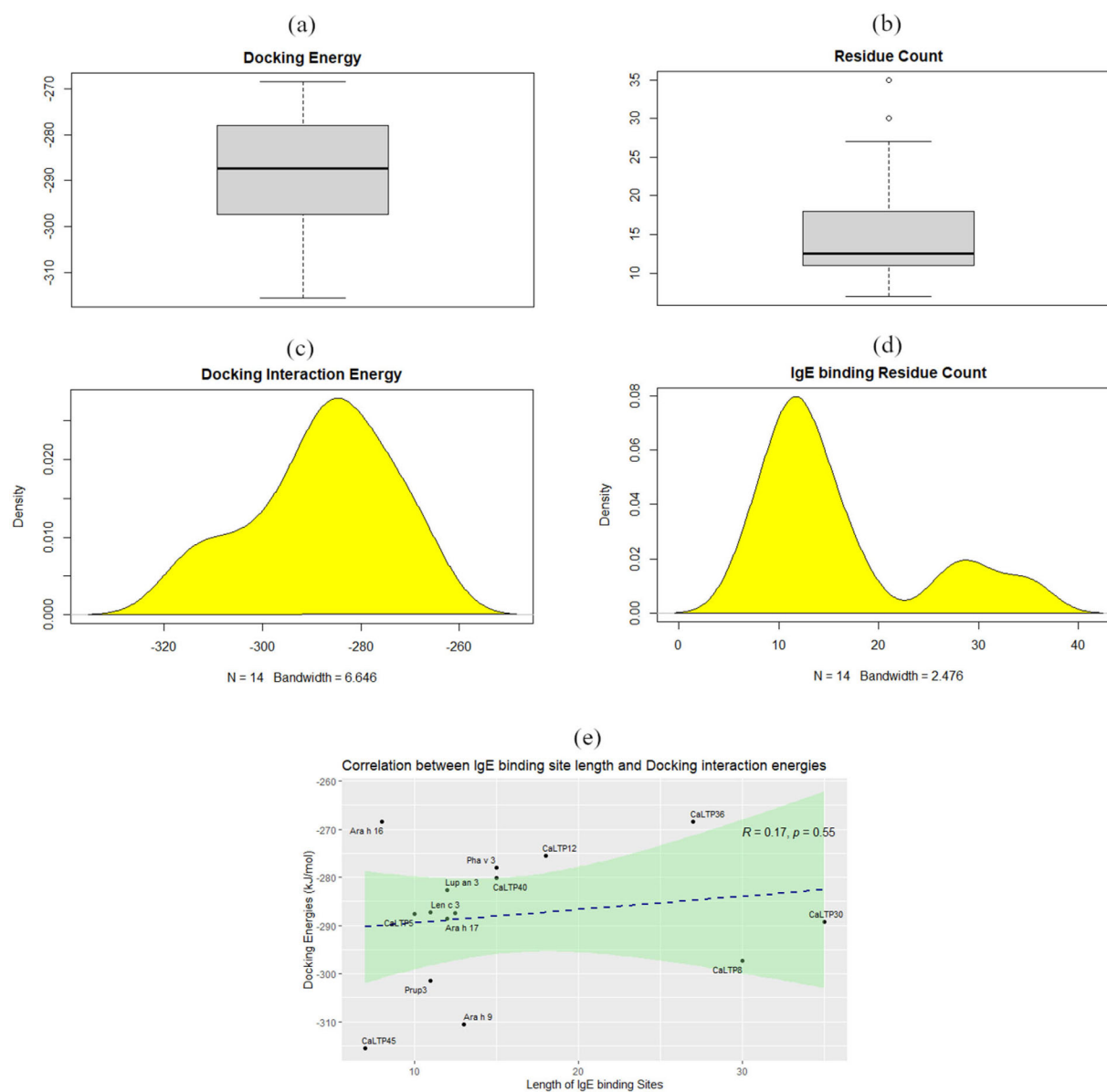


Figure 4. Statistical analysis of the docking results. Box plot and density graphs of docking energies (a, c) and IgE binding residues counts (b, d) in the docked structures of the candidate CaLTps and reference LTPs has been represented. Also, a scatter plot (e) demonstrating the correlation between the IgE binding site length and docking interaction energies is also depicted.

in the compactness of the LTP complex. From our analysis, we see variant fluctuations in the R_g values for most of the LTPs, and they are mostly increasing, indicating decreased compactness of the LTP complexes. However, over a period of 1000ps, the R_g values stabilize for most of the LTPs thereby achieving a compact structure post-simulation. The compactness of the LTP complexes suggests a strong binding interaction between the LTP structures and the IgE molecules. These results strengthen the contribution of LTP-IgE complex binding energies with the help of a molecular dynamics simulation run and depict a relatively stable binding of the IgE epitope with the potential LTP allergens and reference LTP allergens. The similarity in the results during simulation between the CaLTps and reference LTPs provides strong evidence for CaLTps bearing allergen properties, as those of the reference LTPs.

Discussion

The IgE-mediated food allergies are based on the route of sensitization via the gastrointestinal tract (Class-I or food allergies) or through the respiratory tract as inhalants (Class-II or airborne allergies), which results in systematic reactions and oral allergic symptoms, respectively. Complete and incomplete food allergens are known to induce Class-I and Class-II allergies, respectively. One such example of food allergens that is likely to induce systemic reactions is the family of LTPs (Arora et al., 2020; Finkina et al., 2017). LTPs are synthesized in plants when exposed to biotic stresses and are characterized as pathogenesis-related proteins, PR-14. Several LTPs have been identified in different fruits, vegetables, nuts, latex, and pollens, which have been characterized as food allergens (Egger et al., 2010; Wawrzęńczyk et al., 2020). We have performed an *in-silico* analysis on a family of

