Published in final edited form as:

Proteins. 2009 July ; 76(1): 237–243. doi:10.1002/prot.22396.

Structures of Two *Arabidopsis thaliana* Major Latex Proteins Represent Novel Helix-Grip Folds

Betsy L. Lytle^{1,2}, Jikui Song^{1,3}, Norberto B. de la Cruz^{1,2}, Francis C. Peterson^{1,2}, Kenneth A. Johnson^{1,3}, Craig A. Bingman^{1,3}, George N. Phillips Jr.^{1,3}, and Brian F. Volkman^{1,2,*}

¹Center for Eukaryotic Structural Genomics, USA

²Department of Biochemistry, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, Wisconsin 53226, USA

³Department of Biochemistry, University of Wisconsin-Madison, 433 Babcock Drive, Madison, Wisconsin 53706, USA.

Abstract

The major latex proteins (MLP) are a protein family first identified in the latex of opium poppy. They are found only in plants and have 24 identified members in *Arabidopsis* alone as well as in other plants such as peach, strawberry, melon, cucumber, and soybean. While the function of the MLPs is unknown, they have been associated with fruit and flower development and in pathogen defense responses. Based on modest sequence similarity, they have been characterized as members of the Bet v 1 protein superfamily; however, no structures have yet been reported. As part of an ongoing structural genomics effort, we determined the structures of two *Arabidopsis thaliana* MLPs: the solution structure of MLP28 (gene product of At1g70830.1) and the crystal structure of At1g24000.1. The structures revealed distinct differences when compared to one another and to the typical Bet v 1 fold. Nevertheless, NMR titration experiments demonstrated that the characteristic Bet v 1 hydrophobic binding pocket of At1g24000.1 is able to bind a ligand, suggesting that it plays a role in the function of the MLPs. A structure-based sequence analysis identified conserved hydrophobic residues in the long alpha helix that contribute to the binding cavity and may specify preferred ligands for the MLP family.

Keywords

MLP; PR-10; IPR; Bet v 1-like; helix grip fold; NMR; X-ray crystallography; protein structure; structural genomics

INTRODUCTION

The major latex proteins (MLP) are a protein family first identified in the latex of opium poppy (*Papaver somniferum*)¹. They are found only in plants and have 24 identified members in *Arabidopsis* alone as well as in other plants such as peach, strawberry, melon, cucumber, and soybean. While the function of the MLPs is unknown, they have been associated with fruit and flower development and in pathogen defense responses. Based on modest sequence similarity, they have been characterized as members of the Bet v 1 protein superfamily²; however, no MLP structures have yet been reported. Additional sequence comparisons have expanded the Bet v 1 superfamily to include the (S)-norcoclaurine

*To whom correspondence should be addressed: Brian F. Volkman, Medical College of Wisconsin, Department of Biochemistry, 8701 Watertown Plank Road, Milwaukee, Wisconsin 53226, Tel: (414)-456-8400, Fax: (414)-456-6510, bvolkman@mcw.edu.

synthases and cytokinin-specific binding proteins (CSBP) as well as the MLP and intracellular pathogenesis-related class 10 (PR-10) families 3·4. PR-10 proteins, which include tree pollen allergens and major food allergens, are expressed in various tissues and organs in response to pathogen attack as well as environmental stresses such as drought, wounding, ultraviolet radiation, and oxidative stress. The Bet v 1 superfamily is further classified as a member of the Pfam clan, Bet v 1-like (CL0209), which includes the following diverse families outside the plant kingdom: ASHA1, COXG, IP trans, Polyketide cyc, Ring hydroxyl A, and START.

More than 100 Bet v 1 superfamily sequences have now been identified 5, and three-dimensional structures have been determined for at least twelve of these proteins: birch pollen allergen Bet v 1 6, Bet v 11, a hypoallergenic isoform of Bet v 1 7, major cherry allergen Pru av 18, two proteins from yellow lupine subclass LIPR-10.1 9 and two from subclass LIPR-10.2 10·11, *Pachyrrhizus erosus* (jicama) SPE-16 (PDB code 1TW0), major celery allergen Api g 1 12, CSBP from mung bean 13·14, and norcoclaurine synthase from *Thalictrum flavum* (2VNE, 2VQ5).

