

Crystal Structures of Two Homologous Pathogenesis-related Proteins from Yellow Lupine

Jacek Biesiadka¹, Grzegorz Bujacz^{1,2}, Michal M. Sikorski¹ and Mariusz Jaskolski^{1,3*}

¹*Institute of Bioorganic Chemistry, Polish Academy of Sciences, 61-704 Poznan Poland*

²*Technical University of Lodz 90-924 Lodz, Poland*

³*Department of Crystallography Faculty of Chemistry A. Mickiewicz University Grunwaldzka 6, 60-780 Poznan Poland*

Pathogenesis-related class 10 (PR10) proteins are restricted to the plant kingdom where they are coded by multigene families and occur at high levels. In spite of their abundance, their physiological role is obscure although members of a distantly related subclass (cytokinin-specific binding proteins) are known to bind plant hormones. PR10 proteins are of special significance in legume plants where their expression patterns are related to infection by the symbiotic, nitrogen-fixing bacteria. Here we present the first crystal structures of classic PR10 proteins representing two homologues from one subclass in yellow lupine. The general fold is similar and, as in a birch pollen allergen, consists of a seven-stranded β -sheet wrapped around a long C-terminal helix. The mouth of a large pocket formed between the β -sheet and the helix seems a likely site for ligand binding. The shape of the pocket varies because, in variance with the rigid β -sheet, the helix shows unusual conformational variability consisting in bending, disorder, and axial shifting. A surface loop, proximal to the entrance to the internal cavity, shows an unusual structural conservation and rigidity in contrast to the high glycine content in its sequence. The loop is different from the so-called glycine-rich P-loops that bind phosphate groups of nucleotides, but it is very likely that it does play a role in ligand binding in PR10 proteins.

© 2002 Elsevier Science Ltd. All rights reserved

Keywords: protein crystallography; plant pathogenesis-related proteins; plant allergens; cytokinin-binding proteins; RNase activity

*Corresponding author

Introduction

Pathogenesis-related class 10 (PR10) proteins have been detected in nearly all studied plants, both monocotyledonous and dicotyledonous, and also in angiosperms. They were never found in non-plant organisms. The PR10 class includes numerous homologous proteins, first discovered in plant response to pathogen infection,^{1,2} described as intracellular pathogenesis-related (IPR) proteins.^{3,4} On the basis of sequence hom-

ology, two additional, distinct groups of proteins were included in the PR10 class: CSBP⁵ (cytokinin-specific binding proteins) and MLP^{6–8} (major latex proteins). These distant homologues are related to the IPR group by only a 17–20% sequence identity. Although the biological function of PR10 proteins has not been established yet, their expression patterns and biochemical properties are well characterized. In general, PR10 proteins occur in high concentrations in plant roots.^{9–12} They are induced in aerial organs in response to various stressing factors, e.g. wounding,^{13,14} UV radiation,¹⁵ some chemical compounds,¹⁶ or pathogenic infections.^{2,11,15,17–20} In legume plants, expression of a nodule-specific PR10 homologue was also observed at an early stage of symbiosis development in the alfalfa root nodule cortex.²¹ There is a significant variation in PR10 expression patterns between plants, or even between various PR10 homologues in one organism. There are also examples of high, constitutive expression of PR10 proteins in fruits,²² vegetables,²³ various parts of

Abbreviations used: CSBP, cytokinin-specific binding proteins; EDTA, ethylenediamine tetraacetic acid; HDV, hepatitis delta virus; IPR, intracellular pathogenesis-related proteins; LIPR, *Lupinus luteus* pathogenesis-related protein; mlf, maximum-likelihood function; MLP, major latex proteins; rmsd, root-mean-square deviation.

E-mail address of the corresponding author: mariuszj@amu.edu.pl

flowers,^{24–28} or pollen grains,²⁹ in which cases the PR10 proteins act as strong, common allergens, or expression induced by the presence of phytohormones.^{30–33} Connection with response to pathogens has not been demonstrated for CSBP proteins.

Most of the hypotheses explaining the biological role of PR10 proteins focus on the reported RNase activity of some homologues. Two PR10 proteins from ginseng (PgRNS1 and PgRNS2) were first isolated as plant RNases.^{34,35} Nucleolytic activity was also demonstrated *in vitro* for Betv1, the pollen allergen from white birch (*Betula verrucosa*),^{36,37} as well as for PR10 proteins from white³⁸ and yellow lupine (Sikorski, unpublished results). The RNase activity of PR10 proteins suggests a potential role in defense against pathogenic infections. The high level of PR10 proteins in roots may be part of a constitutive defense mechanism in the plant organ that is most exposed to environmental stress. The presence of RNase activity in pollen grains could be related to recognition and self-incompatibility during fertilization.^{37,39} The RNase activity of pollen PR10 proteins could also protect the female reproductive tissue from infection by pathogens during the early phases of pollen tube growth through stigma and style. The low *in vitro* levels of RNase activity could be modulated during the developmental program or after elicitation of plant defense reactions by other factors. However, the high abundance of PR10 proteins in other plant organs, together with various expression models with no relation to pathogen response, make the RNase hypothesis controversial and suggest a more general function of these proteins.^{12,40,41}

Recently, we have identified in yellow lupine two gene families encoding PR10 proteins¹² of LIPR10.1 subclass (155 amino acid residues, three homologues: LIPR10.1A, B, and C) and LIPR10.2 subclass (157 amino acid residues, five homologues: LIPR10.2A, B, C, D, and E). The PR10 proteins of these two subclasses share 59–60% amino acid identity.⁴² Two of these proteins, LIPR10.1A and LIPR10.1B, are very abundant in roots.¹² There is constitutive expression of these proteins in stem, but the protein level is ca 10 times lower. Only LIPR10.1B is constitutively expressed in the petioles and leaves (at a level comparable to stem), while LIPR10.1A is induced in these organs in response to infection with the pathogenic bacteria *Pseudomonas syringae*, or during senescence of leaves.

So far, full structural characterization has been reported only for the Betv1 protein, which was the subject of crystallographic and NMR studies in the native⁴³ and IgG Fab-complexed forms.^{44,45} A preliminary crystallographic study at 2.45 Å resolution has been also reported for the LIPR10.1A protein.⁴⁶ The present paper discusses the structures of two yellow lupine subclass PR10.1 proteins, LIPR10.1A and LIPR10.1B, elucidated by crystallographic methods.

Results and Discussion

Protein expression, activity, crystallization, and data collection

Both proteins were overproduced in *Escherichia coli* cells using the T7 RNA polymerase expression system.⁴⁷ The recombinant proteins were expressed and purified from bacterial lysates according to a procedure described earlier.⁴⁸ The final expression products corresponded exactly to the native protein sequences, devoid of the usual methionine residue introduced at the beginning of the polypeptide chain during the translation initiation step. The excision of the N-terminal methionine was carried out by an endogenous *E. coli* methionyl-aminopeptidase (MAP).⁴⁹

The catalytic efficiency of MAP is inversely proportional to the size of the side-chain of the amino acid in the penultimate position. In both LIPR10.1A and LIPR10.1B this amino acid residue (corresponding to position 1 in the native sequence) is Gly, and therefore the cleavage is almost 98% complete.⁴⁹

The ribonuclease activity of both proteins was tested *in vitro* using ³²P-labeled yeast tRNA^{Phe} and a 73-nucleotide fragment of genomic HDV ribozyme as substrates (M.M.S., unpublished results). No activity was found for LIPR10.1A. For LIPR10.1B, low-activity processing was detected in single-stranded regions of both RNA substrates with no apparent base specificity. The ribonuclease activity of LIPR10.1B was abolished by heat denaturation and proteinase K treatment and substantially decreased in the presence of 6 M urea. The activity did not require divalent cations and was not sensitive to EDTA. The optimum pH was 6–8. Therefore, this RNase activity may be considered as class II, according to the general biochemical classification.⁵⁰

In the present study, the 2.45 Å diffraction data set for LIPR10.1A described earlier⁴⁶ has been superseded by a new data set collected at 100 K using synchrotron radiation. The new data set is of superior quality (Table 1) and extends to a higher resolution (1.95 Å). The crystal used in this new experiment was obtained as described earlier.⁴⁶ The crystallization conditions for LIPR10.1B are different from those for LIPR10.1A (see Methods). The diffraction data were collected at room temperature using conventional X-ray source (Table 1) and therefore the resolution limit is somewhat lower (2.25 Å) than in the previous case.