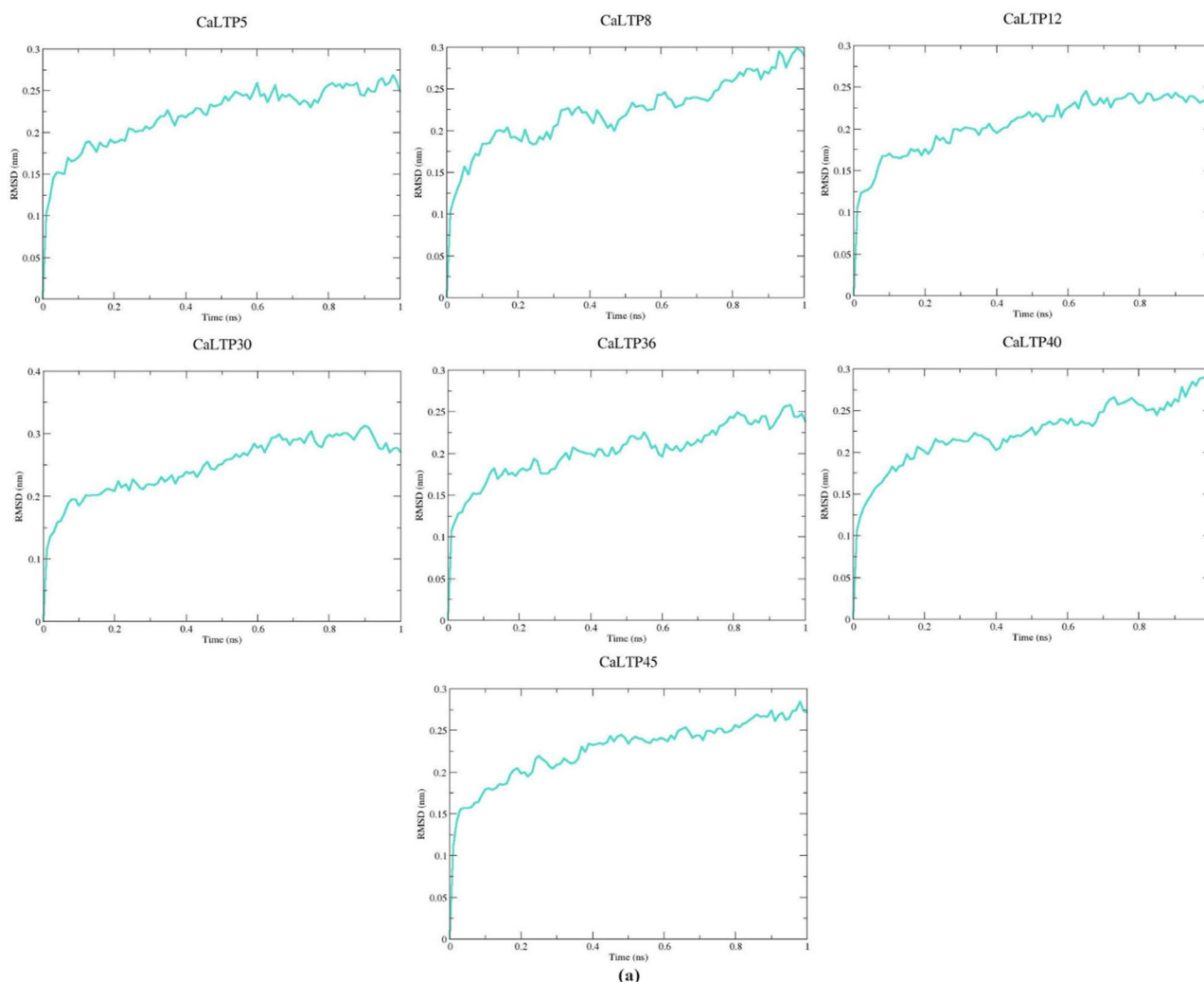


Figure 5. RMSD graphs of simulated structures of 7 CaLTPs (a) and 7 reference LTPs (b), visualized using the software, Grace 5.1.25. The RMSD fluctuations stabilizes over a period of 1 ns between 0.175–0.275 nm.

LTP proteins identified in *Cicer arietinum* L., chickpea, during the attack by a most common pest and polyphagous Arthropod, *Helicoverpa armigera*, to determine their capability as food allergens. Among the 48 LTPs identified in chickpea (unpublished data), only 14 CaLTPs were differentially regulated (unpublished data) during herbivory. Out of these 14 defense-related CaLTPs, seven were defined as allergens, by determining their allergenicity potential, B-cell epitope prediction, three-dimensional structure prediction, docking with IgE epitopes, and molecular dynamic simulation of docked structures. The known allergens from Fabales (Ara h 9, Ara h 16, & Ara h 17 of *Arachis hypogaea*, Len c 3 of *Lens culinaris*, Lup an 3 of *Lupinus angustifolius*, and Pha v 3 of *Phaseolus vulgaris*) and Pru p 3 from *Prunus persica* were used as references for conducting this analysis.

The United Nations Food and Agriculture Organization (FAO) along with the World Health Organization (WHO) have laid down standard guidelines for defining the allergenicity potential of a candidate allergen, which state that a protein can be considered a probable allergen when it is compared to with a known allergen having either similarity with 6–8 consecutive amino acids or sequence similarity greater than 35%, in range of 80 amino acids (FAO/WHO Agriculture and

Consumer Protection, 2001; FAO/WHO Codex Alimentarius Commission, 2003; Stadler & Stadler, 2003). Several bioinformatics tools are easily accessible for determining the allergenicity potential of unknown proteins and are based on different algorithms. We have used three different prediction tools to substantiate the prediction results. The first tool, AlgPred, incorporates various approaches such as prediction based on the composition of protein (amino acids and dipeptides) using support vector machine (SVM); motif-based prediction with the help of MEME/MAST software; prediction based on similarity of segments with allergen representative proteins (ARPs); and segment similarity with known IgE epitopes (Saha & Raghava, 2006b). Likewise, AllerTOP recognizes potential allergens by considering the major chemical properties of the amino acid sequences, while AllergenFP encompasses physical, chemical, and biological characteristics during the prediction (Dimitrov et al., 2014a, 2014b).

