Structural characterization of the Pet c 1.0201 PR-10 protein isolated from roots of *Petroselinum crispum* (Mill.) Fuss

Barbora Stratilová[a,b], Pavel Řehulka[c], Soňa Garajová[a], Helena Řehulková[c], Eva Stratilová[a], Maria Hrmova[d,e], Stanislav Kozmon[a,*]

[a]*Institute of Chemistry, Centre for Glycomics, Slovak Academy of Sciences, Dúbravská cesta 9, SK-84538 Bratislava, Slovakia; *chemsksa@savba.sk, tel. +421259410322, fax +421259410222*

[b]*Faculty of Natural Sciences, Department of Physical and Theoretical Chemistry, Comenius University Bratislava, Mlynská dolina, SK-842 15 Bratislava, Slovakia*

[c]*Department of Molecular Pathology and Biology, Faculty of Military Health Sciences, University of Defence, Třebešská 1575, CZ-50001 Hradec Králové, Czech Republic*

[d]*School of Life Sciences, Huaiyin Normal University, Huai'an 223300, China*

[e]*School of Agriculture, Food and Wine, and Waite Research Institute, Waite Research Precinct, University of Adelaide, Glen Osmond, SA 5064, Australia*

Declaration of conflict of interest: none.

Abstract

The native dimeric *Petroselinum crispum* (Mill.) Fuss protein Pet c 1.0201 and a monomeric xyloglucan endotransglycosylase enzyme (Garajova et al., 2008) isolated from the root cells co-purify and share similar molecular masses and acidic isoelectric points. In this work, we determined the complete primary structure of the parsley Pet c 1.0201 protein, based on tryptic and chymotryptic peptides followed by the manual micro-gradient chromatographic separation coupled with offline MALDI-TOF/TOF mass spectrometry. The bioinformatics approach enabled us to include the parsley protein into the PR-10 family, as it exhibited the highest protein sequence identity with the *Apium graveolens* Api g 1.0201 allergen and the major *Daucus carota* allergen Dau c 1.0201. Hence, we designated the *Petroselinum crispum* protein as Pet c 1.0201 and deposited it in the UniProt Knowledgebase under the accession C0HKF5. 3D protein homology modelling and molecular dynamics simulations of the Pet c 1.0201 dimer confirmed the typical structure of the Bet v 1 family allergens, and the potential of the Pet c 1.0201 protein to dimerize in water. The latter data agreed with those observed for Pet c 1.0201 during protein purification. However, the behavioural properties of Pet c 1.0201 and the celery allergen Api g 1.0101 differed in the presence of salts due to transiently and stably formed dimeric forms of Pet c 1.0201 and Api g 1.0101, respectively.

Keywords

Petroselinum crispum; Apiaceae; parsley; mass spectrometry; 3D structural modelling; molecular dynamics simulations; PR-10 proteins.

1. Introduction

The plant family Apiaceae includes agriculturally important plants such as carrot-*Daucus carota* L., celery-*Apium graveolens* L. and parsley-*Petroselinum crispum* (Mill.) Fuss. All produce proteins classified in the Structural Database of Allergenic Proteins (SDAP), more specifically in the Bet v 1 sub-family (Ivanciuc et al., 2003). Most of well-known pollen allergens cause allergic rhino conjunctivitis and asthma (Gajhede et al., 1996; Taketomi et al., 2006; Asam et al., 2015; Pablos et al., 2016), while food allergens underly oral allergy syndromes (Vanekkrebitz et al., 1995; Beyer et al., 2002; Neudecker et al., 2003). Dramatic expansion of these allergies has led to studies of their origin (Ballmer-Weber et al., 2012), mechanisms of action (Smole et al., 2015; Zulehner et al., 2017) and cross-reactivity (Vieths et al., 2002; Bohle et al., 2003; Bohle, 2007) that lead to disease treatments and prevention (Ballmer-Weber et al., 2005; Hoflehner et al., 2012).

Proteins of the Bet v 1 family fall into the pathogenesis-related protein (PR-proteins) group (Hoffmann-Sommergruber, 2000), as they are synthetized in plants mainly upon pathogen invasions and show defensive roles in plant system. Some of these proteins are expressed due to wounding or environmental stresses such as cold, heat, UV light, drought, flooding, salinity or the exposure to heavy metals (Edreva, 2005; Borad and Sriram, 2008). Compounds implicated in both in plant protection and regulation of their developmental processes such as flowering, fruit ripening, seed germination and embryogenesis were identified (Van Loon and Van Strien, 1999). The precise function of the pathogenesis-related (PR) proteins is unknown but some of them are classified as 1,3-β-D-glucanases, chitinases, thaumatin- and osmotin-like proteins, proteinase inhibitors, endoproteinases, peroxidases, ribonuclease-like proteins, defensins, thionins, lipid transfer proteins, oxalate oxidases and oxalate oxidase-like proteins (Bowles, 1990; Sinha et al., 2014). PR-proteins are divided into 17 families according to their properties, immunologic relationships and structural homologies (Van Loon and Van Strien, 1999; Okushima et al., 2000; Christensen et al., 2002; Sinha et al., 2014).

The Bet v 1 members belong to the PR-protein family 10 (PR-10). The principal representative of this family is a protein from *Petroselinum crispum* (Somssich et al., 1986). The key feature of this family is a high homology in tertiary structures, regardless their low sequence identity (Fernandes et al., 2013). A ubiquitous distribution of the Bet v 1-related proteins among all protein groups suggests that the Bet v 1-like protein was present in the last universal common

ancestor. During evolutionary history, this protein diversified into numerous families with a low sequence similarity but with a versatile scaffold for binding of bulky ligands (Radauer et al., 2008).

The PR-10 family contains small (around 160-residue proteins of 15-18 kDa) and acidic (isoelectric points of 4 - 5) multifunctional proteins (Agarwal and Agarwal, 2014). Although they have a low sequence identity, their secondary structure elements form specific structural arrangements (Fernandes et al., 2013). The tertiary structures of PR-10 proteins appear similar on the surface but show significant differences inside protein folds (Fernandes et al., 2013; Chwastyk et al., 2014) resulting in binding of various types of ligands, *e.g.* phytohormones that play a variety of physiological functions (Mogensen et al., 2002; Sliwiak et al., 2016a, b).

Differences in function of PR-10 proteins are given by variations of several key amino acid residues. The most variable region is localized at the C-terminus containing the conserved glycine-rich loop with the EG(D/N)GG(V/P)G(T/S) sequence (positions 45-52 in parsley *Petroselinum crispum* PR-10.1) preserved even in the most distant PR-10 allergens. This sequence represents a signature motif for PR-10 proteins (Fernandes et al., 2013), in addition to IgE binding residues that are linked to allergenicity (Neudecker et al., 2003; Spangfort et al., 2003). Because of the direct influence on the IgE binding capacity and cross-linking of antigens, leading to a histamine release from mastocytes, oligomerization patterns of these proteins are their key features (Scholl et al., 2005; Rouvinen et al., 2010; Kofler et al., 2014). As mentioned, the sequence identity of PR-10 proteins is low except of certain allergens (*e.g.* 1.0101 and 1.0201), but the 3D structures of the carrot *Daucus carota* and celery *Apium graveolens* allergens show a high structural similarity, although those found in *Petroselinum crispum* are rather different. It is of note that residue variations lead to varying responses of patients to the presence of allergenic isoforms that have a high sequence identity (Wangorsch et al., 2007).

The aim of this work was to provide the molecular and structural characterization of the Pet c 1.0201 protein from *Petroselinum crispum* roots that co-purifies with the xyloglucan endotransglycosylase (XET) enzyme (Garajová et al., 2008) and exhibits the high protein sequence identity with the Api g 1.0201 (Hoffmann-Sommergruber et al., 2000) and Dau c 1.0201 (Vieths et al., 2001) allergens. Our studies were made possible due to previous detailed findings published on the Apiaceae allergens (Hoffmann-Sommergruber et al., 1999; Neudecker et al., 2003; Schirmer et al., 2005; Markovic-Housley et al., 2009). Attention was paid to the dimerization patterns of Pet c 1.0201 observed experimentally in solutions during protein purification.

2. Results and Discussion

2.1. *Protein extraction, purification and electrophoretic characterization*

Proteins extracted from the root cells of *Petroselinum crispum* were purified, following the procedure desribed in Fig. S1, with the aim to obtain the homogeneous XET enzyme (Garajová et al., 2008). The fraction with the XET activity analyzed by IEF-PAGE (Fig. 1A) showed that the purified protein had the isoelectric point of 4.2, while Garajova and co-workers (2008) reported a slightly higher value (isoelectric point 4.6). The molecular mass of the protein visualised in the SDS-PAGE gel corresponding to approximately 17 kDa (Fig. 1B, sample lane – bottom band) was unexpectedly low for XET. In addition to this 17-kDa protein, a second band (Fig. 1B, sample lane – top band) with the molecular mass of approximately 35 kDa was detected. This molecular mass of 35 kDa was in-line with the mass estimation for XET by size-exclusion chromatography on the calibrated Superdex 75 colum (Garajová et al., 2008). However, sonication and boiling of the sample before SDS-PAGE led to a removal of the top 35 kDa band fom the SDS-PAGE gel, while the bottom 17 kDa band remained. The latter observation could be explained by a loss of the quarternary (dimeric) structure of the protein contained in the top band, while the monomeric protein in the bottom band remained unaffected.

2.2. *Mass spectrometry and bioinformatics analyses*

In-gel tryptic or chymotryptic digestion of a bottom protein band, followed by the mass spectrometry analysis combined with *de-novo* sequencing offered the complete primary structure of this protein (Fig. 2, File S1). The proteins from the bottom and upper bands shown on the SDS-PAGE (Fig. 1B) and IEF-PAGE gel (Fig. 1A) gels yielded identical amino acid sequences. On the other hand, the sequence of the XET protein contaminating the top band (Garajova et al., 2008), could not be identified due to its low protein content.

