Molecular Immunology 114 (2019) 19–29

Contents lists available at ScienceDirect

Molecular Immunology

journal homepage: www.elsevier.com/locate/molimm

Comparative structural and thermal stability studies of Cuc m 2.0101, Art v 4.0101 and other allergenic profilins

A. Brenda Kapingidza[a], Sarah E. Pye[a], Noah Hyduke[a], Coleman Dolamore[a], Swanandi Pote[a], Caleb R. Schlachter[a], Scott P. Commins[b], Krzysztof Kowal[c,d], Maksymilian Chruszcz[a,*]

[a] Department of Chemistry and Biochemistry, University of South Carolina, Columbia, SC, 29208, United States
[b] Thurston Arthritis Research Center, University of North Carolina at Chapel Hill, NC, 27599, United States
[c] Department of Allergology and Internal Medicine, Medical University of Bialystok, Bialystok, Poland
[d] Department of Experimental Allergology and Immunology, Medical University of Bialystok, Bialystok, Poland

ARTICLE INFO

Keywords:
Profilin
Allergen
Pollen-food syndrome
Thermal stability
Oligomerization

ABSTRACT

Worldwide, more than one-third of the population suffers from allergies. A significant fraction of officially registered allergens originate from the profilin family of proteins. Profilins are small ubiquitous proteins which are found in plants, viruses and various eukaryotes including mammals. Although they are primarily regarded as minor allergens, profilins are important players in immunoglobulin E (IgE) cross-reactivity. However, in some populations profilins are recognized by IgE from at least 50% of patients allergic to a given allergen source. Cuc m 2.0101 is recognized by IgE in more than 80% of muskmelon-allergic patients. The recombinant isoallergen Cuc m 2.0101 was produced in significant quantities and its X-ray crystal structure was determined. In addition, a new Art v 4.0101 (mugwort profilin) structure was determined. The profilins Cuc m 2.0101 and Art v 4.0101 were compared in terms of their structure and thermal stability. Furthermore, structural similarities and IgE cross-reactivity between profilins from different sources are discussed to explain the molecular basis of various clinical syndromes involving this group of allergens. Special emphasis is placed on discussion of profilins' quaternary structures and their relation to biological function, as well as to protein allergenicity. Moreover, a potential impact of protein purification protocols on the structure of profilins is highlighted.

1. Introduction

Profilins are ubiquitous proteins that were discovered more than 40 years ago in calf thymus (Carlsson et al., 1977). Profilins, which are small (12–15 kDa) and typically eukaryotic proteins, have been proven to be involved in regulating various cellular processes, including membrane trafficking, actin cytoskeletal dynamics, as well as binding to proline-rich regions of proteins (Carlsson et al., 1977; Tanaka and Shibata, 1985). It has also been shown that profilins can interact with phospholipids such as phosphatidylinositol 4,5-bisphosphate and phosphatidylinositol 4-phosphate (Hansson et al., 1988; Lassing and Lindberg, 1985).

Along with proteins like cyclophilins, nonspecific lipid-transfer proteins, polcalcins, and pathogenesis-related proteins 10 (PR-10 s or members of the pathogenesis-related protein 10 family), profilins are also considered to be panallergens (Hauser et al., 2010). These ubiquitous proteins share high sequence identity and have extremely conserved three-dimensional structures (Mari, 2001). This explains why more often high IgE cross-reactivity is observed among panallergens (Chruszcz et al., 2018). However, currently only plants and mites have been identified as sources of allergenic profilins (Chruszcz et al., 2018). Plant profilins that are registered as allergens have high sequence identity (at minimum 65%), even between relatively distant related members (Offermann et al., 2016). The allergenic plant profilins originate from pollens of weeds, trees, and grasses, as well as from fruits, legumes, seeds and vegetables (Hauser et al., 2010). Therefore, profilins can account for the so-called pollen-food syndrome and between 10–30% of patients allergic to pollen are also sensitized to profilins (Tordesillas et al., 2011; van Ree et al., 1995). Profilins have also been indicated as being responsible for food allergy depending on sensitization to cross-reactive pollens in the mugwort-celery-spice syndrome and ragweed-melon allergy (Asero et al., 2011; Ebner et al., 1998).

Abbreviations: DSF, Differential Scanning Fluorimetry; IgE, immunoglobulin E; IPTG, isopropyl β-D-1-thiogalactopyranoside; β-ME, β-mercaptoethanol; PDB, Protein Data Bank

* Corresponding author at: Department of Chemistry and Biochemistry, University of South Carolina, Columbia, SC 29208, USA.
E-mail address: chruszcz@mailbox.sc.edu (M. Chruszcz).

https://doi.org/10.1016/j.molimm.2019.07.004
Received 4 March 2019; Received in revised form 25 June 2019; Accepted 8 July 2019
0161-5890/ © 2019 Elsevier Ltd. All rights reserved.

Furthermore, several studies have implicated profilins to play a role in sensitization to carrot, peach, pear, potato, lychee, tomato, apple, hazelnut, banana and others (Asero et al., 2003). The relatively low potency of profilins as food allergens is mainly correlated to their poor stability; typically they are susceptible to denaturation in conditions encountered during food preparation or the digestive process (Asero et al., 2003). It was also shown that the thermal stability of pollen allergens is relatively low (Offermann et al., 2016).

A majority of the allergenic profilins are considered as minor allergens (Tordesillas et al., 2011). However, Cuc m 2.0101, a profilin from *Cucumis melo* (muskmelon), can elicit IgE production in 80% of muskmelon allergic patients. Therefore, Cuc m 2.0101 is regarded as a major allergen (López-Torrejón et al., 2005a). Muskmelons are found throughout the world and are commercially grown in Turkey, the USA, Spain, Romania, Morocco, India, Mexico, and Egypt (Burger et al., 2010; Ganglberger et al., 2001).

Artemisia vulgaris, known colloquially as mugwort, is another known source of profilins. It is a member of the Asteraceae family found in temperate regions of the northern hemisphere, the Mediterranean basin, and parts of Asia (Asero et al., 2006). Patients with mugwort sensitivity often experience ragweed sensitivity as a result of co-recognition of the major allergens from each plant rather than cross-reactivity of common antigens (Asero et al., 2006). It is noteworthy that *A. vulgaris* is one of the few important allergenic sources causing allergic diseases in 10–14% of pollinosis (pollen allergy) patients (Gadermaier et al., 2004, 2008). Art v 4, an *A. vulgaris* profilin, is a clinically significant weed allergen (Offermann et al., 2016). In patients with mugwort allergy, mugwort pollen is primarily responsible for allergic reactions in late summer and autumn (Offermann et al., 2016). Specifically, Art v 4 has a sensitizing prevalence of 34–36% in Europe (Gadermaier et al., 2014).

In this study, the X-ray crystal structure of Cuc m 2.0101 was determined at 2.4 Å and compared to different structures of Art v 4.0101, including the new structure described here that was determined at 1.9 Å. Additionally, we compared the thermal stability of profilins. Combination of structural and thermal stability studies offers explanations for the molecular basis of various clinical syndromes observed involving profilins. Comparison of profilins provides new insights into the allergenicity of this family of proteins and allergy diagnostics, and may facilitate the development of hypoallergenic molecules that can be exploited in immunotherapy.

2. Materials and methods

2.1. Protein production

Gene coding for Cuc m 2.0101 (131 amino acids) was synthesized and inserted in pJExpress411 plasmid by DNA 2.0 (currently ATUM, Newark, CA). The Cuc m 2 insert was designed with a cleavable N-terminal purification tag MHHHHHHSSGVDLGTENLYFQ/SGSG, where the slash denotes the Tobacco Etch Virus (TEV) protease cleavage site. The plasmid was transformed into *E. coli* BL21 (DE3) pLysS cells. Cultures were grown in Lysogeny Broth with 50 μg/mL kanamycin at 37 °C to an OD_{600} of 0.4. Cultures were then moved to 22 °C and grown to an OD_{600} of 0.8 and induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM, cooled to 16 °C, and induced for 16–18 hours.

Cell pellets were resuspended in lysis buffer (50 mM Tris − HCl, 500 mM NaCl, 2% glycerol, 20 mM β-mercaptoethanol (β-ME), 10 mM imidazole, pH 7.4) and lysed by sonication. The sonicated mixture was centrifuged at 9000 x *g* for 10 min at 4 °C. The supernatant was loaded onto a Bio-Rad Econo-Pac chromatography column containing Thermo-Scientific HisPur Ni-NTA resin equilibrated with lysis buffer (50 mM Tris, 500 mM NaCl, 2% glycerol, 20 mM β-ME, 30 mM imidazole, pH 7.4). The protein was eluted using elution buffer (50 mM Tris − HCl, 50 mM NaCl, 2% glycerol, 20 mM β-ME, 250 mM imidazole, pH 7.4). Elutions containing protein were determined by SDS-PAGE, pooled and dialyzed overnight using Pierce SnakeSkin Dialysis Tubing with a molecular weight cutoff of 3500 Da and dialysis buffer (50 mM Tris − HCl, 150 mM NaCl, pH 7.4). The dialyzed protein was concentrated using a Millipore Amicon Ultra concentrator with a 3000 Da molecular mass cutoff and was further purified by size exclusion chromatography using a GE Healthcare AKTA-Pure FPLC and HiLoad Superdex 200 column. The protein was observed to be predominantly monomeric based on gel filtration results (Supplementary Figure S1). However, a small fraction of the produced protein was present in a dimeric form. Fractions containing the monomeric protein (Supplementary Figure S2) were concentrated and Cuc m 2 concentration was determined spectrophotochemically measuring absorption at 280 nm using a Thermo-Scientific Nanodrop (extinction co-efficient of 19,940 L mol^{-1} cm^{-1}, as determined by ProtParam (Gasteiger et al., 2003)). Due to the small amount of the produced dimeric form, only the monomeric form of the protein was studied further. In addition, we performed experiments in which the monomeric form of the profilin was run through the gel filtration column in presence or absence of 20 mM β-ME (data not shown). In both cases we observed only elution of the monomeric form of the protein without formation of dimers. TEV protease was used to cleave the purification tag from the Cuc m 2 according to the protocol described before (Booth et al., 2018). For crystallization experiments, the protein (stored in 50 mM Tris − HCl, 150 mM NaCl, pH 7.4) was concentrated to 12, 25 and 40 mg/mL. Art v 4 protein was expressed and purified using the protocol developed for Cuc m 2. The concentration of Art v 4 was measured spectrophotochemically at 280 nm using a Thermo-Scientific Nanodrop (extinction co-efficient of 18,450 L mol^{-1} cm^{-1}).