The interaction of the probable allergens with the B-cell epitopes testifies that the residues present on the protein surfaces interact with the antibody molecules. The residues are thus, solvent accessible, hydrophilic, and tend to interact with water molecules in the interacting system. The presence of a large number of epitope binding residues inculcates

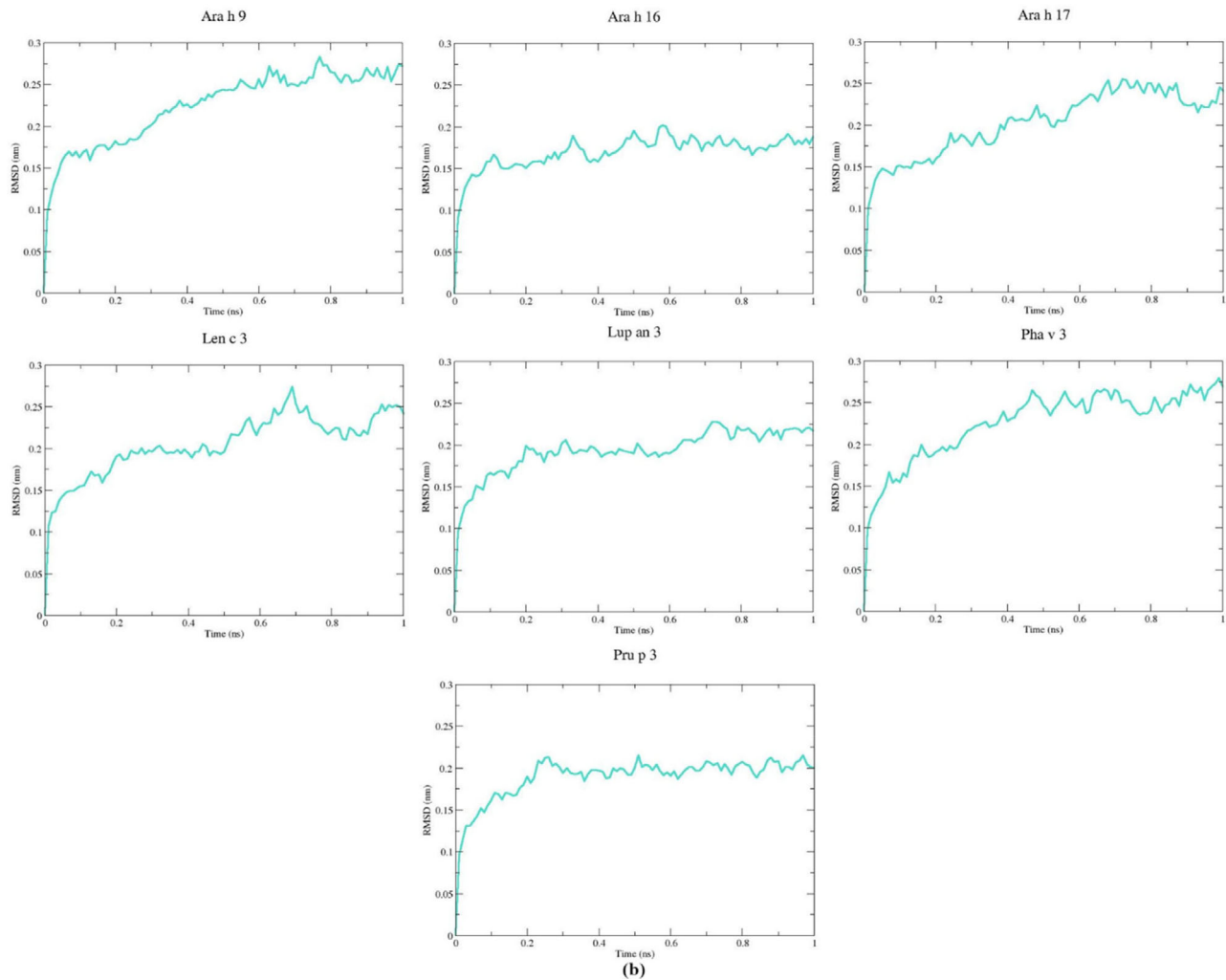


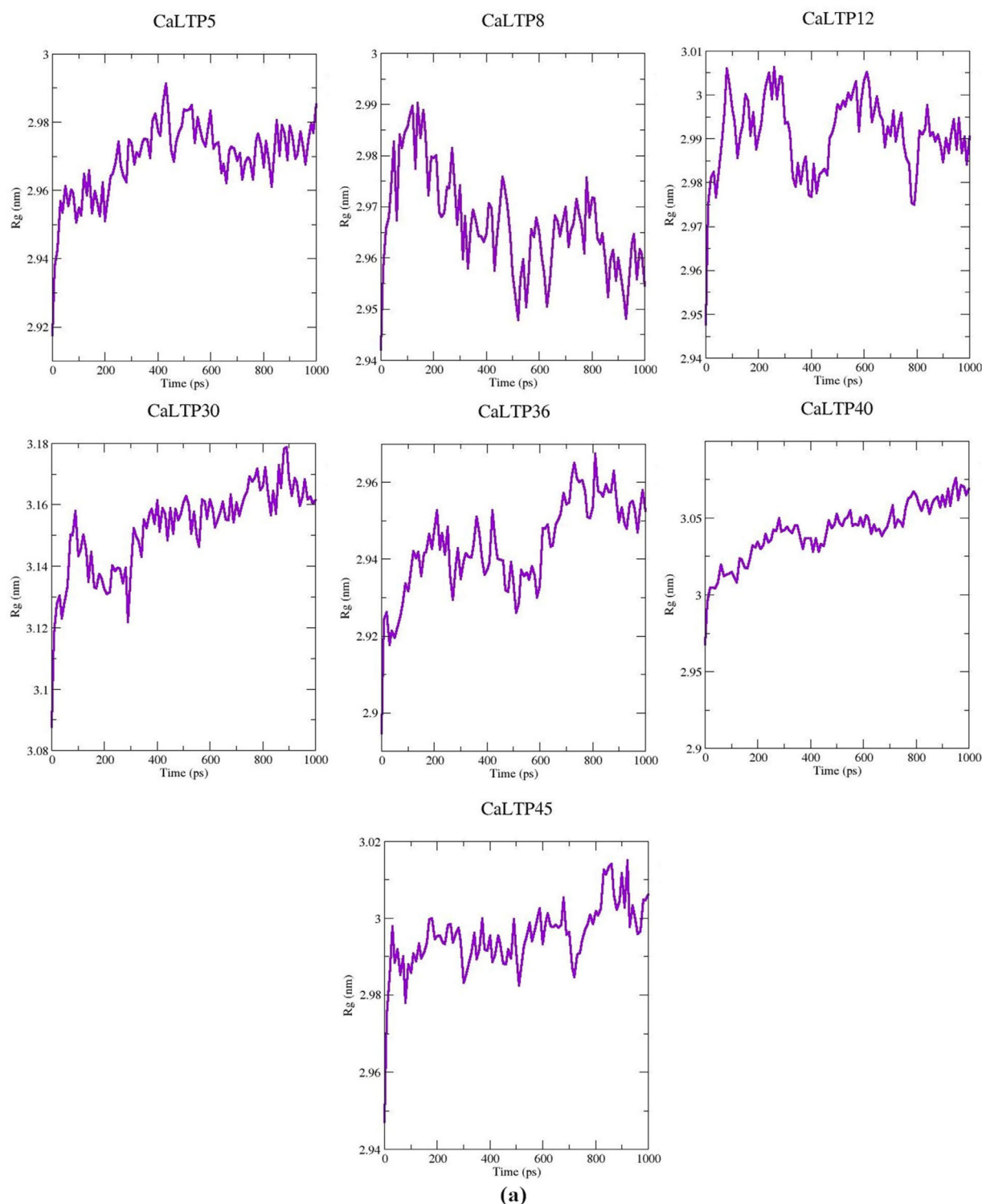
Figure 5. Continued.

stronger interactions with the antibody molecules and thus triggers severe allergic responses (Arora et al., 2020). Therefore, a prerequisite knowledge of the B-cell epitope binding regions present in the potential allergens is useful for a better understanding of the allergic responses caused by these allergens. To assert confidence in the B-cell epitope prediction results, we again conducted the prediction using three different bioinformatics tools and corroborated the final residues' sequences with a consensus sequence derived from all the prediction results. The epitopes from all the 7 LTPs were then aligned together using CLUSTALW algorithm, along with the reference LTPs to check their conservation status. The presence of three conserved domains was observed in each CaLTP similar to the conservation pattern in Pru p 3 and other reference LTPs, providing strong evidence for the similarity in hypersensitivity and allergenicity responses. The alignment of the residues using the WebLogo tool provided insights into the conservation of amino acids in the LTPs, as well as the conservation status of cysteine residues, was prominent (Banerjee et al., 2002; de Weerd et al., 2003).

In addition to this, the molecular interaction between the antigen and antibody during allergic reaction requires