The highest sequence identity of the purified protein (Fig. 1) with sequences in databases was found for Api g 2 (Api g 1.0201) from *Apium graveolens* (Hoffmann-Sommergruber et al., 2000) and Dau c 1 (Dau c 1.0201) from *Daucus carota* (Vieths et al., 2001) proteins (Fig. 3A). The sequence identities of the purified protein with those of Api g 2 and Dau c 1 were 97.4% and 95.4%, respectively. Both Api g 2 and Dau c 1 belonged to type I allergens of the Bet v I family and to the group of PR-10 proteins. For this reason, we suggested to designate the newly identified protein from *Petroselinum crispum* as Pet c 1.0201. The alignment with other known PR-proteins

from *Petroselinum crispum* deposited in the UniProt database indicated the sequence identity of 54.6 to 60.5%, suggesting that until now such protein has not been identified in *Petroselinum* (Fig. 3B). Other comparisons showed a surprisingly high identity of 69.1% of Pet c 1.0201 to ribonucleases (Fig. 3C), whereby in these RNAse proteins the residues that are essential for the RNase activity were present in Lys54, Glu96, Glu148 and Tyr150 key positions. The connection between the RNase activity and antifungal activities was studied by Chadha and Das (2006), who showed that the mutation of Lys54 to Asn54 led to a complete removal of the RNase activity in peanut AhPR-10, concomitant with a loss of antifungal activity. Based on these analyses we concluded that Pet c 1.0201 belongs to the group of PR-10 proteins (Moiseyev et al., 1994; Bantignies et al., 2000; Park et al., 2004; Yan et al., 2008; Zubini et al., 2009).

2.3. 3D Protein structural modelling

The Api g 1 crystal structure (Schirmer et al., 2005) from the Apiaceae family was the first 3D structure of any food allergen solved by X-ray crystallography (PDB accession 2BK0). For this reason, we used the Api g 1 crystal structure as the template for 3D protein structural modelling of Pet c 1.0201, although the sequence identity of Api g 1 and Pet c 1.0201 was only 49% (Fig. 4A). The superposition of the Pet c 1.0201 model with the Api g 1 crystal structure showed that the optimized model had an excellent structural similarity to the Api g 1 crystal structure. Secondary and tertiary structures similarities between PR-10 proteins are generally known to be high (Radauer et al., 2008; Fernandes et al., 2013), in accordance with the root-mean-square deviation (RMSD) value (Maiorov and Crippen, 1994) of 2.5 Å that we detected between the monomeric sub-units of Api g 1 and Pet c 1.0201. We observed slight differences between Api g 1 and Pet c 1.0201 in the positions of the first short α-helix (Fig. 4A, α1) and in the loops linking the N-terminal β-sheet with the first short α-helix (Fig. 4A, L1), and in the positions of the second short α-helix that is linked with the second β-sheet (Fig. 4A, L3), and how those β-sheets were connected through loops (Fig. 4A, L4, L5, L8).

2.4. Molecular dynamic simulations of dimers stability

Theoretical studies investigated the tendency of PR-10 allergen proteins to create dimeric assemblies. The existence of transient or covalently stabilized dimer structures of allergens of the Bet v I family, including Api g 1 was described and directly related to the allergenic effect of these

proteins (Scholl et al., 2005; Rouvinen et al., 2010; Kofler et al., 2014). Our experimental findings, both from SDS-PAGE (Fig. 2B) and MS analyses confirmed the presence of the Pet c 1.0201 dimers. To shed more light on these structural features, we conducted molecular dynamics (MD) simulations to compare the Api g 1.0101 structure with that of the Pet c 1.0201 3D model.

First, we investigated, if Api g 1.0101 and Pet c 1.0201 could dimerize in aqueous solutions lacking NaCl. The visual analysis of inter-monomeric distances in dependence on time during MD simulations (Fig. 5A, B at respective 0 and 200 ns) showed that the monomeric units dimerized mainly through loop L8 situated between β6 and β7 chains, and through the neighbouring part of the β6 chain (Fig. 5A). These dimers were further stabilized *via* the residues on loop L1 that connected the N-terminal β-sheet with the first short α-helix, and through the residues of α1 helix or loop L6 (Fig. 6). Loop L8 was shorter in Api g 1.0101 (Fig 4A) than in Pet c 1.0201 leading to stronger inter-atomic bonds between exposed monomers; thus, no change of a relative position of the Api g 1.0101 monomers was observed during 200 ns simulation (Fig. 5A). The Pet c 1.0201 dimers were stabilized after approximately 50 ns, where mostly hydrophobic interactions were formed at interfaces, resulting in a change of the positions of monomers (Fig. 5B, 6B).

Trajectories generated after 200 ns MD simulations were used for the determination of free interaction energies (E_{int}) using the MM-GBSA analysis (Table 1). Averaged E_{int} for Api g 1.0101 was approximately two-fold lower than that for Pet c 1.0201, suggesting that the stability of the Api g 1.0101 dimer was higher. In accordance with this conclusion, many more hydrogen bonds and van der Waals contacts were observed between the Api g 1.0101 monomers (141 interactions in total) compared to those of the Pet c 1.0201 monomers (90 interactions) (Figs. 6; Tables S1, S2). It was of note that in the region of the second β-sheet of one of the monomers in Pet c 1.0201, these interactions repeatedly disintegrated during MD simulations (Fig. 6C). The theoretical mutation of Ser49 that is unique in Pet c 1.0201 to Gly, which is a conserved residue in this region of PR-10 proteins (Fernandes et al., 2013), led to the re-stabilization of this region. It is questionable, to which extent these interactions disintegrate in the Glu45 region (Fig. 4C), and how this may affect binding of IgE. As for Api g 1.0101, these instabilities in the second β-sheet of one of the monomers were not observed (Fig. 4B).

It was described that some PR-10 allergens could be stabilised by covalent S-S bridges between monomers to stabilize their quaternary structures (Rouvinen et al., 2010; Kofler et al., 2014), however, these S-S bridges were not observed in Api g 1.0101 or Pet c 1.0201. The rationale

for this was that the functional groups of cysteines were distant and oriented inside of the molecules (Fig. S1).

We were further interested in nature of dimerization of Pet c 1.0201 under various salt concentrations with 0.05 M, 0.1 M, 0.2 M, 0.5 M NaCl; these conditions may represent a more natural cellular environment for the PR-10 allergens. In general, the addition of the salt caused destabilization of the Pet c 1.0201 dimers under all studied concentrations, while the stability of the Api g 1.0101 dimers was not affected by NaCl (Table S1). Under 0.05, 0.2 and 0.5 M NaCl concentrations, we observed re-stabilization of the Pet c 1.0201 dimers or secondary stabilization of monomer sub-units after a dimer decay (Fig. 5, Table 1, File S2). As for Pet c 1.0201 stabilization, the monomer sub-units may have adopted differences in mutual positions of monomers, that were like those observed in water, that is lacking NaCl. All these interactions were more unstable compared to those in water, causing dimer dissociation. Thus, the changes observed in solutions with salts can be described by the formation of transiently or partially stable dimers. Trajectories generated through MD simulations were used to determine free interaction energies (E_{int}) using the MM-GBSA analysis (Table 1). The large differences in the estimated E_{int} between monomeric sub-units of Pet c 1.0201 were observed not only under individual NaCl concentrations, but also during a variety of simulations runs with the same salt concentrations. This could be the consequence of the transitional states of dimerization due to destabilization with the added salt.

In the case of Pet c 1.0201, the MD simulation with 0.2 M NaCl identified a relatively stable position of monomers (E_{int} of 5.3 kcal/mol), however, only with the probability of 1:2 during 200 ns simulation (Fig. 5C, 1. run), which suggested that the dimeric form of Pet c 1.0201 could dissociate to a monomeric form. This structure of Pet c 1.0201 at the end of the MD simulation may correspond to one out of two possible states that were revealed for Api g 1 or Dau c 1 in crystal structures (Rouvinen et al., 2010; Markovic-Housley et al., 2009). The contact surface that allowed for these interactions is seemingly greater in a dimer, however, the authors identified only 47 interactions between monomers (Fig. 6C, Table S3), in accordance with a higher stability of the Pet c 1.0201 dimer in water (Fig. 6B, Table S2).

The formation of different homo-dimeric structures of Pet c 1.0201 was seen during two MD simulation runs with 0.5 M NaCl concentration. Here, stable dimers, comparable to those seen during MD simulations of Pet c 1.0201 lacking NaCl, were observed in the first and the third

simulation run. However, the mutual monomer positions of monomers differed from those observed under the conditions of lacking NaCl (Fig. S2). During the third simulation run of Pet c 1.0201 the dimeric structure was similar to that observed for the Pet c 1.0201 MD simulation run under 0.2 M NaCl, while the structure obtained in the first simulation run represented a new dimeric form not seen in any of the previous MD simulations (Fig. S2). In conclusion, MD simulations in the NaCl environment pointed at significant differences between the stability of dimeric forms for the 1.0101 and 1.0201 allergens, due to using a more realistic environment with salts during MD simulations. As it was mentioned above, the higher dimer stability of the Apiaceae isoforms 1.0101 may be one of the factors that lead to their higher allergenicity.