Refinement results and quality of the models

Both structures were solved by molecular replacement with the birch pollen protein (PDB accession code: 1BV1) as a model for LIPR10.1A, and LIPR10.1A—as a model for LIPR10.1B. The refinements were carried out with standard procedures including simulated annealing, positional

Table 1. Statistics of data collection and processing

	LIPR10.1A	LIPR10.1B
Space group	$P2_12_1$	$P2_12_1$
Cell parameters (a, b, c) (Å)	34.97, 56.62, 61.76	64.88, 107.45, 46.88
Molecules in asymmetric unit	1	2
Radiation source	Synchrotron, LURE, beamline D41	Cu K α , rotating anode
Detector type	MarResearch 30 cm image plate	MarResearch 30 cm image plate
Crystal temperature (K)	100	291
No. of measured reflections	89,375	88,882
No. of unique reflections	9446	15,594
Resolution limits (last shell) (Å)	20–1.95 (2.07–1.95)	20–2.25 (2.39–2.25)
R_{int} (last shell)	0.053 (0.127)	0.102 (0.399)
Completeness (last shell) (%)	99.5 (96.8)	97.2 (94.7)
$\langle I/\sigma(I) \rangle$ (last shell)	25.7 (8.5)	13.3 (2.8)
Redundancy (last shell)	9.5 (3.1)	5.6 (3.1)

and B -factor refinement, and manual modeling according to electron density maps. However, the LIPR10.1A model required substantial rebuilding and ca 30% of the structure within variable regions, mainly helix $\alpha 3$ and loops L3, L5 and L7 (for labeling see Figure 1), had to be reconstructed. The asymmetric unit of the LIPR10.1A crystal contains only one molecule. In the crystal structure of LIPR10.1B, two independent molecules are present (A and B) but they do not form any quaternary structure. The quality of both structures is very good (Table 2). The models are characterized by very good conformation of the polypeptide chains, as demonstrated by Ramachandran analysis⁵¹ (Table 2), especially within the β regions.

The structure of LIPR10.1A has been determined to a higher resolution and at low temperature. In spite of the very good refinement results, it displays a puzzling disorder within the C-terminal helix, which may be caused by crystal contacts. The LIPR10.1A crystals are densely packed (Matthews volume⁵² 1.81 Å³/Da, corresponding to

32% solvent content) and this region is involved in interactions with symmetry-related molecules. Although of slightly lower resolution, the structure of LIPR10.1B does not exhibit any problems indicative of disorder, and is characterized by a lower R -factor.

Overall fold

The structures reveal a similar folding scheme: a seven-stranded antiparallel β -sheet wrapped around a long C-terminal α -helix. Two additional, short helices are located near the N terminus (Figure 1). The two principal structural elements, the β -sheet and the long α -helix, are both amphipathic, with polar residues directed toward solvent and with hydrophobic side-chains located mainly at the contact interface and stabilizing the whole structure. Despite the compact fold, a void area exists within the core of the molecule.⁴³ It has an elongated shape and is filled with water molecules. The number of water molecules in each of the present molecules is the same (eight), but their exact positions vary in response to conformational differences.

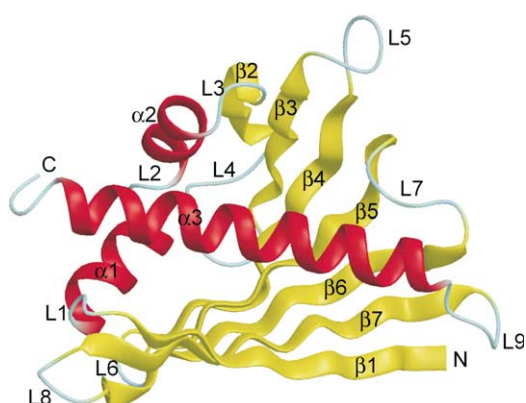


Figure 1. A ribbon diagram of LIPR10.1B (molecule A), with labeling of secondary structure elements, representing the overall fold and topology of PR10 proteins. Note the pocket formed between helix $\alpha 3$ and the β -sheet, viewed here from the direction of helix $\alpha 3$. The mouth of this cavity is formed by the glycine-rich loop L4 and the junction between helices $\alpha 1$ and $\alpha 2$ (loop L2). This and all other structural illustrations have been prepared using DINO (<http://www.dino3d.org>).

Table 2. Refinement statistics

	LIPR10.1A	LIPR10.1B
Program used	CNS 0.9	CNS 0.9
Resolution limits (Å)	12–1.95	12–2.25
No. of reflections	9307	15,594
No. of reflections in test set	930	1539
σ cutoff	0	0
Crystallographic target	mlf	mlf
Bulk solvent correction	Yes	Yes
Protein atoms	1166	2336
Solvent atoms	104	112
R/R_{free}	0.196/0.250	0.186/0.239
$\langle B \rangle$ for main-chain atoms (Å ²)	24.6	32.8
$\langle B \rangle$ for side-chain atoms (Å ²)	28.3	38.3
$\langle B \rangle$ for solvent atoms (Å ²)	31.2	42.9
Rmsd for bond lengths (Å)	0.016	0.015
Rmsd for bond angles (°)	1.90	1.80
Rmsd for dihedral angles (°)	25.3	25.4
Rmsd for improper torsion angles (°)	1.19	1.15
Most favored ϕ/ψ (%)	89.2	93.6

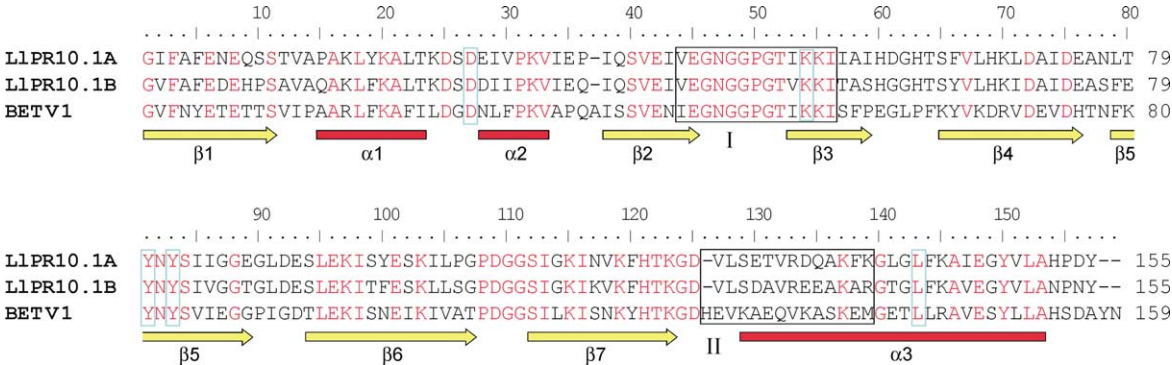


Figure 2. Sequence alignment of PR10 proteins with known three-dimensional structure, calculated in CLUSTAL W.⁶² LLPR10.1A and LLPR10.1B are the yellow lupine proteins (this work), Betv1 is a birch pollen allergen.⁴³ Residues identical in all sequences are indicated in red. Secondary structure elements are indicated according to the LLPR10.1B protein structure (Figure 1). The black boxes indicate the protein regions that are discussed in the text: (I) highly conserved glycine-rich loop L4 and its flanking residues, (II) variable sequence region including the N-terminal portion of helix $\alpha 3$. The cyan boxes indicate conserved residues whose side-chains are directed into the lumen of the putative ligand-binding cavity.