extensive insights into the structure of the potential allergens, and thus, we performed three-dimensional protein structure prediction using the iterative structure assembling and from several threading alignments (Yang & Zhang, 2015b). Ideally, we speculate a correct three-dimensional structure and stable protein conformations for these unknown allergens to elicit the required allergic response (Arora et al., 2020). From our results, the models predicted are relatively stable with the average quality of the models and were considered for docking with the IgE epitopes.

The antigen-antibody docking was performed to mimic the potential allergenicity responses of the LTPs in the presence of IgE epitopes, and the interactions at the atomic level highlighted the interaction of cysteine residues with the epitopes. The interaction of allergens with the IgE molecules is facilitated by changing the protein conformations constantly to acquire the most favorable orientation. The negative interaction energy is inversely related to the stability of the protein (Brenke et al., 2012). The differences in the interaction energies signify variations in the bond strength, i.e. intra- and inter-molecular hydrogen bonding, of each LTP-IgE complex. The number of IgE binding residues in each system was correlated with the docking interaction energies, to



(a)

Figure 6. Radius of gyration (R_g) curves of all the LTPs: (a) 7 CaLTPs and (b) 7 reference LTPs, using the software, Grace 5.1.25. The fluctuations in R_g stabilize over a period of 1000 ps at a value of ~ 2.94 nm.

understand the underlying factors contributing to the stability of these structures. A weak correlation was observed between the IgE binding site length and the docking energies of respective LTP allergens. Not many LTPs were found close to the line of best fit. Among the CaLTPs, only CaLTP5 and CaLTP40 were found to have a relation with the number of IgE binding sites. This suggests there exist some other causes for this deviation or the stability of the proteins.