The link between dimerization and allergenicity of PR-10 proteins (Scholl et al., 2005; Kofler et al., 2014) is known, and our findings now explain the differences in higher allergenicity of the 1.0101 forms compared to those of 1.0201. It was shown by Wangorsch et al. (2007) that the determination of the IgE binding capacity of the 1.0201 forms poised difficulties because only 48% of the *Apium graveolens* allergic patients and 14% of the birch pollen allergic patients reacted to Api g 1.0201; in comparison respective 74% and 41% reacted to the major allergen 1.0101. Because of a lack of patients with a reliably diagnosed allergy to the *Petroselinum crispum* and Apiaceae 1.0201 allergens, our work is largely theoretical in describing the structural properties of the previously uncharacterized *Petroselinum crispum* allergen. The importance of these findings highlights that the differences in structural properties cannot be explained at the primary structural levels, because Glu45 that is known to be responsible for IgE binding, is replaced with lysine in the 1.0101 isoforms while the 1.0201 isoforms, with a lower allergenicity, it retains Glu in the equivalent position. The importance of Glu45 was proven by Spangfort et al. (2003), who showed that the mutation of this residue to Ser evaded binding of IgE. It is of note that Lys45Glu variants of the *Apium graveolens* and *Daucus carota* 1.0101 isoforms showed an increase in IgE binding (Neudecker et al., 2003).

3. Concluding Remarks

The primary structure of the previously uncharacterized member of the PR-10 protein family, Pet c 1.0201 from the *Petroselinum crispum* roots was defined *via* the MALDI-TOF/TOF mass spectrometry. The primary structure was used to construct the Pet c 1.0201 3D protein model, where the Api g 1.0101 crystal structure, with 49% sequence identity to Pet c 1.0201 was utilized

as a template. As Api g 1.0101 formed a dimeric asembly in a crystal form, the structure of the Pet c 1.0201 protein was also modelled in the dimeric form. This was justified by the fact, that the Pet c 1.0201 dimer was observed during protein purification. These dimeric structures were further studied through MD simulations to test their stability in the solutions with various degrees of salinity. The dimer of Api g 1.0101 was stable under all tested conditions contrary to the Pet c 1.0201 dimer that was only stable without NaCl, while the increasing concentrations of NaCl had a destabilizing effect on the Pet c 1.0201 dimer. Nonetheless, certain stabilization was observed with the highest – 0.5 M – concentration of NaCl, although, the mutual dispositions of the monomeric sub-units differed compared to those in the crystal structure of Api g 1.0101. Due to a high sequence identity between Pet c 1.0201 and other Apiaceae 1.0201 allergens, the Pet c 1.0201 protein may serve as a mild allergen for sensitive individuals, with and a generally lower allergenicity for the 1.02 isoforms, compared to those of the 1.01 isoforms. These observations could partially be explained by a lower stability of the 1.02 dimeric forms.

4. Experimental

4.1 Protein extraction and purification

Seeds of *Petroselinum crispum* (Mill.) Fuss (Apiaceae family) were purchased from the SEMO a.s. company (Smržice, Czech Republic). Twenty kg of *Petroselinum crispum* roots were collected in September 2006 from the field in the Danubian Lowland (Slovakia) situated in the area of Žitný Ostrov at DMS (Degrees Minutes Seconds)/latitude-longitude of 48°01'36.96" N and 17°19'20.89" E. At that time, relatively normal or slightly increased humidity and temperatures prevailed in this area except for July and September that were dry and warm. Roots were stored at 5 ºC and processed sequentially for two weeks. The first step of the protein extraction included homogenization in the juice extractor ES-3551 (Severin, Germany), followed by a deep-freezing of the pulp to approximately -20 ºC. Subsequently, proteins in a frozen pulp were extracted at an ambient temperature for 12 h in 0.1 M imidazole (pH 6.0 containing 1 M NaCl). After filtration and centrifugation (23 650 x g, 20 min, 4 °C) the liquid extract was precipitated with ammonium sulphate (Merck, Germany) to 100% saturation for 24 h at 4 °C, centrifuged. After precipitation, the sediment was dissolved in a small amount of the 1 M NaCl solution, dialyzed against distilled water and freeze-dried.

The parsley Pet c 1.0201 protein was co-purified from the liquid extract of the material as described for the XET enzyme (Garajová et al., 2008), using size-exclusion chromatography on the Biogel P-30 (Bio-Rad Laboratories, Hercules, CA, USA), followed by affinity separation on Concanavalin A Sepharose (Pharmacia, Uppsala, Sweden), size-exclusion chromatography on Superdex 75 (GE Healthcare, Chicago, IL, USA), and ion-exchange chromatography on the Mono Q HR 5/50 GL column (GE Healthcare) (Fig. S1). The last two purification steps were conducted on the FPLC instrument (Pharmacia). The Mono Q step was performed in 0.02 M Bis-Tris/HCl buffer, pH 6.4, whereby the protein was released from the column by applying a stepwise gradient of NaCl (0 - 0.7 M). The sample application flow rate was 0.25 ml/min that was altered to 0.5 ml/min before the NaCl gradient was applied. The summary of protein purification steps is shown in Fig. S1. The XET activity was determined using fluorescently labelled acceptor substrates (Stratilová et al., 2010).

4.2 Isoelectric point and molecular mass determination

Isoelectric point determination was performed on the ultrathin-layer isoelectric focusing polyacrylamide gels (IEF-PAGE) in the pH range of 3-10 (Radola et al., 1980) using protein test mixture (Serva, Heidelberg, Germany) for calibration; protein bands in gels were stained with Coomassie Brilliant Blue (Fluka Chemie, Buchs, Germany).

Denaturing 10% (w/v) SDS-PAGE (Laemmli, 1970) gels were run on the Mini-Protean 3 electrophoresis system (Bio-Rad Laboratories) under the reducing conditions. Protein bands were visualized by Coomassie Brilliant Blue (Fluka). Standard PageRuler^{TM} pre-stained protein ladder in the 10-170 kDa range were used (Thermo Fisher Scientific, Waltham, MA, USA) for calibration.

4.3 Mass spectrometry analysis

Coomassie Brilliant blue-stained protein bands from the gels after IEF-PAGE and SDS-PAGE were processed according to the protocol of Shevchenko et al. (2006). Briefly, the excised gel pieces (about 5 x 2 mm) containing separated proteins were dehydrated with 450 µl acetonitrile (5 min) and the solutions removed. Proteins were in-gel reduced with 50 µl 10 mM DTT in 100 mM NH_4HCO_3 at 56 °C for 30 min. The mixture was vortexed after the addition 450 µl acetonitrile and the supernatant removed. Protein alkylation was done with 50 µl 55 mM

316 iodoacetamide in 100 mM NH_4HCO_3 in the dark at ambient temperature for 20 min. The
317 supernatant was removed, gel pieces washed with 450 µl 100 mM NH_4HCO_3 for 5 min and
318 dehydrated with 450 µl acetonitrile for 5 min. The supernatant was removed, and gel pieces were
319 rehydrated for 30 min at 4 °C with the solution of 10 mM NH_4HCO_3 and 10 µg/ml porcine
320 modified sequencing grade trypsin (Promega, Madison, MA, USA) or 20 µl/ml bovine sequencing
321 grade chymotrypsin (Roche, Basel, Switzerland). The digestion solution of 10 mM NH_4HCO_3
322 (about 30-50 µl) was added to cover gel pieces and the digestion was carried out overnight at 37
323 °C. The in-gel digests were handled using manual micro-gradient chromatographic separation
324 (Franc et al., 2012) using the capillary column (length: 30 mm, i.d. 250 µm) packed with 5 µm
325 particles from the Poroshell 300 Extended-C18 column (Agilent Technologies, Santa Clara, CA,
326 USA) or the capillary column (length: 27 mm, i.d. 250 µm) packed with 2.7 µm particles from the
327 Ascentis Express Peptide ES-C18 column (Sigma-Aldrich, St. Louis, MI, USA). After MALDI-
328 TOF/TOF MS analysis, selected peptide fractions were re-dissolved in 60% (v/v) acetonitrile and
329 transferred to 0.5 ml test tubes, dried in a vacuum centrifuge and esterified using 50 µl of the
330 mixture of acetylchloride/ethanol (8:50, v/v) for 2 hours at room temperature. The mixture was
331 dried in a vacuum centrifuge, dissolved in 2 µl 60% (v/v) acetonitrile and analyzed using MALDI-
332 TOF/TOF MS for confirmation of peptide sequences derived from the MALDI-TOF/TOF MS/MS
333 analysis of underivatized peptides.
334 MALDI-TOF/TOF MS measurements in the positive reflectron mode were performed with
335 the Applied Biosystems 4800 Proteomics Analyzer (Applied Biosystems, Foster City, CA, USA).
336 MS spectra were acquired using a dual stage reflectron mirror. Accelerating voltage applied for
337 MS measurements was 20 kV. Raw spectral data were further processed using Data Explorer 4.8
338 software (Applied Biosystems). For mass determination of protein digest, a solution of α-cyano-
339 4-hydroxycinnamic acid (5 mg/ml) in acetonitrile/0.1% (v/v) TFA (3:2, v/v) was used. 0.5 µl of
340 the purified peptide mixture was mixed with 0.5 µl of the matrix solution and dried at ambient
341 temperature. The peak m/z values from MS and MS/MS analyses were submitted to the Mascot
342 protein identification program package (v. 2.3.02) to perform database searching with the
343 following parameters: database - NCBInr (ver. 27.11.2011); taxonomy - all entries; enzyme –
344 trypsin or chymotrypsin; allowed missed cleavages - 1; fixed modifications - carbamidomethyl
345 (C); variable modifications - Oxidation (M), Pyro-cmC (N-term camC), Pyro-glu (N-term E),
346 Pyro-glu (N-term Q); peptide tolerance - 50 ppm; MS/MS tolerance - 250 mmu; peptide charge -

(+1); monoisotopic masses; instrument - MALDI-TOF-PSD. Alternatively, MS/MS spectra of tryptic and chymotryptic peptides were evaluated manually and corresponding peptide fragment ions were identified. These MS/MS spectra were used as the basis for *de-novo* sequencing of the PR10-protein and are presented in Supplementary File S1. The protein sequence data reported in this paper are deposited in the UniProt Knowledgebase under the accession C0HKF5. Similarity based searches were performed using BLAST (Altschul et al., 1997).