The β -sheet

The seven strands of the β -sheet form a twisted shape. The N terminus is included in strand $\beta 1$. Strand $\beta 2$ is located on the opposite edge of the β -sheet and is connected with $\beta 1$ via a right-handed crossover containing loops L1, L2 and L3, and helices $\alpha 1$ and $\alpha 2$. The consecutive strands $\beta 2$ – $\beta 7$ are connected by five hairpin loops bringing strand $\beta 7$ back into hydrogen-bond contact with $\beta 1$. The regularity of the β -sheet is distorted at three places (Leu70–Asp71, Ile84–Ile85, Leu94–Glu95 in LLPR10.1A) where the directions of the carbonyl groups are perturbed resulting in modified hydrogen-bonding patterns and causing β -bulging of the main chain. β -Bulges are characteristic structural elements frequently utilized by protein folds to introduce curvature into β -sheets. The overall geometrical result is the curved shape of the entire β -sheet. These elements are structurally conserved, as they are found in all three copies of the lupine proteins as well as in the birch pollen protein, confirming their importance in maintaining the overall structure of the β -sheet.

The C-terminal helix

The second main structural element, the C-terminal helix $\alpha 3$ is 25 amino acid residues long and can be subdivided in the lupine proteins into two parts according to sequence conservation and structural properties (Figure 2). The N-terminal part (residues Glu128–Lys137 in LLPR10.1A) represents the least conserved region in the PR10 amino acid sequence, and also is among the most variable structural segments. Additionally, in the LLPR10.1A structure, a fragment of this region (residues Lys135–Lys137) is disordered as the corresponding $2F_o - F_c$ density is fragmented even in the main chain area (Figure 3). This region forms intermolecular contacts with the variable loops L3 and L5 of symmetry-related molecules, as well as intramolecular contacts with loop L7, which also is partially disordered. The same region in the two molecules of the LLPR10.1B protein is characterized by very good and contiguous density (Figure 3) and at the same time is not involved in contacts with any disordered elements. It has proper 3.6_{13} helix geometry although it is slightly curved. The C-terminal part of helix $\alpha 3$ (residues Gly138–

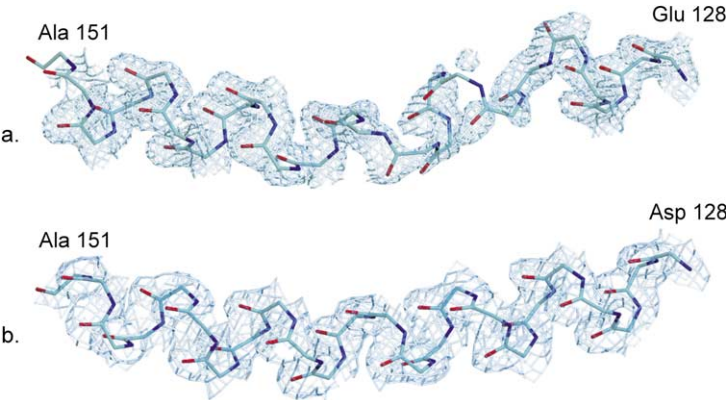


Figure 3. Conformation of helix $\alpha 3$ in LLPR10.1A (a) and LLPR10.1B (molecule A) (b). The helices are drawn in the same orientation, corresponding to the plane of their bending. The $2F_o - F_c$ maps are contoured at the 1σ level.

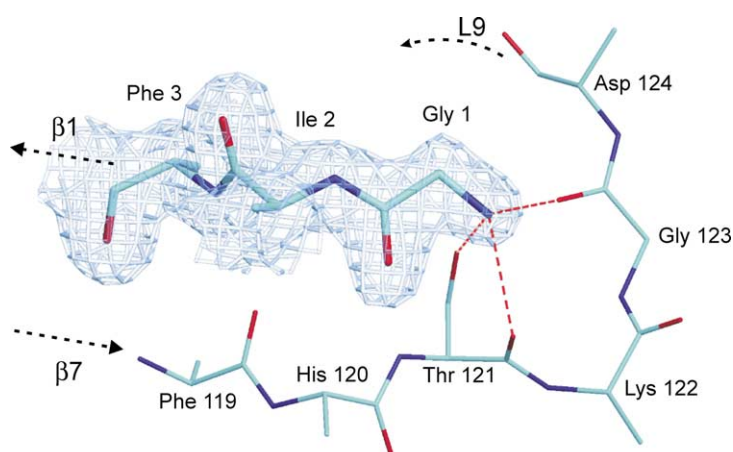


Figure 4. Interactions of the N-terminal glycine with a fragment of loop L9. For clarity, the $2F_o - F_c$ map, contoured at the 1σ level, is displayed only around the N-terminal fragment of strand $\beta 1$. This example is for LIPR10.1A but identical interactions are present in all three LIPR10.1 molecules.

Ala151) has a relatively conserved sequence and its conformation is similar in all three lupine molecules and in 1BV1, although it appears to undergo an axial shift (see below). In contrast to the lupine proteins, the sequence of the birch pollen allergen corresponding to the entire $\alpha 3$ C-terminal region forms a very regular and straight 3.6_{13} helix.⁴³

N and C termini

Both termini of the protein chains of all three molecules are stable and unambiguously defined in the electron density. The N-terminal glycine is a component of strand $\beta 1$ and is stabilized by canonical β -sheet hydrogen bonding. It forms two additional hydrogen bonds at its ammonium $-\text{NH}_3^+$ group, resulting in a very stable structure with clear electron density and low *B*-factors (Figure 4). In LIPR10.1A, the N-terminal $-\text{NH}_3^+$ group of Gly1 forms hydrogen bonds with the carbonyl oxygen atoms of Thr121 and Gly123, and with the hydroxyl group of Thr121. A similar situation is seen in both LIPR10.1B molecules and in 1BV1. These observations exclude the possibility that a disordered N-terminal methionine residue, introduced into the sequence by the bacterial expression machinery, could precede the native sequence. This conclusion is confirmed by mass spectrometry data (not shown) and is explained by the activity of the MAP aminopeptidase in *E. coli* cells.⁴⁹ The significance of the N-terminal interactions is further supported by the high conservation of the amino acid residues Thr121 and Gly123, together with their neighboring residues (Figure 2).