These include the prediction of misfolded structures; incorrect prediction of sheets, helices, and loops; thermodynamic stability of the native structure; exposure of hydrophobic chains to solvents, to name a few. This can be corroborated using NMR spectroscopy or X-ray crystallography to decipher the correct topology and folds of the probable LTP allergens (Arora et al., 2020; A. Singh et al., 2018). We further conducted molecular dynamics simulation to validate the free

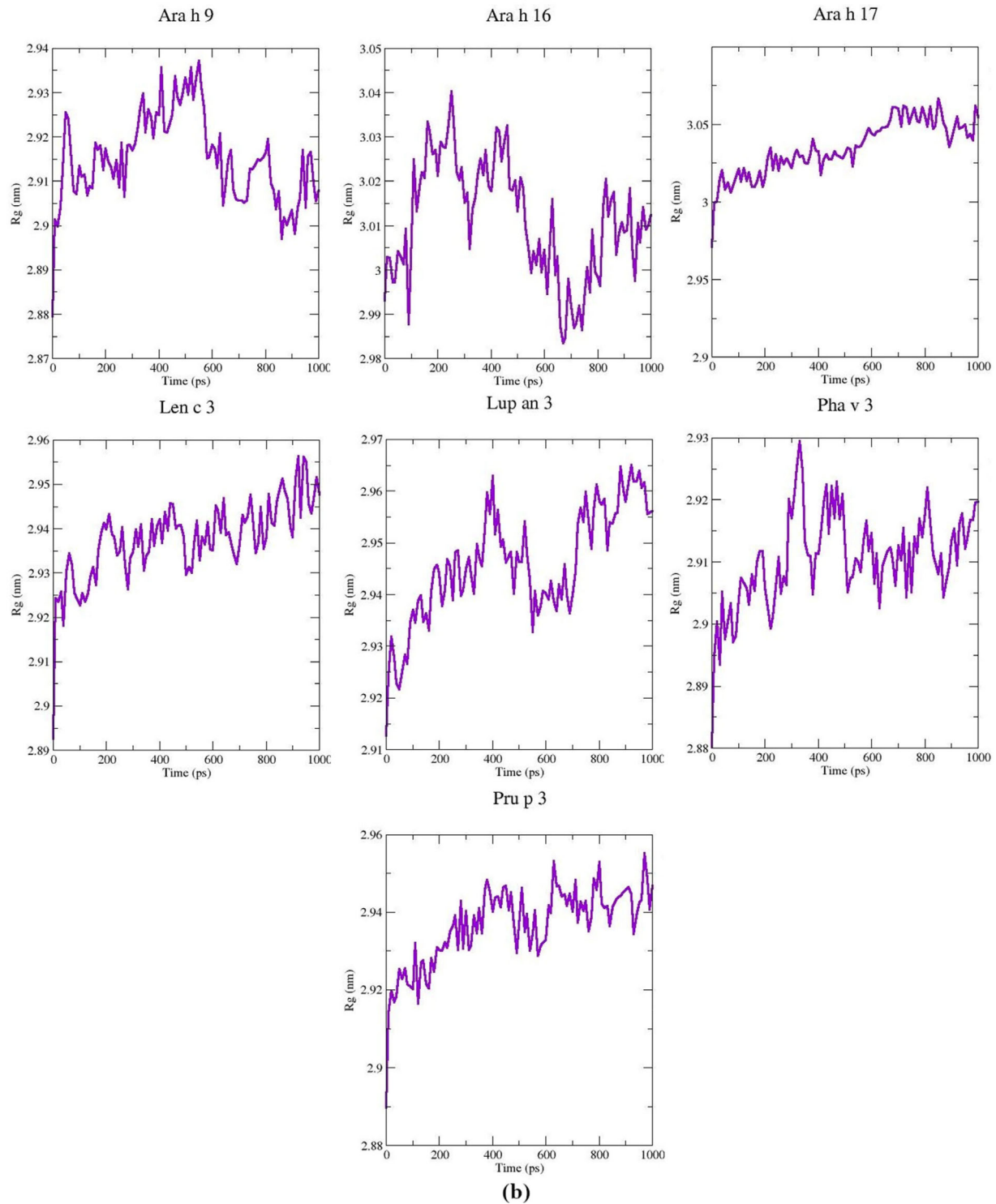


Figure 6. Continued.

movement of docked models in every system. The docking complexes of LTP and IgE molecules were stabilized within 10 ns of subjection to simulation, with stable topology and increasingly compact structures (A. Singh et al., 2018). Notably, our findings have led to a foundation for getting the first insights into the energetics involved in the interaction of these LTPs with IgE molecules, suggesting their potential as allergens.

Conclusion

The process of sensitization to allergic food components depends on the immunological tolerance in response to certain food products. The diminished tolerance results in the synthesis of protein-specific IgE molecules in the human body, which on encountering specific proteins/allergens, can trigger severe immune reactions. The presence of allergens in different food products and their cross-reactivity can be

assessed at the protein level, using *in silico* analyses. In this study, we have identified a family of allergens that get accumulated in the chickpea seeds, when it is attacked by a pod-borer pest, *H. armigera*. This family of allergens has characteristics of pathogenesis-related proteins and is identified as LTPs, which have multifarious roles in plant development, and responses after exposure to environmental stresses. Our findings suggest that these LTPs are potential allergens and can trigger severe allergenic and cross-reactivity responses in humans. Prior knowledge of the binding sites in the B-cell epitopes, and checking their interaction with the candidate allergens, provided strong evidence of the protein-protein interaction, which can serve as a basis for discovering effective drugs in clinical trials to combat the allergenic effects of these LTPs. Further proteomics analysis such as NMR, MS, and X-ray crystallography can validate our results and annotate the function and interaction of these LTPs with different antibody molecules.

Acknowledgements

We sincerely thank Hansraj College, University of Delhi for providing the necessary infrastructure and resources for this study. This study was supported in part by resources and technical expertise from the Georgia Advanced Computing Resource Center, a partnership between the University of Georgia's Office of the Vice President for Research and the Office of the Vice President for Information Technology. We would also like to extend our gratitude to Dr. Anchala Choudhary, Senior Scientist at Arontier, Seoul 06735, South Korea for guiding us in GROMACS.

Disclosure statement

The authors declare no competing financial interests.

Funding

AS acknowledge the funding received from Science and Engineering Research Board (SERB), DST, Ministry of Science and Technology, Government of India, New Delhi, India, for the financial supports given to AS [ECR/2017/002478 (AS)]. Sincere thanks to Prof. Rama, Principal, Hansraj College for her help and support.