2.2. Circular dichroism spectroscopy

Before the crystal structure of Cuc m 2 was determined using X-ray crystallography, the secondary structure of the protein (with purification tag intact) was estimated using circular dichroism spectroscopy (JASCO spectrophotometer). Protein samples for CD experiments were desalted using Sephadex® beads previously equilibrated in 10 mM sodium phosphate buffer that had the pH adjusted to 7.4 using phosphoric acid. The samples were concentrated to 40 μM following the same method as mentioned previously. The instrument was operated according to the manufacturer's protocol at a wavelength range of 220–320 nm. Data processing was done using CAPITO software and the secondary structure percentages were calculated (Wiedemann et al., 2013). For comparison purposes we used the Cuc m 2.0101 structure elucidated here (PDB code 6MBX) and secondary structure definitions assigned by DSSP software (Joosten et al., 2010).

2.3. Thermal stability

Differential Scanning Fluorimetry (DSF) was used to investigate the thermal stability of Cuc m 2.0101 with and without purification tag. DSF experiments were carried out using Bio-Rad CFX96 RT-PCR instrument. All solutions were made as an in-house salt gradient and pH screen. A working concentration of 50 mM for the buffer was used for all the conditions with a pH range of 4.0 − 9.5 (in 0.5 unit increments) and a salt range (sodium chloride) 0 − 1.0 M (no salt, 0.05, 0.10, 0.15, 0.20, 0.25, 0.50, and 1.0 M). All the experiments and steps were done according to the procedures described previously (Booth et al., 2018; Ku et al., 2009). To determine the melting temperatures of Cuc m 2.0101 in each condition, Bio-Rad CFX Manager software was used to decipher the inflection point of the melting curves. Each experiment was performed three times and an average melting temperature (T_m) was calculated.

Theoretical melting temperatures for various profilins that are considered to be allergens were calculated with T_m Predictor (Ku et al., 2009). Sequences of profilins that correspond to the mature forms of the

A

Cuc m 2 (with purification tag)

NaCl (M) \ pH	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5
0.000	44	48	54	52	56	55	56	55	53	54	53	53
0.125	44	50	54	54	56	56	56	56	54	54	55	53
0.250	46	50	54	54	56	56	56	56	54	56	55	53
0.500	46	50	54	54	56	56	56	56	55	55	55	53
0.625	46	50	54	54	56	56	56	56	55	55	55	54
0.750	46	50	54	54	56	56	55	56	55	55	55	53
0.875	48	52	55	54	56	56	56	55	55	56	56	54
1.000	49	52	55	54	56	56	56	56	55	55	55	55

Cuc m 2 (purification tag cleaved)

NaCl (M) \ pH	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5
0.000	46	50	54	53	55	54	55	54	52	54	53	51
0.125	45	50	54	53	54	54	54	54	53	54	54	51
0.250	45	50	54	53	54	54	54	54	52	54	54	51
0.500	45	51	55	53	54	54	55	54	53	54	54	51
0.625	47	51	54	54	55	54	54	54	53	53	54	51
0.750	46	51	54	54	54	54	54	54	52	53	54	50
0.875	47	51	54	54	54	53	54	53	51	53	53	52
1.000	47	51	54	53	54	52	54	52	50	52	52	50

B

Cuc m 2 vs. Art v 4 (with purification tag)

NaCl (M) \ pH	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5
0.000	14	8	14	5	6	3	2	1	-1	0	-1	0
0.250	16	10	12	8	6	3	2	2	0	3	2	0
0.500	16	10	12	6	4	3	2	2	1	2	1	1
1.000	19	12	11	4	4	3	2	2	1	1	2	3

Cuc m 2 vs. Art v 4 (purification tag cleaved)

NaCl (M) \ pH	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5
0.000	10	10	10	3	4	2	3	1	-2	0	-1	-3
0.250	14	10	10	5	2	2	2	1	-1	1	1	-1
0.500	14	11	11	5	2	1	1	1	-1	1	2	-1
1.000	16	11	10	3	2	-1	0	-2	-4	-2	0	-2

Fig. 1. DSF results. A) Average melting temperatures (T_m in °C) for Cuc m 2.0101 with and without purification tag. A gradient of salt (0.0 − 1.0 M NaCl) and pH (4.0 − 9.5) buffers were used. Yellow, green, and blue represent high, average, and low melting temperatures, respectively. **B)** Melting temperature (T_m in °C) differences between Cuc m 2.0101 and Art v 4.0101 for proteins with and without purification tag. Yellow, green, and blue represent high, average, and low melting temperature differences respectively. The standard deviation was less than 1 °C for all experiments. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

proteins were used for these calculations (Supplementary Table S1).

2.4. Crystallization

Crystallization of Cuc m 2.0101 was performed at 277 K. The best quality crystals for both Cuc m 2.0101 (with purification tag) and Art v 4.0101 (cleaved purification tag) were grown using the sitting drop vapor diffusion method. The best diffracting Cuc m 2.0101 crystals were obtained when 40 mg/mL protein sample was mixed 1:1 with well solution containing 2.8 M sodium acetate, pH 7.0. Prior to data collection, Cuc m 2.0101 crystals were cryo-protected using 50% glycerol solution, then immediately cryo-cooled in liquid nitrogen. Art v 4.0101 was crystallized at 293 K in MRC 2-drop 96-well crystallization plates (Molecular Dimensions) by mixing 0.49 M sodium phosphate monobasic pH 6.9 (Hampton Index Screen) with recombinant protein in a 1:1 ratio.

2.5. Data collection and structure determination

Data for the new crystal form of Art v 4.0101 were collected remotely using the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source (APS), Argonne National Laboratory. Data for Cuc m 2.0101 were collected at the 19-BM beamline of the Structural Biology Center, APS. Diffraction images of each protein were processed with HKL-2000 (Otwinowski and Minor, 1997). The structures of Cuc m 2.0101 and Art v 4.0101 were solved by molecular replacement using MOLREP (Vagin and Teplyakov, 1997) and HKL-3000 (Minor et al., 2006) with the previously published structure of Art v 4.0101 (PDB code 5EM0) as the starting model (Offermann et al., 2016). The models were rebuilt using BUCCANEER (Cowtan, 2006) and COOT (Emsley and Cowtan, 2004). The structures were refined using REFMAC (Murshudov et al., 2011) and HKL-3000. Various programs from the CCP4 package (Winn et al., 2011) were used for data analysis. COOT and MOLPROBITY (Davis et al., 2007) were used for validation of models. Both structures together with structure factors were deposited into the Protein Data Bank (Berman et al., 2000) with accession codes 6B6J and 6MBX for Art v 4.0101 and Cuc m 2.0101 models, respectively.

2.6. Various computational approaches

A search was performed using the Allergome and AllFam software to identify the various profilins (Mari and Riccioli, 2004; Radauer et al., 2008). Allergen.org was used to identify officially registered allergens (Chapman, 2008), and Clustal Omega (Sievers et al., 2011) with the SIAS Server (http://imed.med.ucm.es/Tools/sias.html) were used to determine their sequence similarity and identity to Cuc m 2.0101. Figures showing sequence alignment were prepared using ESPript (Gouet et al., 2003). Figures and structure alignments were done with PyMOL (DeLano, 2002). Analysis of the quaternary structure was performed using PDBePISA (Krissinel and Henrick, 2007). DALI (Holm et al., 2006) and PDBeFOLD (Krissinel and Henrick, 2004) were used to search PDB for structural homologs of Art v 4.0101 and Cuc m 2.0101.

3. Results

3.1. Protein production and CD spectroscopy

Recombinant Cuc m 2.0101 was produced in *E. coli* and maximum yields of up to 40 mg/L culture was obtained. The observed protein recovery after cleavage of the purification tag by TEV protease was about 80%. Recombinant Art v 4.0101 had comparable quantities to Cuc m 2.0101 (Offermann et al., 2016).

Secondary structure was determined by CAPITO software based on CD spectroscopy data. Cuc m 2.0101 (with purification tag) secondary structure was estimated as follows: α-helices 0–27%, β-strands 30–48% and random coils (irregular) 43–58% (Wiedemann et al., 2013). These results are in agreement with the secondary structure calculated using the crystal structure of Cuc m 2.0101 reported here (α-helices 24%, β-strands 35% and random coils (irregular) 41%).

3.2. Thermal stability

The results for the DSF experiments performed for Cuc m 2.0101

(Fig. 1A) show that the maximum melting temperature (56 °C) is similar to those observed for previously characterized Amb a 8.0101 (profilin from ragweed), Art v 4.0101 and Bet v 2 (profilin from birch) (Booth et al., 2018). From the melting temperatures acquired, Cuc m 2.0101 with the purification tag was slightly more stable than with the cleaved tag. The salt concentration does not have a major impact on Cuc m 2.0101 thermal stability, and the only exemption is seen at pH 4.0 for the protein with the purification tag. In comparison with Amb a 8.0101, Art v 4.0101 and Bet v 2.0101, the muskmelon profilin is significantly more stable in solutions with lower pH values. For example, when Cuc m 2.0101 is compared with Art v 4.0101, the muskmelon protein is clearly more stable in the 4.0–6.0 pH range, and the differences are especially pronounced in solutions with pH between 4.0 and 5.0. In these conditions, Cuc m 2.0101 has melting temperatures 8–19 °C higher in comparison with Art v 4.0101 (Fig. 1B), Amb a 8.0101, and Bet v 2.0101. The biggest differences are observed for solutions with pH 4.0.

To compare the thermal stability of profilins, a computational approach was used as described by Ku et al. (Ku et al., 2009). This approach allowed for the division of the allergenic profilins into three categories (Supplementary Table S1). The first category includes proteins that have a predicted melting temperature above 65 °C, and only one profilin (Vit v 4, grapevine) was predicted to belong to this group. The second category includes proteins melting in 55–65 °C range, and there are several allergenic profilins predicted to be in this group (Cap a 2 (bell pepper), Citr l 2 (watermelon), Cor a 2 (hazel), Cuc m 2 (muskmelon), Gly m 3 (soybean), Mal d 4 (apple), Par *j* 3.0201 (pellitory of the wall), Pha v 5 (bean), Pru av 4 (cherry), Pru du 4 (almond), Pru p 4 (peach), Sola l 1 (tomato), Sola t 8 (potato) and Tyr p 36 (storage mite)). The third and most numerous group includes profilins that have melting temperature predicted to be below 55 °C. The predicted melting temperatures agree with the experimental results for Amb a 8.0101, Art v 4.0101, Bet v 2.0101 and Cuc m 2.0101 (Booth et al., 2018) (Fig. 1). At the same time, the predictions are in agreement with various other reports that indicated profilins as labile proteins (Amnuaycheewa and de Mejia, 2010; Ballmer-Weber et al., 2002; Sirvent et al., 2012).