The Bet v 1 fold consists of a curved seven-stranded β -sheet wrapped around a long C-terminal helix, α 3, and has been classified as a type of “helix-grip” fold 15. Between these two structural elements is a large, Y-shaped hydrophobic cavity, which is closed at one end by two short helices, α 1 and α 2, that connect strands β 1 and β 2. The role of this hydrophobic cavity as a ligand-binding pocket was first suggested based on structural similarities between the hydrophobic cavity of Bet v 1 and the cholesterol-binding pocket of the steroidogenic acute regulatory (StAR)-related lipid transfer (START) domain of the human protein MLN64 15·16. Based on this hypothesis, NMR titration experiments of Pru av 1 binding to homocastasterone were conducted and provided the first evidence that this cavity could bind plant steroids 8. More recently, structures of Bet v 11 in complex with deoxycholate 7, a compound structurally similar to the plant hormones, brassinosteroids, and CSBP in complex with the cytokinin, zeatin, 13 enabled the determination of specific binding interactions within the hydrophobic pockets. In both cases, the hydrophobic binding pocket could accommodate two ligand molecules. In addition to the hydrophobic pocket, the Bet v 1 fold is characterized by a rigid and highly conserved glycine-rich loop, or so-called “P-loop” motif, which has the sequence GxGGxGT. In studies of the complex between Bet v 1 and an IgG Fab' fragment from mouse (human IgE mAbs are difficult to obtain in sufficient quantities from allergic patients), the P-loop sequence was shown to be contained within the binding surface, suggesting that its rigid conformation helps define the IgE epitope 17·18.

Here we report the first structures of two MLP proteins, which display unique structural differences from the canonical Bet v 1 fold described above. MLP28 (SwissProt/TrEMBL ID Q9SSK9), the product of gene At1g70830.1, and the At1g24000.1 gene product (SwissProt/TrEMBL ID P0C0B0/Q93VR4), proteins which share 32% sequence identity, were independently selected as fold-space targets by the Center for Eukaryotic Structural Genomics. The structure of a single domain (residues 17–173) of MLP28 was solved by NMR spectroscopy, while the full-length At1g24000.1 structure was determined by X-ray crystallography. MLP28 displays greater than 30% sequence identity to at least eight MLPs from other species. For example, the MLP28 sequence shares 64% identity to peach Pp-MLP119 and 55% identity to cucumber Csf2 20. In contrast, the At1g24000.1 sequence is highly divergent (Fig. 1), containing a gap of 33 amino acids when compared to all other known MLPs. Even when the gap is excluded, the sequence identity with MLPs from other species is less than 30%. Unlike some of the MLPs from other species, none of the *A. thaliana* MLPs have been characterized biochemically. We show by NMR chemical shift mapping that At1g24000.1 binds progesterone, demonstrating that despite its sequence

dissimilarity, the hydrophobic binding pocket is conserved and, therefore, may play a role in its biological function and that of the MLP family in general.

MATERIALS AND METHODS

Cloning, protein expression and purification

A protein construct corresponding to residues 17–173 of *A. thaliana* gene At1g70830.1 (MLP28) was prepared in [U - ^{13}C , ^{15}N]-labeled form according to CESC wheat germ cell-free protocols as previously described 21. Briefly, the protein was expressed with an N-terminal His₆ fusion tag in wheat germ extract supplemented with [U - ^{13}C , ^{15}N] amino acids (Cambridge Isotope Labs) and purified by Ni-NTA affinity chromatography, followed by size-exclusion chromatography. At1g24000.1 was prepared by heterologous expression in *E. coli* as previously described 22,23.