The C terminus is not involved in any distinct secondary structure, but as a result of numerous interactions with other fragments of the molecule, most of the atoms have very clear definition in the electron density maps. The intramolecular interactions are different in the three lupine molecules, although in all cases, the hydroxyl group of Tyr155 forms a short hydrogen bond with the side chain of Asp25 at the junction of helices $\alpha 1$ and $\alpha 2$. The intermolecular interactions stabilizing the

C terminus are different in each case. In LIPR10.1A, the stabilization results from interactions with the side-chains in strand $\beta 1$ (and also in loop L5) in a symmetry-related molecule. In the LIPR10.1B structure, the intermolecular contacts are fewer and are only observed in molecule B. The position of the side-chain of the corresponding Tyr158 in 1BV1 resembles that of Tyr155 in the lupine proteins but the general conformation of the C-terminal segment is different. The C-terminal group of 1BV1 is stabilized by an intramolecular salt bridge with Arg17.⁴³

Loop regions

The fold of the peptide chain of the LIPR10 molecules includes nine loops. Five of them are β -hairpin loops connecting the consecutive β -strands of the β -sheet. Most of the loop regions have higher temperature factors than the well-defined secondary structure elements, but only one of them, L5, residues His59-Thr63, shows symptoms of higher mobility or disorder resulting in broken $2F_o - F_c$ density. The analogous loop in the 1BV1 structure was reported in double conformation and with high temperature factors.⁴³ Nevertheless, the entire L5 loop in molecule B of LIPR10.1B could be traced in contiguous main chain density at the 1σ level. Since this loop is not involved in many interactions and since its disorder is observed in some degree in all analyzed structures, it is likely that its flexibility is an intrinsic property of PR10 proteins.

Glycine-rich loop L4 (Gly45-Thr51)

The glycine-rich loop L4 (Gly45-Thr51) is part of the most conserved fragment of the PR10 sequence (Figure 2). This sequence motif, xEGNG-GPGTxKKI, was earlier described as similar to the phosphate-binding motif identified in several other proteins, where it is often referred to as the P-loop.^{53,54} In the PR10 motif, the glycine residues are preserved even in the most distant homologues, with overall sequence identity below 20%.

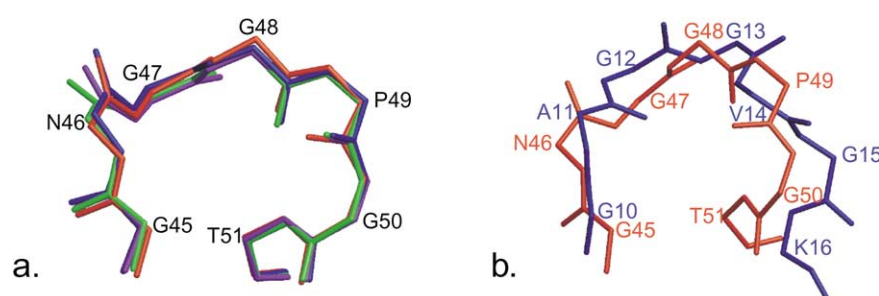


Figure 5. Conformation of the glycine-rich loop L4 in various PR10 structures: LIPR10.1A (red), LIPR10.1B (molecule A, blue; molecule B, purple), and Betv1 (PDB code 1BV1, green) (a). Superposition of the main-chain atoms was done in ALIGN.⁵⁵ The glycine-rich phosphate binding motif (P-loop) present in the mononucleotide-binding p21 H-ras oncogene protein⁶³ (PDB code 5P21, blue) is conformationally different from the L4 loop of PR10 proteins (LIPR10.1A, red) (b).

Comparison of the known PR10 structures reveals that this fragment also shows unusual structural conservation. It has a nearly identical conformation in the three LIPR10 molecules reported here (Figure 5(a)). However, the loop as a whole appears to be shifted between the structures when the molecules are superimposed using the structurally conserved main chain atoms (ALIGN⁵⁵). In the most dramatic case of LIPR10.1B (chain A *versus* chain B), the shift amounts to 1 Å. However, if only the loop areas are superposed, the agreement is excellent, as illustrated by the small rms deviation between the C α atoms of even the most distant analogs, for instance, 0.19 Å for nine C α atoms of the LIPR10.1A/1BV1 pair. Such perfect structural conservation of an element containing as many as four glycine residues is difficult to explain. Some of the L4 loop residues form hydrogen bonds with other protein atoms, located mainly in the conserved neighborhood of the loop (Gly47N...Ile52O, Gly50N...Leu70O, in LIPR10.1A), providing some stabilizing interactions. These interactions are of main-chain...main-chain type and, theoretically, should be sequence-independent. Also, several intramolecular hydrogen bonds formed by the hydroxyl group of Thr51 seem to contribute to the increased rigidity of the loop. All four glycine residues have unusual main-chain torsion angles, with positive ϕ values in three cases. Two of them fall in the left-handed helix area. The remaining loop L4 residues are in the generous β region of the Ramachandran plot.

In view of the unusual structural conservation and rigidity of the L4 loop it is tempting to implicate it for a physiological role of PR10 proteins. However, the earlier suggestions,⁴³ referring to its glycine-rich sequence, about its similarity to the phosphate binding motif, are not supported by the structural data as the conformation of loop L4 is different from the canonical P-loop structure⁵³ (Figure 5(b)).

Putative ligand-binding site

A putative ligand-binding site can be identified between the glycine-rich loop and the junction of helices α 1 and α 2 (Figure 1), at the entrance to a

cavity described earlier in the birch pollen allergen molecule.⁴³ The site is large enough to hold a 100–200 Da ligand molecule and is lined with potential hydrogen bond donor and acceptor groups (Table 3). Several side-chains directed into this part of the cavity are conserved as well (Figure 2). Close to the entrance to the cavity, there are side chains of the strictly conserved residues Lys53 and Asp27. Other conserved residues (Tyr80, Tyr82, Leu141) are located at the bottom of the pocket. Some less-conserved residues, however, are also found there (Ser26, His68 in the lupine proteins).

The cavity harbors in each lupine protein as many as eight internal water molecules. They are all defined by very good electron density and most of them are characterized by relatively low temperature factors. In all cases, they are interconnected by hydrogen bonds and form hydrogen-bond contacts with hydrophilic residues on the cavity walls. However, the networks of connections are not the same, reflecting the conformational differences among the protein molecules, i.e. the rotation of the β -sheet and α 3 helix translation and bending, as described below. The internal water molecules are connected to the outside of the protein molecules by networks of hydrogen bonds involving residues at the putative ligand-binding site. Only four internal water molecules were reported in the cavity of the Betv1 protein.⁴³

Table 3. Hydrogen bonding of residues forming the putative ligand-binding site in the LIPR10.1A structure

H-bonded partners		Distance (Å)
Ser26 OG	Thr23 O	3.06
Ser26 OG	Asp27 N	3.09
Ser26 OG	HOH 1046 O	3.09
Asn46 OD1	HOH 1084 O	2.52
Lys53 NZ	Asp27 OD1	2.69
Lys53 NZ	Asp27 OD2	3.14
Lys53 NZ	HOH 1049 O	3.18
Tyr80 OH	HOH 1005 O	2.59
Tyr80 OH	HOH 1008 O	2.35
Tyr82 OH	HOH 1001 O	2.81
Tyr82 OH	HOH 1003 O	2.44
HOH 1091 O	Thr51 O	2.63
HOH 1091 O	His68 NE2	3.17
HOH 1091 O	HOH 1049 O	3.03