3.3. Structural Analysis of Cuc m 2.0101 and Art v 4.0101

The Cuc m 2 crystal structure was determined at 2.4Å (PDB code 6MBX). Cuc m 2.0101 (155 amino acid residues – including the N-terminal purification tag) crystallized in the C2 space group with three molecules in the asymmetric unit (Table 1). The crystal structure could be also solved in $P3_221$ space group (one molecule in the asymmetric unit); however, in this case, the refinement and model update were not able to reduce the R_{free} value below 35%. Residues corresponding to the purification tag were not visible in the electron density map and absent from the final model. Similar to other allergenic profilins, like Art v 4.0101 (Fig. 2), Cuc m 2.0101 also exhibits a crystal structure with three α-helices and seven-stranded antiparallel β-sheets. Interestingly, Cuc m 2.0101 with the cleaved tag crystallized as well, and the crystals were isomorphous (had very similar unit cell parameters and the same symmetry) to those formed by protein with the purification tag. However, their diffraction quality was poor and their structural analysis was not pursued.

The structure of a new crystal form of Art v 4.0101 was determined at 1.9 Å (PDB code 6B6J). The protein crystallized in the $P3_1$ space group with one molecule in the asymmetric unit. The previously determined Art v 4.0101 structure (PDB code 5EM0) was derived from an orthorhombic form of the protein ($P2_12_12$ space group). Both Art v 4.0101 structures superpose with RMSD value of 0.5 Å over 133 aligned C_α carbons which illustrates very similar conformations of the protein main chain. The overall structure is very similar despite the fact that the protein structures were obtained from crystallization conditions that differed in pH (8.5 for 5EM0 and 6.9 for 6B6J reported here). Both Art v 4.0101 and Cuc m 2.0101 used for crystallization were monomeric in solution and in their crystal forms. On the other hand, however, previous studies have shown that natural and recombinant Art v 4 can form dimers and tetramers that exhibit increased allergenicity and are stabilized by intermolecular disulfide bonds or ionic interactions (Wopfner et al., 2002). Furthermore, Amb a 8, Hev b 8, Zea m 12, and several others are reported to be present in monomeric or dimeric states (Mares-Mejía et al., 2016).

Table 1
Data collection, processing and refinement statistics. Parameters for the highest resolution shells are in parentheses.

Protein/PDB code	Art v 4.0101/6B6J	Cuc m 2.0101/ 6MBX
Diffraction source	APS 22-ID	APS 19-BM
Wavelength (Å)	1.000	0.979
Temperature (K)	100	100
Space group	$P3_1$	C2
a, *b*, *c* (Å)	32.7, 32.7, 81.8	102.6, 59.2, 82.9
α, β, γ (°)	90.00, 90.00, 120.00	90.00, 90.02, 90.00
Resolution range (Å)	40.00-1.90 (1.93-1.90)	40.00-2.40 (2.44-2.40)
Solvent content (%)	46	51
No. of unique reflections	7348 (374)	19,535 (962)
Completeness (%)	95.5 (97.7)	99.2 (100.0)
Redundancy	2.3 (2.2)	2.5 (2.5)
$\langle I/\sigma(I)\rangle$	28.0 (3.5)	31.8 (2.3)
$R_{measure}$	0.065 (0.368)	0.042 (0.607)
$R_{p.i.m.}$	0.040 (0.233)	0.026 (0.368)
Overall *B* factor from Wilson plot (Å^2)	34.4	63.5
$CC_{1/2}$	0.787	0.876
R_{cryst}	0.165 (0.224)	0.231 (0.361)
R_{free}	0.211 (0.291)	0.261 (0.390)
R.m.s. deviations		
Bonds (Å)	0.013	0.010
Angles (°)	1.7	1.3
Average B factors (Å^2)	43.6	94.2
Ramachandran plot		
Favored (%)	97.7	96.6
Allowed (%)	2.3	3.4
Outliers (%)	0.0	0.0

In the case of Hev b 8 and Zea m 12, the dimer formation is also associated with creation of an intermolecular disulfide bond. This covalent bond is formed between Cys12 from two different molecules (Mares-Mejía et al., 2016). Art v 4 has three cysteine residues (Cys13, Cys95 and Cys117) and structural data indicates formation of an internal disulfide (SS) bridge between Cys95 and Cys117 (Figs. 2 and 3). Such a SS bridge is not observed in Cuc m 2, despite having two cysteines (Cys13 and Cys115). Interestingly, two different crystal structures with Cys13 and Cys117 oxidized or reduced were reported for Bet v 2 (Soh et al., 2017).

3.4. Structurally similar profilins to Cuc m 2 and Art v 4

Currently, in the PDB there are structures of profilins originating from 18 different organisms, eight of which are plants. Among the eight plant profilins, seven are reported to be allergens (Fig. 3). These seven allergenic profilins have 14 crystal structures associated with them: Amb a 8 - 5EM1, 5EV0, 5EVE; Ara h 5 - 4ESP; Art v 4 – 5EM0, 6B6J; Bet v 2 - 1CQA, 5NZB, 5NZC; Cuc m 2 - 6MBX; Hev b 8 - 1G5U, 5FEG, 5FDS; Zea m 12 - 5FEF). Allergenic plant profilins share at least 70% identical and 80% similar residues when compared to Cuc m 2.0101 (Fig. 4). This figure also shows that watermelon (Citr l 2.0101) and strawberry (Fra a 4.0101) profilins are the most similar in terms of sequence to Cuc m 2.0101, while Amb a 8.0101 and Art v 4.0101 are the most different when compared to the muskmelon profilin. The high level of primary structure similarity also translates into very similar structures. For

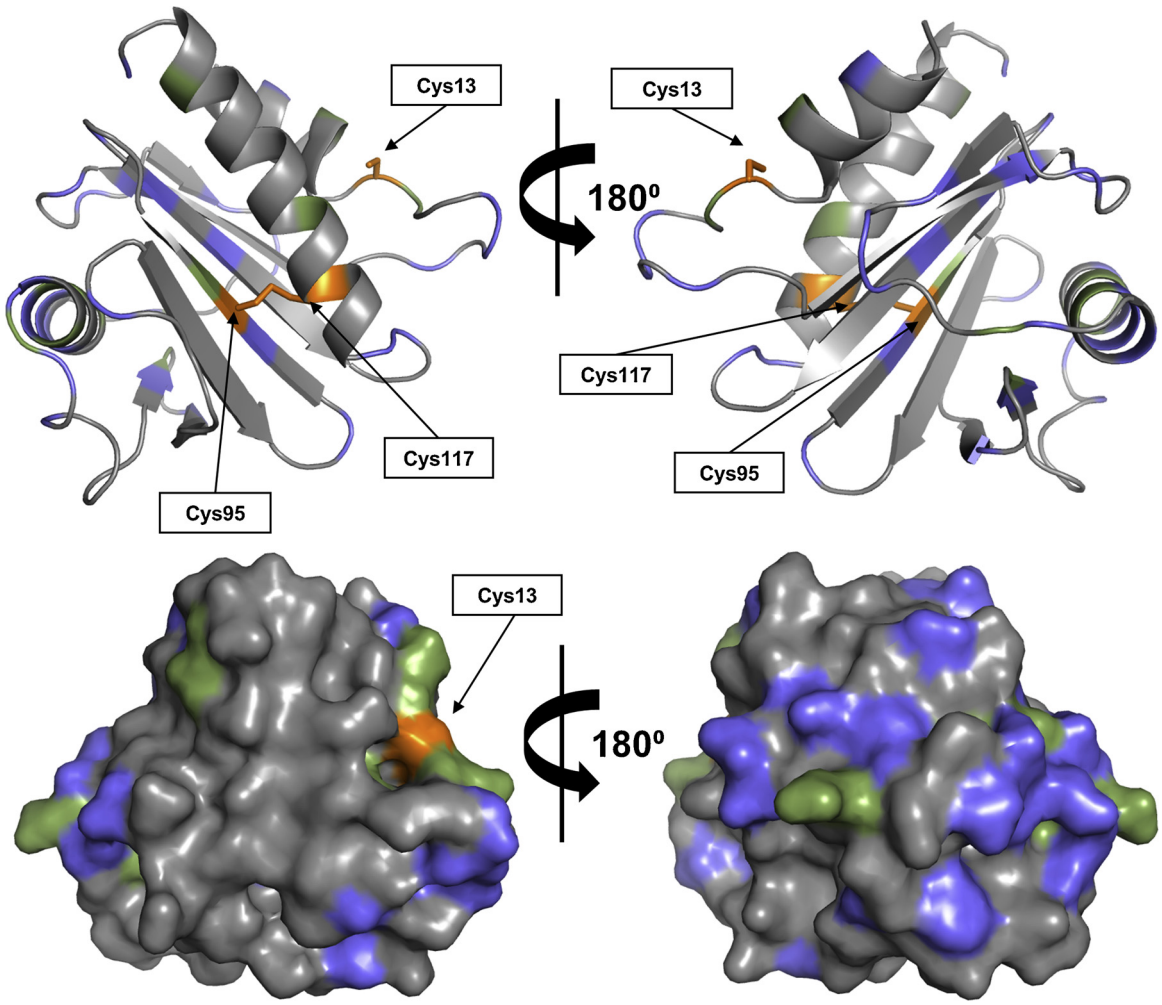


Fig. 2. Structure of Art v 4.0101. Sequence identity and similarity between Art v 4.0101 and Cuc m 2.0101 shown on the Art v 4.0101 structure (PDB code 6B6J). The top images are cartoon representation with the cysteine residues shown in stick representation while the bottom images show surface representation. In Cuc m 2 there is no cysteine that corresponds to Cys95 of Art v 4. This amino acid is replaced by Thr residue and therefore no internal disulfide bridge is present in Cuc m 2. Grey color indicates identical residues, and green color is used to marked residues that are similar, but not identical. Regions displayed in purple display neither identity nor similarity. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

example, superposition of Hev b 8.0101 (PDB code 5FDS) with Cuc m 2 (PDB code 6MBX; 87% sequence similarity and 81% sequence identity) results in an RMSD value of 0.6 Å over 130 aligned C_α atoms. Similarly, superposition of Hev b 8.0101 and Art v 4.0101 results in RMSD value of 0.9 Å over 129 aligned C_α atoms (72% sequence identity and 82% sequence similarity). This shows that even when comparing one profilin allergen from a fruit source (Cuc m 2), one from a pollen source (Art v 4), and one from a latex source (Hev b 8), they are still extremely similar in terms of their primary, secondary and tertiary structures. It also explains why profilins are considered panallergens.