4.4 Bioinformatics analysis

The primary structure of the *Petroselinum crispum* protein Pet c 1.0201 was analyzed using BLAST and alignment tools in the UniProtKB database (UniProt C, 2017). Alignments were generated in Clustal Omega (Sievers et al., 2011).

4.5 3D Protein structure modelling

The 3D protein homology model of the MS-determined primary structure of the *Petroselinum crispum* Pet c 1.0201 protein was constructed based on the food allergen from the family Apiaceae (Schirmer et al., 2005) as the template structure (PDB accession 2BK0), using Modeller9v6 (Sali and Blundell, 1993). The 2BK0 structure is a dimer, therefore, the 3D model of the parsley *Petroselinum crispum* Pet c 1.0201 protein was also prepared in the dimeric form.

4.6 Molecular dynamic simulations of dimer stability

The 3D protein homology model and the template structure 2BK0 of dimeric forms were protonated using the Protein Preparation Wizard included in the Schrodinger Suite 2015.2 (Schrodinger 2015, USA) for the neutral pH and the positions of the added protons were optimized according their neighbourhood. Protonated structures were optimized using the molecular mechanics force field Amber99sb in Amber14 (Case et al., 2014). Optimized structures were solvated by the water molecules (TIP3P) in a cubic box with at least a 17 Å-thick water layer. Charge of the model systems was neutralized by adding of Na^+ counter ions. Additional Na^+ and Cl^- ions were applied into the box for simulations in various environmental salinities (0.05 M, 0.1 M, 0.2 M, 0.5 M for the 3D protein homology model and 0.2 M for Api g 1.0101). Prepared structures were equilibrated under periodic boundary conditions to the temperature of 300 K. The simulation box was heated to 300 K over the course of a 100 ps constant volume simulation (NVE)

applying the 50 kcal/mol restraint to all solute atoms, followed by a 300 ps 1 bar constant pressure simulation. Next, a series of constant pressure steps 10 ps in length were carried out, in which the strength of the restraint was reduced gradually to 25 kcal/mol and 10 kcal/mol, which was followed by a 50 ps simulation with a restraint strength of 5 kcal/mol and 70 ps simulation with a restraint strength of 2.5 kcal/mol. The last two steps were 300 ps simulations with the protein backbone held by a 1 kcal/mol restraint and a 300 ps restraint-free simulation. Based on the equilibrated structure production, we defined MD simulations such that the production simulation run under the periodic boundary conditions using NPT ensemble lasted for 200 ns. Snapshots from the simulation ware taken every 10 ps. MD simulations were performed in the Amber14 program using the Amber99sb force filed parameters. Simulations for each salt concentration were run in triplicates; standard deviations (σ) were calculated from these simulations. The Virtual Molecular Dynamics (VMD) program was used for visualization of trajectories (Humphrey et al., 1996). The backbone root-mean-square-deviation **(**RMSD) values and inter-monomeric distances were calculated *via* the *cpptraj* AmberTools14 utility (Case et al., 2014).

Trajectories were used to analyse interaction energies between monomer molecules. Interaction energies were calculated using the MM-GBSA analysis implemented in the Amber14 program using the MMPBSA.py script from AmberTools14 (Case et al., 2014). Every 10^{th} snapshot from the production trajectory was taken for the analysis.

Acknowledgement

This work was financially supported by the VEGA grant No. 2/0058/16, Slovakia (E.S.), and by the Ministry of Defence of the Czech Republic - Long-term organization development plan of the Faculty of Military Health Sciences, University of Defence (P.Ř.). Research was supported by the funding from Huaiyin Normal University and the Australian Research Council Linkage Project (DP120100900) to M.H., the program SASPRO (ArIDARuM, 0005/01/02 - SK) co-funded by the People Programme (Marie Curie Actions 7FP, grant agreement REA no. 609427 - SK) and co-financed by the Slovak Academy of Sciences (S.K.). The authors are grateful to Dr Zuzana Firáková and Mrs Helena Čigašová for technical assistance.

Table 1. Free interaction energies E_{int} of the Pet c 1.0201 3D model based on MM-GBSA and for the Api g 1 crystal structure under various concentrations of NaCl.

		1. run[a]		2. run[a]		3. run[a]		Average[a]
Protein	Concentration NaCl [M]	E_{int} [kcal/mol]	± σ	E_{int} [kcal/mol]	± σ	E_{int} [kcal/mol]	± σ	E_{int} [kcal/mol]
2BK0	0	-13.50	3.20	-12.60	3.98	-14.65	4.79	-13.58
	0.2	-17.00	8.45	-18.56	5.68	-20.30	3.73	-18.62
Pet c 1.0201	0	-8.25	4.77	-6.90	4.24	-6.23	3.70	-7.12
	0.05	-0.19	6.26	-7.78	8.61	-0.05	2.74	-2.67
	0.1	-2.10	6.20	-1.59	5.55	-0.48	3.47	-1.39
	0.2	-5.30	9.55	3.47	6.55	0.69	2.37	-0.38
	0.5	-6.56	10.97	0.52	3.72	-13.75	6.47	-6.59

[a] Average value is based on three individual runs, standard deviations (σ) were calculated from each MM-GBSA calculation.

References

Agarwal, P., Agarwal, P.K., 2014. Pathogenesis related-10 proteins are small, structurally similar but with diverse role in stress signaling. Mol. Biol. Rep. 41, 599-611. doi: 10.1007/s11033-013-2897-4.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389-3402. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC146917/pdf/253389.pdf.

Asam, C., Hofer, H., Wolf, M., Aglas, L., Wallner, M., 2015. Tree pollen allergens-an update from a molecular perspective. Allergy 70, 1201-1211. doi: 10.1111/all.12696.

Ballmer-Weber, B.K., Wangorsch, A., Bohle, B., Kaul, S., Kundig, T., Fotisch, K., van Ree, R., Vieths, S., 2005. Component-resolved in vitro diagnosis in carrot allergy: Does the use of recombinant carrot allergens improve the reliability of the diagnostic procedure? Clin. Exp. Allergy 35, 970-978. doi: 10.1111/j.1365-2222.2005.02294.x.

Ballmer-Weber, B.K., Hansen, K.S., Sastre, J., Andersson, K., Batscher, I., Ostling, J., Dahl, L., Hanschmann, K.M., Holzhauser, T., Poulsen, L.K., et al., 2012. Component-resolved in vitro diagnosis of carrot allergy in three different regions of Europe. Allergy 67, 758-766. doi: 10.1111/j.1398-9995.2012.02827.x.

Bantignies, B., Seguin, J., Muzac, I., Dedaldechamp, F., Gulick, P., Ibrahim, R., 2000. Direct evidence for ribonucleolytic activity of a PR-10-like protein from white lupin roots. Plant Mol. Biol. 42, 871-881. doi: 10.1023/A:1006475303115.

Beyer, K., Bardina, L., Grishina, G., Sampson, H.A., 2002. Identification of sesame seed allergens by 2-dimensional proteomics and Edman sequencing: Seed storage proteins as common food allergens. J. Allergy Clin. Immunol. 110, 154-159. doi: 10.1067/mai.2002.125487.

Bohle, B., Radakovics, A., Jahn-Schmid, B., Hoffmann-Sommergruber, K., Fischer, G.F., Ebner, C., 2003. Bet v 1, the major birch pollen allergen, initiates sensitization to Api g 1, the major allergen in celery: evidence at the T cell level. Eur. J. Immunol. 33, 3303-3310. doi: 0.1002/eji.200324321.

Bohle, B., 2007. The impact of pollen-related food allergens on pollen allergy. Allergy 62, 3-10. doi:10.1111/j.1398-9995.2006.01258.x

Bowles, D.J., 1990. Defense-related proteins in higher plants. Annu. Rev. Biochem. 59, 873-907. doi: 10.1146/annurev.bi.59.070190.004301.

Case, D.A., Babin, V., Berryman, J.T., Betz, R.M., Cai, Q., Cerutti, D.S., Cheatham, T.E., Darden, T.A., Duke, R.E., Gohlke, H., et al., 2014. AMBER 14. University of California, San Francisco; 2014. http://ambermd.org/.

Chadha, P, Das, R.H., 2006. A pathogenesis related protein, AhPR10 from peanut: an insight of its mode of antifungal activity. Planta 225, 213-222. doi:10.1007/s00425-006-0344-7.

Christensen, A.B., Cho, B.H., Naesby, M., Gregersen, P.L, Brandt, J., Madriz-Ordenana, K., Collinge, D.B., Thordal-Christensen, H., 2002. The molecular characterization of two barley proteins establishes the novel PR-17 family of pathogenesis-related proteins. Mol. Plant Pathol. 3, 135-144. doi: 10.1046/j.1364-3703.2002.00105.x.

Chwastyk, M., Jaskolski, M., Cieplak, M., 2014. Structure-based analysis of thermodynamic and mechanical properties of cavity-containing proteins-case study of plant pathogenesis-related proteins of class 10. FEBS J. 281, 416-429. doi: 10.1111/febs.12611.

Edreva, A., 2005. Pathogenesis-related proteins: research progress in the last 15 years. Gen. Appl. Plant Physiol. 31, 105-124. https://pdfs.semanticscholar.org/251a/8c7bfee5a062f4e336474c87338e26882af2.pdf

Fernandes, H., Michalska, K., Sikorski, M., Jaskolski, M., 2013. Structural and functional aspects of PR-10 proteins. FEBS J. 280, 1169-1199. doi:10.1111/febs.12114.