NMR spectroscopy

The sample used for NMR contained ~0.7 mM [U - ^{13}C , ^{15}N] MLP28, 10 mM deuterated bis-Tris (pH 7.0), and 5 mM dithiothreitol in 95% H₂O and 5% $^2\text{H}_2\text{O}$. All NMR data were acquired at 25°C on a Bruker 600 MHz spectrometer equipped with a triple-resonance CryoProbe™ and processed with NMR Pipe software (Delaglio et al. 1995). Chemical shift assignments and 3D structure determination of MLP28, and chemical shift mapping of progesterone binding to At1g24000.1 were performed as detailed in Supplemental Methods.

Crystal structure determination and refinement

The At1g24000.1 protein was subjected to the UW192 crystallization screen and the structure was solved using SAD methodology as described in Supplemental Methods.

Deposition of NMR data and atomic coordinates

Coordinates and restraints have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.pdb.org/>) under PDB code 2I9Y (MLP28) and PDB code 1VJH (At1g24000.1). All time-domain NMR data and chemical shift assignments have been deposited in BioMagResBank (<http://www.bmrb.wisc.edu/>) under BMRB accession number 15720 for MLP28.

RESULTS AND DISCUSSION

Construct design and structure determination

The sequence of MLP28 consists of two domains which share 93% sequence identity with each other 24. Only one other *A. thaliana* MLP family member, MLP34, contains two duplicated domains. This gene is considered to be a paralog of MLP28, along with 3 other genes, predicted to have arisen from tandem duplications within the *A. thaliana* genome in an analysis of phylogenetic relationships and genomic positional data 24. We determined the solution structure of one domain of MLP28 (referred to herein simply as MLP28), corresponding to residues 17–173, after unsuccessful attempts to crystallize the entire protein. Under the solvent conditions used, MLP28 was a stable, folded monomer. However, increased ionic strength led to a tendency for the protein to aggregate into higher order oligomers (data not shown). An automated procedure for iterative NOE assignment 25 was used to calculate the structure, and molecular dynamics calculations in explicit solvent were used in the final refinement 26. The final conformers were generated from 1766 non-redundant NOE derived distance restraints and 197 dihedral angle constraints (Supplemental Table 1).

Structure determination efforts on full-length At1g24000.1 were conducted in parallel by both NMR and X-ray crystallography. Nearly complete ^1H , ^{15}N and ^{13}C resonance assignments were determined, but further work on the solution structure was abandoned once the crystal structure was solved. The At1g24000.1 crystal belongs to the space group $P2_1$ with unit cell parameters of $a = 45.68 \text{ \AA}$, $b = 34.26 \text{ \AA}$, $c = 78.56 \text{ \AA}$, $\alpha = 90.00^\circ$, $\beta = 90.01^\circ$, and $\gamma = 90.00^\circ$. The asymmetric unit contains two protein molecules. It is unknown whether this dimer is biologically relevant.

Structure descriptions

The MLP28 and At1g24000.1 structures both display the helix-grip fold¹⁵, albeit with significant differences from the typical Bet v 1 fold. The MLP28 structure consists of a six-stranded β -sheet enfolding a 21-amino acid C-terminal α -helix (Figure 2A and B). A short α -helix is positioned at one end of a large hydrophobic cavity and a single turn of 3_{10} helix, in the loop between strands $\beta 4$ and $\beta 5$, is found at the other end.

