Biosci. Rep. (2012) / **32** / 567–575 (Printed in Great Britain) / doi 10.1042/BSR20120058

Original Paper

Bioscience Reports

Solution structure of the strawberry allergen Fra a 1

Christian SEUTTER VON LOETZEN*, Kristian SCHWEIMER*, Wilfried SCHWAB†, Paul RÖSCH* and Olivia HARTL-SPIEGELHAUER*[1]

*Lehrstuhl Biopolymere und Forschungszentrum für Bio-Makromoleküle, Universität Bayreuth, Universitätsstr. 30, 95447 Bayreuth, Germany, and †Biotechnology of Natural Products, Technische Universität München, Liesel-Beckmann-Str. 1, 85354 Freising, Germany

Synopsis

The PR10 family protein Fra a 1E from strawberry (*Fragaria x ananassa*) is down-regulated in white strawberry mutants, and transient RNAi (RNA interference)-mediated silencing experiments confirmed that Fra a 1 is involved in fruit pigment synthesis. In the present study, we determined the solution structure of Fra a 1E. The protein fold is identical with that of other members of the PR10 protein family and consists of a seven-stranded antiparallel β-sheet, two short V-shaped α-helices and a long C-terminal α-helix that encompass a hydrophobic pocket. Whereas Fra a 1E contains the glycine-rich loop that is highly conserved throughout the protein family, the volume of the hydrophobic pocket and the size of its entrance are much larger than expected. The three-dimensional structure may shed some light on its physiological function and may help to further understand the role of PR10 proteins in plants.

Key words: Bet v 1 superfamily, Fra a 1, *Fragaria x ananassa*, NMR structure, pathogenesis-related protein

Cite this article as: Seutter von Loetzen, C., Schweimer, K., Schwab, W., Rösch, P. and Hartl-Spiegelhauer, O. (2012) Solution structure of the strawberry allergen Fra a 1. Biosci. Rep. **32**, 567–575

www.bioscirep.org

Bioscience Reports

INTRODUCTION

In North and Central Europe as well as in North America approximately 15–20 % of the population are allergic to pollen, with birch pollen being the main cause of those allergies [1,2]. Over 90 % of the birch pollen allergic patients show sensitization to the major birch pollen allergen Bet v 1 [3]. In total 50–90 % of the patients who are sensitized to Bet v 1 show IgE-mediated allergic cross reactions to food allergens such as those contained in fruits, vegetables and nuts [4,5]. These cross reactions are often based on the similarity of the three-dimensional allergen structures [6–9].

A series of Bet v 1 homologous food allergens have been identified and summarized in the superfamily of Bet v 1-like proteins, such as Api g 1 from celery [10], Mal d 1 from apple [11], Pru av 1 from cherry [12] and Gly m 4 from soya bean [13]. The members of the Bet v 1 superfamily are 153–160 amino acid proteins with molecular mass ~17 kDa and theoretical isoelectric points between pH 4.4 and 6.1. The sequence identities with Bet v 1 are 40–70 %, and all these proteins show similar conformations with identical arrangements of secondary structure elements that correlate with the cross-reactivity.

The pollen and food allergens belong to the class 10 of pathogenesis-related proteins, protein family PR10, a subdivision of the Bet v 1-like superfamily. The PR10 proteins are conserved and ubiquitous in the plant kingdom, with roles in plant development [17] and immune defence induced by microbial attack [18], fungal elicitors [18,19], wounding [20] or stress stimuli [21]. The PR10 family contains, for example, the yellow lupine proteins LIPR-10.2A and LIPR-10.2B [14,15] and the St. John's Wort protein Hyp1 [16]. The classification of this protein group is based on the 3D (three-dimensional) structure. All structurally studied proteins show the same overall fold. However, their biological function is still not well recognized.

As an additional member of the Bet v 1 superfamily, a strawberry (*Fragaria x ananassa*) protein was detected in red fruited genotypes and named Fra a 1 [22]. A hypoallergenic white fruited genotype was almost devoid of the protein. Subsequently, five potential Fra a 1 isoforms (Fra a 1A–Fra a 1E) were identified with only seven variable amino acid positions, and it could be shown that Fra a 1A is recognized by IgE antibodies from strawberry allergic patients [23,24]. Two additional Fra a isoforms with greater variability were identified and named Fra a 2 and Fra a 3 [25]. Whereas *Fra a 1e* transcripts are expressed predominantly in roots with decreasing expression levels

Abbreviations used: LB, Luria–Bertani; NOESY, nuclear Overhauser enhancement spectroscopy; rmsd, root mean square deviation; RNAi, RNA interference.

[1] To whom correspondence should be addressed (email Olivia.Hartl-Spiegelhauer@uni-bayreuth.de).

© 2012 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

during fruit ripening, *Fra a 2* expression increases in ripe fruits and *Fra a 3* expression is constant in all stages. Transient RNAi (RNA interference)-mediated silencing of the *Fra a* genes in strawberries leads to reduced levels of anthocyanins and upstream metabolites as well as down-regulation of chalcone synthase (*FaCHS*) and *PAL* (phenylalanine ammonia lyase) genes [25], clearly indicating a key role of Fra a isoforms in the pigment synthesis pathway.

These different appearances of the members of the Bet v 1 superfamily of proteins, combined with the observation of divergent expression patterns in different species and different isoforms within a single species, renders it difficult to assign clear physiological functions to these proteins. One of the possible approaches to clarify the role of the Bet v 1 proteins in Nature is the study and comparison of their conformations in detail, an approach we here continue with the presentation of the solution structure of Fra a 1 in atomic resolution.

MATERIALS AND METHODS

Expression of recombinant *Fra a 1e*

The *Fra a 1e* gene was cloned into the expression vector pQE70 (Qiagen) by using the restriction sites SphI and BamHI. The protein contained a His_6 tag at the C-terminus. The recombinant plasmid was transformed together with the vector pREP4 (Qiagen) into *Escherichia coli* BL21 (DE3). *Fra a 1e* was expressed in the host strain in LB (Luria–Bertani) medium at 20°C over 4 h. To obtain ^{15}N- and $^{15}N/^{13}C$-labelled protein overexpressing *E. coli* were grown on M9 minimal medium with ^{15}N enriched $(^{15}NH_4)_2SO_4$ or ^{15}N enriched $(^{15}NH_4)_2SO_4$/[U-^{13}C]glucose respectively.

Protein purification

Crude cell extracts were prepared in 50 mM Hepes buffer (pH 7.0) with 300 mM NaCl for purification. Soluble His-tagged Fra a 1E was purified by immobilized metal-ion affinity chromatography [Ni-NTA (Ni^{2+}-nitrilotriacetate) column with a bed volume of 1 ml; Qiagen] using a step gradient from 20 to 300 mM imidazole. Fractions were analysed by SDS/19 % PAGE, and pure protein fractions were pooled and dialysed against 20 mM Hepes buffer, pH 7.0, containing 150 mM NaCl. This buffer was used as equilibration and elution buffer of a size-exclusion chromatography to remove residual impurities (Superdex S75 GL 10/300 column, total bed volume: 24 ml; GE Healthcare). Two litres of cell culture from LB and M9 minimal medium yielded approximately 11.5 and 7.6 mg of protein respectively. The purity of recombinant Fra a 1E exceeded 95 % as estimated from SDS/PAGE on 19 % gels. The DC protein assay (BioRad) as well as a molar absorption coefficient of $\varepsilon_{280} = 14900\ M^{-1} \cdot cm^{-1}$ were used to calculate protein concentrations. Samples of Bet v 1.0101 were prepared as described previously [26].

