Crystal structure of timothy grass allergen Phl p 12.0101 reveals an unusual profilin dimer*

Andrea O’Malley1, A. Brenda Kapingidza1, Noah Hyduke1, Coleman Dolamore1, Krzysztof Kowal123 and Maksymilian Chruszcz1,2

1Department of Chemistry and Biochemistry, University of South Carolina, Columbia, SC 29208, USA; 2Department of Allergology and Internal Medicine, Medical University of Białystok, Białystok 15-276, Poland; 3Department of Experimental Allergology and Immunology, Medical University of Białystok, Białystok 15-276, Poland

INTRODUCTION

Timothy (Phleum pratense) is a perennial grass native to most of Europe and found in most of the United States. Timothy grass grows well in wet, humid areas and is quite winter-hardy (Ogle et al., 2011). It is grown commercially as a common animal fodder, especially for horses, other grazing animals like cattle, and for domesticated pets such as guinea pigs. It can also be used as erosion control in some regions (Ogle et al., 2011). In America, timothy is mainly cultivated in the Pacific Northwest (humid areas around the Puget Sound), the Northeast, and the Midwest (Hitchcock & Chase, 1951). Due to this ubiquitous nature of timothy grass, it is a potent source of allergens for many patients sensitized to the grass (Scarparotta et al., 2013; Rossi et al., 2008; Sekerkova et al., 2012; Almeida et al., 2019). Sensitization and reaction to pollen from timothy grass is a contributor to allergic rhinitis and other symptoms such as rhinoconjunctivitis. Due to high pollen counts and long pollen seasons, grasses are the most important seasonal allergen source in Europe and the USA (Kowal, 2020; Durham et al., 2012; Maloney et al., 2014). Moreover, high cumulative exposure to grass pollens each season leads to development of a robust IgE response to a wide range of grass proteins in allergic patients (Hatzler et al., 2012).

Currently, ten P. pratense allergens have been registered by the World Health Organization and International Union of Immunological Societies (WHO/IUIS) Allergen Nomenclature Sub-committee (www.allergen.org) (Radauer et al., 2014). Among them, Phl p 1 and Phl p 5 are considered the most important, as a majority of IgE in timothy grass-allergic individuals is directed against these molecules (Rossi et al., 2008). A significant fraction of the P. pratense allergens were characterized in terms of their molecular and structural properties, with structures determined for Phl p 1, Phl p 2, Phl p 3 Phl p 4, Phl p 5, Phl p 6, and Phl p 7 (Zafred et al., 2013; Padavattan et al., 2009; Henzl et al., 2013; Gohl et al., 2017; Winter et al., 2020; Mitropoulou et al., 2018; Schwieger et al., 2008; De Marino et al., 1999).

The profilin from timothy grass, Phl p 12, has been registered as a minor allergen. In different populations of grass pollen allergic patients, up to 35% show IgE binding to Phl p 12 (Cudowska et al., 2020; Kowal et al., 2020). Although Phl p 12 is regarded as a minor allergen, it displays T cell response prevalence and strength comparable to Phl p 1, a major allergen from P. pratense (Lund et al., 2018). Phl p 12 has three registered isoallergens, and only one of them (Phl p 12.0101) was characterized at the molecular level (Cudowska et al., 2020).

Profilins are small, ubiquitous proteins found extensively in plants. Their primary function involves binding actin in plant cytoskeletons and controlling its polymeri-
zation and communication (Jimenez-Lopez et al., 2012; Jimenez-Lopez et al., 2013). Profilins are included in the panallergen classification due to their wide range of functions, high likelihood of IgE cross-reactivity, and conserved sequences and structures (McKenna et al., 2016). Profilins are particularly noted for their marked similarity between protein sequences, with 57 registered plant allergens ranging from 65 to 92% sequence identity. As a protein family, profilins are primarily minor allergens and are known for a high likelihood of co-sensitization as well as cross-reactivity situations caused by their sequence and structure similarity (Chruszcz et al., 2018; Offermann et al., 2016).

Here we present a crystal structure of Phl p 12.0101, which reveals that this allergen may form an unusual dimeric structure that was not previously observed among any profilins. The structure of Phl p 12.0101 is discussed in the context of allergic sensitization, IgE binding, and allergy diagnostics, as well as in production of recombinant allergens.