The MLP28 structure deviates significantly from the conserved Bet v 1 fold in that the $\beta 2$ strand and the second short α -helix between strands $\beta 1$ and $\beta 2$ of the Bet v 1 fold are replaced by a long, flexible loop comprising residues 45–65. Heteronuclear NOE values of less than 0.6 for residues 45–60 confirm that the loop is mobile and highly unstructured (Fig. 2C). The amide resonances of residues 62–65 were undetected; therefore, these ^{15}N - ^1H NOE values could not be determined. The glycine-rich loop (residues 66–72), which is adjacent to the unstructured loop, is somewhat rigid as indicated by ^{15}N - ^1H NOE values of 0.6–0.8 and lower rmsds than residues 45–65. Strands $\beta 2$ and $\beta 3$ (residues 76–96), which stretch out across the top of the molecule perpendicular to the C-terminal helix, display higher rmsd values. However, ^{15}N - ^1H NOE values similar to the rest of the molecule suggest that the strands fluctuate on a slower timescale (millisecond to microsecond) as a rigid body. A potential salt bridge between His81 (in the loop between $\beta 2$ and $\beta 3$) and Glu160 (in the long α -helix) may serve to stabilize the strands, which serve as a lid over the largest hydrophobic cavity opening. In the NMR structure of Pru av 1, the equivalent loop is much shorter and shows significant internal flexibility as indicated by low ^{15}N - ^1H NOE values⁸. This loop is also poorly-defined in both the solution and crystal structures of Bet v 1⁶. In contrast, the crystal structure of yellow lupine LIPR-10.2B in complex with zeatin contains a salt bridge in the same position as that of MLP28 that similarly serves to “gate” the cavity opening¹¹.

The At1g24000.1 crystal structure displays a fold similar to that of MLP28, consisting of a five-stranded β -sheet and a 21-amino acid C-terminal α -helix (Fig. 2D). Like MLP28, it also has a single short α -helix ($\alpha 1$) and an equivalent 3_{10} helix. The two structures can be superimposed with an r.m.s.d of $1.6 \text{ \AA}$ for 102 aligned C^α atoms (Fig. 2E). However, as noted earlier, the At1g24000.1 structure differs significantly from MLP28, as well as the Bet v 1 fold, due to the deletion of 33 amino acids following the $\alpha 1$ helix. Compared to MLP28, the majority of the flexible loop, the glycine-rich loop, and the $\beta 2$ strand are absent. As a result, the hydrophobic cavity is noticeably smaller than that of MLP28. Also, there is only a small loop between $\alpha 1$ and the equivalent $\beta 3$ strand, rather than a long flap covering the cavity. Additional electron density observed in the cavity was ascribed to low molecular weight polyethylene glycol.

Structural homologs of MLP28 and sequence analysis

A search for structures homologous to MLP28 using the Secondary Structure Matching (SSM) server (<http://www.ebi.ac.uk/msd-srv/ssm/>) yielded At1g24000.1 as the top match (Q-score = 0.43). Alignment statistics for the next closest structural neighbors, CSBP from mung bean in complex with zeatin, SPE-16, yellow lupine LIPR10.1B and LIPR10.2B, and

Bet v 1 are summarized in Supplemental Table 3, and representative homologs are depicted in Fig. 3A. MLP28 and At1g24000.1 display low sequence homology to the most similar Bet v 1 superfamily structures (20% identity or less). A structure-based sequence alignment of representative MLPs and the homologous structures highlights some interesting sequence differences (Fig. 3B). In particular, a conserved cluster of residues is found at the closed end of the hydrophobic pocket consisting of Tyr115, Lys116, Ser117, Phe118, Glu142, and Lys143 (MLP28 numbering), which is not conserved in the PR-10 structures. CSBP differs from the other two structures in that Ser117 is conserved and Phe118, Glu142 and Lys143 are conservatively substituted with Tyr, Asp, and His, respectively. A number of other residues in the MLP28 cavity, such as Trp77, Glu90, Ile102, and Phe118, are conserved or conservatively substituted in CSBP but not in the others. Notably, the residues in CSBP corresponding to Trp77 and Glu90 are ligand-binding residues.

These patterns of sequence conservation suggest that the MLPs are more closely related to CSBP than the other PR-10 proteins are. It is interesting, however, that none of these four residues are strictly conserved in At1g24000.1. Additional residues that are conserved in the MLPs, including At1g24000.1, are found just before and within the long α helix (Pro151, Leu155, and Ile165). The proline likely serves a structural role in determining the turn of the main chain and thereby positioning the helix within the groove of the β -sheet. Conservation of the two hydrophobic residues appears to be significant, especially because the helix is the most variable region, both in structure and sequence, among the PR-10 proteins. This suggests that these residues are important in determining the binding specificity of the MLPs.