Franc, V., Sebela, M., Rehulka, P., Koncitikova, R., Lenobel, R., Madzak, C., Kopecny, D., 2012. Analysis of N-glycosylation in maize cytokinin oxidase/dehydrogenase 1 using a manual microgradient chromatographic separation coupled offline to MALDI-TOF/TOF mass spectrometry. J. Prot. 75, 4027-4037. doi: 10.1016/j.jprot.2012.05.013.

Gajhede, M., Osmark, P., Poulsen, F.M., Ipsen, H., Larsen, J.N., Joost van Neerven, R.J., Schou, C., Lowenstein, H., Spangfort, M.D., 1996. X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy. Nat. Struct. Biol. 3, 1040-1045. doi:10.1038/nsb1296-1040.

Garajova, S., Flodrova, D., Ait-Mohand, F., Farkas, V., Stratilova, E., 2008. Characterization of two partially purified xyloglucan endotransglycosylases from parsley (*Petroselinum crispum*) roots. Biologia 63, 313-319. doi: 10.2478/s11756-008-0067-2.

Hoffmann-Sommergruber, K., O'Riordain, G., Ahorn, H, Ebner, C., Laimer da Camara Machado, M., Puhringer, H., Scheiner, O., Breiteneder, H., 1999. Molecular characterization of Dau c 1,

the Bet v 1 homologous protein from carrot and its cross-reactivity with Bet v 1 and Api g 1. Clin Exp Allergy 29, 840-847. doi: 10.1046/j.1365-2222.1999.00529.x.

Hoffmann-Sommergruber, K., 2000. Plant allergens and pathogenesis-related proteins - What do they have in common? Int. Arch. Allergy Immunol. 122, 155-166. doi:10.1159/000024392.

Borad, V., Sriram, S., 2008. Pathogenesis-Related Proteins for the Plant Protection. Asian J. Exp. Sci. 22, 189-196. doi: 10.1.1.494.8921.

Hoffmann-Sommergruber, K., Ferris, R., Pec, M., Radauer, C., O'Riordain, G., Laimer da Camara Machado, M., Scheiner, O., Breiteneder, H., 2000. Characterization of Api g 1.0201, a new member of the Api g 1 family of celery allergens. Int. Arch. Allergy Immunol. 122, 115-123. doi:10.1159/000024367.

Hoflehner, E., Hufnagl, K., Schabussova, I., Jasinska, J., Hoffmann-Sommergruber, K., Bohle, B., Maizels, R.M., Wiedermann, U., 2012. Prevention of Birch Pollen-Related Food Allergy by Mucosal Treatment with Multi-Allergen-Chimers in Mice. PLoS One 7, e39409. doi: 10.1371/journal.pone.0039409

Humphrey, W., Dalke, A., Schulten, K., 1996. VMD: visual molecular dynamics. J. Mol. Graph. 14, 33-38, 27-38. doi: 10.1016/0263-7855(96)00018-5.

Ivanciuc, O., Schein, C.H., Braun, W., 2003. SDAP: database and computational tools for allergenic proteins. Nucleic Acids Research 31, 359-362. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC165457/

Kofler, S., Ackaert, C., Samonig, M., Asam, C., Briza, P., Horejs-Hoeck, J., Cabrele, C., Ferreira, F., Duschl, A., Huber, C., Brandstetter, H., 2014. Stabilization of the Dimeric Birch Pollen Allergen Bet v 1 Impacts Its Immunological Properties. J. Biol. Chem. 289, 540-551. doi: 10.1074/jbc.M113.518795.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-685. doi:10.1038/227680a0.

Maiorov, V.N., Crippen, G.M., 1994. Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins. J. Mol. Biol. 235, 625-634. doi: 10.1006/jmbi.1994.1017.

Markovic-Housley, Z., Basle, A., Padavattan, S., Maderegger, B., Schirmer, T., Hoffmann-Sommergruber, K., 2009. Structure of the major carrot allergen Dau c 1. Acta Cryst. D 65,1206-1212. doi: 10.1107/S0907444909034854.

504 Mogensen, J.E., Wimmer, R., Larsen, J.N., Spangfort, M.D., Otzen, D.E., 2002. The major birch
505 allergen, Bet v 1, shows affinity for a broad spectrum of physiological ligands. J. Biol. Chem.
506 277, 23684-23692. doi: 10.1074/jbc.M202065200.
507 Moiseyev, G.P., Beintema, J.J., Fedoreyeva, L.I., Yakovlev, G.I., 1994. High sequence similarity
508 between a ribonuclease from ginseng calluses and fungus-elicited proteins from parsley
509 indicates that intracellular pathogenesis-related proteins are Ribonucleases. Planta 193, 470-
510 472. doi: 10.1007/BF00201828.
511 Neudecker, P., Lehmann, K., Nerkamp, J., Haase, T., Wangorsch, A., Fotisch, K., Hoffmann, S.,
512 Rosch, P., Vieths, S., Scheurer, S., 2003. Mutational epitope analysis of Pru av 1 and Api g 1,
513 the major allergens of cherry (*Prunus avium*) and celery (*Apium graveolens*): correlating IgE
514 reactivity with three-dimensional structure. Biochem. J. 376, 97-107. *doi*: 10.1042/bj20031057.
515 Okushima, Y., Koizumi, N., Kusano, T., Sano, H., 2000. Secreted proteins of tobacco cultured
516 BY2 cells: identification of a new member of pathogenesis-related proteins. Plant Mol. Biol.
517 42, 479-488. doi: 10.1023/A:100639332698.
518 Pablos, I., Wildner, S., Asam, C., Wallner, M., Gadermaier, G., 1016 Pollen Allergens for
519 Molecular Diagnosis. Curr. Allergy Asthma Rep. 16, 31. doi: 10.1007/s11882-016-0603-z.
520 Park, C.J., Kim, K.J., Shin, R., Park, J.M., Shin, Y.C., Paek, K.H., 2004. Pathogenesis-related
521 protein 10 isolated from hot pepper functions as a ribonuclease in an antiviral pathway. Plant J.
522 37, 186-198. doi: 10.1046/j.1365-313X.2003.01951.x.
523 Radauer, C., Lackner, P., Breiteneder, H., 2008. The Bet v 1 fold: an ancient, versatile scaffold
524 for binding of large, hydrophobic ligands. BMC Evol. Biol. 8, 286. doi:10.1186/1471-2148-8-
525 286.
526 Radola, B.J., 1980. Ultrathin-layer isoelectric focusing in 50–100 μm polyacrylamide gels on
527 silanized glass plates or polyester films. Electrophoresis 1, 43-56.
528 doi:10.1002/elps.1150010109.
529 Rouvinen, J., Janis, J., Laukkanen, M.L., Jylha, S., Niemi, M., Paivinen, T., Makinen-Kiljunen,
530 S., Haahtela, T., Soderlund, H., Takkinen, K., 2010. Transient Dimers of Allergens. Plos One
531 5, e9037. doi: 10.1371/journal.pone.0009037.
532 Sali, A., Blundell, T.L., 1993. Comparative protein modelling by satisfaction of spatial restraints.
533 J. Mol. Biol. 234, 779-815. doi: 10.1006/jmbi.1993.1626.

534 Schirmer, T., Hoffmann-Sommergruber, K., Susani, M., Breiteneder, H., Markovic-Housley, Z.,
535 2005. Crystal structure of the major celery allergen Api g 1: Molecular analysis of cross-
536 reactivity. J. Mol. Biol. 351, 1101-1109. doi:10.1016/jmb.2005.06.054.
537 Scholl, I., Kalkura, N., Shedziankova, Y., Bergmann, A., Verdino, P., Knittelfelder, R., Kopp, T.,
538 Hantusch, B., Betzel, C., Dierks, K., et al., 2005. Dimerization of the major birch pollen allergen
539 Bet v 1 is important for its in vivo IgE-cross-linking potential in mice. J. Immunol. 175, 6645-
540 6650. doi: 10.4049/jimmunol.175.10.6645.
541 Schrödinger-Inc., 2015. Schrödinger Suite 2015.2. 2015-2 edition. New York, NY: Schrödinger,
542 LLC.
543 Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., Mann, M., 2006. In-gel digestion for mass
544 spectrometric characterization of proteins and proteomes. Nat. Protoc. 1, 2856-2860. doi:
545 10.1038/nprot.2006.468.
546 Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H.,
547 Remmert, M., Soding, J., et al., 2011. Fast, scalable generation of high-quality protein multiple
548 sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539. doi: 10.1038/msb.2011.75.
549 Sinha, M., Singh, R.P., Kushwaha, G.S., Iqbal, N., Singh, A., Kaushik, S., Kaur, P., Sharma, S.,
550 Singh, T.P., 2014. Current Overview of Allergens of Plant Pathogenesis Related Protein
551 Families. Sci. World J., art.ID 543195. doi.org/10.1155/2014/543195.
552 Sliwiak, J., Dauter, Z., Jaskolski, M., 2016a. Crystal Structure of Hyp-1, a Hypericum perforatum
553 PR-10 Protein, in Complex with Melatonin. Front. Plant Sci. 7, 10. doi:
554 10.3389/fpls.2016.00668
555 Sliwiak, J., Dolot, R., Michalska, K., Szpotkowski, K., Bujacz, G., Sikorski, M., Jaskolski, M.,
556 2016b. Crystallographic and CD probing of ligand-induced conformational changes in a plant
557 PR-10 protein. J. Struct. Biol. 193, 55-66. doi.org/10.1016/j.jsb.2015.11.008.
558 Smole, U., Radauer, C., Lengger, N., Svoboda, M., Rigby, N., Bublin, M., Gaier, S., Hoffmann-
559 Sommergruber, K., Jensen-Jarolim, E., Mechtcheriakova, D., Breiteneder, H., 2015. The major
560 birch pollen allergen Bet v 1 induces different responses in dendritic cells of birch pollen
561 allergic and healthy individuals. PLoS One 10, e0117904. doi:10.1371/journal.pone.0117904.
562 Somssich, I.E., Schmelzer, E, Bollmann, J., Hahlbrock K., 1986. Rapid activation by fungal
563 elicitor of genes encoding pathogenesis-related proteins in cultured parsley cells. Proc. Natl.