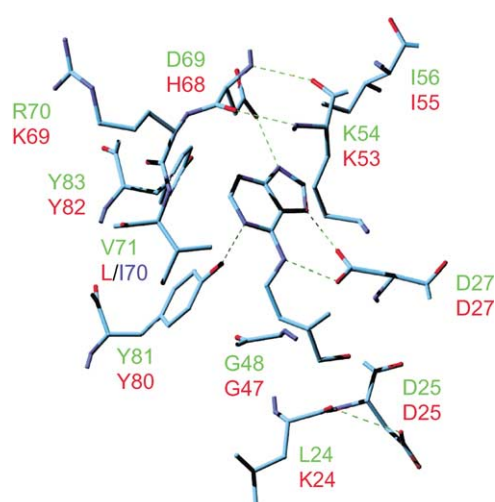


Figure 6. The results of zeatin docking in the putative binding site of PR10 proteins. The modeling was done using the GOLD program⁵⁶ (see text). The Figure illustrates the conservation of the key residues in LIPR10.1A (red), LIPR10.1B (blue, given if different), and Betv1 (green). The marked interactions correspond to Betv1.⁴³

The main goal of the current research on PR10 proteins is to identify their physiological function. In spite of the very high levels at which these proteins are present in plant cells, and in spite of the modulation of their levels in pathogenesis-related situations, their function remains obscure. On the other hand, the CSBP proteins, which are distant PR10 homologues, show very high affinity for plant hormones such as zeatin. To test if a molecule like zeatin could be bound in the PR10 binding site, we have performed a computer simulation experiment using the genetic algorithm protein-ligand docking program GOLD.⁵⁶ Repeated calculations for different starting points converged at the same geometry (characterized by the fitness parameter in the range 48–52%) with the zeatin purine moiety directed towards the binding cavity and anchored by four hydrogen bonds. In the sequence of LIPR10.1A, the binding residues correspond to Asp27, His68, and Tyr80 (Figure 6). Since zeatin (6-[4-hydroxy-3-methylbut-2-enylamino]purine) can be considered a nucleoside analog, the successful docking experiment may indicate, that the putative PR10 binding site is compatible with binding of a nucleoside-like

substrate. This is consistent with the reported ribonuclease activity of PR10 proteins (see Introduction).

Structural comparisons of PR10 proteins

Comparison of the two molecules of LIPR10.1B reveals several differences. The most evident difference involves a fragment containing strands $\beta 2$, $\beta 3$, $\beta 4$ together with two connecting loops, the glycine-rich loop L4 and the disordered loop L5 (59–63), which is rotated around an axis that is approximately perpendicular to the surface of the β -sheet at the midpoint of strand $\beta 3$.

In Table 4, which summarizes all pairwise comparisons of the existing PR10 models, the main chains of the two copies of LIPR10.1B superpose with a relatively low overall root-mean-square deviation (rmsd) value of 0.32 Å. In this light, superpositions of different proteins show rather high rmsd values, about 1 Å. This illustrates that there is an appreciable degree of conformational freedom among the PR10 proteins in spite of their common folding canon, and that even very close homologues (like the present LIPR10.1A/LIPR10.1B pair) can show significant structural differences.

Detailed comparisons show high structural conservation and stability within the β -sheet. In contrast, the C-terminal helix $\alpha 3$ appears to be labile, as its location and conformation are different in the three crystal structures compared here. It is straight in 1BV1, slightly bent in LIPR10.1B, and strongly bent in LIPR10.1A. The bending deformations take place in a plane that is normal to the β -sheet face with the central curved part indented towards the sheet, and the helical ends lifted away from it. In addition, in LIPR10.1A the central (kinked) part of the helix is disordered (Figure 3). Structural superpositions based on the most conserved C $^{\alpha}$ atoms within the β -sheet reveal a lack of overlap of the C-terminal helices. The same results are obtained in superpositions based on automatically selected common atoms, done in ALIGN⁵⁵ (Table 4, Figure 7), because the C-terminal helix segments are excluded from the calculations as too divergent.

Two diverse structural elements of the PR10 protein fold are thus discernible. One is the stable and rigid β -sheet, acting as a folding scaffold, twisted to create an elongated groove. In this

Table 4. Rmsd (Å) of main-chain atom pairs (rmsd/maximum distance (no. of pairs)) after least-squares superpositions calculated in ALIGN⁵⁵ (auto mode)

	LIPR10.1B (A)	LIPR10.1B (B)	1BV1
LIPR10.1A	0.90/2.33 (543)	1.02/2.54 (536)	1.14/2.90 (511)
LIPR10.1B (A)		0.32/0.83 (568)	0.99/2.53 (491)
LIPR10.1B (B)			1.13/2.95 (507)

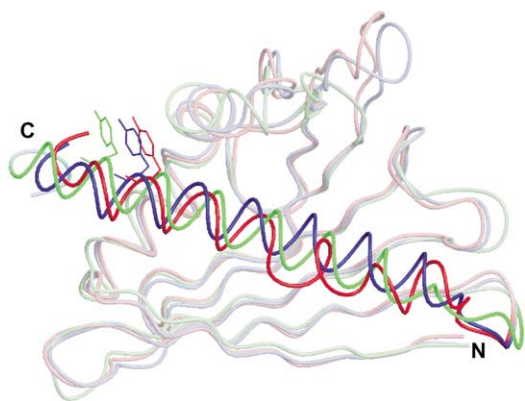


Figure 7. Main-chain traces of LIPR10.1A (red), LIPR10.1B (molecule A, blue) and Betv1 (PDB code 1BV1, green) in a structural superposition calculated by ALIGN.⁵⁵ The side-chain of the conserved Tyr148 residue (Tyr150 in 1BV1) is shown to visualize the axial shift of helix $\alpha 3$.

groove, the long and elastic C-terminal helix $\alpha 3$ is situated. In addition to displaying various degree of bending deformation, the $\alpha 3$ helix is also not precisely fixed in the groove as it can slide along its general direction. For example, in the LIPR10.1A/1BV1 pair, the helix is shifted by about one full coil along the helical axis, and this translation can be correlated with the one-amino-acid insertion present in loop L9 of Betv1 (Figure 2). The translation of the $\alpha 3$ helix in the two copies of the LIPR10.1B protein is the same. In general, helix $\alpha 3$ is stabilized in the groove by hydrophobic interactions. Despite the high conservation of the hydrophobic residues on the concave face of the β -sheet, the exact interactions with the helix are different in each structure. This reflects both the sequence variability in the N-terminal part of the helix and its variable translation along the groove axis.

The two termini of the protein chain seem to play a role in anchoring or positioning of the flexible $\alpha 3$ helix within the core β -structure. The N-terminal glycine stabilizes the conformation of loop L9, which connects helix $\alpha 3$ with the rest of the protein. This interaction determines the turn of the main chain and directs the helix towards the groove of the β -sheet. Together with the hydrophobic interactions it could play a significant role during the protein folding process. The C terminus seems to be relatively stable, providing additional stabilizing interaction, although neither its sequence nor its structure is strictly conserved.

Implications for physiological function of PR10 proteins

Several hypotheses proposed to explain the function of PR10 proteins have focused on their RNase activity. The biological significance of these observations is controversial because of the low level of the hydrolytic activity and of the high abundance,

unusual for enzymes, of the PR10 proteins in many plant organs. It is also puzzling that one of the present proteins (LIPR10.1B) displayed some (low) RNase activity, while the close LIPR10.1A homologue none at all (Sikorski, unpublished results). This raises the possibility that the reported RNase activities may be accidental and of no biological significance. The previously postulated role of the glycine-rich motif in phosphate binding should be revised. The glycine carbonyl groups are located in close proximity to the pocket and are likely to accept hydrogen bonds from a donor-rich molecule. A phosphate anion, even in a partly protonated form, is unlikely in this role.