All but one (Tyr p 36.0101) allergenic profilins officially registered by the World Health Organization and International Union of Immunological Societies (WHO/IUIS) Allergen Nomenclature Subcommittee originate from plants. Tyr p 36.0101 originates from storage mite (*Tyrophagus putrescentiae*) and shares only 33% and 34% of identical residues with Cuc m 2.0101 and Art v 4.0101, respectively. It is also worth mentioning that, human profilins, while similar in structure to Cuc m 2.0101 and Art v 4.0101 as indicated by similarity searches performed using DALI and PDBeFOLD, share less than 25% sequence identity with these plant profilins (Chruszcz et al., 2018).

Among over 60 allergenic proteins analyzed (Supplementary Table S1 and Figure S3), only thirteen of them (Aca f 2.0101, Ama r 2.0101, Ama v 2.01, Ama v 2.02, Amb a 8.0101, Art v 4.0101, Art v 4.0201, Hel a 2.0101, Koc s 2.0101, Pro *j* 2.0101, Sal k 4.0101, Sal k 4.0201 and Vit v 4) have conserved Cys95 and Cys117 (based on Art v 4.0101 residue positions). We predict that all these profilins most likely will have an intramolecular disulfide bridge formed by the Cys 95 and Cys117. Interestingly, all of the analyzed allergenic profilins have conserved Cys13 and Cys117 (Supplementary Figure S3). We also presume that all these allergenic profilins can possibly form dimeric assemblies that were described for Hev b 8, which involve formation of a disulfide bridge between two Cys13 residues from different molecules (Mares-Mejía et al., 2016)

Notably, comparison of plant profilin sequences and structures highlighted a region that connects α-helix 1 with β-strand 1 (Fig. 3 and Supplementary Figure S3). This solvent exposed loop region has a short 1–3 amino acid insertion in 18 of the plant profilins analyzed here including Bet v 2.0101, Art v 4.0101, and Amb a 8.0101. In the case of Amb a 8.0101, it was shown that Gly17 present in this loop changes conformation upon binding of polyproline (Offermann et al., 2016).

3.5. IgE epitopes

Profilins represent one of the most numerous allergen families that are associated with many clinical syndromes. Some of them, (Cit s 2, Citr l 2.0101, Cuc m 2 and Gly m 3) being major allergens, have instigated many researchers to investigate IgE epitopes associated with

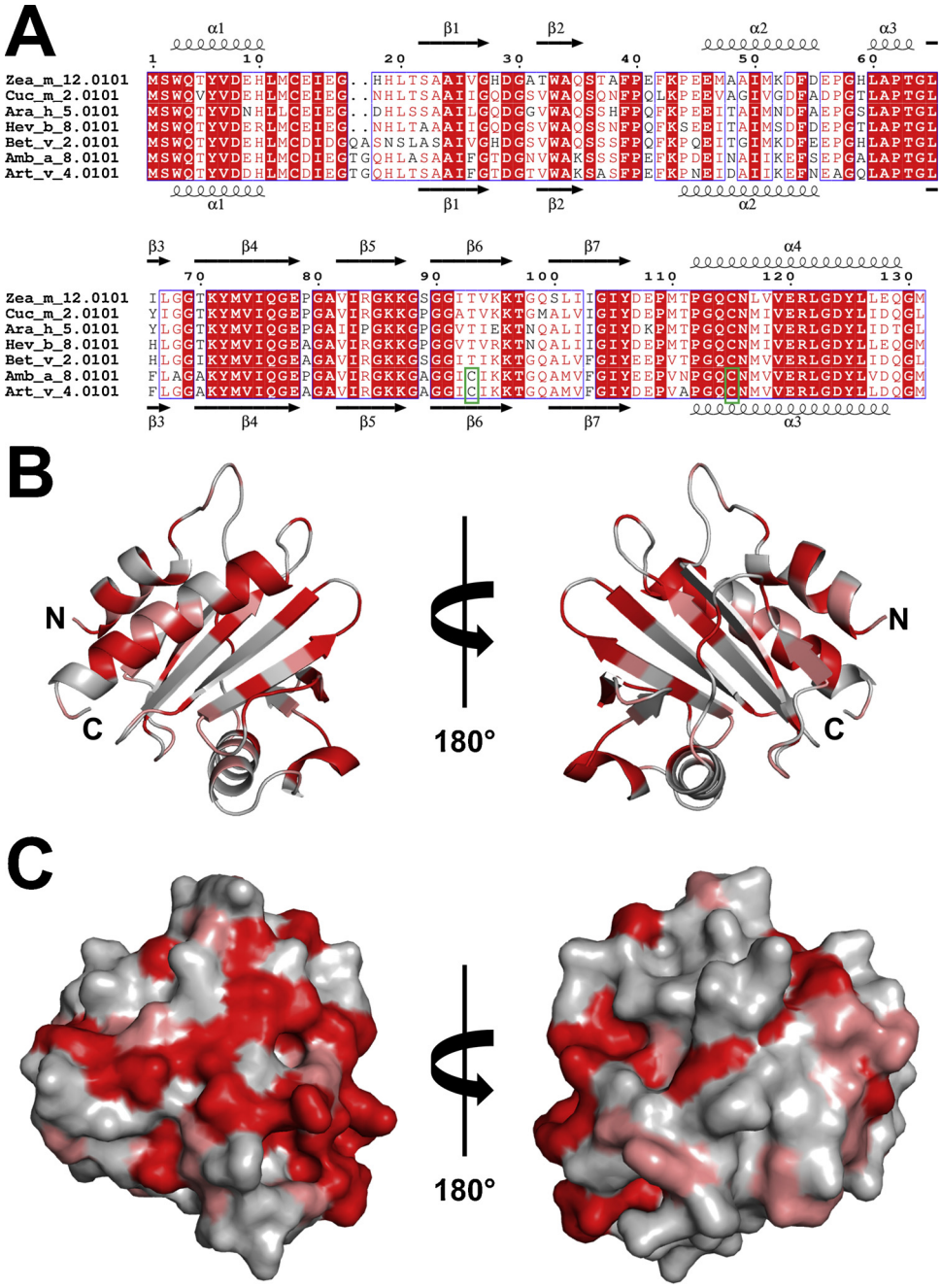


Fig. 3. Sequence conservation. A) Sequence comparison of all allergenic profilins reported in the PDB. Secondary structure elements are marked for Zea m 12.0101 (PDB code 5FEF) and Art v 4.0101 (PDB code 6B6J), respectively. Cysteine residues forming disulfide bridges in Amb a 8.0101 and Art v 4.0101 are marked using green boxes. Initiator methionine residues that are not present in mature forms of the profilins are included. **B)** Conservation of profilin sequences mapped on structure of Cuc m 2.0101 presented in cartoon representation. Residues that are conserved in all analyzed allergenic profilins herein (Supplementary Table S1 and Figure S3) are marked in red. Residues from profilins that have their structures determined and are completely conserved are shown in red and salmon. It implicates that salmon corresponds to residues that are conserved for profilins shown in this figure, but which are not completely conserved among all analyzed profilins presented in Supplementary Fig. 3S. N- and C-terminal residues are marked with corresponding letter. **C)** Conservation of sequences mapped on molecular surface of Cuc m 2.0101. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

these proteins (Cases et al., 2010; Jeong et al., 2013; López-Torrejón et al., 2005a, b). In a study to characterize IgE epitopes of Cuc m 2 by López-Torrejón and colleagues, four IgE epitopes of Cuc m 2 were described (Fig. 5) (López-Torrejón et al., 2007). Two of these epitopes were defined as major (E1 – residues 66–75 and 81–93; E2 – residues 95–99 and 122–131), because they showed the strongest IgE-binding capacity, and two additional epitopes (E3 – residues 2–10; E4 - residues 35–45) showed low IgE-binding capacity. Later, another sequence was identified - $S_2W_3A_5Y_6D_9H_{10}T_{111}P_{112}G_{113}Q1_{14}N_{116}M_{117}R_{121}L_{122}$ as forming a Cuc m 2 mimotope (Fig. 5) (Tordesillas et al., 2010). These studies were used to design Cuc m 2 mutants with reduced IgE-binding capacities (Tordesillas et al., 2011). Interestingly, the mimotope overlaps mainly with epitope E3 (Fig. 5) and a significant area of the mimotope does not correspond to any of the identified IgE epitopes. Our structural studies allow mapping of these epitopes and the mimotope on the experimental Cuc m 2 model and demonstrate that a potential formation of a dimeric structure would interfere with IgE biding to epitopes E2 and E3, as well as to the mimotope (Fig. 6).