564 Acad. Sci. U.S.A. 83, 2427-2430. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC323310/
565 pdf/pnas00312-0153.pdf.
566 Spangfort, M.D., Mirza, O., Ipsen, H., van Neerven, R.J.J., Gajhede, M., Larsen, J.N., 2003.
567 Dominating IgE-binding epitope of Bet v 1, the major allergen of birch pollen, characterized
568 by X-ray crystallography and site-directed mutagenesis. J. Immunol. 171, 3084-3090. doi:
569 10.4049/jimmunol.171.6.3084.
570 Stratilova, E., Ait-Mohand, F., Rehulka, P., Garajova, S., Flodrova, D., Rehulkova, H., Farkas, V.,
571 2010. Xyloglucan endotransglycosylases (XETs) from germinating nasturtium (*Tropaeolum*
572 *majus*) seeds: Isolation and characterization of the major form. Plant Physiol. Biochem. 48,
573 207-215. doi:10.1002/elps.1150010109.
574 Taketomi, E.A., Sopelete, M.C., de Sousa Moreira P.F., de Assis Machado Vieira, F., 2006. Pollen
575 allergic disease: pollens and its major allergens. Braz. J. Otorhinolaryngol. 72, 562-567. doi:
576 10.1016/S1808-8694(15)31005-3.
577 The UniProt C, 2017. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 45, D158-
578 D169. https://doi.org/10.1093/nar/gkw1099.
579 Vanek-Krebitz, M., Hoffmann-Sommergruber, K., Laimer da Camara Machado, M., Susani, M.,
580 Ebner, C., Kraft, D., Scheiner, O., Breiteneder, H., 1995. Cloning and Sequencing of Mal-D-1,
581 the Major Allergen from Apple (*Malus-Domestica*), and Its Immunological Relationship to Bet-
582 V-1, the Major Birch Pollen Allergen. Biochem. Biophys. Res. Commun. 214, 538-551.
583 doi:10.1006/bbrc.1995.2320.
584 Van Loon, L.C., Van Strien, E.A., 1999. The families of pathogenesis-related proteins, their
585 activities, and comparative analysis of PR-1 type proteins. Physiol. Mol. Plant Pathol. 55, 85-
586 97. https://www.uu.nl/sites/default/files/pmpp-vanloon-1999.pdf.
587 Vieths, S., Wangorsch, A., Ballmer-Weber, B.K., 2001. Molecular characterization of the major
588 carrot allergen Dau c 1.0201, a homolog of the celery allergen Api g 1.0201.
589 EMBL/GenBank/DDBJ databases. Submitted (DEC-2001).
590 Vieths, S., Scheurer, S., Ballmer-Weber, B., 2002. Current understanding of cross-reactivity of
591 food allergens and pollen, In: Fu, T.J., Gendel, S.M. (Eds.), Genetically Engineered Foods
592 Assessing Potential Allergenicity 964. New York Acad. Sci., New York, pp. 47-68.
593 doi: 10.1111/j.1749-6632.2002.tb04132.x.

Wangorsch, A., Ballmer-Weber, B.K., Rösch, P., Holzhauser, T., Vieths, S., 2007. Mutational epitope analysis and cross-reactivity of two isoforms of Api g 1, the major celery allergen. Mol. Immunol. 44, 2518–2527. doi:10.1016/j.molimm.2006.12.023.

Yan, Q., Qi, X., Jiang, Z., Yang, S., Han, L., 2008. Characterization of a pathogenesis-related class 10 protein (PR-10) from *Astragalus mongholicus* with ribonuclease activity. Plant Physiol. Biochem. 46, 93-99. doi: 10.1016/j.plaphy.2007.10.002.

Zubini, P., Zambelli, B., Musiani, F., Ciurli, S., Bertolini, P., Baraldi, E., 2009. The RNA hydrolysis and the cytokinin binding activities of PR-10 proteins are differently performed by two isoforms of the Pru p 1 peach major allergen and are possibly functionally related. Plant Physiol. 150, 1235-1247. doi: 10.1104/pp.109.139543.

Zulehner, N., Nagl, B., Briza, P., Roulias, A., Ballmer-Weber, B., Zlabinger, G.J., Ferreira, F., Bohle, B., 2017. Characterization of the T-cell response to Dau c 1, the Bet v 1-homolog in carrot. Allergy 72, 244-251. doi: 10.1111/all.12938.

Figures and Legends

Fig. 1. IEF-PAGE (A) and SDS-PAGE (B) of the Coomassie Brilliant blue stained proteins purified from the *Petroselinum crispum* roots.

Fig. 2. Peptide sequences of the Pet c 1.021 protein purified from the *Petroselinum crispum* roots, where identified tryptic (red types) and chymotryptic (blue types) peptides are highlighted.

Fig. 3. Multiple sequence alignment of Pet c 1.021 (C0HKF5) purified from the *Petroselinum crispum* roots with: A – proteins with the highest sequence identity to Api g 2 - P92918 and Dau c 1 - AAL76932. B – PR-proteins previously found in *Petroselinum crispum* (PR1_PETCR - Q40795, PR2_PETCR - P27538, PR11_PETCR - P19417, PR13_PETCR - P19418), C – ribonuclease 1 (P80889) and ribonuclease 2 (P80890) from *Panax ginseng* with key amino acid residues implicated in RNase activity (Chadha and Das, 2006) shown in black boxes.

Fig. 4. Superpositions of optimized tertiary structures of the Pet c 1.0201 3D homology model and the Api g 1.0101 crystal structure (PDB accession 2BK0). A: blue – the template structure of Api g 1, red – the 3D model of Pet c 1.0201, both with annotated secondary structure elements: α – α-

helix, β – β-sheet, L – loop, and N- and C-terminals. B: monomers of Api g 1.0101 and C: monomers of Pet c 1.0201 after 200 ns MD simulations. Red indicates the position of Glu45 responsible for IgE binding.

Fig. 5. Inter-monomeric distances dependent on time of MD simulation with relative dispositions of monomeric units at the beginning (0 ns) and the end (200 ns) of each simulation. A: Api g 1.0101 in water; B: Pet c 1.0201 in water; C: Pet c 1.0201 in 0.2 M NaCl. Relative positions at 0 ns and after 200 ns for respective black and red trajectories are visualized; the blue trajectory shows a breakdown of a dimeric form into monomer units.

Fig. 6. Vizualization of amino acid residues (yellow) that stabilize dimers and are responsible for IgE binding (red) in A: Api g 1.0101 (template 2BK0); B: Pet c 1.0201 in water; C: Pet c 1.0201 in 0.2 M salt.

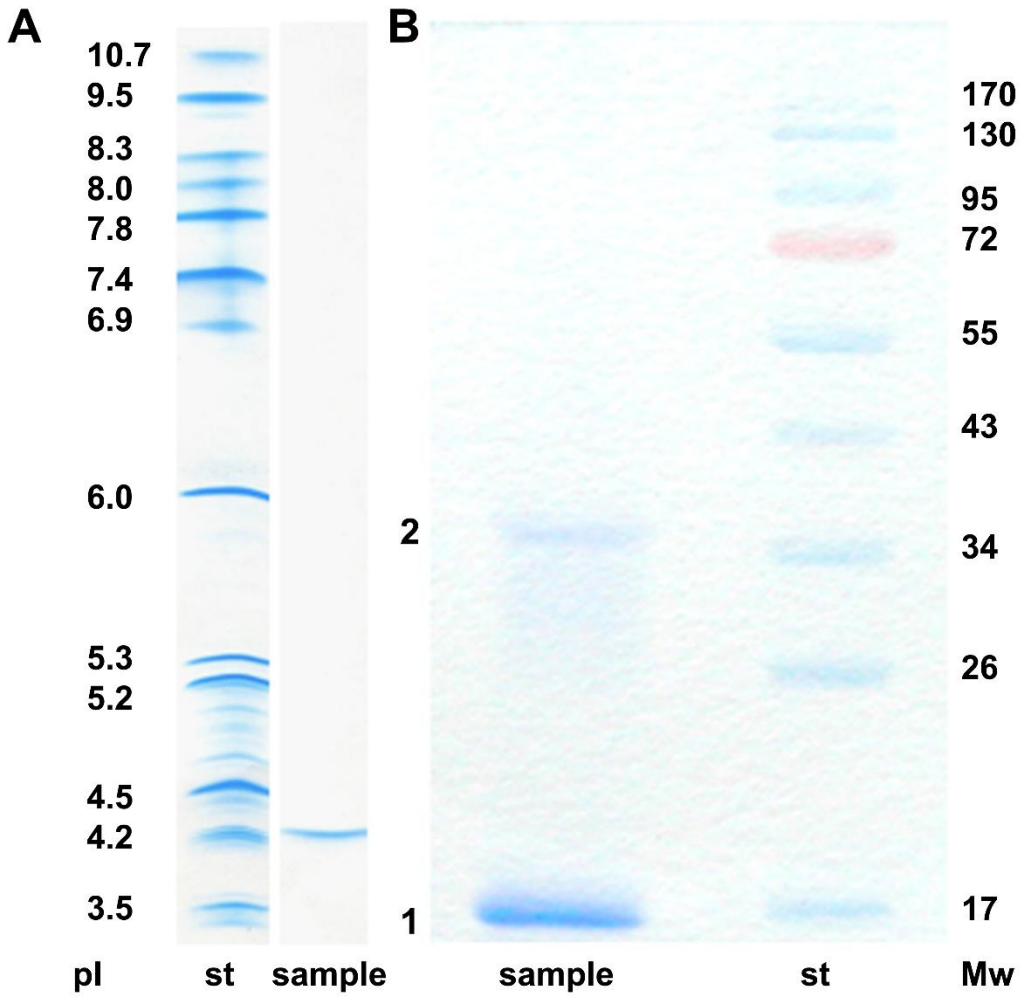
A
10.7
9.5
8.3
8.0
7.8
7.4
6.9
6.0
5.3
5.2
4.5
4.2
3.5
pI
st
sample
B
2
1
sample
170
130
95
72
55
43
34
26
17
st
Mw

Fig. 1.