Size-exclusion chromatography

In order to determine the oligomerization state and the molecular mass of recombinant Fra a 1E, size-exclusion chromatography was performed with a Superdex S75 GL 10/300 column (total bed volume: 24 ml; GE Healthcare). Potassium phosphate buffer, 10 mM, pH 7.0, containing 150 mM NaCl was used as sample buffer as well as for column equilibration and for isocratic protein elution. Molecular mass calibration of the column was performed with albumin (67.0 kDa), ovalbumin (43.0 kDa), carbonic anhydrase (29.0 kDa) and cytochrome *c* (12.4 kDa).

CD spectroscopy

Far ultraviolet CD measurements were recorded on a Jasco J-810 spectropolarimeter (Japan Spectroscopic) equipped with a 1-mm quartz cuvette (Hellma). Spectra of 10 μM Fra a 1E and 7.0 μM Bet v 1.0101 were obtained at 20°C in 1 mM Hepes buffer, pH 7.0, containing 5 mM NaCl. At least 10 scans in the range between 260 and 190 nm with 50 nm/min scanning speed were averaged for each measurement, and the resulting spectrum was smoothed and normalized using the following formula to calculate the mean residue molar ellipticity:

$$[\Theta]_{MRW} = \Theta / c \cdot d \cdot N$$

where Θ is the measured ellipticity, MRW is the mean residue mass, c is the protein concentration, d is the path length and N is the number of amino acids.

Computational methods

Multiple and pairwise sequence alignments were performed with ClustalW (EMBL-EBI) [27] and LALIGN (EMBL-EBI) [28], and homologous protein structures were aligned with the program Coot [29]. The cavity volumes of the Bet v 1 homologous proteins were calculated with CASTp with default parameters [30]. For NMR structures, the cavity volumes were determined for every single structure of the NMR bundle. The quoted volumes are means ± S.D. Figures of protein structures were generated using the program PyMOL v1.0 (2009, DeLano Scientific LLC).

NMR spectroscopy

All NMR experiments were performed at 298 K on a Bruker Avance 800 MHz spectrometer with cryogenically cooled triple resonance probes equipped with pulsed field-gradient capabilities, and standard double- and triple-resonance through-bond correlation. NMR spectra to assign chemical shifts [31] were obtained with a 700 μM ^{15}N, ^{13}C-labelled Fra a 1E sample in 20 mM Hepes, pH 7.0, 150 mM NaCl and 10 % ^2H_2O. Three-dimensional ^{13}C and ^{15}N edited NOESY (nuclear Overhauser enhancement spectroscopy) experiments (mixing times 120 ms) were recorded for derivation of distance restraints. NMR data were processed using in-house software and visualized with NMRView [32].

© 2012 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

Table 1 Comparison of Bet v 1 homologous proteins

Protein	PDB code (method)	Sequence identity to Fra a 1E (%)	Sequence similarity to Fra a 1E (%)	Backbone rmsd values (Å)	Backbone rmsd values for secondary structures (Å)	Cavity size (Å^3)	Reference
Fra a 1E	2LPX (NMR)	100	100	–	–	2468 ± 440	Present study
Pru av 1.0101	1E09 (NMR)	79.4	95.0	1.82	1.04	1643 ± 280	[6]
Bet v 1.0101	1BTV (NMR)	53.8	81.2	2.33	1.71	1633 ± 152	[45]
	1BV1 (X-ray)	53.8	81.2	1.85	1.32	1527	[45]
Gly m 4.0101	2K7H (X-ray)	52.5	78.1	2.32	1.71	1621 ± 242	[9]
LIPR10.2B	2QIM (X-ray)	51.9	80.6	1.24	1.93	2181.7	[41]
Bet v 1.1001	1FM4 (X-ray)	51.2	80.6	1.28	1.93	2359	[38]
LIPR10.2A	1XDF (X-ray)	50.9	78.6	1.42	1.80	329.3	[46]

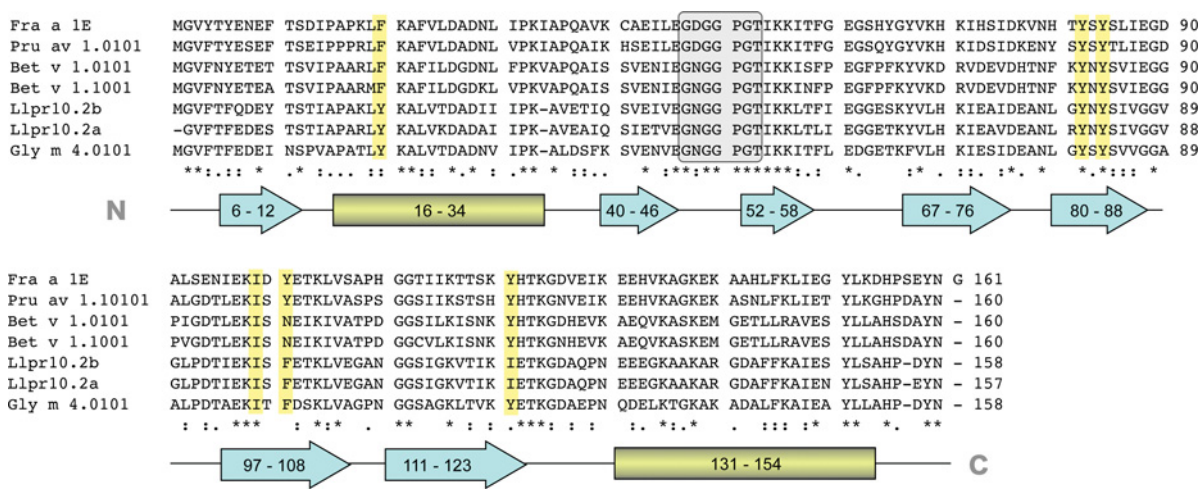

Figure 1 Sequence alignment of Bet v 1 homologous proteins with known structure
Multiple sequence alignment of Bet v 1 homologous proteins as performed with ClustalW: Fra a 1E (*Fragaria x ananassa*, strawberry), Pru av 1 (*Prunus avium*, cherry), Bet v 1 (*Betula verrucosa*, birch), LIPR-10.2A (*Lupinus luteus*) and Gly m 4 (*Glycine max*, soya bean). Amino acids are marked with asterisks (identical), colons (conserved) and dots (semi-conserved). The highly conserved glycine-rich loop region is highlighted by a grey box. Conserved cavity residues are highlighted in yellow. Secondary structure elements are indicated as present in Fra a 1E (bottom line: β-sheets, blue; α-helices, green).

Structure calculation

NOESY cross peaks were classified according to their relative intensities and converted into distance restraints with upper limits of 3.0 Å (strong), 4.0 Å (medium), 5.0 Å (weak) and 6.0 Å (very weak) (where 1 Å = 0.1 nm). For ambiguous distance restraints, the r^{-6} summation over all assigned possibilities defined the upper limit. Dihedral restraints were taken from analysis of chemical shifts by the TALOS software package [33]. Hydrogen bonds were included for backbone amide protons in regular secondary structure, if the amide proton did not show a water exchange cross peak in the ^{15}N-edited NOESY spectrum. Structures were calculated with the program XPLOR-NIH 1.2.1 [34] using a three-step simulated annealing protocol with floating assignment of prochiral groups including a conformational database potential.

The 20 structures showing the lowest values of the target function excluding the database potential were further analysed with XPLOR-NIH 1.2.1 [34], MolMol [35] and PROCHECK 3.5.4 [36]. The structure coordinates were deposited in the protein data bank (pdb accession number 2LPX), and chemical shift assignments were deposited in the BioMagResBank (BMRB accession code 18281).