4. Discussion

Profilins are small eukaryotic proteins which are considered panallergens and play a role in the regulation of various cellular processes (Carlsson et al., 1977; Tanaka and Shibata, 1985). They share high sequence similarity and identity as well as high 3D structure conservation (Mari, 2001). This holds true even among distantly related members and profilins originating from different sources (Chruszcz et al., 2018; Offermann et al., 2016). Furthermore, high sequence homology between profilins plays a major role in IgE cross-reactivity with other allergens (Hauser et al., 2010). Although profilins are mainly regarded as minor allergens, they are important in cross-reactive reactions. Cuc m 2, however, is considered a major allergen, which

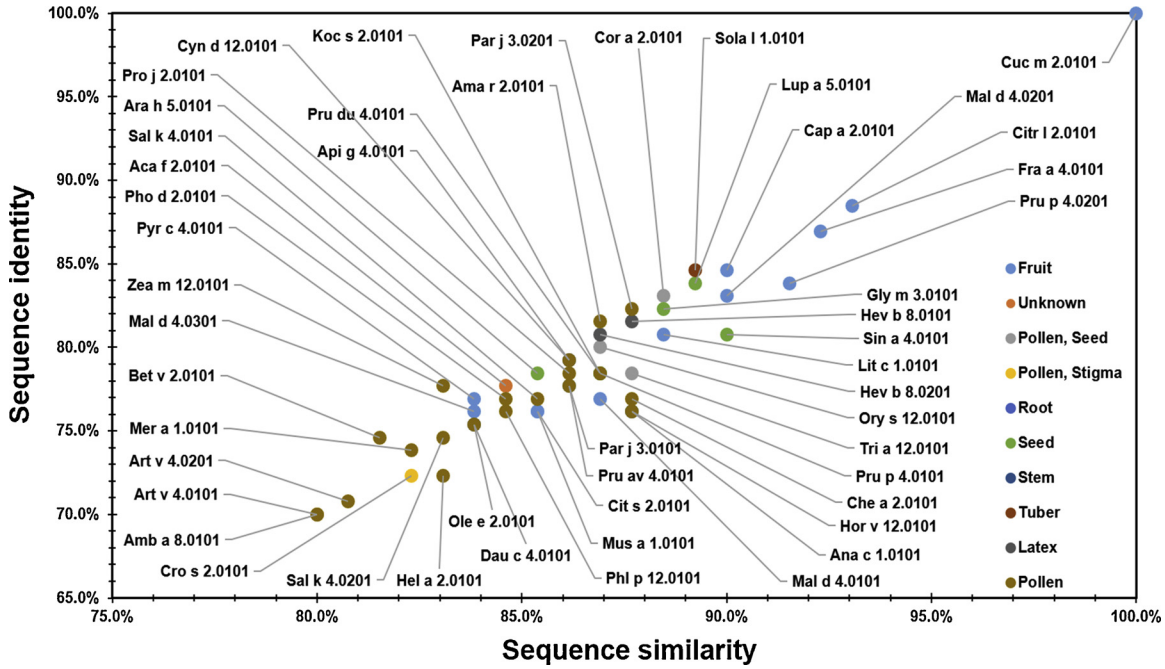


Fig. 4. Sequence similarities and identities between Cuc m 2.0101 and various plant profilins that are officially registered as allergens by the World Health Organization and International Union of Immunological Societies (WHO/IUIS) Allergen Nomenclature Sub-committee (www. allergen.org). Proteins are colored according to allergen source. Sequence comparison was performed using Clustal Omega (Sievers et al., 2011) and SIAS (http://imed.med.ucm.es/Tools/sias.html).

induces an IgE response in more than 80% of patients allergic to melon (López-Torrejón et al., 2005a). Our DSF results indicate a significantly higher thermal stability of Cuc m 2 in solutions with low pH (Fig. 1) when this allergen is compared with profilins such as Amb a 8, Art v 4 and Bet v 2 that originate from pollens. Interestingly, the maximum thermal stabilities of the aforementioned pollen profilins are not different from the stabilities observed for Cuc m 2. As Cuc m 2 is a food allergen, it is possible that its increased resistance to low pH conditions combined with the fact that muskmelons are ingested in raw form are the reasons that renders this profilin a major allergen.

Results of the maximum melting temperature predictions for the allergenic profilins are consistent with our experimental data, and indeed they show that almost all profilins are thermally labile (Supplementary Table S1). Predicted, relatively high melting temperature (> 65 °C) for Vit v 4 will require an experimental confirmation. The results of thermal stability predictions also show that there may be a subgroup of allergenic profilins that are more stable (melting in the 55–65 °C range). This subgroup includes not only Cuc m 2.0101, but also Citr l 2.0101 and Gly m 3.0101 that are major allergens originating from watermelon and soybean, respectively. Interestingly, both DSF data and melting temperature predictions indicate that the presence of internal disulfide bridge is not critical for profilin stability. This is supported by our experimental results showing that Art v 4.0101 and Cuc m 2.0101 have a very similar maximum thermal stability, despite the fact that T_m Predictor classifies the profilin from mugwort as less stable.

Structural studies reveal that Art v 4.0101 and Cuc m 2.0101 have very similar main chain conformations, which reflects 80% sequence similarity and 70% sequence identity. These two experimental structures, together with structures reported for Amb a 8, Ara h 5, Bet v 2, Hev b 8 and Zea m 12, provide insights into molecular properties of allergenic profilins. The intramolecular disulfide bridge is not critical for maintaining an overall fold for the profilins, and that there are no significant changes in the structure of these molecules in the presence or absence of the disulfide bridge. Analysis of the recent profilin structures provides a more elaborate picture on possible quaternary assemblies that may be present in this group of proteins. In the case of recombinant Hev b 8, the formation of the dimeric assembly resulted in increase of IgE-mediated degranulation of rat basophilic leukemia cells (Mares-Mejía et al., 2016). Comparison of monomeric structure of Cuc m 2 (presented here) and the dimeric structure of Hev b 8 highlights some implications of the dimer formation (Fig. 6). For example, it is clear that the dimer formation would partially block residues forming epitopes E2 and E3 that were determined by López-Torrejón et al. (Fig. 6A), and completely block the mimotope found by Tordesillas et al. (Fig. 6D). However, the studies of Cuc m 2 epitopes and the mimotope show that there are IgE antibodies that can recognize the residues that are part of a putative Cuc m 2.0101 dimerization interface. As alluded to by the Hev b 8 example above, the oligomerization state of a profilin seems to play a major role in IgE response. This prompts a question about the oligomerization state of natural profilins and their quaternary form(s) to which the human body is exposed to, and what differences in IgE response do these states elicit. It is also worth to note that the residues forming the 674 $Å^2$ dimer interface in Hev b 8 are almost completely conserved in Cuc m 2.0101 (Fig. 6 D), and this is also true for Art v 4.0101.

The conserved Cys13 residue that is important for formation of the dimeric assembly of recombinant Hev b 8 may change its oxidation state in various profilins. The oxidation state of this residue was shown to have a pronounced effect on Bet v 2 resistance for cathepsin S digestion and therefore the endolysosomal stability of this profilin, which impacts the allergenicity (Soh et al., 2017). Our previous (Offermann et al., 2016) and current studies of Art v 4.0101 structures reveal that Cys12 in this profilin may also change its oxidation state. In the case of the first reported structure (PDB code 5EM0), the Cys12 is reduced as it forms a covalent intermediate with β-ME which was used during protein purification. In the case of the structure described in this manuscript (PDB code 6B6J), the thiol is in the reduced form. The presence or absence of the covalent modification of cysteine in our Art v 4.0101 crystal structure most likely is related to the different crystallization conditions that were used (pH 8.5 for 5EM0 and pH 6.9 for 6B6J). This is consistent with the fact that the formation of disulfide bridges is usually favored in alkaline solutions (Buchanan et al., 2013). Therefore, it is clear that conditions used during purification and/or crystallization

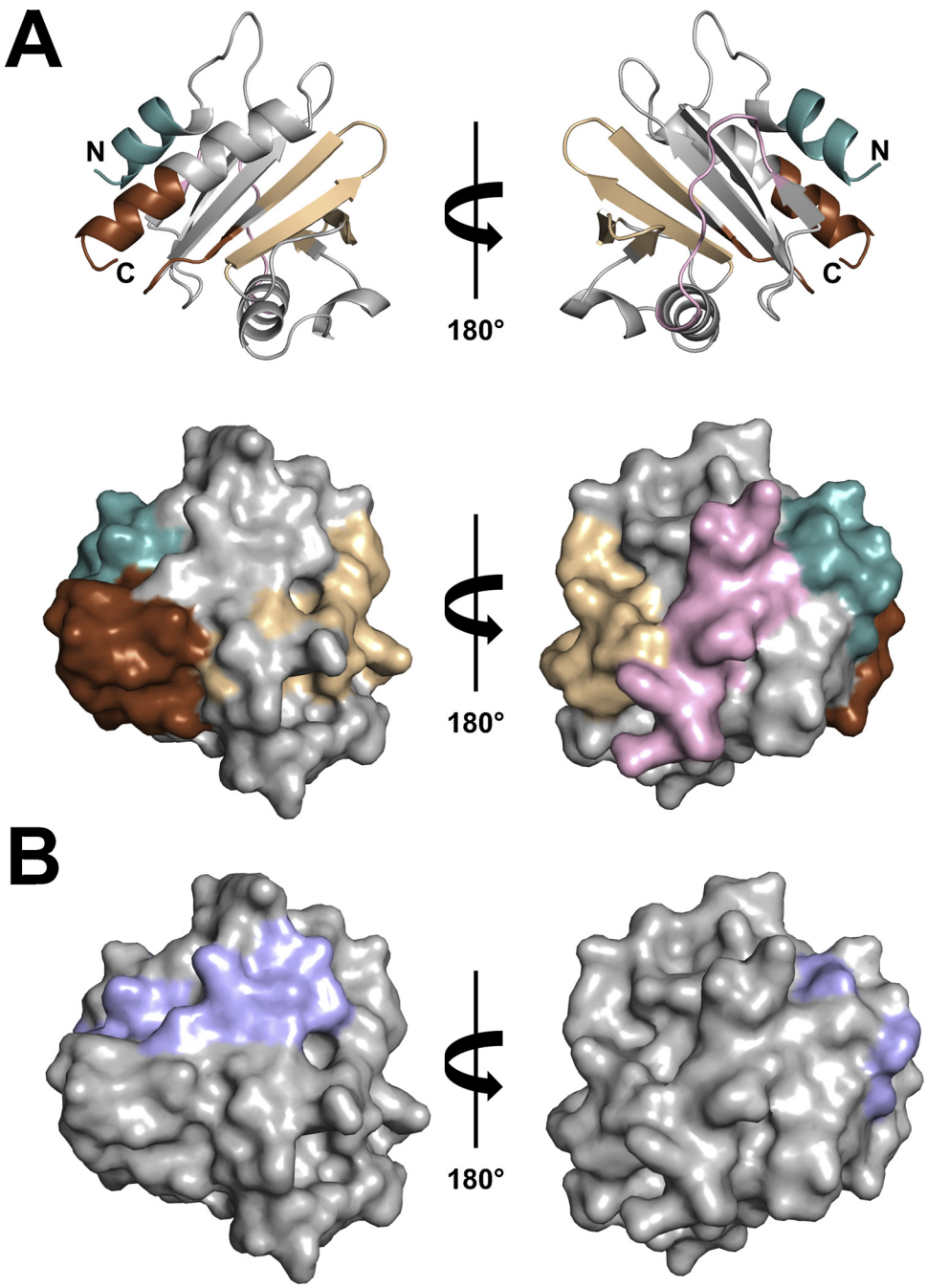


Fig. 5. Identified epitopes and mimotope. A) Cartoon (top) and surface (bottom) representations of Cuc m 2.0101 with mapped IgE epitopes as determined by Lopez-Torrejon et al. (2007). N- and C-terminal residues are marked with corresponding letters. Epitope 1, consisting of residues 66–75 and 81–93, in wheat, epitope 2 (residues 95–99 and 122–131) in brown, epitope 3 (residues 2–10) in light teal, and epitope 4 (residues 35–45) in light pink. **B)** Mimotope determined by Tordesillas et al. and consisting of residues 2–3, 5–6, 9–10, 111–114, 116–117, as well as 121–122 is shown in light blue (Tordesillas et al., 2010). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

will affect the final form of the protein and/or protein structure. It also suggests that the presence of the dimeric structure of recombinant Hev b 8 and other oligomeric assemblies reported for profilins may be an artefact of protein production (Mares-Mejía et al., 2016). Quite often profilins are purified in denaturing conditions and refolded, which may result in formation of the intermolecular disulfide bridges. For example, this purification protocol was followed during isolation of the natural Art v 4 from mugwort pollen (Wopfner et al., 2002).