ORCID

Harshita Negi  <http://orcid.org/0000-0002-9418-5098>
 Harshita Saxena  <http://orcid.org/0000-0002-7982-6563>
 Indrakant K. Singh  <http://orcid.org/0000-0002-3211-8471>
 Archana Singh  <http://orcid.org/0000-0003-3916-7513>

References

Arora, R., Kumar, A., Singh, I. K., & Singh, A. (2020). Pathogenesis related proteins: A defensin for plants but an allergen for humans. *International journal of Biological Macromolecules*, 157, 659–672. <https://doi.org/10.1016/j.ijbiomac.2019.11.223>

Banerjee, B., Kurup, V. P., Greenberger, P. A., Kelly, K. J., & Fink, J. N. (2002). C-terminal cysteine residues determine the IgE binding of *Aspergillus fumigatus* allergen Asp f 2. *The Journal of Immunology*, 169(9), 5137–5144. <https://doi.org/10.4049/jimmunol.169.9.5137>

Bauer, P., Hess, B., & Lindahl, E. (2022). *GROMACS 2022.2 manual*. <https://doi.org/10.5281/ZENODO.6637572>

Benkert, P., Biasini, M., & Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics (Oxford, England)*, 27(3), 343–350. <https://doi.org/10.1093/bioinformatics/btq662>

Borges, J.-P., Barre, A., Culerrier, R., Granier, C., Didier, A., & Rougé, P. (2008). Lipid transfer proteins from Rosaceae fruits share consensus epitopes responsible for their IgE-binding cross-reactivity. *Biochemical and Biophysical Research Communications*, 365(4), 685–690. <https://doi.org/10.1016/j.bbrc.2007.11.046>

Bowie, J. U., Lüthy, R., & Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science (New York, N.Y.)*, 253(5016), 164–170. <https://doi.org/10.1126/science.1853201>

Brenke, R., Hall, D. R., Chuang, G.-Y., Comeau, S. R., Bohnuud, T., Beglov, D., Schueler-Furman, O., Vajda, S., & Kozakov, D. (2012). Application of asymmetric statistical potentials to antibody–protein docking. *Bioinformatics (Oxford, England)*, 28(20), 2608–2614. <https://doi.org/10.1093/bioinformatics/bts493>

Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Science*, 2(9), 1511–1519. <https://doi.org/10.1002/pro.5560020916>

Crooks, G. E., Hon, G., Chandonia, J.-M., & Brenner, S. E. (2004). WebLogo: A sequence logo generator. *Genome Research*, 14(6), 1188–1190. <https://doi.org/10.1101/gr.849004>

de Weerd, N., Bhalla, P. L., & Singh, M. B. (2003). Effect of cysteine mutagenesis on human IgE reactivity of recombinant forms of the major rye grass pollen allergen Lol p 1. *Allergology International*, 52(4), 183–190. <https://doi.org/10.1046/j.1323-8930.2003.00300.x>

Desta, I. T., Porter, K. A., Xia, B., Kozakov, D., & Vajda, S. (2020). Performance and its limits in rigid body protein-protein docking. *Structure (London, England: 1993)*, 28(9), 1071–1081.e3. <https://doi.org/10.1016/j.str.2020.06.006>

Dimitrov, I., Bangov, I., Flower, D. R., & Doytchinova, I. (2014a). AllerTOP v.2—A server for in silico prediction of allergens. *Journal of Molecular Modeling*, 20(6), 2278. <https://doi.org/10.1007/s00894-014-2278-5>

Dimitrov, I., Naneva, L., Doytchinova, I., & Bangov, I. (2014b). AllergenFP: Allergenicity prediction by descriptor fingerprints. *Bioinformatics (Oxford, England)*, 30(6), 846–851. <https://doi.org/10.1093/bioinformatics/btt619>

Egger, M., Hauser, M., Mari, A., Ferreira, F., & Gadermaier, G. (2010). The role of lipid transfer proteins in allergic diseases. *Current Allergy and Asthma Reports*, 10(5), 326–335. <https://doi.org/10.1007/s11882-010-0128-9>

Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., & Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *The Journal of Chemical Physics*, 103(19), 8577–8593. <https://doi.org/10.1063/1.470117>

FAO/WHO Agriculture and Consumer Protection. (2001). *Evaluation of allergenicity of genetically modified foods: Report of a Joint FAO/WHO Expert Consultation on Allergenicity of Foods Derived from Biotechnology*, 22–25 January 2001 (WHO/FOS/2001.01; Joint FAO/WHO Expert Consultation on Allergenicity of Foods Derived from Biotechnology). Food and Agriculture Organization of the United Nations. <https://apps.who.int/iris/handle/10665/340572>

FAO/WHO Codex Alimentarius Commission. (2003). *Codex principles and guidelines on foods derived from biotechnology*. Joint FAO/WHO Food Standards Programme.