MATERIALS AND METHODS

Protein purification and crystallization

Recombinant Phl p 12.0101 was produced and purified as previously described (Cudowska et al., 2020). Briefly, the Phl p 12.0101 construct with a cleavable N-terminal polyhistidine tag was produced in E. coli BL21 (DE3) cells. The protein was purified using a combination of immobilized metal affinity chromatography and size exclusion chromatography. 20 mM β-mercaptoethanol (β-ME) was used during protein purification; however, it was not included in the final size exclusion chromatography step. Cleavage of the purification tag was performed according to the previously described protocol (Booth et al., 2018). Purified protein was stored in buffer containing 50 mM Tris-HCl, 150 mM NaCl, pH 7.4.

Protein was crystallized at 277 K over the course of one year. Using the vapor diffusion method in a 4 µL drop, the best-diffraction crystals were obtained when 20 mg/mL protein solution was mixed 1:1 with crystallization solution containing 0.5 M sodium citrate at pH 6.5. Prior to data collection, crystals were cryocooled in liquid nitrogen.

Data collection, processing and structure determination

The South East Regional Collaborative Access Team beamline 22ID at the Advanced Photon Source (Argonne National Laboratory, Lemont, IL) was used to perform diffraction experiments. HKL-2000 software package was used for data processing (Otwinowski, 1997), and data collection statistics are reported in Table 1.

Phl p 12.0101 structure was solved by molecular replacement using MOLREP (Vagin & Teplyakov, 1997) integrated with HKL-3000 (Minor et al., 2006). The structure of maize profilin (Zea m 12; PDB code: 5FEF) was used as a search model. Initial model was rebuilt with Buccaneer (Cowtan, 2006), COOT (Emsley & Cowtan, 2004), and HKL-3000. REFMAC (Murshudov et al., 2011) and COOT were used for structure refinement. Translation-Libration-Screw (TLS) Motion Determination server (Painter & Merritt, 2006) was used to divide the protein model into segments, and TLS parameters were implemented in the final stages of the refinement process. COOT and MOLPROBITY (Davis et al., 2007) were used for validation of the Phl p 12.0101 model. Refinement and validation statistics are summarized in Table 1. The Phl p 12.0101 structure together with its structure factors was deposited to the PDB with accession code 7KYW. Diffraction images were deposited in the Integrated Resource for Reproducibility in Macromolecular Crystallography (https://proteindiffraction.org; Grabowski et al., 2019; Grabowski et al., 2016; with https://doi.org/10.18430/m3.irrmc.5715).

Various computational approaches

PDBePISA was used to investigate an oligomeric form of Phl p 12.0101 in crystal state (Krissinel & Henrick, 2007). PDBeFOLD (Krissinel & Henrick, 2004) and DALI (Holm & Rosenstrøm, 2010) servers were used to find proteins that have similar structures to Phl p 12.0101. Sequence alignment was prepared using Jalview (Waterhouse et al., 2009) and MUSCLE (Edgar, 2004). Figures showing protein structures were prepared with PYMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.). APBS implemented in PyMOL was used for electrostatics calculations (Jurrus et al., 2018).

Table 1. Data collection and refinement statistics. Values in parentheses are for the highest resolution shell.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Phl p 12.0101</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB accession code</td>
<td>7KYW</td>
</tr>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
</tr>
<tr>
<td>Diffraction source</td>
<td>APS, 22ID</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.0000</td>
</tr>
<tr>
<td>Space group</td>
<td>P3121</td>
</tr>
<tr>
<td>a, c (Å)</td>
<td>56.3, 94.4</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>40.00-2.30 (2.34-2.30)</td>
</tr>
<tr>
<td>No. of unique reflections</td>
<td>8119 (405)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.8 (100.0)</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.066 (0.766)</td>
</tr>
<tr>
<td>Rp.lim</td>
<td>0.021 (0.825)</td>
</tr>
<tr>
<td>CC1/2</td>
<td>0.942 (0.825)</td>
</tr>
</tbody>
</table>

| **Refinement** | |
| Resolution range (Å) | 33.94-2.30 (2.36-2.30) |
| Completeness (%) | 99.7 (99.0) |
| No. of reflections, working set | 7700 (556) |
| No. of reflections, test set | 395 (29) |
| Final Rcryst | 0.197 (0.302) |
| Final Rfree | 0.242 (0.400) |
| Rmsd bonds (Å) | 0.014 |
| Rmsd angles (°) | 1.5 |
| Ramachandran Plot | |
| Allowed regions (%) | 98.0 |
| Favored regions (%) | 100.0 |
RESULTS AND DISCUSSION