We concur that the oligomerization of profilins is very intriguing and may have various biological implications. For example, it was suggested that human, birch and yeast profilins can be ligands for themselves (Mittermann et al., 1998). Moreover, profilins are involved in many molecular interactions, and some of these interactions involve the binding to proline-rich regions of various proteins (Witke, 2004). Comparison of the dimeric structure of Hev b 8 with model of Amb a 8 in complex with polyproline (Fig. 6B) implicates that the dimerization of a profilin will also limit the ability of the protein to interact with the proline-rich peptides because the oligomerization interface and the peptide binding sites partially overlap. Conversely, the presence of both monomeric and dimeric forms of a profilin could be a meticulous solution that can be used in cells to modify affinity of the protein to various binding partners. Currently, it is also not known how the structure of the tetrameric assemblies reported for Art v 4 looks like (Wopfner et al., 2002). It is possible that the tetramer may be considered as a dimer of covalently linked dimers, like it was determined for Hev b 8.

Comparison of the allergenic plant profilins in terms of their amino acid similarity and identity clearly illustrates the molecular basis of cross-reactivity (Fig. 4 and Supplementary Figure S3). Moreover, the likelihood of the cross-reactivity is high as predicted using A-RISC approach (Chruszcz et al., 2018). Using the A-RISC indexes, (Fig. 4) the likelihood of cross-reactivity between Cuc m 2.0101 is the greatest for

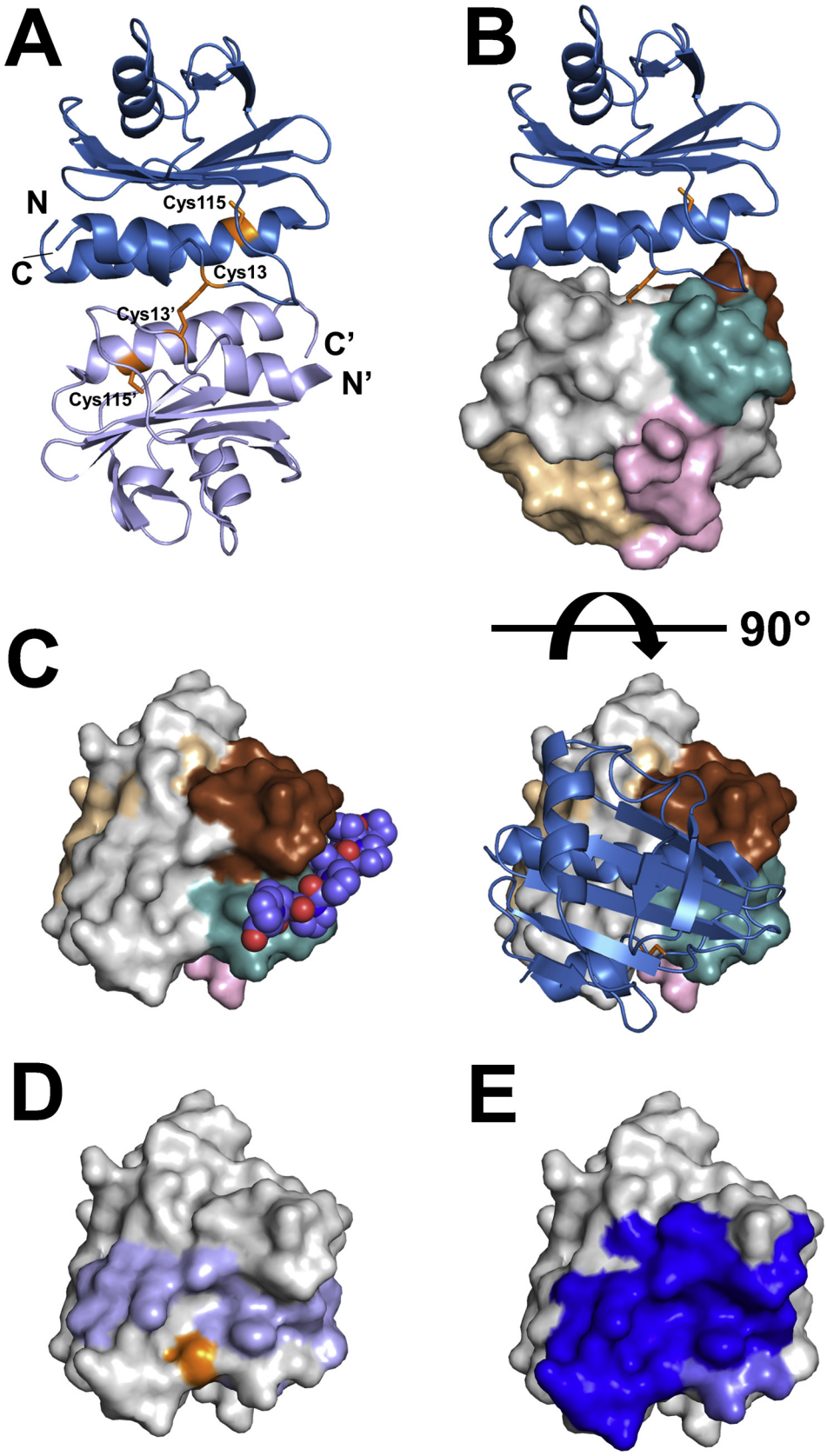


Fig. 6. Dimeric and monomeric profilins. A) Dimeric structure of Hev b 8.0101. **B)** Cuc m 2.0101 (surface representations) superposed on one of the molecules forming Hev b 8.0101 dimer. Epitope 1 is marked in wheat, epitope 2 in brown, epitope 3 in light teal and epitope 4 in light pink. **C)** Cuc m 2.0101 with modeled polyproline peptide (shown in sphere representation). The peptide binding site is formed by residues forming epitopes 2 and 3 identified by Lopez-Torrejon et al. (2007). **D)** Mimotope determined by Tordesillas et al. (2010) (light blue). Cys residues that may be responsible for formation of a dimeric structure is marked in orange. **E)** Putative Cuc m 2.0101 dimer interface based on structure homology of Hev b 8 (PDB code 5FEG). Residues marked in blue are completely conserved, while residues shown in purple are not conserved. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

the following allergens: Citr l 2.0101 (watermelon), Fra a 4.0101 (strawberry), Pru p 4.0101 (peach), Cap a 2.0101 (bell pepper), Lup a 5.0101 (white lupine) and Sola l 1.0101 (tomato). In addition, Fig. 4 suggests that the likelihood of a cross-reactivity reaction between Cuc m 2 and weed profilins like Amb a 8 and Art v 4 is somewhat less likely, but still falling into the highest risk category according to the A-RISC model. This is also well illustrated in Fig. 2 as the solvent exposed and conserved residues between Cuc m 2.0101 and Art v 4.0101 form both

continuous and discontinuous regions which may be possibly cross-reactive IgE epitopes. This suggests that in the case of profilins, despite their similarity, one should expect presence of both cross-reactive and species-specific antibodies.

In summary, we have shown that structures of profilins reveal various molecular details that are related to the biological functions of this group of proteins, as well as to their allergenic properties. In addition, it should be stressed that the conditions used during protein purification may impact some molecular properties of proteins (both natural and recombinant), and that these properties may have a significant impact on down-stream experiments performed with the proteins.

Declaration of Competing Interest

Nothing declared.

Acknowledgments

Structural results shown in this report are derived from data collected at Southeast Regional Collaborative Access Team (SER-CAT; 22 ID and 19 BM) beamlines at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at www.ser-cat.org/members.html. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract Nos. DE-AC02-06CH11357 and W-31-109-Eng-38. This work was partially supported by ASPIRE III grant from the Office of the Vice President of Research at the University of South Carolina. MC and ABK were partially supported by grant R01AI077653 from National Institute of Allergy and Infectious Diseases.