Another indication of the role of PR10 proteins comes from the fact that CSBP are also classified as members of this group if generous sequence similarity criteria are accepted.⁵ The biological role of CSBP proteins is much better recognized as they have been found to strongly bind both purine- and urea-type cytokinin analogs.⁵ In view of the limited sequence homology, it will be first necessary to determine the structure of typical cytokinin-binding proteins to confirm their structural similarity to other PR10 proteins and to validate their PR10 classification. A positive answer to this question is suggested by the successful docking of zeatin in the putative PR10 active site near the glycine-rich loop (above). In a more general interpretation, the simulation results suggest that the physiological role of PR10 proteins might involve binding nucleosides or their analogs. Crystallographic studies of CSBP–cytokinin complexes should provide new insights about the binding site and the nature of cytokinin binding and, hopefully, about the physiological role of other PR10 proteins as well. It will be interesting to see if all PR10 proteins share the same structural elements with CSBP proteins for binding their targets. The present structural data indicate, as the most probable, the pocket adjacent to loop L4, as this region has very high sequence conservation in both protein groups.

Concluding remarks

The two crystal structures presented here provide valuable information about the pathogenesis-related PR10 proteins. The structures confirm the PR10 folding canon established by the structure of the birch pollen allergen protein Betv1⁴³ (1BV1). It consists of a seven-stranded antiparallel β -sheet wrapped around a long C-terminal helix ($\alpha 3$). Strands $\beta 1$ and $\beta 2$ form the edges of the sheet and are separated by two short helices ($\alpha 1$, $\alpha 2$). There are nine loops connecting the secondary structure elements. Loop L5 (between strands $\beta 3$ and $\beta 4$) shows a variable degree of disorder. In contrast, loop L4, formed by a sequence containing as many as four glycine residues, is extremely rigid and has identical conformation in all known PR10 molecules. In spite of its conformational stability, its precise positioning is not absolutely fixed and

appears to be correlated with visible rigid-body movements of some structural elements of the proteins. The glycine-rich loop L4 is, together with helices $\alpha 1$ and $\alpha 2$, at the mouth of an unusual pocket formed within the protein core between the concave face of the β -sheet and helix $\alpha 3$. The shape and surface of this pocket are compatible with binding a medium-size organic ligand, such as a nucleotide or a cytokinin molecule. In the apo LIPR10.1A and LIPR10.1B structures the pocket is filled with eight well-ordered water molecules. The helix $\alpha 3$ shows a surprising degree of sequence and structural variability, particularly in its N-terminal part that is responsible for its proper location and orientation in the β -sheet groove. In 1BV1, this helix is straight and ordered, in the two molecules of LIPR10.1B it is kinked in the middle, while in LIPR10.1A it is strongly bent and disordered in the central part. In addition, it displays significant axial translations, amounting to as much as a full coil shift. In consequence, the conserved points of the β - α interactions utilize different elements from the helical component in different structures. The N terminus seems to be important for controlling the final conformation of helix $\alpha 3$. Gly1, which is conserved in all known PR10 structures, interacts with the loop through which the helix enters the β -groove. The C-terminal residues anchor the other end of the helix but there is sequence variability seen in this region. On the basis of low sequence homology (20% identity), another group of plant proteins, known as CSBP, are also included in the PR10 class. The CSBP proteins bind plant hormones (cytokinins) with very high affinity and thus, in contrast to other PR10 proteins, have a better defined physiological function. It will be interesting to see if structural studies confirm the classification of CSBP proteins in the PR10 family, and if yes, how this will help to solve the puzzle of the role of PR10 proteins in plants.

Methods

LIPR10.1A

Originally, the structure was solved using room temperature 2.45 Å diffraction data obtained for small crystals, grown from *E. coli*-expressed protein samples, as described previously.^{46,48} A molecular replacement solution was obtained⁴⁶ using the birch-pollen allergen protein 1BV1⁴³ as a probe. Subsequently, a new data set extending to 1.95 Å was collected at low temperature using beamline D41 of the LURE synchrotron and a MarResearch image plate detector. A crystal with dimensions $0.05 \times 0.1 \times 0.3$ mm³ was frozen using mineral oil as cryoprotectant according to the procedure described by Teng.⁵⁷ A total of 89,375 measured reflections were reduced to a set of 9446 unique data characterized by $R_{\text{int}} = 0.053$ and 99.5% completeness. Processing of the images and scaling of the data was done using programs DENZO and SCALEPACK from the HKL package.⁵⁸ A summary of the data processing statistics is presented in Table 1.

The initial model, obtained with the first, 2.45 Å, dataset⁴⁶ was used as the starting point for the refinements. After truncation of the regions that were poorly defined in the initial maps, the refinement started from R/R_{free} of 0.47/0.47. The refinement was done using CNS 0.9⁵⁹ and included bulk solvent correction. The maximum likelihood function was used for the refinement, which included all reflections within the 12–1.95 Å resolution range without any σ cutoff. After several cycles of simulated annealing and modeling in “O”,⁶⁰ all missing fragments of the model were reconstructed ($R/R_{\text{free}} = 0.27/0.33$). In subsequent rounds of the refinement, 104 water molecules were gradually added and rebuilding sessions alternated with positional and B -factor refinement. The refinement converged with R/R_{free} of 0.196/0.250 and produced a model with very good geometry as summarized in Table 2. The refined atomic coordinates of the LIPR10.1A structure have been deposited with the Protein Data Bank, accession number 1ICX.

LIPR10.1B

The protein sample was prepared by *E. coli* expression and purified according to a procedure similar to that described for LIPR10.1A.⁴⁸ The protein was crystallized at a concentration of 5.6 mg/ml in the presence of 0.9 M sodium citrate and 0.1 M Hepes, pH 8.0. Room-temperature diffraction data were collected for a plate-like crystal measuring $0.05 \times 0.2 \times 0.4$ mm³ using Cu K α radiation and a 300 mm MarResearch image plate detector. In the resolution range 15–2.25 Å 88,882 reflections were measured and after data reduction resulted in a 97.2% complete data set consisting of 15,594 unique reflections characterized by $R_{\text{int}} = 0.102$ (Table 1). The diffraction images were integrated using DENZO and merged using SCALEPACK from the HKL package.⁵⁸

Molecular replacement calculations were done using the EPMR program,⁶¹ with the LIPR10.1A model as a probe. The program located two protein molecules in the asymmetric unit (A and B), corresponding to a Matthews volume⁵² of 2.45 Å³/Da and 49.5% solvent content. The solution was characterized by a high correlation coefficient (0.578), but also by a high R -factor (0.525). Structure factor refinement was done in CNS 0.9 in a way analogous to that described for LIPR10.1A. Although the initial R -value was high, the first three steps of simulated annealing, performed with increasing resolution range, resulted in rapid decrease of R/R_{free} from 0.510/0.542 to 0.307/0.383. The subsequent steps of refinement included simulated annealing, positional refinement, B -factor refinement and manual modeling according to electron density maps, and addition of water molecules to a total of 112. The final R/R_{free} converged at 0.186/0.239 (Table 2). The refined atomic coordinates of the LIPR10.1B structure have been deposited with the Protein Data Bank, accession number 1IFV.