Another key sequence difference between the MLPs and the other Bet v 1 proteins is the lack of conservation of the glycine-rich loop (GxGGxGT). In all the MLPs, except At1g24000.1 in which the loop is absent, this sequence is GxxxxxG, with the third amino acid of the motif usually Trp. Since the MLPs have no known allergenicity, this loop is not expected to serve as a IgE epitope, as with the PR-10 proteins¹⁹. However, it is interesting that the loop has retained its rigid conformation. It remains to be determined whether the modified loop serves a functional role.

NMR analysis of ligand binding

The ability of At1g24000.1 to bind a hydrophobic ligand was investigated by an HSQC titration of [U - ^{13}C , ^{15}N] At1g24000.1 with progesterone. While progesterone is a human hormone not found in plants, it shares enough structural similarity to the plant steroid, brassinolide, to serve as a representative ligand. Combined $^1\text{H}/^{15}\text{N}$ chemical shift perturbations are shown in Fig. 4A. Significant changes (> 0.4 p.p.m.) are clustered in three regions, Ala23-Glu41, Ile52-Ile59, and Ile102-Leu113 (Fig. 4B). It can be seen that the affected residues are located at the opening as well as in the bottom of the cavity, providing evidence that this cavity can accommodate a steroid molecule. Included within these regions are the conserved hydrophobic residues of the long α helix, which were identified above as potential hydrophobic binding ligands on the basis of sequence conservation among the MLPs. It is therefore likely that the cavity in MLP28 and other MLPs can bind a ligand in a similar manner, but further studies are needed to determine the specific mode of binding and whether more than one ligand can bind in the pocket.

Functional roles of MLPs

Although the function of the major latex proteins is unknown, gene expression of *A. thaliana* MLPs has been shown to be affected by a variety of environmental stimuli or stresses. One such stimulus is gravity, which has a pronounced effect on plant root growth. A study involving whole-genome microarray analysis of *Arabidopsis* root apices after

gravity and mechanical stimulation showed that transcription of the At1g70830 (MLP28) gene, as well as the highly similar At1g70890 (MLP43) gene, was up-regulated within 15 min of gravity stimulation. Another MLP (At4g23670) gene was among a cluster of five genes upregulated by more than 3-fold within 2 min. All three of these genes were then subsequently down-regulated below control levels within 60 min²⁹. Based on the rapid and transient response to the stimulus, MLP gene expression can be considered a specific response to the gravity stimulation, rather than being a secondary response. It is possible that these MLPs play a role in gravitropism, the response of a plant to gravity, through binding of hormones in the hydrophobic pocket. The hormone auxin, in conjunction with brassinosteroids, is known to play a central role in gravitropism^{30,31} along with other plant hormones, such as cytokinin, ethylene, and gibberellic acid²⁹. Therefore, there is a wide range of putative MLP binding ligands.

In addition to their role in gravitropism, the MLPs are differentially expressed during plant development. For example, At4g23670 is expressed during germination and seedling development only, and is found only in the flower³². On the other hand, expression of At4g23780, which shows 79% sequence identity, continues throughout development and is found in all tissues such as leaves, stems, flowers, and siliques (fruit). MLP28 shares the highest sequence homology with peach Pp-MLP1 and cucumber CSF-2, both of which are highly expressed during early fruit growth¹⁹. To our knowledge soybean is the only plant known thus far to contain both an MLP (MSG protein) and PR-10 protein (SAM22)³³. For this reason, it seems probable that the functional role of MLPs would be similar to that of the PR-10 proteins. It is thus likely that the MLPs carry out their function in plant growth and development through the binding of hormones or other metabolites in a conserved hydrophobic cavity.