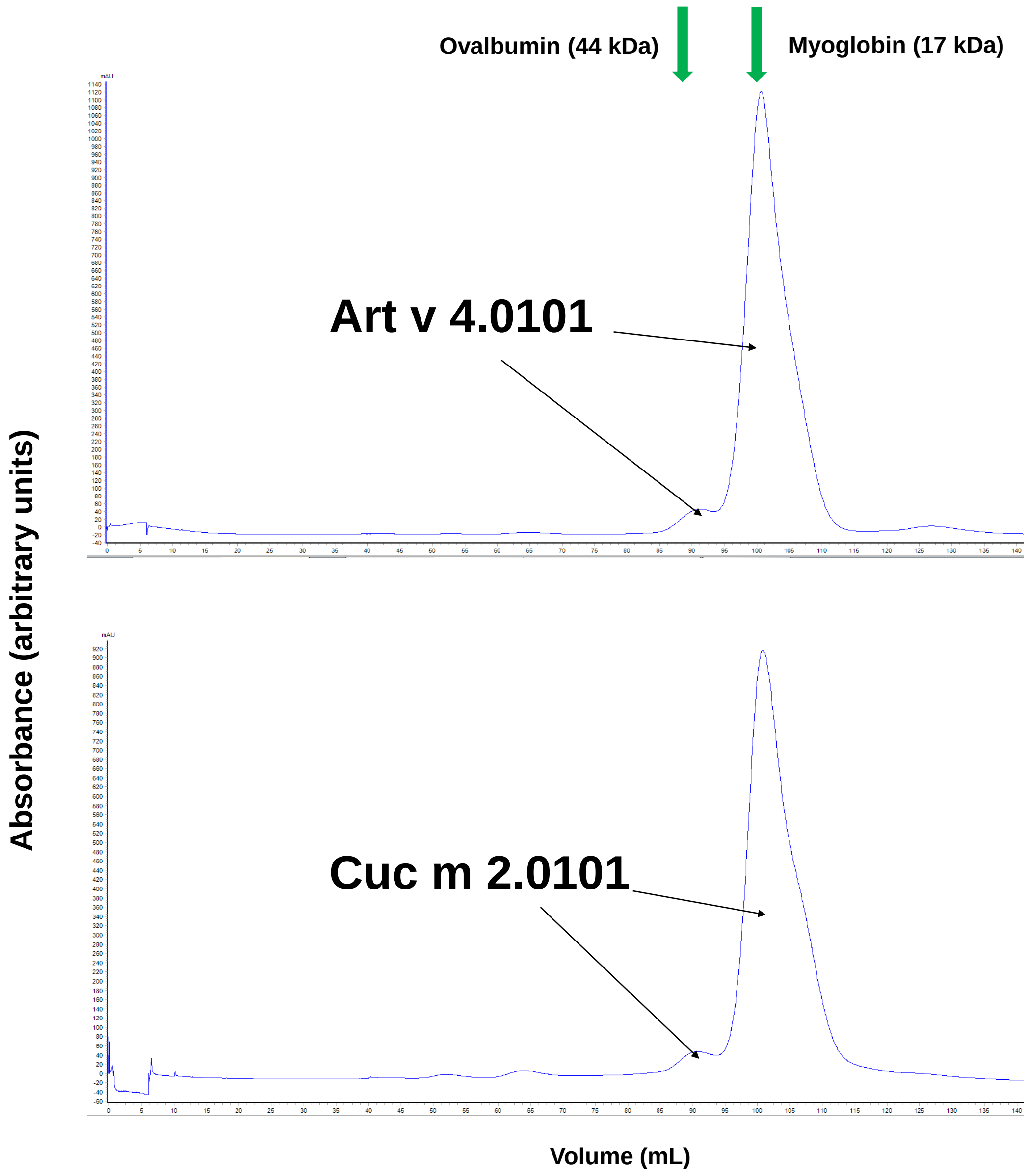
Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.molimm.2019.07.004.

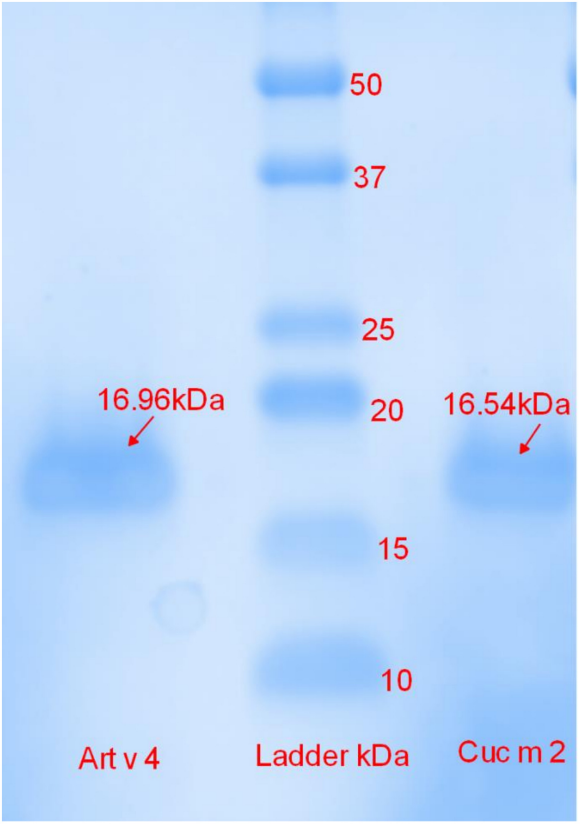
References

Amnuaycheewa, P., de Mejia, E.G., 2010. Purification, characterisation, and quantification of the soy allergen profilin (Gly m 3) in soy products. Food Chem. 119, 1671–1680.
Asero, R., Mistrello, G., Amato, S., 2011. The nature of melon allergy in ragweed-allergic subjects: a study of 1000 patients. Allergy Asthma Proc. 32, 64–67.
Asero, R., Mistrello, G., Roncarolo, D., Amato, S., Zanoni, D., Barocci, F., Caldironi, G., 2003. Detection of clinical markers of sensitization to profilin in patients allergic to plant-derived foods. J. Allergy Clin. Immunol. 112, 427–432.
Asero, R., Wopfner, N., Gruber, P., Gadermaier, G., Ferreira, F., 2006. Artemisia and ambrosia hypersensitivity: co-sensitization or co-recognition? Clin. Exp. Allergy 36, 658–665.
Ballmer-Weber, B., Hoffmann, A., Wüthrich, B., Lüttkopf, D., Pompei, C., Wangorsch, A., Kästner, M., Vieths, S., 2002. Influence of food processing on the allergenicity of celery: DBPCFC with celery spice and cooked celery in patients with celery allergy. Allergy 57, 228–235.
Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The protein data bank. Nucleic Acids Res. 28, 235–242.
Booth, W.T., Schlachter, C.R., Pote, S., Ussin, N., Mank, N.J., Klapper, V., Offermann, L.R., Tang, C., Hurlburt, B.K., Chruszcz, M., 2018. Impact of an N-terminal polyhistidine tag on protein thermal stability. ACS Omega 3, 760–768.
Buchanan, A., Clementel, V., Woods, R., Harn, N., Bowen, M.A., Mo, W., Popovic, B., Bishop, S.M., Dall'Acqua, W., Minter, R., 2013. Engineering a therapeutic IgG molecule to address cysteinylation, aggregation and enhance thermal stability and expression. mAbs 5, 255–262.
Burger, Y., Paris, H.S., Cohen, R., Katzir, N., Tadmor, Y., Lewinsohn, E., Schaffer, A.A., 2010. Genetic diversity of Cucumis melo. Hortic. Rev. 36, 165–198.
Carlsson, L., Nyström, L.-E., Sundkvist, I., Markey, F., Lindberg, U., 1977. Actin polymerizability is influenced by profilin, a low molecular weight protein in non-muscle cells. J. Mol. Biol. 115, 465–483.
Cases, B., Pastor-Vargas, C., Dones, F.G., Perez-Gordo, M., Maroto, A.S., De Las Heras, M., Vivanco, F., Cuesta-Herranz, J., 2010. Watermelon profilin: characterization of a major allergen as a model for plant-derived food profilins. Int. Arch. Allergy Immunol. 153, 215–222.
Chapman, M.D., 2008. Allergen nomenclature. Clin. Allergy Immunol. 18, 51–64.
Chruszcz, M., Kapingidza, A.B., Dolamore, C., Kowal, K., 2018. A robust method for the estimation and visualization of IgE cross-reactivity likelihood between allergens belonging to the same protein family. PLoS One 13, e0208276.
Cowtan, K., 2006. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr. Sect. D: Biol. Crystallogr. 62, 1002–1011.
Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall III, W.B., Snoeyink, J., Richardson, J.S., 2007. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. 35, W375–W383.
DeLano, W.L., 2002. Pymol: an open-source molecular graphics tool. CCP4 Newsletter On Protein Crystallography 40, 82–92.
Ebner, C., Jensen-Jarolim, E., Leitner, A., Breiteneder, H., 1998. Characterization of allergens in plant-derived spices: apiaceae spices, pepper (Piperaceae), and paprika (bell peppers, Solanaceae). Allergy 53, 52–54.
Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132.
Gadermaier, G., Dedic, A., Obermeyer, G., Frank, S., Himly, M., Ferreira, F., 2004. Biology of weed pollen allergens. Curr. Allergy Asthma Rep. 4, 391–400.
Gadermaier, G., Hauser, M., Ferreira, F., 2014. Allergens of weed pollen: an overview on recombinant and natural molecules. Methods 66, 55–66.
Gadermaier, G., Wopfner, N., Wallner, M., Egger, M., Didierlaurent, A., Regl, G., Aberger, F., Lang, R., Ferreira, F., Hawranek, T., 2008. Array-based profiling of ragweed and mugwort pollen allergens. Allergy 63, 1543–1549.
Ganglberger, E., Radauer, C., Wagner, S., Ríordáin, G.Ó., Beezhold, D.H., Brehler, R., Niggemann, B., Scheiner, O., Jensen-Jarolim, E., Breiteneder, H., 2001. Hev b 8, the Hevea brasiliensis latex profilin, is a cross-reactive allergen of latex, plant foods and pollen. Int. Arch. Allergy Immunol. 125, 216–227.
Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., Bairoch, A., 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 31, 3784–3788.
Gouet, P., Robert, X., Courcelle, E., 2003. ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res. 31, 3320–3323.
Hansson, A., Skoglund, G., Lassing, I., Lindberg, U., Ingelman-Sundberg, M., 1988. Protein kinase C-dependent phosphorylation of profilin is specifically stimulated by phosphatidylinositol bisphosphate (PIP2). Biochem. Biophys. Res. Commun. 150, 526–531.
Hauser, M., Roulias, A., Ferreira, F., Egger, M., 2010. Panallergens and their impact on the allergic patient. Allergy Asthma Clin. Immunol. 6, 1–14.
Holm, L., Kääriäinen, S., Wilton, C., Plewczynski, D., 2006. Using Dali for structural comparison of proteins. Curr. Protoc. Bioinformatics 14, 1–24.
Jeong, K.-H., Choi, M.-S., Lee, S.-K., Seo, M.-J., Hwang, T.-Y., Yun, H.-T., Kim, H.-S., Kim, J.-T., Kwon, Y.-U., Kim, Y.-H., 2013. Development of low Gly m Bd 30K (P34) allergen breeding lines using molecular marker in soybean [Glycine max (L.) Merr.]. Plant Breed. Biotechnol. 1, 298–306.
Joosten, R.P., Te Beek, T.A., Krieger, E., Hekkelman, M.L., Hooft, R.W., Schneider, R., Sander, C., Vriend, G., 2010. A series of PDB related databases for everyday needs. Nucleic Acids Res. 39, D411–D419.
Krissinel, E., Henrick, K., 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. Sect. D 60, 2256–2268.
Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797.
Ku, T., Lu, P., Chan, C., Wang, T., Lai, S., Lyu, P., Hsiao, N., 2009. Predicting melting temperature directly from protein sequences. Comput. Biol. Chem. 33, 445–450.
Lassing, I., Lindberg, U., 1985. Specific interaction between phosphatidylinositol 4, 5-bisphosphate and profilactin. Nature 314, 472–474.
López-Torrejón, G., Díaz-Perales, A., Rodríguez, J., Sánchez-Monge, R., Crespo, J.F., Salcedo, G., Pacios, L.F., 2007. An experimental and modeling-based approach to locate IgE epitopes of plant profilin allergens. J. Allergy Clin. Immunol. 119, 1481–1488.
López-Torrejón, G., Crespo, J., Sánchez-Monge, R., Sánchez-Jiménez, M., Alvarez, J., Rodriguez, J., Salcedo, G., 2005a. Allergenic reactivity of the melon profilin Cuc m 2 and its identification as major allergen. Clin. Exp. Allergy 35, 1065–1072.
López-Torrejón, G., Ibanez, M., Ahrazem, O., Sánchez-Monge, R., Sastre, J., Lombardero, M., Barber, D., Salcedo, G., 2005b. Isolation, cloning and allergenic reactivity of natural profilin Cit s 2, a major orange allergen. Allergy 60, 1424–1429.
Mares-Mejía, I., Martínez-Caballero, S., Garay-Canales, C., Cano-Sánchez, P., Torres-Larios, A., Lara-González, S., Ortega, E., Rodríguez-Romero, A., 2016. Structural insights into the IgE mediated responses induced by the allergens Hev b 8 and Zea m 12 in their dimeric forms. Sci. Rep. 6, 1–13.
Mari, A., 2001. Multiple pollen sensitization: a molecular approach to the diagnosis. Int. Arch. Allergy Immunol. 125, 57–65.
Mari, A., Riccioli, D., 2004. The Allergome web site-a database of allergenic molecules. Aim, structure, and data of a web-based resource. J. Allergy Clin. Immunol. 113, S301.
Minor, W., Cymborowski, M., Otwinowski, Z., Chruszcz, M., 2006. HKL-3000: the integration of data reduction and structure solution–from diffraction images to an initial model in minutes. Acta Crystallogr. D Biol. Crystallogr. 62, 859–866.
Mittermann, I., Fetrow, J.S., Schaak, D.L., Almo, S.C., Kraft, D., Heberle-Bors, E., Valenta, R., 1998. Oligomerization of profilins from birch, man and yeast. Profilin, a ligand for itself? Sex. Plant Reprod. 11, 183–191.
Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., Vagin, A.A., 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. 67, 355–367.
Offermann, L.R., Schlachter, C.R., Perdue, M.L., Majorek, K.A., He, J.Z., Booth, W.T., Garrett, J., Kowal, K., Chruszcz, M., 2016. Structural, functional, and immunological