Finkina, E. I., Melnikova, D. N., Bogdanov, I. V., & Ovchinnikova, T. V. (2017). Plant pathogenesis-related proteins PR-10 and PR-14 as components of innate immunity system and ubiquitous allergens. *Current Medicinal Chemistry*, 24(17), 1772–1787. <https://doi.org/10.2174/0929867323666161026154111>

García-Casado, G., Pacios, L. F., Díaz-Perales, A., Sánchez-Monge, R., Lombardero, M., García-Selles, F. J., Polo, F., Barber, D., & Salcedo, G. (2003). Identification of IgE-binding epitopes of the major peach allergen Pru p 3. *The Journal of Allergy and Clinical Immunology*, 112(3), 599–605. [https://doi.org/10.1016/S0091-6749\(03\)01605-1](https://doi.org/10.1016/S0091-6749(03)01605-1)

Garino, C., Coisson, J. D., & Arlorio, M. (2016). In silico allergenicity prediction of several lipid transfer proteins. *Computational Biology and Chemistry*, 60, 32–42. <https://doi.org/10.1016/j.compbiolchem.2015.11.006>

- Gautam, B. (2020). Energy minimization. In R. Trindade Maia, R. Maciel de Moraes Filho, & M. Campos (Eds.), *Homology molecular modeling—Perspectives and applications*. IntechOpen. <https://doi.org/10.5772/intechopen.94809>
- Giangrieco, I., Alessandri, C., Rafeiani, C., Santoro, M., Zuzzi, S., Tuppo, L., Tamburrini, M., D'Avino, R., Ciardiello, M. A., & Mari, A. (2015). Structural features, IgE binding and preliminary clinical findings of the 7kDa Lipid Transfer Protein from tomato seeds. *Molecular Immunology*, 66(2), 154–163. <https://doi.org/10.1016/j.molimm.2015.02.025>
- Hauser, M., Roulias, A., Ferreira, F., & Egger, M. (2010). Panallergens and their impact on the allergic patient. *Allergy, Asthma, and Clinical Immunology*, 6(1), 1. <https://doi.org/10.1186/1710-1492-6-1>
- Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINC: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12), 1463–1472. [https://doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:12<1463::AID-JCC4>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H)
- Hoffmann-Sommergruber, K. (2000). Plant allergens and pathogenesis-related proteins. *International Archives of Allergy and Immunology*, 122(3), 155–166. <https://doi.org/10.1159/000024392>
- Jespersen, M. C., Peters, B., Nielsen, M., & Marcatili, P. (2017). BepiPred-2.0: Improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Research*, 45(W1), W24–W29. <https://doi.org/10.1093/nar/gkx346>
- Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118(45), 11225–11236. <https://doi.org/10.1021/ja9621760>
- Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., & Jorgensen, W. L. (2001). Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *The Journal of Physical Chemistry B*, 105(28), 6474–6487. <https://doi.org/10.1021/jp003919d>
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, 10(6), 845–858. <https://doi.org/10.1038/nprot.2015.053>
- Kerchev, P. I., Fenton, B., Foyer, C. H., & Hancock, R. D. (2012). Plant responses to insect herbivory: Interactions between photosynthesis, reactive oxygen species and hormonal signalling pathways. *Plant, Cell & Environment*, 35(2), 441–453. <https://doi.org/10.1111/j.1365-3040.2011.02399.x>
- Kozakov, D., Beglov, D., Bohnuud, T., Mottarella, S. E., Xia, B., Hall, D. R., & Vajda, S. (2013). How good is automated protein docking?: Automated Protein Docking. *Proteins*, 81(12), 2159–2166. <https://doi.org/10.1002/prot.24403>
- Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., Beglov, D., & Vajda, S. (2017). The ClusPro web server for protein–protein docking. *Nature Protocols*, 12(2), 255–278. <https://doi.org/10.1038/nprot.2016.169>
- Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26(2), 283–291. <https://doi.org/10.1107/S0021889892009944>
- Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR*, 8(4), 477–486. <https://doi.org/10.1007/BF00228148>
- Lleonart, R., Cisteró, A., Carreira, J., Batista, A., & Moscoso del Prado, J. (1992). Food allergy: Identification of the major IgE-binding component of peach (*Prunus persica*). *Annals of Allergy*, 69(2), 128–130.
- Lüthy, R., Bowie, J. U., & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature*, 356(6364), 83–85. <https://doi.org/10.1038/356083a0>
- Mark, P., & Nilsson, L. (2001). Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *The Journal of Physical Chemistry A*, 105(43), 9954–9960. <https://doi.org/10.1021/jp003020w>
- Morales, M., López-Matas, M. Á., Moya, R., & Carnés, J. (2014). Cross-reactivity among non-specific lipid-transfer proteins from food and pollen allergenic sources. *Food Chemistry*, 165, 397–402. <https://doi.org/10.1016/j.foodchem.2014.05.101>
- Pacios, L. F., Tordesillas, L., Cuesta-Herranz, J., Compes, E., Sánchez-Monge, R., Palacín, A., Salcedo, G., & Díaz-Perales, A. (2008). Mimotope mapping as a complementary strategy to define allergen IgE-epitopes: Peach Pru p 3 allergen as a model. *Molecular Immunology*, 45(8), 2269–2276. <https://doi.org/10.1016/j.molimm.2007.11.022>
- Pastorello, E. A., Pravettoni, V., Trambaioli, C., Pompei, C., Brenna, O., Farioli, L., & Conti, A. (2001). Lipid transfer proteins and 2S albumins as allergens. *Allergy*, 56(s67), 45–47. <https://doi.org/10.1034/j.1398-9995.2001.00914.x>
- Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: A unified platform for automated protein structure and function prediction. *Nature Protocols*, 5(4), 725–738. <https://doi.org/10.1038/nprot.2010.5>
- RStudio Team. (2020). *RStudio: Integrated development for R*. RStudio. <http://www.rstudio.com/>
- Saha, S., & Raghava, G. P. S. (2004). BcePred: Prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. In G. Nicosia, V. Cutello, P. J. Bentley, & J. Timmis (Eds.), *Artificial immune systems* (Vol. 3239, pp. 197–204). Springer. https://doi.org/10.1007/978-3-540-30220-9_16
- Saha, S., & Raghava, G. P. S. (2006a). Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins*, 65(1), 40–48. <https://doi.org/10.1002/prot.21078>
- Saha, S., & Raghava, G. P. S. (2006b). AlgPred: Prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic acids Research*, 34(Web Server issue), W202–W209. <https://doi.org/10.1093/nar/gkl343>
- Saha, S., & Raghava, G. P. S. (2007). Prediction methods for B-cell epitopes. In D. R. Flower (Ed.), *Immunoinformatics* (Vol. 409, pp. 387–394). Humana Press. https://doi.org/10.1007/978-1-60327-118-9_29
- Salcedo, G., Sánchez-Monge, R., Barber, D., & Díaz-Perales, A. (2007). Plant non-specific lipid transfer proteins: An interface between plant defence and human allergy. *Biochimica et Biophysica Acta*, 1771(6), 781–791. <https://doi.org/10.1016/j.bbalip.2007.01.001>
- Salcedo, G., Sanchez-Monge, R., Diaz-Perales, A., Garcia-Casado, G., & Barber, D. (2004). Plant non-specific lipid transfer proteins as food and pollen allergens. *Clinical and Experimental Allergy*, 34(9), 1336–1341. <https://doi.org/10.1111/j.1365-2222.2004.02018.x>
- Schneider, T. D., & Stephens, R. M. (1990). Sequence logos: A new way to display consensus sequences. *Nucleic acids Research*, 18(20), 6097–6100. <https://doi.org/10.1093/nar/18.20.6097>
- Schrödinger, L., & DeLano, W. (2020). *The PyMOL molecular graphics system* (2.0). <https://pymol.org/2/>
- Sharma, P., Gaur, S. N., & Arora, N. (2013). In silico identification of IgE-binding epitopes of osmotin protein. *PLoS One*, 8(1), e54755. <https://doi.org/10.1371/journal.pone.0054755>
- Singh, A., Singh, I. K., & Verma, P. K. (2008). Differential transcript accumulation in *Cicer arietinum* L. in response to a chewing insect *Helicoverpa armigera* and defence regulators correlate with reduced insect performance. *Journal of Experimental Botany*, 59(9), 2379–2392. <https://doi.org/10.1093/jxb/ern111>
- Singh, A., Tyagi, C., Nath, O., & Singh, I. K. (2018). *Helicoverpa*-inducible Thioredoxin h from *Cicer arietinum*: Structural modeling and potential targets. *International Journal of Biological Macromolecules*, 109, 231–243. <https://doi.org/10.1016/j.ijbiomac.2017.12.079>
- Singh, I. K., Kumar, S., Singh, S., & Singh, A. (2017). Expression profiling of mitogen-activated protein kinase genes from chickpea (*Cicer arietinum* L.) in response to *Helicoverpa armigera*, wounding and signaling compounds. *Journal of Asia-Pacific Entomology*, 20(3), 942–948. <https://doi.org/10.1016/j.aspen.2017.07.003>
- Singh, I. K., & Singh, A. (Eds.). (2021). *Plant-pest interactions: From molecular mechanisms to chemical ecology: Chemical ecology*. Springer Singapore. <https://doi.org/10.1007/978-981-15-2467-7>
- Stadler, M. B., & Stadler, B. M. (2003). Allergenicity prediction by protein sequence. *The FASEB Journal*, 17(9), 1141–1143. <https://doi.org/10.1096/fj.02-1052fe>
- Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, 38(7), 3022–3027. <https://doi.org/10.1093/molbev/msab120>