Conclusions

The first two structures solved for members of the MLP family reveal a novel helix-grip fold that differs significantly from the Bet v 1 fold. Despite low sequence conservation, MLP28 and At1g24000.1 are structurally similar and share a number of unique conserved residues when compared to other Bet v 1 superfamily proteins. However, the deletion of 33 residues in the At1g24000.1 sequence results in a more compact structure with a small hydrophobic cavity. Despite the structural differences, we demonstrated that At1g24000.1 is able to bind a representative hormone, progesterone, indicating that At1g24000.1, and in all likelihood the MLP family, can participate in a wide variety of biochemical processes through plant hormone-mediated cell signaling.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by the NIH Protein Structure Initiative grants P50 GM64598 and U54 GM074901 (J. L. Markley, P.I.). We are grateful to Kent Brodie of the MCW Human and Molecular Genetics Center for Linux cluster implementation of structure calculation software. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract No. W-31-109-Eng-38. Use of the BioCARS Sector 14 was supported by the National Institutes of Health, National Center for Research Resources, under grant number RR007707.

REFERENCES

1. Nessler CL, Allen RD, Galewsky S. Identification and Characterization of Latex-Specific Proteins in Opium Poppy. *Plant Physiol.* 1985; 79(2):499–504. [PubMed: 16664439]

2. Osmark P, Boyle B, Brisson N. Sequential and structural homology between intracellular pathogenesis-related proteins and a group of latex proteins. *Plant Mol Biol.* 1998; 38(6):1243–1246. [PubMed: 9869429]
3. Liscombe DK, MacLeod BP, Loukanina N, Nandi OI, Facchini PJ. Evidence for the monophyletic evolution of benzyloquinoline alkaloid biosynthesis in angiosperms. *Phytochemistry.* 2005; 66:2501–2520. [PubMed: 16342378]
4. Radauer C, Breiteneder H. Evolutionary biology of plant food allergens. *Journal of Allergy and Clinical Immunology.* 2007; 120(3):518–525. [PubMed: 17689599]
5. Liu J-J, Ekramoddoullah AKM. The family 10 of plant pathogenesis-related proteins: Their structure, regulation, and function in response to biotic and abiotic stresses. *Physiol Mol Plant Pathology.* 2006; 68:3–13.
6. Gajhede M, Osmark P, Poulsen FM, Ipsen H, Larsen JN, Joost van Neerven RJ, Schou C, Lowenstein H, Spangfort MD. X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy. *Nat Struct Biol.* 1996; 3(12):1040–1045. [PubMed: 8946858]
7. Markovic-Housley Z, Degano M, Lamba D, von Roepenack-Lahaye E, Clemens S, Susani M, Ferreira F, Scheiner O, Breiteneder H. Crystal structure of a hypoallergenic isoform of the major birch pollen allergen Bet v 1 and its likely biological function as a plant steroid carrier. *J Mol Biol.* 2003; 325(1):123–133. [PubMed: 12473456]
8. Neudecker P, Schweimer K, Nerkamp J, Scheurer S, Vieths S, Sticht H, Rosch P. Allergic cross-reactivity made visible: solution structure of the major cherry allergen Pru av 1. *J Biol Chem.* 2001; 276(25):22756–22763. [PubMed: 11287426]
9. Biesiadka J, Bujacz G, Sikorski MM, Jaskolski M. Crystal structures of two homologous pathogenesis-related proteins from yellow lupine. *J Mol Biol.* 2002; 319(5):1223–1234. [PubMed: 12079359]
10. Pasternak O, Biesiadka J, Dolot R, Handschuh L, Bujacz G, Sikorski MM, Jaskolski M. Structure of a yellow lupin pathogenesis-related PR-10 protein belonging to a novel subclass. *Acta Crystallogr D Biol Crystallogr.* 2005; 61(Pt 1):99–107. [PubMed: 15608381]
11. Fernandes H, Pasternak O, Bujacz G, Bujacz A, Sikorski MM, Jaskolski M. Lupinus luteus Pathogenesis-Related Protein as a Reservoir for Cytokinin. *J Mol Biol.* 2008; 378:1040–1051. [PubMed: 18406424]
12. Schirmer T, Hoffmann-Sommergrube K, Susani M, Breiteneder H, Markovic-Housley Z. Crystal structure of the major celery allergen Api g 1: molecular analysis of cross-reactivity. *J Mol Biol.* 2005; 351(5):1101–1109. [PubMed: 16051263]
13. Pasternak O, Bujacz GD, Fujimoto Y, Hashimoto Y, Jelen F, Otlewski J, Sikorski MM, Jaskolski M. Crystal structure of Vigna radiata cytokinin-specific binding protein in complex with zeatin. *Plant Cell.* 2006; 18(10):2622–2634. [PubMed: 16998071]
14. Pasternak O, Bujacz A, Biesiadka J, Bujacz G, Sikorska M, Jaskolski M. MAD phasing using the (Ta6Br12)²⁺ cluster: a retrospective study. *Acta Cryst.* 2008; D64:595–606.
15. Iyer LM, Koonin EV, Aravind L. Adaptations of the helix-grip fold for ligand binding and catalysis in the START domain superfamily. *Proteins.* 2001; 43(2):134–144. [PubMed: 11276083]
16. Tsujishita Y, Hurley JH. Structure and lipid transport mechanism of a StAR-related domain. *Nat Struct Biol.* 2000; 7(5):408–414. [PubMed: 10802740]
17. Mirza O, Henriksen A, Ipsen H, Larsen JN, Wissenbach M, Spangfort MD, Gajhede M. Dominant epitopes and allergic cross-reactivity: complex formation between a Fab fragment of a monoclonal murine IgG antibody and the major allergen from birch pollen Bet v 1. *J Immunol.* 2000; 165(1):331–338. [PubMed: 10861069]
18. Spangfort MD, Mirza O, Ipsen H, Van Neerven RJ, Gajhede M, Larsen JN. Dominating IgE-binding epitope of Bet v 1, the major allergen of birch pollen, characterized by X-ray crystallography and site-directed mutagenesis. *J Immunol.* 2003; 171(6):3084–3090. [PubMed: 12960334]
19. Ruperti B, Bonghi C, Ziliotto F, Pagni S, Rasori A, Varotto S, Tunutti P, Giovannoni J, Ramina A. Characterization of a major latex protein (MLP) gene down-regulated by ethylene during peach fruitlet abscission. *Plant Science.* 2002; 163:265–272.