Phl p 12.0101 – overall structure

Recombinant Phl p 12.0101 crystallized in P3,21 space group with one molecule in the asymmetric unit, and the protein model was refined at 2.3 Å resolution. The model contains all residues forming the mature form of the protein (residues 2–131), as well as glycine (residue 1) that corresponds to a fragment of the purification tag. The overall fold of the protein chain is very similar to that observed in other profilins that have their structures determined. The structure reported here superposes very well (over the whole sequence; rmsd values ~1 Å) with other models of plant profilins such as profilins 1–3 from Arabidopsis thaliana (Qiao et al., 2019, Thorn et al., 1997), Hevea brasiliensis (Hev b 8.0102) (Galicía et al., 2015; Mares-Mejia et al., 2016), muskmelon (Cuc m 2.0101) (Kapingidza et al., 2019), mugwort (Art v 4.0101), and ragweed (Amb a 8.0101) (Offermann et al., 2016). The core of the molecule includes an antiparallel beta sheet composed of five strands (β2↑β1↓β7↑β6↓β5↑), which is flanked from one side by two helices (H1 and H3), and from the other side by another helix (H2) and a hairpin motif (β384) (Fig. 1). However, helix H3 in Phl p 12.0101 is shorter in comparison to the equivalent helices from plant profilins, such as that observed in mugwort profilin Art v 4.0101 (Fig. 1B). In Phl p 12.0101, an initial fragment of helix H3 is unwound and Cys115, which in other plant profilin structures is pointing toward the protein core, is facing toward the protein surface.

The structural analysis revealed that both Cys13 and Cys115 form intermolecular disulfide bridges (Cys13-Cys115’ and Cys115-Cys13; Fig. 2) that are responsible for Phl p 12.0101 dimer formation (Fig. 3). Symmetry of the dimer coincides in this case with crystal symmetry. The dimer interface has an area of 709 Å², and this molecular assembly, in addition to the two covalent bonds, is stabilized by a mixture of hydrophobic and hydrogen bonding interactions. Each molecule forming the dimer contributes approximately 15 residues to the interface.
and most of these residues are highly conserved among profilins that are reported as allergens. However, the presence of dimerization in the crystal structure of Phl p 12.0101 was unexpected, as initial stages of protein purification were performed in reducing conditions, and the protein was determined to be monomeric prior to crystallization (Cudowska et al., 2020). Most likely, oxidation of the protein during the crystallization process, which took several months, resulted in dimer formation.

Phl p 12 has three officially registered isoallergens that have the same length and vary by just a few residues. Phl p 12.0102 can be considered a Phl p 12.0101 mutant in which arginine that is present in the 12.0101 isoallergen is substituted by alanine. Phl p 12.0103 has four differing residues from Phl p 12.0101 (97% sequence identity), including R81A as seen in Phl p 12.0102. Phl p 12.0101 and Phl p 12.0103 also feature differences between leucine and phenylalanine (L26F and F54L) and one change from glycine to alanine (G69A) in comparison to Phl p 12.0101 and Phl p 12.0102. Both cysteines (C13 and C115) found in Phl p 12.0101 are found in Phl p 12.0102 and Phl p 12.0103, as well as other residues forming the dimer interface, indicating that these isoallergens may form dimers identical to the one observed in Phl p 12.0101 crystal.

Comparison of Phl p 12.0101 with other profilins

As already mentioned, the overall fold of Phl p 12.0101 is very similar to that observed for other profilins. However, the observed conformation of Phl p 12.0101 fragment containing Cys115 is very unusual and was not reported previously (Fig. 1). Similarly, the participation of the conserved Cys115 (Fig. 2) in formation of an oligomeric assembly in plant profilins was also not observed. At the same time, profilins have been shown to exist as monomers or higher-order oligomers (Mares-Mejia et al., 2016; Mittermann et al., 1998; Wopfner et al., 2002; Willerroider et al., 2003). It was suggested that this may be accomplished through disulfide bond formation between two or more identical profilin molecules. Phl p 12.0101 as presented here is the second profilin allergen with its structure solved as a homodimer, following Hev b 8.0102 (Mares-Mejia et al., 2016). However, there are significant differences between the dimers formed by Phl p 12.0101 and Hev b 8.0102. For example, while the Phl p 12.0101 dimer is stabilized by two disulfide bridges, only one such covalent link is present in Hev b 8.0102 (Fig. 3). In addition, despite the fact that most of the residues forming the dimer interfaces originate from the N- and C-terminal fragments of the protein chains (including helices H1 and H3), the arrangements of the molecules in the dimers are different (Fig. 3). At the same time, residues that mediate dimer formation in Hev b 8.0102 and Phl p 12.0101 are highly conserved, suggesting that other plant profilins may form such oligomeric assemblies.