RESULTS AND DISCUSSION

General comparison, CD spectroscopy and size-exclusion chromatography

According to sequence alignment with PR10 proteins with known structure, Fra a 1E is most similar to the major cherry allergen Pru av 1 with a sequence identity of 79.4 % [6] (Figure 1 and Table 1). The glycine-rich loop motif, residues 46–51, is highly conserved throughout the Bet v 1 superfamily.

© 2012 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

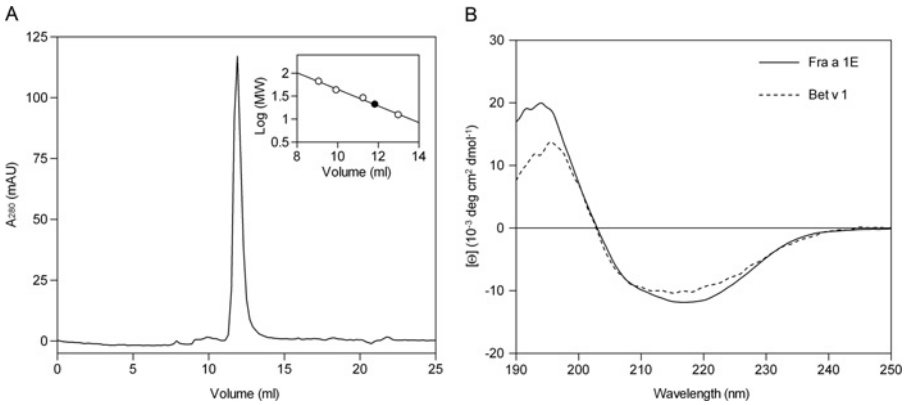

Figure 2 Biophysical properties of Fra a 1E
(**A**) Size-exclusion chromatography of Fra a 1E with a Superdex S75 GL 10/300 column. The run was performed with 0.35 mg of Fra a 1E in 10 mM potassium phosphate buffer, pH 7.0 and 150 mM NaCl. The fit to the data obtained for the molecular masses of the standard proteins (inset, open circles) was used for the determination of the molecular mass of Fra a 1E (closed circle). (**B**) Far UV CD spectra of Fra a 1E (continuous line) and Bet v 1.0101 (dotted line) were acquired at 20 °C with a bandwith of 1 nm, a sensitivity of 100 mdeg and a data density of 10 points/nm in a 1 mm quartz cuvette with 10 μM Fra a 1E in 1 mM Hepes buffer, pH 7.0, containing 5 mM NaCl, and 7 μM Bet v 1.0101 in identical buffer.

Size-exclusion chromatography of Fra a 1E revealed a single peak corresponding to a molecular mass of 20.8 kDa, in good agreement with the theoretical mass of 19.0 kDa (Figure 2A). No additional peaks could be detected, indicating that under the present conditions Fra a 1E is monomeric.

The CD spectrum of Fra a 1E displays characteristic features of a folded protein (Figure 2B), but differs from the Bet v 1 CD spectrum in several respects: the minima wavelengths (Fra a 1E: 217 nm; Bet v 1: 215 nm) as well as the maxima wavelengths (Fra a 1E: 194 nm; Bet v 1: 196 nm) are slightly different, the Bet v 1 CD 196 nm has lower amplitude in maximum as well as minimum ellipticity.

NMR spectroscopy

The NMR spectra of Fra a 1E show signal dispersion characteristics of a well-folded protein, and the application of double- and triple-resonance NMR experiments yielded nearly complete assignments of all ^{1}H, ^{13}C and ^{15}N resonances (see Supplementary Figures S1 and S2 at http://www.bioscirep.org/bsr/032/bsr0320567add.htm). In total 1601 experimentally derived restraints were obtained for the structure calculation. The resulting ensemble of 20 structures shows good stereochemical properties and only low violation of experimental and geometrical restraints, it is well defined with an overall backbone rmsd (root mean square deviation) of 0.92 Å for residues Gly^{2}–Asp^{154} and 0.54 Å for the regions with defined secondary structure only. More than 90 % of the residues are located in the most favoured regions of the Ramachandran map (see Table 2).

Description of the structure

The solution structure (Figure 3A) of Fra a 1E consists of a seven-stranded antiparallel β-sheet (residues Tyr^{6}–Ser^{12}, Lys^{40}–Glu^{46}, Gly^{52}–Thr^{58}, Tyr^{67}–Asp^{76}, His^{80}–Glu^{88}, Glu^{97}–Ala^{108} and

Table 2 Statistics for structure calculations
NOE, nuclear Overhauser effect.

Experimental restraints	Value
Distance restraints	
NOE	1270
Intraresidual	52
Sequential	287
Medium range	340
Long range	591
Hydrogen bonds	122
Dihedral restraints	209
Restraint violation	
Average distance restraint violation	0.0027 ± 0.001 Å
Maximum distance restraint violation	0.13 Å
Average dihedral restraint violation	0.15 ± 0.07°
Maximum dihedral restraint violation	3.5°
Deviation from ideal geometry	
Bond length	0.00053 ± 0.00002 Å
Bond angle	0.11 ± 0.004°
Co-ordinate precision*	
Val^{2}–Asp^{154}	
Backbone heavy atoms	0.91 Å
All heavy atoms	1.42 Å
Defined secondary structure†	
Backbone heavy atoms	0.54 Å
All heavy atoms	1.06 Å
Ramachandran plot statistics‡	90.3 %, 7.9 %, 0.7 %, 1.1%

*The precision of the coordinates is defined as the average atomic rmsd between the accepted simulated annealing structures and the corresponding mean structure calculated for the given sequence regions.
†The region of defined secondary structure is composed of Val^{2}–Ser^{11}; Ala^{16}–Ile^{34}; Lys^{40}–Glu^{46}; Gly^{53}–Thr^{58}; Tyr^{67}–Asp^{76}; His^{80}–Glu^{88}; Glu^{97}–Ala^{108}; Gly^{111}–Thr^{123}; Lys^{130}–Lys^{153}.
‡Ramachandran plot statistics are determined by PROCHECK and noted by most favoured/additionally allowed/generously allowed/disallowed.

© 2012 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

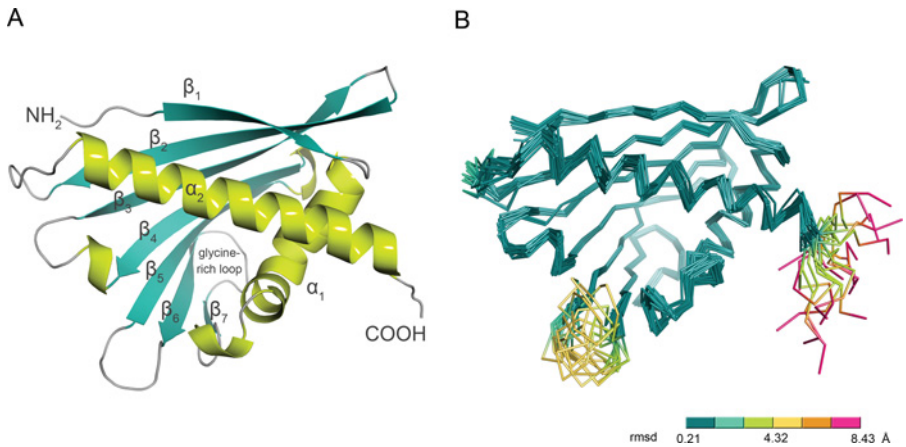

Figure 3 The overall structure of Fra a 1E
(**A**) Cartoon representation of the lowest energy solution structure of Fra a 1E (α-helices, green; β-sheets, blue; loop-regions, grey). (**B**) Backbone overlay of the 20 lowest energy solution structures of Fra a 1E, with backbone rmsd values (0.21–1.58 Å, blue; 1.58–2.95 Å, dark green; 2.95–4.32 Å, light green; 4.32–5.69 Å, yellow; 5.69–7.06 Å, orange; 7.06–8.43 Å, pink). Higher rmsd values generally correspond to higher structural flexibility.