- Tordesillas, L., Pacios, L. F., Palacin, A., Quirce, S., Armentia, A., Barber, D., Salcedo, G., & Diaz-Perales, A. (2009). Molecular basis of allergen cross-reactivity: Non-specific lipid transfer proteins from wheat flour and peach fruit as models. *Molecular Immunology*, 47(2–3), 534–540. <https://doi.org/10.1016/j.molimm.2009.07.028>
- Turner, P. J. (2015). *XMGRACE (5.1. 25)*. Center for coastal and land-margin research. Oregon Graduate Institute of Science and Technology.
- Vajda, S., Yueh, C., Beglov, D., Bohnuud, T., Mottarella, S. E., Xia, B., Hall, D. R., & Kozakov, D. (2017). New additions to the ClusPro server motivated by CAPRI. *Proteins: Structure, Function, and Bioinformatics*, 85(3), 435–444. <https://doi.org/10.1002/prot.25219>
- van Ree, R. (2002). Clinical importance of non-specific lipid transfer proteins as food allergens. *Biochemical Society Transactions*, 30(Pt 6), 910–913. <https://doi.org/10.1042/bst0300910>
- Wawrzęńczyk, A., Żbikowska-Gotz, M., Wawrzęńczyk, A., & Bartuzi, Z. (2020). Sensitisation to lipid transfer proteins in pollen-allergic adults with food allergy. *Postępy dermatologii i Alergologii*, 37(4), 508–512. <https://doi.org/10.5114/ada.2020.98278>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (2nd ed.). Springer. <https://doi.org/10.1007/978-3-319-24277-4>
- Yadav, M., Pandey, J., Chakraborty, A., Hassan, M. I., Kundu, J. K., Roy, A., Singh, I. K., & Singh, A. (2022). A comprehensive analysis of calmodulin-like proteins of glycine max indicates their role in calcium signaling and plant defense against insect attack. *Frontiers in Plant Science*, 13, 817950. <https://doi.org/10.3389/fpls.2022.817950>
- Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, 12(1), 7–8. <https://doi.org/10.1038/nmeth.3213>
- Yang, J., & Zhang, Y. (2015a). I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Research*, 43(W1), W174–W181. <https://doi.org/10.1093/nar/gkv342>
- Yang, J., & Zhang, Y. (2015b). Protein structure and function prediction using I-TASSER. *Current Protocols in Bioinformatics*, 52, 5.8.1–5.815. <https://doi.org/10.1002/0471250953.bi0508s52>
- Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, 9(1), 40. <https://doi.org/10.1186/1471-2105-9-40>
- Zheng, W., Zhang, C., Li, Y., Pearce, R., Bell, E. W., & Zhang, Y. (2021). Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations. *Cell Reports Methods*, 1(3), 100014. <https://doi.org/10.1016/j.crmeth.2021.100014>