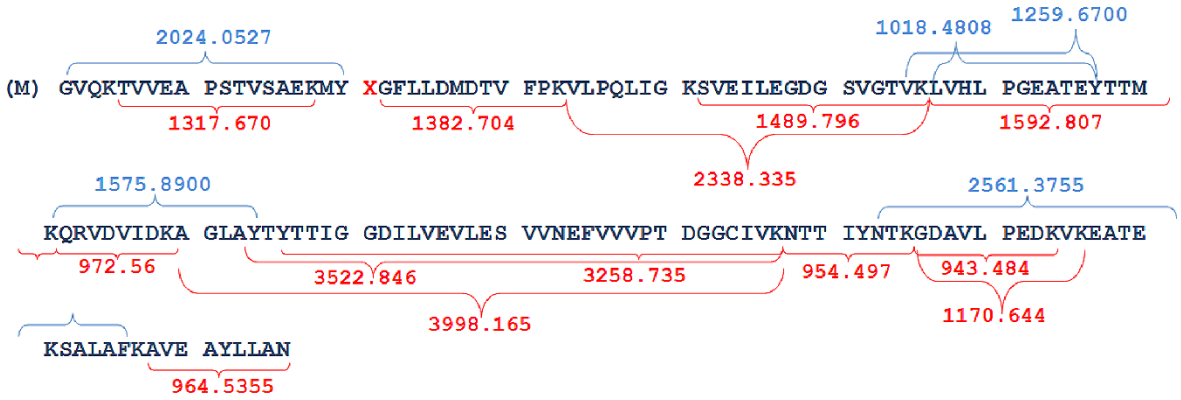
2024.0527
1018.4808
1259.6700
(M) GVQKTVVEA PSTVSAEKMY XGFLLDMDTV FPKVLPQLIG KSVEILEGDG SVGTVKLVHL PGEATEYTTM
1317.670
1382.704
1489.796
1592.807
2338.335
1575.8900
2561.3755
KQRVDVIDKA GLAYTYTTIG GDILVEVLES VVNEFVVVPT DGGCIVKNTT IYNTKGDAVL PEDKVKEATE
972.56
3522.846
3258.735
954.497
943.484
3998.165
1170.644
KSALAFKAVE AYLLAN
964.5355

Fig. 2.


A
C0HKF5 GVQKTVVEAPSTVSAEKMY-GFLLDMDTVFPKVLPQLIGKSVEILEGDGSVGTVKLVHLGEATEYTTMKQRVDVIDKAGLAYTYTTIGGDILVEVLESVVNEF 102
P92918 MGVQKTVVEAPSTVSAEKMYQGFLLDMDTVFPKVLPQLI-KSVEILEGDGGVGTVKLVHLGEATEYTTMKQKVDVIDKAGLAYTYTTIGGDILVDVLESVVNEF 103
AAL76932 MGVQKTEVEAPSTVSAEKMYQGFLLDMDTVFPKVLPQLI-KSVEILEGDGGVGTVRLVHLGEATEYTTMKQKVDVIDKAGLGYTYTTIGGDILVEGLESVVNQF 103
C0HKF5 VVVPTDGGCIVKNTTIYNTKGDAVLPEDKVKEATEKSALAFKAVEAYLLAN----- 153
P92918 VVVPTDGGCIVKNTTIYNTKGDAVLPEDKIKEATEKSALAFKAVEAYLLANLQFLA 159
AAL76932 VVVPTDGGCIVKNTTIYNTKGDAVLPEDKVKEATEKSALAFKAVEAYLLAN----- 154
B
P19417 MGVQKSEVETTSSVSAEKLFKGLCLDIDTLLPQVLPGAI-KSSETLEGDGGVGTVKLVHLGDASPFKTMKQKVDAIDKATFTYSYSIIDGDILLGFIESINNHF 103
P19418 MGVQKSEVEATSSVSAEKLFKGLCLDIDTLLPRVLPGAI-KSSETLEGDGGVGTVKLVHLGDASPFKTMKQKVDAIDKATFTYSYSIIDGDILLGFIESINNHF 103
Q40795 MGVQKSEVVITSSVPAAKLFKALCLDIDTLLPQVLPGAI-KGGEILEGDGGVGTVKLVTLGDASPYKTMKQKIDAIDKEAFTFSYSIIDGDILLGYIDSINNHL 103
C0HKF5 GVQKTVVEAPSTVSAEKMY-GFLLDMDTVFPKVLPQLIGKSVEILEGDGSVGTVKLVHLGEATEYTTMKQRVDVIDKAGLAYTYTTIGGDILVEVLESVVNEF 102
P27538 MGAVTTDVEVASSVPAQTIYKGFLLDMDNIIPKVLPQAI-KSIEIISGDGGAGTIKKVTLGEVSQFTVVKQRIDEIDAEALKYSYSIIEGDLLLGIIESITSKF 103
P19417 TAVPNADGGCTVKSTIIFNTKGDAVVPEENIKFANDQNLTIFKAVEAYLIAN---- 155
P19418 TAVPNADGGCTVKSTIIFNTKGDAVVPEENIKFANDQNLTIFKAVEAYLIAN---- 155
Q40795 SFVPTADGGCTATSTAVFNTKGDAVVPEENIKFANDQNNLIFKAVEAYLLAN---- 155
C0HKF5 VVVP-TDGGCIVKNTTIYNTKGDAVLPEDKVKEATEKSALAFKAVEAYLLAN---- 153
P27538 TVVP-TDGGCIVKNTTIYTPIGDAVIPEENVKEATEQSGMVFKAIEAYLLANPGAY 158
C
P80889 MGVQKTEVEATSTVPAQKLYAGLLLDIDDILPKAFPQAI-KSSEIIEGDGGVGTVKLVTLGEASQFNTMKQRIDAIDKDALTYTYSIIGGDILLDIIEKSIVNH 102
P80890 MGVQKTETQAISPVPAEKLFKGSFLDMDTVVPKAFPEGI-KSVQVLEGNGGVGTIKNVTLGDATPFNTMKTRIDAIDEHAFTYTYTIIGGDILLDIIEKSIENH 102
C0HKF5 GVQKTVVEAPSTVSAEKMY-GFLLDMDTVFPKVLPQLIGKSVEILEGDGSVGTVKLVHLGEATEYTTMKQRVDVIDKAGLAYTYTTIGGDILVEVIEKSVVNE 101
P80889 FTIVPTPDGGSIVKNTTIYNTIGDAVIPEENIKDATEKAGLIFKAVEAVLLAN 154
P80890 FKIVPT-DGGSTITQTTIYNTIGDAVIPEENIKDATDKSIQLFKAVEAVLLAN 153
C0HKF5 FVVVPT-DGGCIVKNTTIYNTKGDAVLPEDKVKEATEKSALAFKAVEAVLLAN 153

Fig. 3.