Plant profilins that are reported as allergens have at least two cysteine residues. Until now, all structural studies indicated that only residues equivalent to Cys13 of Phl p 12.0101 were exposed to solvent. The solvent-exposed thiol group was shown to be prone to modifications and disulfide bond formation, like in the case of Hev b 8.0102 (Mares-Mejia et al., 2016). Moreover, reducing agents such as 2-mercaptoethanol (β-ME) and dithiothreitol (DTT) were used in purification of these proteins and may have an effect on profilin stability (Soh et al., 2017; Mittermann et al., 1998). Under reducing conditions, profilin exists nearly exclusively in the monomeric form (Cudowska et al., 2020; Kapingidza et al., 2019; Offermann et al., 2016). However, when the profilin is under non-reducing conditions (in buffer without the reducing agent), the profilin has the potential to oligomerize. Bet v 2, the profilin from birch, was shown in SDS-PAGE to form dimers and trimers, with monomers in higher concentration than the dimeric or trimeric form (Mittermann et al., 1998). In non-reducing conditions, Hev b 8 forms primarily monomers and dimers, with the suggestion that it forms transient oligomers during gel filtration (Mares-Mejia et al., 2016). Art v 4 was shown to form dimers, trimers, and tetramers under non-reducing conditions (Wopfner et al., 2002). Art v 4 has an extra cysteine residue (C95) in addition to the conserved C13 and C115 observed in Phl p 12 (Fig. 1), Hev b 8, and Bet v 2, possibly making oligomerization easier for...
Art v 4 in comparison to the other profilins. Similar to Hev b 8, varying oligomerization was observed when Art v 4 tetramers were converted back to stable dimers upon the addition of the reducing agent DTT (Wopfner et al., 2002). Interestingly, in contrast to other profilins that seem to exist primarily as monomers, Cap a 2 (Capsicum annuum, various peppers) was found almost exclusively in dimeric form with size exclusion chromatography, and the addition of DTT acted only partially to convert it to a monomer (Willerroider et al., 2003).

**Profilin oligomerization in the context of biological function**

Profilins interact with many proteins, and some of these interactions are mediated by binding to sequences that include proline-rich regions (Witke, 2004). It was shown in the dimeric structure of Hev b 8.0102 that the regions responsible for binding of proline-rich peptides partially overlap with the oligomerization interface (Kapingidza et al., 2019), which is also true for the Phl p 12.0101 dimer. Superposition of the profilin dimer structures with a structure of Amb a 8 in complex with polyproline (PDB code: 5EVO) and with a complex of human profilin-1-poly-Pro-actin (PDB code: 3CHW; Fig. 4) suggests that the dimers may potentially interact with relatively short proline-rich peptides (~ six residues) through the part of the binding site available in monomeric forms of the proteins. However, there is currently no information on the ability of the profilin dimers to interact with proline-rich peptides. If profilins can oligomerize in cells (Babich et al., 1996), it is plausible that the different forms of these proteins may have modified affinity towards binding partners (Kapingidza et al., 2019). Figure 4 also shows that dimeric assemblies of Hev b 8.0102 and Phl p 12.0101 are most likely not able to bind actin. In this case, more potential clashes are observed for the Hev b 8 dimer, due to its bent structure (Fig. 5).

Plant profilins are encoded by multigene families (Kovar et al., 2000). Therefore, several different isoforms of this protein may be present in the same organism, and the isoforms are not functionally alike. Phl p 12.0101 belongs to the profilin-1 class, while Hev b 8.0102 is a member of the profilin-2 class (Kovar et al., 2000). Therefore, one can speculate that the different dimeric assemblies observed for these proteins may be representative of different profilin classes. Despite the structural differences, there are large patches of surface that are negatively charged in both dimers (Fig. 3) which can be potentially involved in interactions with positively charged molecules. Similarly, positively charged surface areas may participate in interactions with negatively charged molecules. For example, it was demonstrated that profilins can interact with phosphatidylinositol 4,5-bisphosphate (PIP2) and inhibit hydrolysis of this molecule by phospholipase C (Goldschmidt-Clermont et al., 1990). According to UniProt (The UniProt, 2017), the Phl p 12.0101 region responsible for PIP2 binding corresponds to the Arg81-Thr97 fragment that contains several positively charged residues. Analysis of the profilin dimers’ structures indicates that the putative PIP2 binding site is relatively far from the dimerization interfaces, and therefore the dimeric Hev b 8.0102 and Phl p 12.0101 may interact with the lipid.