20. Suyama T, Yamada K, Mori H, Takeno K, Yamaki S. Cloning cDNAs from genes preferentially expressed during fruit growth in cucumber. *J Am Soc Hort Sci.* 1999; 124:136–139.
21. Vinarov DA, Lytle BL, Peterson FC, Tyler EM, Volkman BF, Markley JL. Cell-free protein production and labeling protocol for NMR-based structural proteomics. *Nat Methods.* 2004; 1(2): 149–153. [PubMed: 15782178]
22. Sreenath HK, Bingman CA, Buchan BW, Seder KD, Burns BT, Geetha HV, Jeon WB, Vojtik FC, Aceti DJ, Frederick RO, Phillips GN Jr, Fox BG. Protocols for production of selenomethionine-labeled proteins in 2-L polyethylene terephthalate bottles using auto-induction medium. *Protein Expr Purif.* 2005; 40(2):256–267. [PubMed: 15766867]
23. Jeon WB, Aceti DJ, Bingman CA, Vojtik FC, Olson AC, Ellefson JM, McCombs JE, Sreenath HK, Blommel PG, Seder KD, Burns BT, Geetha HV, Harms AC, Sabat G, Sussman MR, Fox BG, Phillips GN Jr. High-throughput purification and quality assurance of *Arabidopsis thaliana* proteins for eukaryotic structural genomics. *J Struct Funct Genomics.* 2005; 6((2–3)):143–147. [PubMed: 16211511]
24. Cannon SB, Young ND. OrthoParaMap: distinguishing orthologs from paralogs by integrating comparative genome data and gene phylogenies. *BMC Bioinformatics.* 2003; 4:35. [PubMed: 12952558]
25. Herrmann T, Guntert P, Wuthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol.* 2002; 319(1):209–227. [PubMed: 12051947]
26. Linge JP, Williams MA, Spronk CA, Bonvin AM, Nilges M. Refinement of protein structures in explicit solvent. *Proteins.* 2003; 50(3):496–506. [PubMed: 12557191]
27. Song J, Zhao Q, Lee MS, Markley JL. ¹H, ¹⁵N and ¹³C resonance assignments of the putative Bet v 1 family protein Atlg24000.1 from *Arabidopsis thaliana*. *J Biomol NMR.* 2005; 32(4):335. [PubMed: 16211487]
28. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr.* 2004; 60((Pt 12 Pt 1)):2256–2268. [PubMed: 15572779]
29. Kimbrough JM, Salinas-Mondragon R, Boss WF, Brown CS, Sederoff HW. The fast and transient transcriptional network of gravity and mechanical stimulation in the *Arabidopsis* root apex. *Plant Physiol.* 2004; 136(1):2790–2805. [PubMed: 15347791]
30. Li L, Xu J, Xu ZH, Xue HW. Brassinosteroids stimulate plant tropisms through modulation of polar auxin transport in *Brassica* and *Arabidopsis*. *Plant Cell.* 2005; 17(10):2738–2753. [PubMed: 16141452]
31. Blancaflor EB, Masson PH. Plant gravitropism. Unraveling the ups and downs of a complex process. *Plant Physiol.* 2003; 133(4):1677–1690. [PubMed: 14681531]
32. Wu F-Z, Lu T-C, Shen Z, Wang B-C, Wang H-X. N-Terminal acetylation of two major latex proteins from *Arabidopsis thaliana* using electrospray ionization tandem mass spectrometry. *Plant Mol Biol Rep.* 2008; 26:88–97.
33. Stromvik MV, Sundararaman VP, Vodkin LO. A novel promoter from soybean that is active in a complex developmental pattern with and without its proximal 650 base pairs. *Plant Mol Biol.* 1999; 41(2):217–231. [PubMed: 10579489]
34. Delano, WL. The PyMOL Molecular Graphics System. San Carlos, CA, USA: DeLano Scientific; 2002.