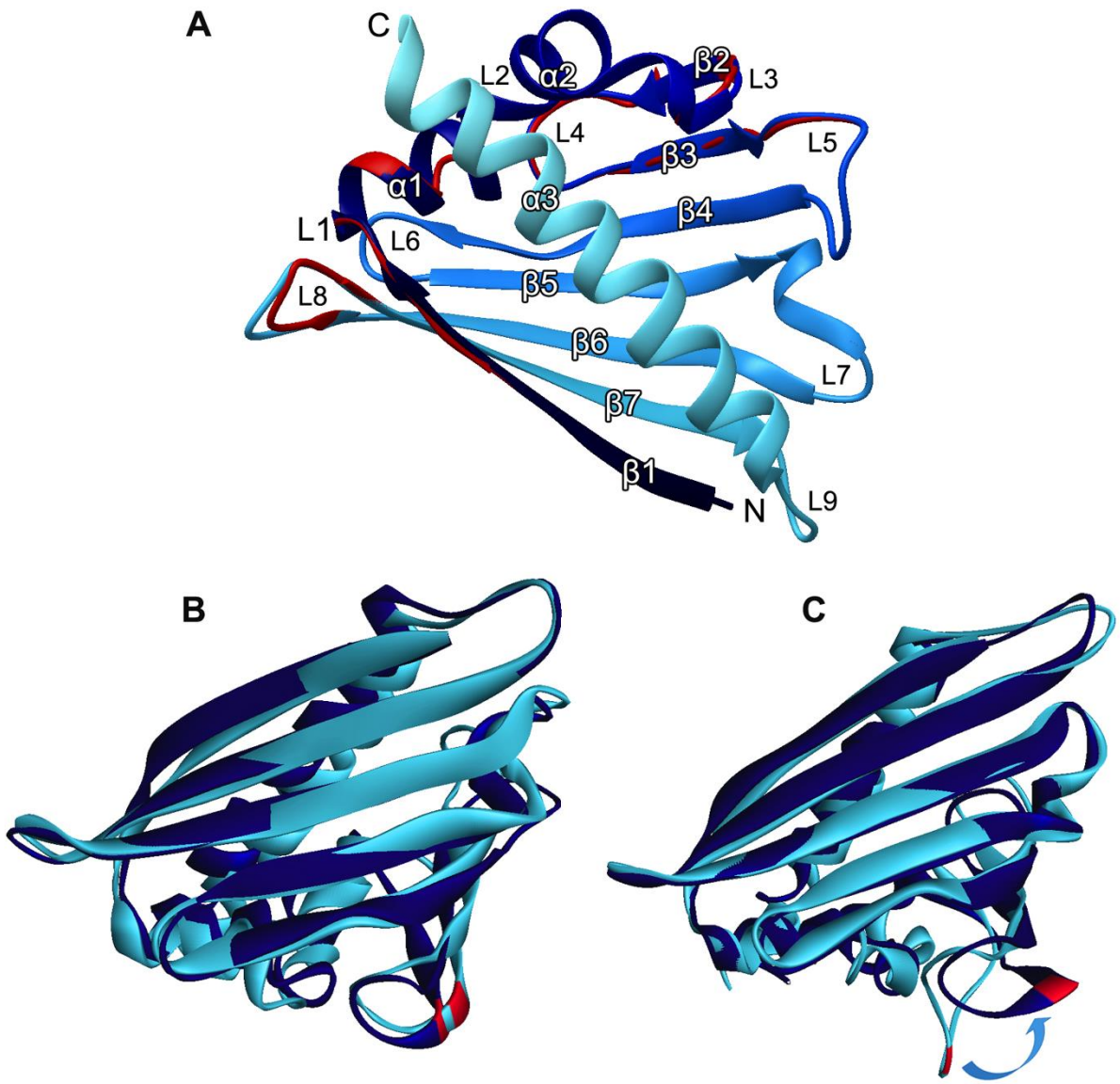

Fig. 4.

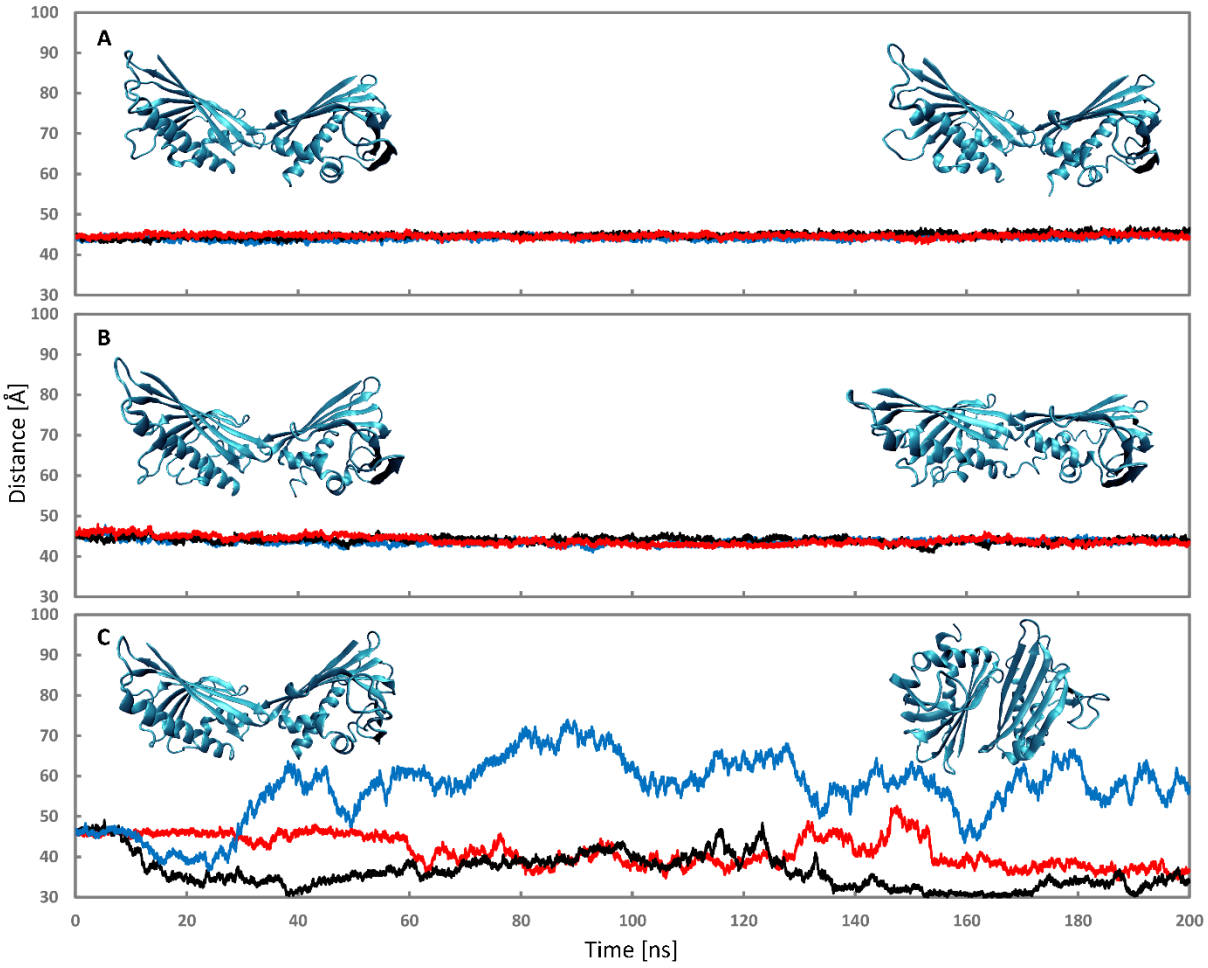
A
B
C
Distance [Å]
Time [ns]
100
90
80
70
60
50
40
30
0
20
40
60
80
100
120
140
160
180
200

Fig. 5.

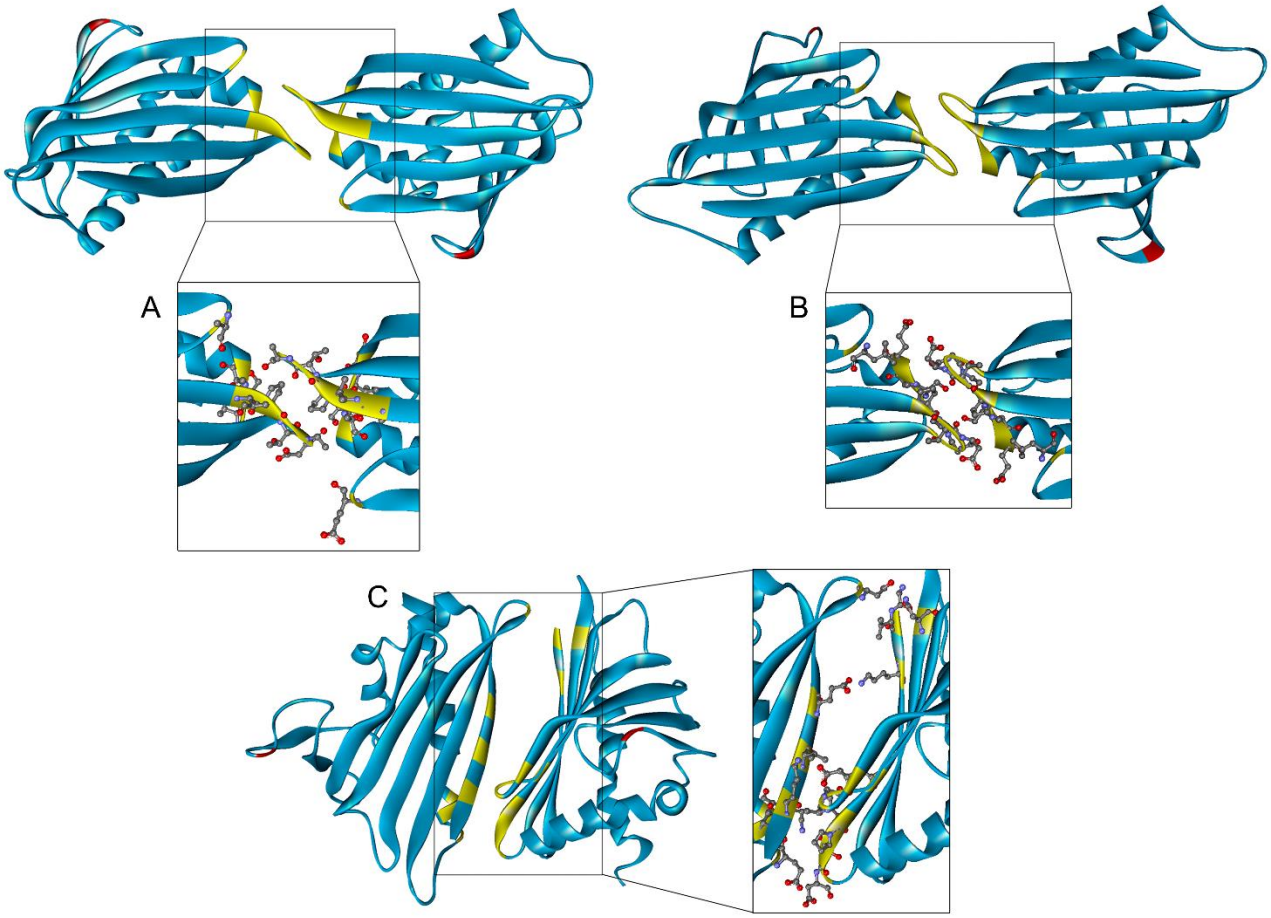

Fig. 6.