**Structure of Phl p 12.0101 in relation to allergy**

Although plant profilins recognized as allergens are considered to exist mainly as monomeric proteins, it was demonstrated that oligomerization has an impact on their allergenicity. For example, it was shown that Hev b 8.0102 dimers and monomers were cross-reactive when tested with an IgE mAb (Mares-Mejia et al., 2020); however, higher basophil degranulation in a rat basophilic leukemia cell line was observed with the dimeric form of the allergen as compared to the monomer due to increased cross-linking between bound mAb and FceRI receptors (Mares-Mejia et al., 2020). This may be related to the structure of FcεRI-IgE complexes and the distance between cross-linked IgE molecules required for optimal mast cell activation (Knol, 2006). Moreover, it has recently been demonstrated for polcalcins that multivalency of oligomers can overcome the need for high-affinity interactions necessary for effective mast cell activation (Bucai et al., 2019). Dimer formation of Hev b 8.0102 also exposed known IgE-binding epitopes on the surface of the homodimer (Mares-Mejia et al., 2016), leaving the dimer able to bind more IgE than the monomer. A similar situation with changed IgE-binding epitopes may be possible with other allergenic profilins that can form dimers, such as Phl p 12, Cuc m 2 and Amb a 8.

While there is experimental evidence pointing to differences between IgE binding of monomeric and di-
meric forms of profilins, it is still not clear whether the human body is exposed to the monomeric form of these allergens or a mixture of monomeric and oligomeric forms. Interestingly, reported clinical significance of sensitization to profilins differs in individual patients, ranging from profilin acting as a molecule of no clinical relevance to an active aeroallergen and to a trigger of a systemic anaphylaxis (Rodriguez Del Río et al., 2018). The variable clinical response in patients sensitized to profilins may depend on co-factors which can possibly affect the ability of those molecules to be recognized by IgE on mast cells and basophils (Rodriguez Del Río et al., 2018). Moreover, it must be stressed that the process of profilin purification can introduce a significant bias and result in samples that do not well represent the relative abundance of different oligomeric forms observed in the starting source material. In fact, profilins are often purified from natural sources using protocols that involve denaturing conditions and refolding. Such protocols may lead to formation of oligomeric assemblies that are mediated by intermolecular disulfide bridges. On the other hand, use of affinity columns that contain polyl-Pro may lead to selection of profilin forms that have high affinity toward the proline-rich peptides. Additional complications can be related to the presence or absence of various reducing agents that can clearly shift equilibrium between monomeric and oligomeric forms of profilins. Therefore, purification and storage conditions of natural and recombinant profilins that are used in allergy research or diagnostics have to be taken into consideration.

Analysis of 69 plant profilins that are registered as allergens by WHO/IUIS shows that all of these proteins have two conserved Cys residues (Cys13 and Cys115 in Aca f 2.0101). In 15 of these proteins, there are three Cys residues (Ama r 2.0101, Amb a 8.0101, Amb a 8.0102, Art v 4.0101, Art v 4.0201, Hel a 2.0101, Hor v 12.0101, Koc s 2.0101, Pro j 2.0101, Sal k 4.0101 and Sal k 4.0201, as they all have an equivalent of Cys95. Most likely this group of profilins does not form the dimeric assembly that was revealed for Phl p 12.0101, because formation of an intermolecular disulfide bond. In the case of Hor v 12.0101 and the Tri a 12 isoallergens, the additional cysteine residue precedes Cys13, and there is no experimental data on its involvement in formation of disulfide bridges. However, its localization suggests that the thiol group of this residue should be solvent-exposed and prone to modification.

Aca f 2.0101 has four cysteine residues, which follow the general pattern observed in Amb a 8.0101 and Art v 4.0101 with additional N-terminal Cys. Again, in this case it is not known whether the N-terminal cysteine participates in Aca f 2.0101 oligomerization. Similarly, it is not known whether the four cysteine residues present in the profilin from banana (Mus a 1.0101) participate in formation of intramolecular or intermolecular disulfide bridges.

In summary, the oxidative state of cysteine residues in profilins that were identified as allergens requires more study, as it may be important from the perspective of recombinant protein production and purification of profilins from natural sources, as well as for our understanding of allergy. Moreover, as recombinant proteins are used extensively in allergy diagnostics, the conditions of their storage should prevent unwanted modifications. In addition, characterization of allergenic profilins originating from natural sources is needed to provide information on postranslational modifications and oligomeric states that are present in this group of proteins and that may presumably have impact on their interactions with the human immune system.

Conflict of Interest

A.O., A.B.K., N.H, C.D., and M.C. declare no conflict of interest. K.K. received lecture fees from: ALK Abello,
Astra Zeneca, Berlin Chemie, Chiesi, Emma, Hal Allergy, Meda Pharma, Orion Pharma, and royalties from UpToDate.

REFERENCES


The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC