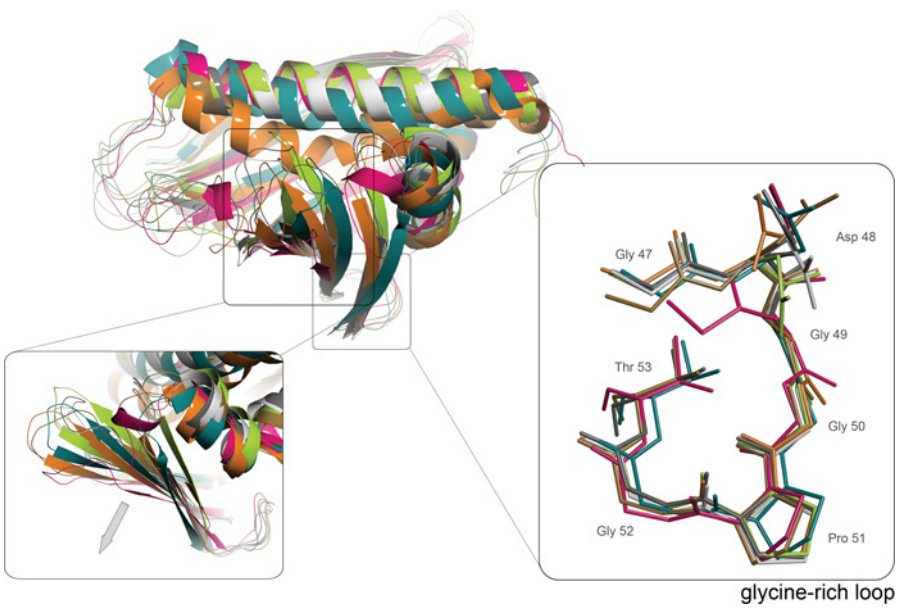

Figure 4 Comparison of Fra a 1E with homologous proteins
Overlay of the average Fra a 1E structure (2LPX, pink) with the structures of Pru av 1 (1E09, brown), Bet v 1.0101 (1BV1, dark grey), the hypoallergenic Bet v 1 isoform Bet v 1.1001 (1FM4, light grey), Gly m 4 (2K7H, dark green), LIPR-10.2A (1XDF, orange) and LIPR-10.2B (2QIM, light green). The structures are displayed in cartoon representation. The insets show the glycine-rich loop region, residues Gly^{47}–Thr^{53}, in stick representation and the region around β-strands 5 and 6 in cartoon representation, colours as above.

Gly^{111}–Thr^{123}) that is closed by a long C-terminal (residues Glu^{131}–Asp^{154}) and a short V-shaped α-helix (residues Ala^{16}–Ile^{34}) with a kink at Asp^{26}. The helices and the β-strands encompass a very large hydrophobic cavity, as is typical for Bet v 1 homologous proteins. A comparison of the secondary structure elements of Fra a 1E with the structures listed in Table 1 reveals backbone atomic rmsd values between 1.04 and 1.93 Å, with the most closely related structure being that of Pru av 1 [6] which also has the highest sequence homology to Fra a 1E. Whereas these proteins show the same overall fold, there detailed structures contain subtle differences.

For example, strands β5 and β6 are bent away from the hydrophobic pocket (Figure 4, inset, grey arrow) in Fra a 1E, and the orientation of the C-terminal helix differs between the homologous proteins; both variations influence size and shape of the hydrophobic cavity.

© 2012 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

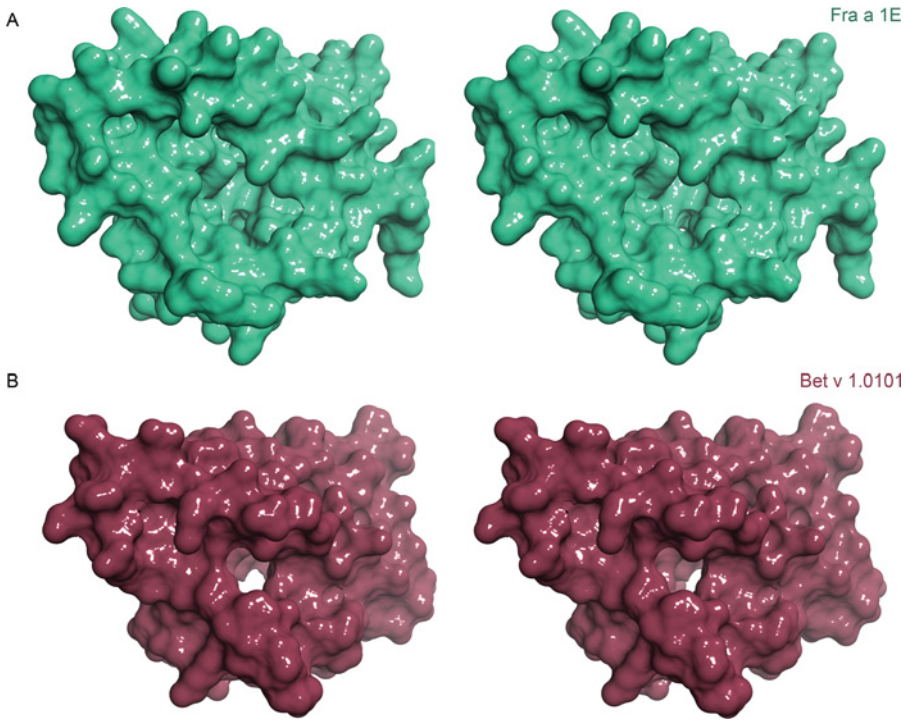

Figure 5 The hydrophobic pocket
Stereoscopic view of surface representations of (**A**) Fra a 1E (green) and (**B**) Bet v 1 (red), with orientation close to Figure 3 displaying the large entrance to the hydrophobic cavity.

Several studies were focused on ligand binding to this hydrophobic pocket. Binding, transport or storage are obvious possibilities for the physiological function of these proteins [6,37–40]. A range of ligands that might be of physiologic relevance could be identified in an ANS (8-anilinonaphthalene-1-sulfonic acid) displacement assay with Bet v 1.0101, and compounds such as flavonoids, fatty acids and cytokinins bind with low micromolar affinities to the protein [37]. Interestingly, the cytokinin kinetin seems to have a binding site alternative to the hydrophobic cavity. Structural studies revealed that two or three ligand molecules can bind inside the pocket, mainly stabilized by hydrophobic and a few hydrogen-bonding interactions [38,41,42].

The structural differences are also reflected in the differences of the CD spectra. Fra a 1E contains 32 % α-helical, 39 % β-sheet and 29 % unstructured regions, while Bet v1 shows 24 % α-helical, 39 % β-sheet and 37 % unstructured regions.

The hydrophobic cavity

Like in other family members, the cavity of Fra a 1E is accessible via three structural gaps, between the long C-terminal α-helix and β1, close to the glycine-rich loop, and between C-terminal α-helix and the loop from Thr^{58}–Tyr^{67}.

The first gap is not present in all homologues as its presence depends on exact shape and arrangement of the C-terminal helix. The second gap is quite small, and its size of approximately 4–9 Å depends on the nature of the amino acid side chains in different homologues (in Fra a 1E: 5 Å). The glycine-rich loop between β2 and β3, from Gly^{47}–Thr^{53}, has the consensus sequence GXGGXGT which is highly conserved throughout the Bet v 1 superfamily (rmsd of 0.35 Å for the alignment of all loop atoms, Figure 4, inset). This loop was suggested to be a binding region for kinetin [39] and to play an essential role in IgG and IgE binding [6,7,43,44], making it a general IgE binding epitope responsible for cross reactions between Bet v 1 homologous proteins [4].

The third gap, between the C-terminal α-helix and the loop from Thr^{58}–Tyr^{67}, is the largest. Compared with Bet v 1 and Pru av 1 [6,45], this loop is less well defined in Fra a 1E (Figure 3B). The loop flexibility is strongly related to the cavity volume, as can be judged from the volume difference of approximately 200 $Å^3$ between the Fra a 1E structures with the highest rmsd difference of the loops.

The major difference between other members of the Bet v 1 superfamily and Fra a 1E is the large volume of its hydrophobic pocket of roughly 2500 $Å^3$. The cavities of Bet v 1, Pru av 1 or Gly m 4 are only approximately 1600 Å, and the cavity of LIPR-10.2A is even smaller at 330 $Å^3$, due to a strong inward kink of the C-terminal helix. The flexible orientation of the loop Thr^{58}–Tyr^{67} as well as the displacements of strands β5 and β6 in Fra a 1E also contributes to the remarkable size of its cavity. Compared with the homologous structures, the main cavity entrance is much larger in Fra a 1E (Figures 5A and 5B) due to the different orientation of the C-terminal helix and the rearrangement of amino acids in

© 2012 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

the N-terminus of the helix. This helix aligns well with those of Pru av 1 and LIPR-10.2B (Figure 4) and, in addition, in Fra a 1E the helix residues Glu^{131}–Glu^{139} are shifted away from the hydrophobic core, moving the larger amino acids such as Lys^{130}, Glu^{131}, His^{132}, Lys^{135} and Glu^{139} away from the cavity entrance. The increased size of this gap is also due to the movement of loop Pro^{36}–Ala^{38} away from the entrance on the opposite side of the C-terminal helix, so that the protein backbone and Gln^{37} are not restrictive.

The residues inside the Fra a 1E cavity are very similar to those of homologous proteins, particularly Pru av 1 ([6], Figure 1). The hydrophobic residues Phe^{20}, Tyr^{82}, Tyr^{84}, Ile^{99}, Tyr^{101} and Tyr^{121} are widely conserved and adopt similar conformations in the different structures, pointing to similar functions. Indeed, these amino acids were suggested to establish van der Waals interactions with ligands [6,38,40].

Structural relatives

In spite of the overall folds being highly similar, the structural alignment of the homologous proteins reveals several differences, and, although Fra a 1E may well play a role in substrate binding, storage, or transport, the substrate class may be different from the other Bet v 1 proteins. LIPR-10.2A and LIPR-10.2B from yellow lupine are rather flexible proteins that can undergo conformational adaptations upon ligand binding. This might explain the large variations of the cavity volumes observed for LIPR-10.2A, LIPR-10.2B, Bet v 1.0101 and Bet v 1.1001 (Table 1). LIPR-10.2B binding of three *trans*-zeatin molecules inside the pocket leads to an increase of the cavity volume by 1800 $Å^3$ [41,42,46] when compared with LIPR-10.2A in the empty state. The Bet v 1.1001 cavity volume increases by 700 $Å^3$ when two molecules of deoxycholate are bound [38]. Hence, the 2500 $Å^3$ cavity of Fra a 1E provides for accommodation of more copies of a single ligand or, alternatively, a larger ligand.

As Fra a 1E is expressed in different plant compartments, it may fulfil various functions [25]. In addition to all the ligand-binding studies it was observed that Bet v 1 is able to bind and permeabilize membranes, and ligand binding may facilitate translocation [47]. One may thus speculate that Fra a 1E is involved in the transport of hydrophobic ligands into plant vacuoles for storage, or that it may bind hydrophobic compounds necessary for cell growth, such as plant hormones or flavonoids. To contribute to ending these speculations, the current focus of our studies is the identification of physiological ligands of the Bet v 1 superfamily proteins.

AUTHOR CONTRIBUTION

Olivia Hartl-Spiegelhauer designed the experiments and, together with Paul Rösch, co-ordinated the study. Christian Seutter von Loetzen performed the majority of the experiments together with Kristian Schweimer, who designed the NMR experiments. Wilfried Schwab provided reagents and conceptual input for the completion of the paper. Paul Rösch and Olivia Hartl-Spiegelhauer wrote the paper. All authors read and approved the manuscript.

FUNDING

This work was supported by Sonderforschungsbereich (SFB) [grant number 924 (to W.S.)].

REFERENCES

1 Wuthrich, B., Schindler, C., Leuenberger, P. and Ackermann-Liebrich, U. (1995) Prevalence of atopy and pollinosis in the adult population of Switzerland (SAPALDIA study). Swiss Study on Air Pollution and Lung Diseases in Adults. Int. Arch. Allergy Immunol. **106**, 149–156

2 Vieths, S., Scheurer, S. and Ballmer-Weber, B. (2002) Current understanding of cross-reactivity of food allergens and pollen. Ann. N. Y. Acad.Sci. **964**, 47–68

3 Ipsen, H. and Lowenstein, H. (1983) Isolation and immunochemical characterization of the major allergen of birch pollen (*Betula verrucosa*). J. Allergy Clin. Immunol. **72**, 150–159

4 Dreborg, S. (1988) Food allergy in pollen-sensitive patients. Ann. Allergy **61**, 41–46

5 Geroldinger-Simic, M., Zelniker, T., Aberer, W., Ebner, C., Egger, C., Greiderer, A., Prem, N., Lidholm, J., Ballmer-Weber, B. K., Vieths, S. and Bohle, B. (2011) Birch pollen-related food allergy: clinical aspects and the role of allergen-specific IgE and IgG4 antibodies. J. Allergy Clin. Immunol. **127**, 616–622.e1

6 Neudecker, P., Schweimer, K., Nerkamp, J., Scheurer, S., Vieths, S., Sticht, H. and Rösch, P. (2001) Allergic cross-reactivity made visible: solution structure of the major cherry allergen Pru av 1. J. Biol. Chem. **276**, 22756–22763

7 Neudecker, P., Lehmann, K., Nerkamp, J., Haase, T., Wangorsch, A., Fotisch, K., Hoffmann, S., Rösch, P., Vieths, S. and Scheurer, S. (2003) Mutational epitope analysis of Pru av 1 and Api g 1, the major allergens of cherry (*Prunus avium*) and celery (*Apium graveolens*): correlating IgE reactivity with three-dimensional structure. Biochem. J. **376**, 97–107

8 Schirmer, T., Hoffimann-Sommergrube, K., Susani, M., Breiteneder, H. and Markovic-Housley, Z. (2005) Crystal structure of the major celery allergen Api g 1: molecular analysis of cross-reactivity. J. Mol. Biol. **351**, 1101–1109

9 Berkner, H., Neudecker, P., Mittag, D., Ballmer-Weber, B. K., Schweimer, K., Vieths, S. and Rösch, P. (2009) Cross-reactivity of pollen and food allergens: soybean Gly m 4 is a member of the Bet v 1 superfamily and closely resembles yellow lupine proteins. Biosci. Rep. **29**, 183–192

10 Breiteneder, H., Hoffmann-Sommergruber, K., O'Riordain, G., Susani, M., Ahorn, H., Ebner, C., Kraft, D. and Scheiner, O. (1995) Molecular characterization 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. **233**, 484–489

11 Vanek-Krebitz, M., Hoffmann-Sommergruber, K., Laimer da Camara Machado, M., Susani, M., Ebner, C., Kraft, D., Scheiner, O. and Breiteneder, H. (1995) Cloning and sequencing of Mal d 1, the major allergen from apple (*Malus domestica*), and its immunological relationship to Bet v 1, the major birch pollen allergen. Biochem. Biophys. Res. Commun. **214**, 538–551

12 Scheurer, S., Metzner, K., Haustein, D. and Vieths, S. (1997) Molecular cloning, expression and characterization of Pru a 1, the major cherry allergen. Mol. Immunol. **34**, 619–629

13 Kleine–Tebbe, J., Vogel, L., Crowell, D. N., Haustein, U. F. and Vieths, S. (2002) Severe oral allergy syndrome and anaphylactic reactions caused by a Bet v 1- related PR-10 protein in soybean, SAM22. J. Allergy Clin. Immunol. **110**, 797–804

© 2012 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

14 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. Pol. **44**, 565–578

15 Handschuh, L., Femiak, I., Kasperska, A., Figlerowicz, M. and Sikorski, M. M. (2007) Structural and functional characteristics of two novel members of pathogenesis-related multigene family of class 10 from yellow lupine. Acta Biochim. Pol. **54**, 783–796

16 Bais, H. P., Vepachedu, R., Lawrence, C. B., Stermitz, F. R. and Vivanco, J. M. (2003) Molecular and biochemical characterization of an enzyme responsible for the formation of hypericin in St. John's wort (*Hypericum perforatum* L.). J. Biol. Chem. **278**, 32413–32422

17 Hoffmann-Sommergruber, K. (2000) Plant allergens and pathogenesis-related proteins. What do they have in common? Int. Arch. Allergy Immunol. **122**, 155–166

18 Somssich, I. E., Schmelzer, E., Bollmann, J. and Hahlbrock, K. (1986) Rapid activation by fungal elicitor of genes encoding "pathogenesis-related" proteins in cultured parsley cells. Proc. Natl. Acad. Sci. U.S.A. **83**, 2427–2430

19 Walter, M. H., Liu, J. W., Grand, C., Lamb, C. J. and Hess, D. (1990) Bean pathogenesis-related (PR) proteins deduced from elicitor-induced transcripts are members of a ubiquitous new class of conserved PR proteins including pollen allergens. Mol. Gen. Genet. **222**, 353–360

20 Warner, S. A. J., Scott, R. and Draper, J. (1992) Characterisation of a wound-induced transcript from the monocot asparagus that shares similarity with a class of intracellular pathogenesis-related (PR) proteins. Plant Mol. Biol. **19**, 555–561

21 Walter, M. H., Liu, J. W., Wunn, J. and Hess, D. (1996) Bean ribonuclease-like pathogenesis-related protein genes (Ypr10) display complex patterns of developmental, dark-induced and exogenous-stimulus-dependent expression. Eur. J. Biochem. **239**, 281–293

22 Karlsson, A. L., Alm, R., Ekstrand, B., Fjelkner-Modig, S., Schiott, A., Bengtsson, U., Bjork, L., Hjerno, K., Roepstorff, P. and Emanuelsson, C. S. (2004) Bet v 1 homologues in strawberry identified as IgE-binding proteins and presumptive allergens. Allergy **59**, 1277–1284

23 Hjerno, K., Alm, R., Canback, B., Matthiesen, R., Trajkovski, K., Bjork, L., Roepstorff, P. and Emanuelsson, C. (2006) Down-regulation of the strawberry Bet v 1-homologous allergen in concert with the flavonoid biosynthesis pathway in colorless strawberry mutant. Proteomics **6**, 1574–1587

24 Musidlowska-Persson, A., Alm, R. and Emanuelsson, C. (2007) Cloning and sequencing of the Bet v 1-homologous allergen Fra a 1 in strawberry (*Fragaria ananassa*) shows the presence of an intron and little variability in amino acid sequence. Mol. Immunol. **44**, 1245–1252

25 Munoz, C., Hoffmann, T., Escobar, N. M., Ludemann, F., Botella, M. A., Valpuesta, V. and Schwab, W. (2010) The strawberry fruit Fra a allergen functions in flavonoid biosynthesis. Mol. Plant. **3**, 113–124

26 Boehm, M. and Rösch, P. (1997) Expression in *Escherichia coli*, purification, and spectroscopic characterization of two mutant Bet v 1 proteins. Biol. Chem. **378**, 687–695

27 Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. and Thompson, J. D. (2003) Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. **31**, 3497–3500

28 Huang, X. and Miller, W. (1991) A time-efficient, linear-space local similarity algorithm. Adv. Appl. Math. **12**, 337–357

29 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D: Biol. Crystallogr. **60**, 2126–2132

30 Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y. and Liang, J. (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res. **34**, W116–W118

31 Sattler, M., Schleucher, J. and Griesinger, C. (1999) Herteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog. NMR Spectrosc. **34**, 93–158

32 Johnson, B. A. and Blevins, R. A. (1994) NMRView: A computer program for the visualization and analysis of NMR data. J. Biomol. NMR **4**, 603–614

33 Cornilescu, G., Delaglio, F. and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J. Biomol. NMR **13**, 289–302

34 Schwieters, C. D., Kuszewski, J. J., Tjandra, N. and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package. J. Magn. Reson. **160**, 66–74

35 Koradi, R., Billeter, M. and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. J. Mol. Graph. **14**, 51–55

36 Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR **8**, 477–486

37 Mogensen, J. E., Wimmer, R., Larsen, J. N., Spangfort, M. D. and Otzen, D. E. (2002) The major birch allergen, Bet v 1, shows affinity for a broad spectrum of physiological ligands. J. Biol. Chem. **277**, 23684–23692

38 MarkovicHousley, Z., Degano, M., Lamba, D., von Roepenack-Lahaye, E., Clemens, S., Susani, M., Ferreira, F., Scheiner, O. and Breiteneder, H. (2003) 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. **325**, 123–133

39 Koistinen, K. M., Soininen, P., Venalainen, T. A., Hayrinen, J., Laatikainen, R., Perakyla, M., Tervahauta, A. I. and Karenlampi, S. O. (2005) Birch PR-10c interacts with several biologically important ligands. Phytochemistry **66**, 2524–2533

40 Pasternak, O., Bujacz, G. D., Fujimoto, Y., Hashimoto, Y., Jelen, F., Otlewski, J., Sikorski, M. M. and Jaskolski, M. (2006) Crystal structure of Vigna radiata cytokinin-specific binding protein in complex with zeatin. Plant Cell **18**, 2622–2634

41 Fernandes, H., Pasternak, O., Bujacz, G., Bujacz, A., Sikorski, M. M. and Jaskolski, M. (2008) Lupinus luteus pathogenesis-related protein as a reservoir for cytokinin. J. Mol. Biol. **378**, 1040–1051

42 Fernandes, H., Bujacz, A., Bujacz, G., Jelen, F., Jasinski, M., Kachlicki, P., Otlewski, J., Sikorski, M. M. and Jaskolski, M. (2009) Cytokinin-induced structural adaptability of a Lupinus luteus PR-10 protein. FEBS J. **276**, 1596–1609

43 Mirza, O., Henriksen, A., Ipsen, H., Larsen, J. N., Wissenbach, M., Spangfort, M. D. and 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 Bet v 1. J. Immunol. **165**, 331–338

44 Spangfort, M. D., Mirza, O., Ipsen, H., Joost van Neerven, R. J., Gajhede, M. and Larsen, J. N. (2003) 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. **171**, 3084–3090

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

© 2012 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

46 Pasternak, O., Biesiadka, J., Dolot, R., Handschuh, L., Bujacz, G., Sikorski, M. M. and Jaskolski, M. (2005) Structure of a yellow lupin pathogenesis-related PR-10 protein belonging to a novel subclass. Acta Crystallogr. D: Biol. Crystallogr. **61**, 99–107

47 Mogensen, J. E., Ferreras, M., Wimmer, R., Petersen, S. V., Enghild, J. J. and Otzen, D. E. (2007) The major allergen from birch tree pollen, Bet v 1, binds and permeabilizes membranes. Biochemistry **46**, 3356–3365

Received 20 June 2012/12 July 2012; accepted 13 July 2012
Published as Immediate Publication 22 August 2012, doi 10.1042/BSR20120058

© 2012 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

Biosci. Rep. (2012) / **32** / 567–575 (Printed in Great Britain) / doi 10.1042/BSR20120058

Original Paper

SUPPLEMENTARY ONLINE DATA

Solution structure of the strawberry allergen Fra a 1

Christian SEUTTER VON LOETZEN*, Kristian SCHWEIMER*, Wilfried SCHWAB†, Paul RÖSCH* and Olivia HARTL-SPIEGELHAUER*[1]

*Lehrstuhl Biopolymere und Forschungszentrum für Bio-Makromoleküle, Universität Bayreuth, Universitätsstr. 30, 95447 Bayreuth, Germany, and †Biotechnology of Natural Products, Technische Universität München, Liesel-Beckmann-Str. 1, 85354 Freising, Germany

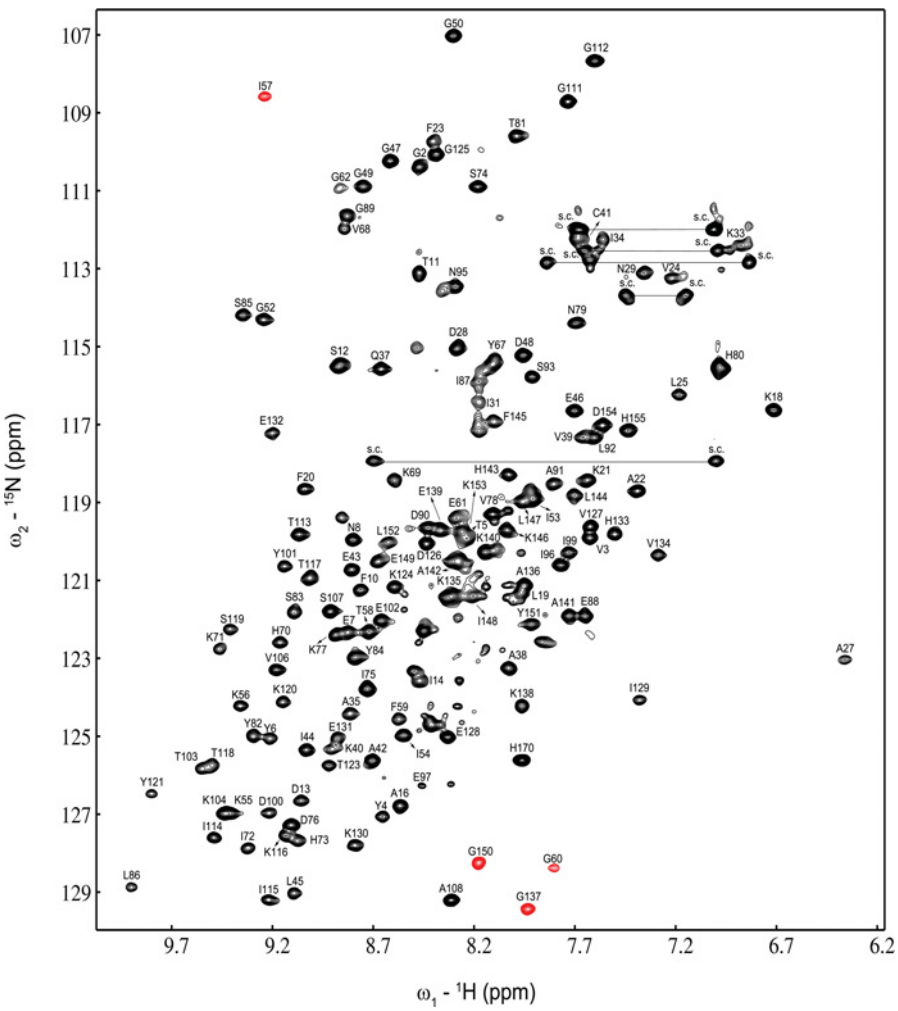

Figure S1 Backbone assignment of Fra a 1E
[^{15}N]-HSQC NMR spectrum of ^{15}N, ^{13}C-labelled Fra a 1E (400 μM), measured at 298 K and 800 MHz. Assigned backbone amide resonances are labelled with amino acid type in one letter code and residue number; s.c., side-chain amide resonance.

[1] To whom correspondence should be addressed (email Olivia.Hartl-Spiegelhauer@uni-bayreuth.de).

© 2012 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

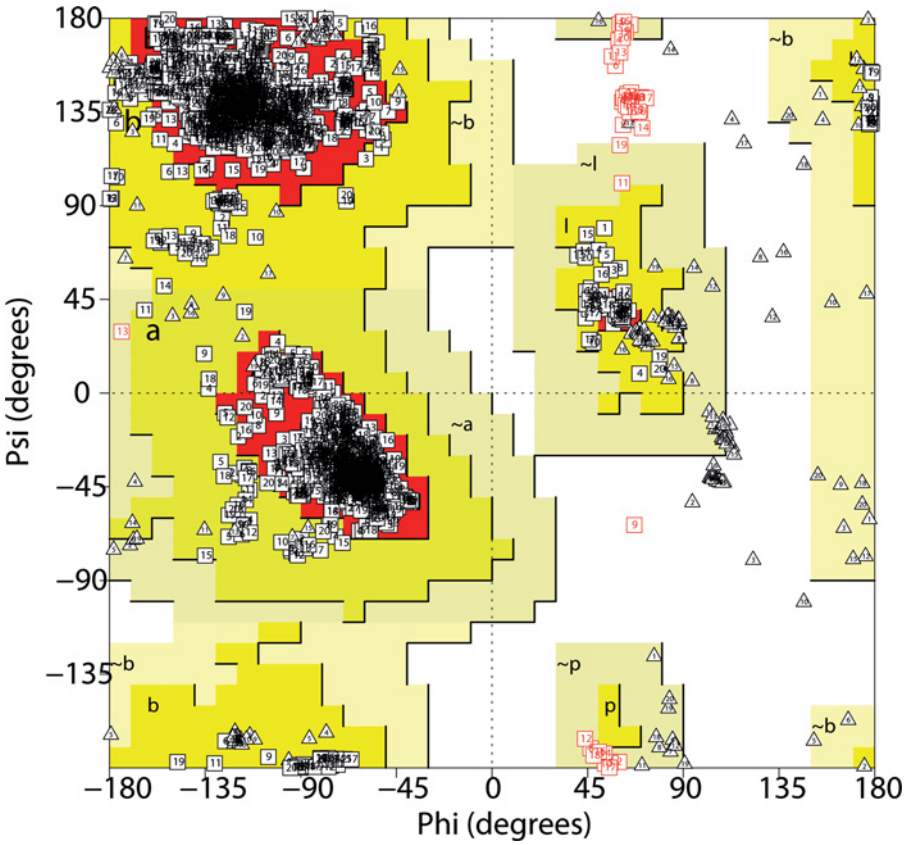

Figure S2 Ramachandran plot of the 20 NMR solution structures
Squares, all residues except glycine and proline residues; triangles, glycine residues. Model numbers are shown inside each data point. PROCHECK-NMR analysis showed that 90.3% of the residues are located in the most favoured regions (A, B, L), 7.9% are located in additional allowed regions (a, b, l, p), 0.7% are located in generously allowed regions (~a, ~b, ~l, ~p) and only 1.1% are located in disallowed regions of the Ramachandran map.