Acknowledgments

We would like to thank Alina Kasperska for skillful help with protein purification. Support from the International Centre for Genetic Engineering and Biotechnology (ICGEB) is gratefully acknowledged (grant CRP/PL00-06). Some of the calculations were performed

using the facilities of the Poznan Supercomputing and Networking Center.

References

- Somssich, I. E., Schmelzer, E., Bollmann, J. & Hahlbrock, K. (1986). Rapid activation by fungal elicitor of genes encoding "pathogenesis-related" proteins in cultured parsley cells. *Proc. Natl Acad. Sci. USA*, **83**, 2427–2430.
- Somssich, I. E., Schmelzer, E., Kawalleck, P. & Hahlbrock, K. (1988). Gene structure and *in situ* transcript localization of pathogenesis-related protein 1 in parsley. *Mol. Gen. Genet.* **213**, 93–98.
- van Loon, L. C., Pierpoint, W. S., Boller, T. & Conejero, V. (1994). Recommendations for naming plant pathogenesis-related proteins. *Plant Mol. Biol. Rep.* **12**, 245–264.
- van Loon, L. C. & van Strien, E. A. (1999). The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Phys. Mol. Plant Pathol.* **55**, 85–97.
- Fujimoto, Y., Nagata, R., Fukasawa, H., Yano, K., Azuma, M., Iida, A. *et al.* (1998). Purification and cloning of cytokinin-specific binding protein from mung bean (*Vigna radiata*). *Eur. J. Biochem.* **258**, 794–802.
- Nessler, C. L. & Burnett, R. J. (1992). Organization of the major latex protein gene family in opium poppy. *Plant Mol. Biol.* **20**, 749–752.
- Pozueta-Romero, J., Klein, M., Houlne, G., Schantz, M.-L., Meyer, B. & Schantz, R. (1995). Characterization of a family of genes encoding a fruit-specific wound-stimulated protein of bell pepper (*Capsicum annuum*): identification of a new family of transposable elements. *Plant Mol. Biol.* **28**, 1011–1025.
- Osmark, P., Boyle, B. & Brisson, N. (1998). Sequential and structural homology between intracellular pathogenesis-related proteins and a group of latex proteins. *Plant Mol. Biol.* **38**, 1243–1246.
- Crowell, D. N., John, M. E., Russel, D. & Amasino, R. M. (1992). Characterisation of stress-induced, developmentally regulated gene family from soybean. *Plant Mol. Biol.* **18**, 459–466.
- Mylona, P., Moerman, M., Yang, W.-C., Gloudemans, T., van De Kerckhove, J., van Kammen, A. *et al.* (1994). The root epidermis-specific pea gene RH2 is homologous to a pathogenesis-related gene. *Plant Mol. Biol.* **26**, 39–50.
- Breda, C., Sallaud, C., El-Turk, J., Buffard, D., de Kozak, I., Esnault, R. & Kondorosi, A. (1996). Defence reaction in *Medicago sativa*: a gene encoding a class 10 PR protein is expressed in vascular bundles. *Mol. Plant-Microbe Interact.* **9**, 713–719.
- Sikorski, M. M., Biesiadka, J., Kasperska, A. E., Kopcinska, J., Lotocka, B., Golinowski, W. & Legocki, A. B. (1999). Expression of genes encoding PR10 class pathogenesis-related proteins is inhibited in yellow lupine root nodules. *Plant Sci.* **149**, 125–137.
- Warner, S. A. J., Scott, R. & Draper, J. (1992). Characterisation of a wound-induced transcript from the monocot asparagus that shares similarity with a class of intracellular pathogenesis-related (PR10) proteins. *Plant Mol. Biol.* **19**, 555–561.
- Warner, S. A. J., Scott, R. & Draper, J. (1993). Isolation of an asparagus intracellular PR gene (AoPR1) wound-responsive promoter by the inverse polymerase chain reaction and its characterization in transgenic tobacco. *Plant J.* **3**, 191–201.
- Pinto, P. M. & Ricardo, C. P. P. (1995). *Lupinus albus* L. pathogenesis-related proteins that show similarity to PR10 proteins. *Plant Physiol.* **109**, 1345–1351.
- Utriainen, M., Kokko, H., Auriola, S., Sarrazin, O. & Karenlampi, S. (1998). PR-10 protein is induced by copper stress in roots and leaves of a Cu/Zn tolerant clone of birch, *Betula pendula*. *Plant Cell Environ.* **21**, 821–828.
- Fristensky, B., Horovitz, D. & Hadwiger, L. A. (1988). cDNA sequences for pea disease resistance response. *Plant Mol. Biol.* **11**, 713–715.
- Schmelzer, E., Kruger-Lebus, S. & Hahlbrock, K. (1989). Temporal and spatial patterns of gene expression around sites of attempted fungal infection in parsley leaves. *Plant Cell*, **1**, 993–1001.
- Matton, D. P. & Brisson, N. (1989). Cloning, expression, and sequence conservation of pathogenesis-related gene transcripts of potato. *Mol. Plant-Microbe Interact.* **2**, 325–331.
- McGee, J. D., Hamer, J. E. & Hodges, T. K. (2001). Characterization of a PR-10 pathogenesis-related gene family induced in rice during infection with *Magnaporthe grisea*. *Mol. Plant-Microbe Interact.* **14**, 877–886.
- Gamas, P., de Billy, F. & Truchet, G. (1998). Symbiosis-specific expression of two *Medicago truncatula* nodulin genes, *MtN1* and *MtN13*, encoding products homologous to plant defense proteins. *Mol. Plant-Microbe Interact.* **11**, 393–403.
- Vanek-Krebitz, M., Hoffman-Sommergruber, K., Machado, M. L. D., Susani, M., Ebner, C., Kraft, D. *et al.* (1995). Cloning and sequencing of *Mal d 1*, the major allergen from apple (*Malus domestica*), and its immunological relationship to Betv 1, the major birch pollen allergen. *Biochem. Biophys. Res. Commun.* **214**, 538–551.
- Breiteneder, H., Hoffman-Sommergruber, K., Oriordan, G., Susani, M., Ahorn, H., Ebner, C. *et al.* (1995). Molecular characterisation of *Api g 1*, the major allergen of celery (*Apium graveolens*), and its immunological and structural relationships to a group of 17-kDa tree pollen allergens. *Eur. J. Biochem.* **223**, 484–489.
- Breiteneder, H., Pettenburger, K., Bito, A., Valenta, R., Kraft, D., Rumpold, H. *et al.* (1989). The gene coding for the major birch pollen allergen *Betv 1* is highly homologous to a pea disease resistance response gene. *EMBO J.* **8**, 1935–1938.
- Warner, S. A. J., Gill, A. & Draper, J. (1994). The developmental expression of the asparagus intracellular PR protein (AoPR1) gene correlates with sites of phenylpropanoid biosynthesis. *Plant J.* **6**, 31–43.
- Swoboda, I., Scheiner, O., Kraft, D., Breitenbach, M., Heberle-Bors, E. & Vicente, O. (1994). A birch gene family encoding pollen allergens and pathogenesis-related proteins. *Biochim. Biophys. Acta*, **1219**, 457–464.
- Constabel, C. P. & Brisson, N. (1995). Stigma- and vascular-specific expression of the PR10a gene of potato: a novel expression of a pathogenesis-related gene. *Mol. Plant-Microbe Interact.* **8**, 104–113.
- Huang, J. C., Chang, F. C. & Wang, C. S. (1997). Characterization of a lily tapetal transcript that shares sequence similarity with a class of intracellular pathogenesis-related (IPR) proteins. *Plant Mol. Biol.* **34**, 681–686.

29. Apold, J., Florvaag, E. & Elsayed, S. (1981). Comparative studies on tree-pollen allergens. I. Isolation and partial characterization of a major allergen from birch pollen (*Betula verrucosa*). *Int. Arch. Allergy Appl. Immunol.* **64**, 439–447.
30. Barratt, D. H. P. & Clark, J. A. (1993). Proteins arising during the late stages of embryogenesis in *Pisum sativum* L. *Planta*, **184**, 14–23.
31. Iturriaga, E. A., Leech, M. J., Barrat, D. H. P. & Wang, T. L. (1994). Two ABA-responsive proteins from pea (*Pisum sativum* L.) are closely related to intracellular pathogenesis-related proteins. *Plant Mol. Biol.* **24**, 235–240.
32. Broderick, K., Pittock, C., Arioli, T., Creaser, E. H., Weinman, J. J. & Rolfe, B. G. (1997). Pathogenesis-related proteins in *Trifolium subterraneum*: a general survey and subsequent characterisation of a protein inducible by ethephon and redlegged earth mite attack. *Aust. J. Plant Physiol.* **24**, 819–829.
33. Carpin, S., Laffer, S., Schoentgen, F., Valenta, R., Chénieux, J.-C., Rideau, M. & Hamdi, S. (1998). Molecular characterization of a cytokinin-inducible periwinkle protein showing sequence homology with pathogenesis-related proteins and the Betv 1 allergen family. *Plant Mol. Biol.* **36**, 791–798.
34. Moiseyev, G. P., Beintema, J. J., Fedoreyeva, L. I. & Yakovlev, G. I. (1994). High sequence similarity between a ribonuclease from ginseng calluses and fungus-elicited proteins from parsley indicates that intracellular pathogenesis-related proteins are ribonucleases. *Planta*, **193**, 470–472.
35. Moiseyev, G. P., Fedoreyeva, L. I., Zhuravlev, Y. N., Yasnetskaya, E., Jekel, P. A. & Beintema, J. J. (1997). Primary structure of two ribonucleases from ginseng callus. *FEBS Letters*, **407**, 207–210.
36. Bufer, A., Spangfort, M. D., Kahlert, H., Schlaak, M. & Becker, W.-M. (1996). The major birch pollen allergen, Betv 1, shows ribonuclease activity. *Planta*, **199**, 413–415.
37. Swoboda, I., Hoffman-Sommergruber, K., O'Riordain, G., Scheiner, O., Heberle-Bors, E. & Vicente, O. (1996). Betv 1 proteins, the major birch pollen allergens and members of a family of conserved pathogenesis-related proteins, show ribonuclease activity *in vitro*. *Physiol. Plant.* **96**, 433–438.
38. 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.
39. Huang, S., Lee, H. S., Karunanandaa, B. & Kao, T. H. (1994). Ribonuclease activity of *Petunia inflata* S proteins is essential for rejection of self-pollen. *Plant Cell*, **6**, 1021–1028.
40. Sikorski, M. M., Szybiak-Strozycka, U., Strozycki, P., Golinska, B., Madrzak, C. J., Kamp, R. M. *et al.* (1989). Coordinated synthesis of leghemoglobin and root protein R18 in yellow lupine. *Acta Biochim. Polon.* **36**, 63–72.
41. Walter, M. H., Liu, J.-W., Wunn, J. & Hess, D. (1996). Bean ribonuclease-like pathogenesis-related protein genes display complex patterns of developmental, dark-induced and exogenous-stimulus-dependent expression. *Eur. J. Biochem.* **239**, 281–293.
42. Sikorski, M. M., Handschuh, L., Biesiadka, J. & Legocki, A. B. (2000). Two subclasses of yellow lupine PR10 proteins and their possible function during the symbiosis development. In *Nitrogen Fixation: From Molecules to Crop Productivity. Current Plant Science and Biotechnology in Agriculture* (Pedrosa, F. O., Hungria, M., Yates, M. G. & Newton, W. E., eds), vol. 38, pp. 319–322, Kluwer Academic Publishers, Dordrecht.
43. Gajhede, M., Osmark, P., Poulsen, F. M., Ipsen, H., Larsen, J. N., van Neerven, R. J. J. *et al.* (1996). X-ray and NMR structure of Betv 1, the origin of birch pollen allergy. *Nature Struct. Biol.* **3**, 1040–1045.
44. Spangfort, M. D., Mirza, O., Svensson, L. A., Larsen, J. N., Gajhede, M. & Ipsen, H. (1999). Crystallization and preliminary X-ray analysis of birch-pollen allergen Betv1 in complex with a murine monoclonal IgG Fab' fragment. *Acta Crystallog. sect. D*, **55**, 2035–2036.
45. Mirza, O., Henriksen, A., Ipsen, H., Larsen, J. N., Wissenbach, M., Spangfort, M. D. & Gajhede, M. (2000). 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 Betv1. *J. Immunol.* **165**, 331–338.
46. Biesiadka, J., Sikorski, M. M., Bujacz, G. & Jaskolski, M. (1999). Crystallization and preliminary X-ray structure determination of *Lupinus luteus* PR10 protein. *Acta Crystallog. sect. D*, **55**, 1925–1927.
47. Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). Use of T7 polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89.
48. Sikorski, M. M. (1997). Expression of *Lupinus luteus* cDNA coding for PR10 protein in *Escherichia coli*: purification of the recombinant protein for structural and functional studies. *Acta Biochim. Polon.* **44**, 565–578.
49. Hirel, Ph.-H., Schmitter, J.-M., Dessen, P., Fayat, G. & Blanquet, S. (1989). Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. *Proc. Natl Acad. Sci. USA*, **86**, 8247–8251.
50. Bariola, P. A. & Green, P. J. (1997). Ribonucleases. In *Plant Ribonucleases: Structure and Function* (D'Alessio, G. & Riordan, J. F., eds), pp. 63–190, Academic Press, New York.
51. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95–99.
52. Matthews, B. W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497.
53. Saraste, M., Sibbald, P. R. & Wittinghofer, A. (1990). The P-loop—a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.* **15**, 430–434.
54. Bossemeyer, D. (1994). The glycine-rich sequence of protein kinases: a multifunctional element. *Trends Biochem. Sci.* **19**, 201–205.
55. Cohen, G. R. (1997). ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. *J. Appl. Crystallog.* **30**, 1160–1161.
56. Jones, G., Willett, P., Glen, R. C., Leach, A. R. & Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **267**, 727–748.
57. Teng, T. Y. (1990). Mounting of crystals for macromolecular crystallography in free-standing thin film. *J. Appl. Crystallog.* **23**, 387–391.
58. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
59. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gross, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system (CNS): a new

- software system for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
60. Jones, T. A., Cowan, S., Zou, J. Y. & Kjeldgaard, M. (1994). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
61. Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). Rapid automated molecular replacement by evolutionary search. *Acta Crystallog. sect. D*, **55**, 484–491.
62. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.
63. Pai, E. F., Krengel, U., Petsko, G. A., Goody, R. S., Kabsch, W. & Wittinghofer, A. (1990). Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. *EMBO J.* **9**, 2351–2359.

Edited by R. Huber

(Received 2 January 2002; received in revised form 15 April 2002; accepted 15 April 2002)