Figure 1. Phylogenetic tree showing the relationships between representative members of the major latex protein (MLP) family.

Figure 2.

(A) Structure of the MLP28 domain. Backbone trace of the ensemble of 20 NMR structures is shown as a stereo image, with α -helices in violet and β -strands in green, produced using the program PyMOL 34. (B) Ribbon diagram of the lowest energy MLP28 conformer. (C) Backbone r.m.s.d. values and ^1H - ^{15}N heteronuclear NOEs plotted for each residue of MLP28. (D) Ribbon diagram of At1g24000.1 (E) Superposition of MLP28 (green) and At1g24000.1 (orange).

Figure 3.

(A) Ribbon diagrams of MLP28, At1g24000.1 and representative structural homologs CSBP from mung bean (2FLH), birch pollen allergen Bet v 1 (1BV1), and jicama SPE-16 (1TW0). (B) Sequence alignment of selected MLP family proteins as well as the three structural homologs shown in (A). Strictly conserved residues in > 70% of the sequences are highlighted in black and those that are conservatively substituted are highlighted gray. Stars identify residues of CSBP that interact with the ligands via hydrogen bonds, and crosses identify residues that interact via van der Waals interactions.

Figure 4.

(A) Combined $^1\text{H}/^{15}\text{N}$ chemical shift perturbations for each backbone amide group of At1g24000.1 upon binding of 5 mM progesterone plotted as a function of At1g24000.1 residue. Residues which displayed perturbations higher than 0.4 ppm are colored green. Along with the Pro residues, the amides of Ser1, Thr2, Thr28, Lys34, Thr38, His100, and Ser101 were unassigned. (B) The residues showing the highest chemical shift perturbations upon progesterone binding are highlighted on the structure of At1g24000.1, shown in the same orientation as in Fig 2D and after a 180° rotation about the vertical axis.