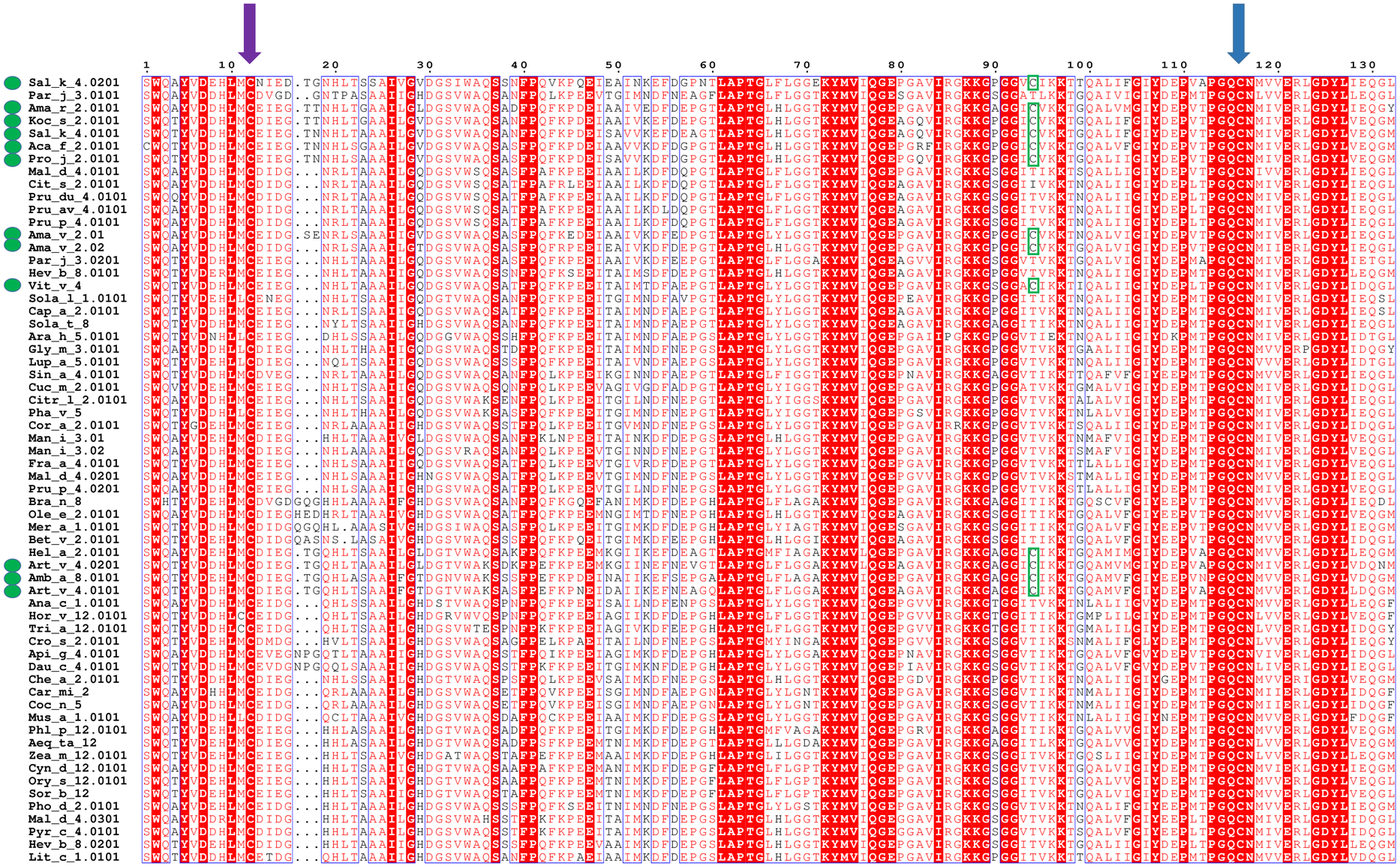
characterization of profilin panallergens Amb a 8, Art v 4, and Bet v 2. J. Biol. Chem. 291, 15447–15459.
Otwinowski, Z., Minor, W., 1997. Processing of X-ray diffraction data collected in oscillation mode. Meth. Enzymol. 276, 307–326.
Radauer, C., Bublin, M., Wagner, S., Mari, A., Breiteneder, H., 2008. Allergens are distributed into few protein families and possess a restricted number of biochemical functions. J. Allergy Clin. Immunol. 121, 847–852.
Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539.
Sirvent, S., Palomares, O., Cuesta-Herranz, J., Villalba, M., Rodríguez, R., 2012. Analysis of the structural and immunological stability of 2S albumin, nonspecific lipid transfer protein, and profilin allergens from mustard seeds. J. Agric. Food Chem. 60, 6011–6018.
Soh, W.T., Briza, P., Dall, E., Asam, C., Schubert, M., Huber, S., Aglas, L., Bohle, B., Ferreira, F., Brandstetter, H., 2017. Two distinct conformations in Bet v 2 determine its proteolytic resistance to cathepsin S. Int. J. Mol. Sci. 18, 2156–2170.
Tanaka, M., Shibata, H., 1985. Poly (l-proline)-binding proteins from chick embryos are a profilin and a profilactin. Eur. J. Biochem. 151, 291–297.
Tordesillas, L., Gamboa, P., Sanz, M.L., Palacin, A., Gomez-Casado, C., Cuesta-Herranz, J., Pacios, L.F., Salcedo, G., Diaz-Perales, A., 2011. A mutant of the major melon allergen, Cuc m 2, with reduced IgE binding capacity is a good candidate for specific immunotherapy. Mol. Immunol. 49, 504–511.
Tordesillas, L., Pacios, L.F., Palacin, A., Cuesta-Herranz, J., Madero, M., Diaz-Perales, A., 2010. Characterization of IgE epitopes of Cuc m 2, the major melon allergen, and their role in cross-reactivity with pollen profilins. Clin. Exp. Allergy 40, 174–181.
Vagin, A., Teplyakov, A., 1997. MOLREP: an automated program for molecular replacement. J. Appl. Crystallogr. 30, 1022–1025.
van Ree, R., Fernández-Rivas, M., Cuevas, M., van Wijngaarden, M., Aalberse, R.C., 1995. Pollen-related allergy to peach and apple: an important role for profilin. J. Allergy Clin. Immunol. 95, 726–734.
Wiedemann, C., Bellstedt, P., Görlach, M., 2013. CAPITO—a web server-based analysis and plotting tool for circular dichroism data. Bioinformatics 29, 1750–1757.
Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., 2011. Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67, 235–242.
Witke, W., 2004. The role of profilin complexes in cell motility and other cellular processes. Trends Cell Biol. 14, 461–469.
Wopfner, N., Willeroider, M., Hebenstreit, D., van Ree, R., Aalbers, M., Briza, P., Thalhamer, J., Ebner, C., Richter, K., Ferreira, F., 2002. Molecular and immunological characterization of profilin from mugwort pollen. Biol. Chem. 383, 1779–1789.



Supplementary Figure S1. Gel filtration results for Art v 4.0101 (top) and Cuc m 2.0101 (bottom). Green arrows mark elution volumes corresponding to standards (Myoglobin and Ovalbumin). Presence of the profilins in the marked peaks was confirmed using SDS-PAGE (data not shown). The gel filtration results indicate that both monomeric and dimeric forms of Art v 4.0101 and Cuc m 2.0101 are generated during production of these recombinant proteins.



Supplementary Figure S2. SDS-PAGE showing purity of Art v 4.0101 and Cuc m 2.0101 samples. Protein versions with His-tag arepresented.



Supplementary Figure S3. Sequence alignment of all reported allergenic plant profilins. The conserved Cys residues are marked with arrows. Cys residues that may participate in dimer formation are marked with the purple arrow. Cys residues that are involved in formation of intramolecular disulfide bridge are using the blue arrow and green boxes. Names of the profilins in which the formation of the intramolecular disulfide bridge is possible are marked with green circles.