Received 20 June 2012/12 July 2012; accepted 13 July 2012
Published as Immediate Publication 22 August 2012, doi 10.1042/BSR20120058

© 2012 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

Biosci. Rep. (2012) / **32** / 567–575 (Printed in Great Britain) / doi 10.1042/BSR20120058

Original Paper

SUPPLEMENTARY ONLINE DATA

Solution structure of the strawberry allergen Fra a 1

Christian SEUTTER VON LOETZEN*, Kristian SCHWEIMER*, Wilfried SCHWAB†, Paul RÖSCH* and Olivia HARTL-SPIEGELHAUER*[1]

*Lehrstuhl Biopolymere und Forschungszentrum für Bio-Makromoleküle, Universität Bayreuth, Universitätsstr. 30, 95447 Bayreuth, Germany, and †Biotechnology of Natural Products, Technische Universität München, Liesel-Beckmann-Str. 1, 85354 Freising, Germany

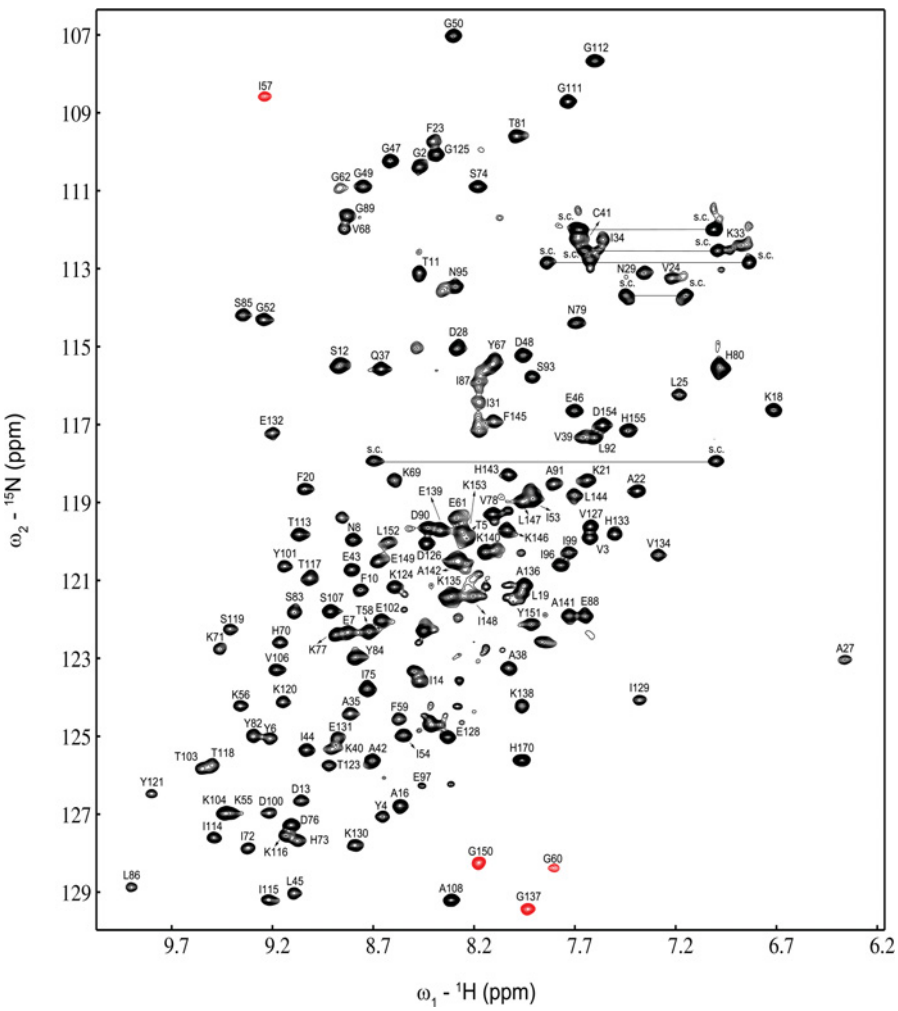

Figure S1 **Backbone assignment of Fra a 1E**
[^{15}N]-HSQC NMR spectrum of ^{15}N, ^{13}C-labelled Fra a 1E (400 μM), measured at 298 K and 800 MHz. Assigned backbone amide resonances are labelled with amino acid type in one letter code and residue number; s.c., side-chain amide resonance.

[1] To whom correspondence should be addressed (email Olivia.Hartl-Spiegelhauer@uni-bayreuth.de).

© 2012 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.

Bioscience Reports www.bioscirep.org

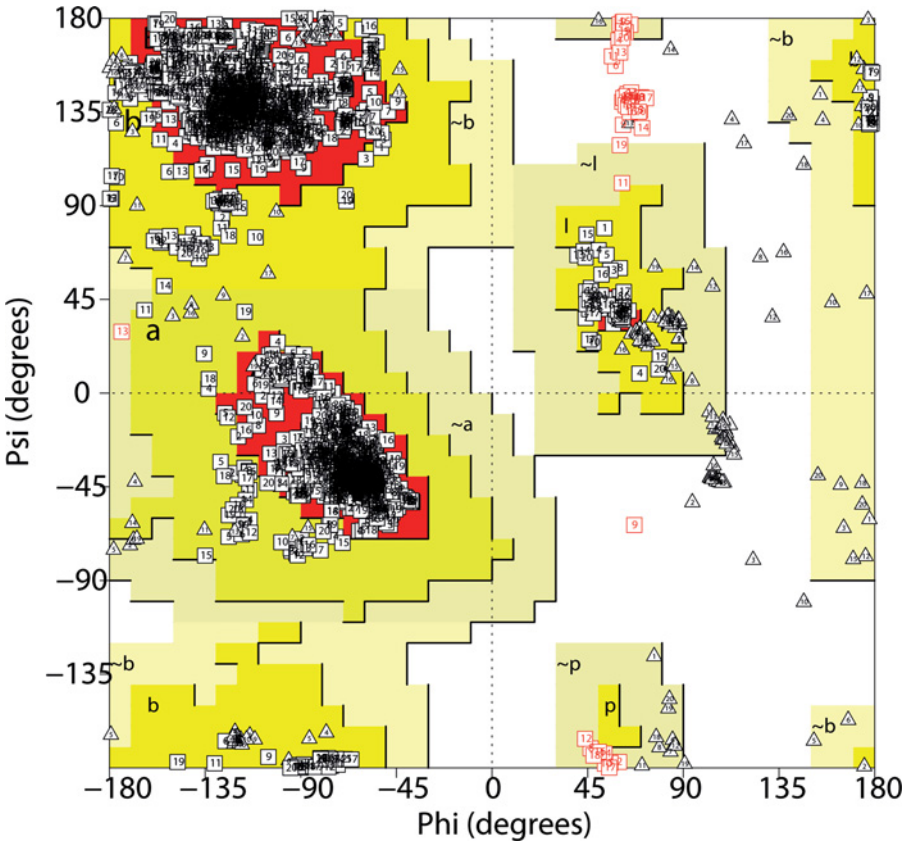

Figure S2 Ramachandran plot of the 20 NMR solution structures
Squares, all residues except glycine and proline residues; triangles, glycine residues. Model numbers are shown inside each data point. PROCHECK-NMR analysis showed that 90.3% of the residues are located in the most favoured regions (A, B, L), 7.9% are located in additional allowed regions (a, b, l, p), 0.7% are located in generously allowed regions (~a, ~b, ~l, ~p) and only 1.1% are located in disallowed regions of the Ramachandran map.

Received 20 June 2012/12 July 2012; accepted 13 July 2012
Published as Immediate Publication 22 August 2012, doi 10.1042/BSR20120058

© 2012 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial Licence (http://creativecommons.org/licenses/by-nc/2.5/) which permits unrestricted non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited.