

Purification and structural analysis of the novel glycoprotein allergen Cyn d 24, a pathogenesis-related protein PR-1, from Bermuda grass pollen

Lu-Ping Chow^{1,5}, Li-Li Chiu¹, Kay-Hooi Khoo², Ho-Jen Peng³, Sue-Yee Yang³, Shih-Wen Huang⁴ and Song-Nan Su³

1 Graduate Institute of Biochemistry and Molecular Biology, College of Medicine, National Taiwan University, Taipei, Taiwan

2 Institute of Biochemistry, Academia Sinica, Taipei, Taiwan

3 Department of Medical Research and Education, Veterans General Hospital-Taipei, Taipei, Taiwan

4 Department of Pediatric, Division of Immunology and Allergy, University of Florida, Gainesville, FL, USA

5 Department of Medical Genetics, National Taiwan University Hospital, Taipei, Taiwan

Keywords

allergen; Bermuda grass pollen; glycoprotein; pathogenesis-related proteins; purification

Correspondence

S.-N. Su, Department of Medical Research and Education, Taipei Veterans General Hospital, Taipei, Taiwan 112

Fax: +886 22875 1562

Tel: +886 22871 2121 ext. 3379

E-mail: snsu@vghtpe.gov.tw

(Received 9 June 2005, revised 27 September 2005, accepted 3 October 2005)

doi:10.1111/j.1742-4658.2005.05000.x

Bermuda grass pollen (BGP) contains a very complex mixture of allergens, but only a few have been characterized. One of the allergens, with an apparent molecular mass of 21 kDa, has been shown to bind serum IgE from 29% of patients with BGP allergy. A combination of chromatographic techniques (ion exchange and reverse phase HPLC) was used to purify the 21 kDa allergen. Immunoblotting was performed to investigate its IgE binding and lectin-binding activities, and the Lysyl-C endopeptidase digested peptides were determined by N-terminal sequencing. The cDNA sequence was analyzed by RACE PCR-based cloning. The protein mass and the putative glycan structure were further elucidated using MALDI-TOF mass spectrometry. The purified 21 kDa allergen was designated Cyn d 24 according to the protocol of International Union of Immunological Societies (IUIS). It has a molecular mass of 18 411 Da by MALDI-TOF analysis and a pI of 5.9. The cDNA encoding Cyn d 24 was predicted to produce a 153 amino acid mature protein containing two conserved sequences seen in the pathogen-related protein family. Carbohydrate analysis showed that the most abundant N-linked glycan is a α (3)-fucosylated pauci-mannose ($\text{Man}_3\text{GlcNAc}_2$) structure, without a Xyl β -(1,2)-linked to the branching β -Man. Thus, Cyn d 24 is a glycoprotein and the results of the sequence alignment indicate that this novel allergen is a pathogenesis-related protein 1. To the best of our knowledge, this is the first study to identify any grass pollen allergen as a pathogenesis-related protein 1.

Type I (IgE-mediated) allergy is a clinical disorder that affects about 20% of the population in developed countries. Pollen is a major contributor to outdoor airborne allergens that cause type I allergic reactions. Bermuda grass (*Cynodon dactylon*) pollen (BGP) is one of the major causes of respiratory allergy in warm climates [1–3] and more than 12 IgE-binding BGP proteins have been reported [4,5]. Many attempts have been made to

study these allergens, but only few have been well characterized. Of these allergens, Cyn d 1 is the most important allergen and more than 96% of individuals allergic to BGP are hypersensitive to Cyn d 1 [3,6–9]. The function of the Cyn d 1 is not known but the results of a sequence search from databank showed some sequence similarity with β -expansin. Another allergen, BG60, is a metalloflavoprotein with three or four

Abbreviations

BGP, Bermuda grass pollen; CM, carboxymethyl; ELISA, enzyme-linked immunosorbent assay; MAb, monoclonal antibody; MALDI-TOF MS, matrix assisted laser desorption ionization-time of flight mass spectrometry; PAS, periodic acid-Schiff's stain; PSD, post source decay; RP-HPLC, high performance liquid chromatography.

isoforms and these have high pI values ranging from 9.5 to 10.5 [10–15]. Similarly, Cyn d 7 is a calcium-binding protein containing two Ca^{2+} , and its allergenicity and cross-reactivity have been investigated in other pollens [16,17]. Cyn d 12 is a profilin (actin-binding protein) that is involved in about 20% of the cross-reactivity found among pollen and food allergic patients [18]. Finally and recently, a 46 kDa allergen has been reported and its internal peptide sequences shown to have some sequence similarity to cytochrome c oxidase III from corn grass pollen [19].

Pollen extracts have been used for desensitization treatment of allergic patients, but such extracts contain a very complex mixture of poorly characterized proteins. Our recent studies using 2D gel and immunoblot techniques indicated that BGP may contain up to 230 proteins and about 65 of these are IgE binding proteins [20]; however, some of these are known to be either isoforms or the degraded products of allergens. Furthermore, allergic patients may show differential immune responses to different allergens in the pollen. As a consequence, individuals allergic to BGP require individual diagnosis and therapy, and an understanding of the structure of these allergens is essential to the improvement of diagnosis and the design of adequate therapeutic treatment. We previously identified a BGP protein with an apparent molecular mass of 21 kDa on SDS/PAGE and this protein was able to bind serum IgE from about 29% of patients with BGP allergy [5]. Our preliminary study showed that peptides of the 21 kDa protein had sequence similarity to pathogenesis-related proteins. These proteins are induced by stresses such as fungal and bacterial infections, flooding, freezing temperature, or chemicals such as ethylene and salicylic acid [21]. Many plant allergens from food and pollen have been found to be pathogenesis-related proteins and these have been grouped into one of 14 families [22]. In this study, we describe the purification, characterization and cDNA cloning of this allergen from Bermuda grass pollen. Results from sequence alignment indicate that this newly identified allergen is a pathogenesis-related protein 1 (PR-1).

Results

Purification of allergen Cyn d 24

Cyn d 24 was purified from BGP by two chromatographic steps using a CM-TSK column and reverse phase HPLC. When fraction AS1 in starting buffer was applied to the CM-TSK column, one major peak (C1) and two minor peaks (C2 and C3) were eluted using the starting buffer (Fig. 1A). Peak C3, the only

peak to contain a protein with an apparent molecular mass of about 21 kDa on SDS/PAGE, contained a major protein (> 90% pure) which bound serum IgE from allergic patients on immunoblots (data not shown). When the C3 fraction was chromatographed on a reverse phase HPLC column, 2 peaks were seen (Fig. 1B). On SDS/PAGE, the major peak contained a single protein with an apparent molecular mass of 21 kDa (Fig. 1B, inset), which bound human serum IgE, and which was designated Cyn d 24 according to recommendations of the International Union of Immunological Societies (IUIS) nomenclature subcommittee

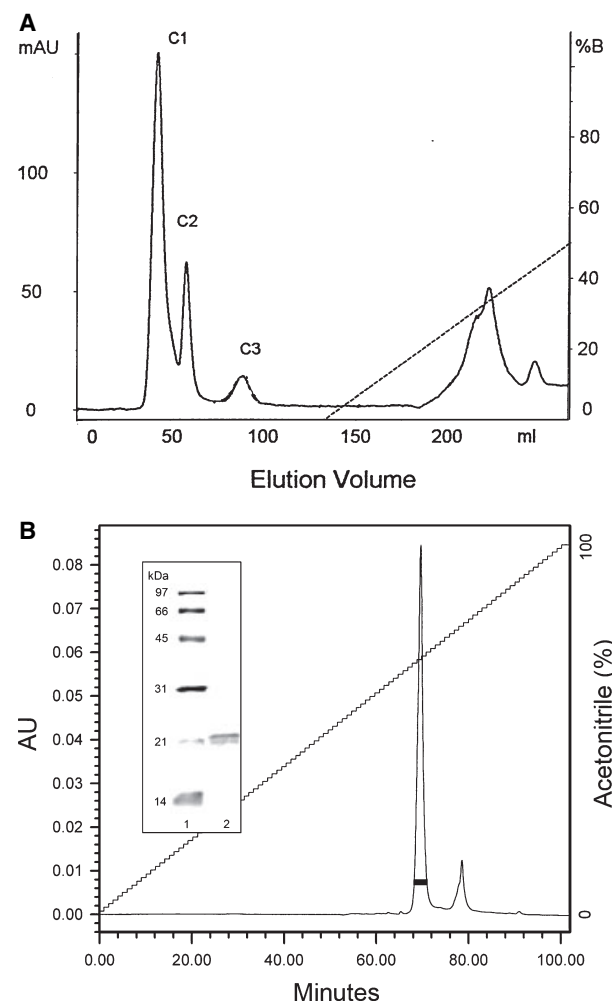


Fig. 1. (A) Chromatography of fraction AS1 (15 mg) on a CM-TSK column (35 × 1.6 cm). The fractions containing a protein with a molecular mass of 21 kDa on SDS/PAGE were pooled as indicated by the black bar (fraction C3). (B) Chromatography of fraction C3 (1 mg) on a semipreparative RP-HPLC column (C4, 10 × 200 mm). A sample (3 µg) of the major peak (indicated by the black bar) was loaded onto an SDS/PAGE (12.5%) and stained with Coomassie blue (inset, lane 2). Six standard proteins were included (inset, lane 1).

[23]. MALDI-TOF mass spectrometry gave a molecular mass of m/z 18 411 Da. The pI value was estimated to be 5.9. The protein gave a pink color on PAS staining (data not shown), indicating it was a glycoprotein. The yield of purified Cyn d 24 was < 0.1% of total soluble BGP protein.

Lectin-binding activity of Cyn d 24

As Cyn d 24 was found to be a glycoprotein, its carbohydrate moieties were analyzed using seven lectins, *Glycine max*, *Dolichos biflorus*, *Helix pomatia*, *Triticum vulgaris*, *Maclura pomifera*, *Tetragonolobus purpuria* and *Canvalia ensiformis*. Of these, only that from *Canvalia ensiformis* (Con A) was found to bind to Cyn d 24 (data not shown). These results demonstrate that Cyn d 24 contains a carbohydrate moiety with free terminal α -D-mannopyranoside or α -D-glucopyranoside residues.

Allergenicity of native Cyn d 24

Human IgE antibodies reacting with Cyn d 24 were demonstrated by ELISA using allergic sera. Serum samples from 35 allergic patients, with a value of at least three for their skin responses to BGP crude extract, were tested. When tested for IgE reactive with Cyn d 24, 12 showed no reactivity, one showed low reactivity, nine showed medium reactivity, six showed medium-high reactivity, and seven showed high reactivity. The prevalence of Cyn d 24 immunoreactivity in this study was about 65%, with 23 out of 35 of patients being allergic to BGP, which is much higher than the previous report of about 29% prevalence.

Partial amino acid sequencing of Cyn d 24

To obtain peptide fragments of Cyn d 24 that would allow us to design oligonucleotide primers for cloning the specific gene, the allergen was subjected to proteolytic treatment. The peptides were separated by C_{18} column (Fig. 2A), and their sequences were determined by Edman degradation (Fig. 2B). The alignment of these sequences with those contained in the GeneBank/EMBL database revealed significant similarity with PR-1.

Cloning and sequencing of cDNA encoding Cyn d 24

Cyn d 24-specific cDNA was obtained by cDNA synthesis and PCR amplification from total RNA isolated from BGP. A sense primer, designed on the basis of

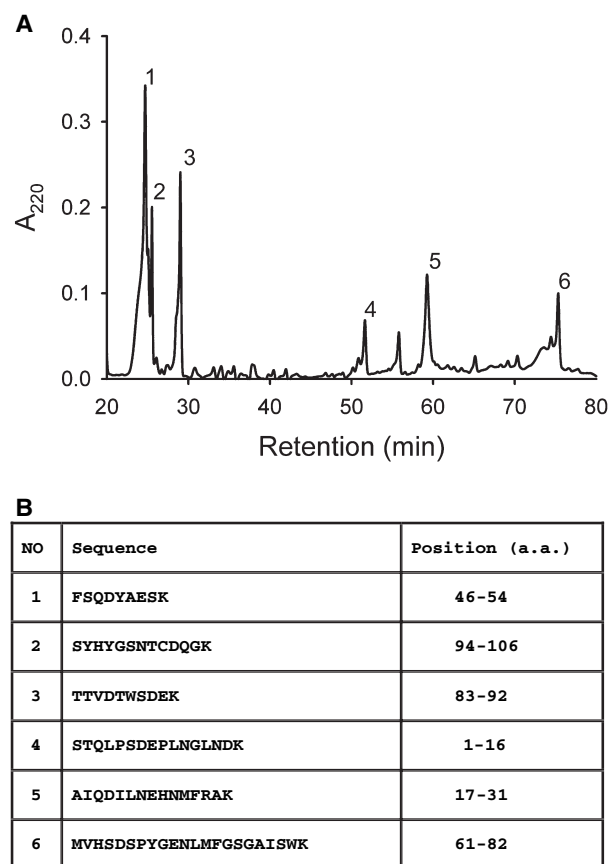


Fig. 2. Determination of internal sequences of Cyn d 24. (A) Reverse phase HPLC separation of peptides obtained from nCyn d 24 by treatment with Lys-C endopeptidase, with 5–60% acetonitrile gradients in 0.06% trifluoroacetic acid. (B) Amino acid sequences of the peptides, the position in the complete sequence of Cyn d 24 is indicated.

the amino acids of number 4 and number 5 peptide sequences (Fig. 2B), and an antisense primer, designed on the basis of the amino acids of the number 2 peptide sequence (Fig. 2B), were used in the first step of PCR cloning. This experiment resulted in a cDNA fragment with an estimated size of 240 bp, which corresponds to the N-terminal portion of the allergen. Based on this partial sequence, two specific primers were designed and used together with the anchor primer AP2 to obtain the 5' and 3' portions of the Cyn d 24 cDNA. The nucleotide sequence of the Cyn d 24 gene is shown in Fig. 3A. The cDNA contains an open reading frame of 750 nucleotides encoding 250 amino acids. The N-terminal sequence of the mature protein is predicted to begin at Ser98 (Fig. 3A). The molecular mass of the amino acid sequence deduced from the clone (17 374 Da) was notably lower than that obtained by MALDI-TOF MS for the

A

1										M	V	D	L	Q	A	A	A	L	V	I	L	
	TCC	ACG	CGT	TGG	GAG	CTC	TCC	CAT	ATG	GTC	GAC	CTG	CAG	GCG	GCC	GCA	CTA	GTG	ATT	CTC	36	
13	I	C	I	C	L	L	F	A	G	G	H	L	A	A	A	S	K	S	F	G		
	ATC	TGC	ATC	TGC	CTG	CTC	TTC	GCC	GGC	GGC	CAC	CTC	GCC	GCG	GCT	AGC	AAG	AGC	TTC	GGC	96	
33	G	G	G	G	Y	G	G	E	G	S	A	A	A	Q	E	V	Q	T	A	A		
	GGC	GGT	GGA	GGC	TAT	GGC	GGA	GAG	GGA	TCA	GCA	GCC	GCC	CAG	GAG	GTC	CAG	ACC	GCC	GCC	156	
53	Q	E	A	V	E	G	A	E	Q	V	A	S	E	S	A	S	L	T	T	P		
	CAG	GAG	GCG	GTA	GAG	GGC	GCC	GAG	CAG	GTA	GCG	TCC	GAG	TCA	GCC	TCC	CTC	ACA	ACA	CCA	216	
73	T	T	R	E	E	Q	P	A	A	E	A	A	A	S	T	A	G	G	S	Q		
	ACC	ACC	AGG	GAA	GAA	CAA	CCG	GCG	GCA	GAG	GCC	GCG	GCG	TCC	ACC	GCT	GGC	GGT	AGC	CAA	276	
93	Q	E	G	Y	G	S	T	Q	L	P	S	D	E	P	L	N	G	L	N	D		
	CAA	GAA	GGA	TAT	GGC	AGC	ACC	CAA	CTT	CCA	TCG	GAC	GAG	CCA	TTG	AAC	GGG	CTC	AAC	GAC	336	
113	K	A	I	Q	D	I	L	N	E	H	N	M	F	R	A	K	E	R	V	P		
	AAG	GCC	ATA	CAG	GAC	ATC	CTC	AAC	GAG	CAC	AAC	ATG	TTC	CGC	GCC	AAG	GAG	CGC	GTC	CCG	396	
133	P	L	T	W	N	T	T	L	A	K	F	S	Q	D	Y	A	E	S	K	L		
	CCG	CTC	ACG	TGG	AAC	ACG	ACG	CTT	GCC	AAG	TTC	TCG	CAG	GAC	TAC	GCG	GAG	TCG	AAG	CTG	456	
153	K	K	D	C	K	M	V	H	S	D	S	P	Y	G	E	N	L	M	F	G		
	AAG	AAG	GAC	TGC	AAG	ATG	GTG	CAC	TCG	GAC	TCG	CCC	TAC	GGG	GAG	AAC	CTG	ATG	TTC	GGC	516	
173	S	G	A	I	S	W	K	T	T	V	D	T	W	S	D	E	K	K	S	Y		
	TCC	GGC	GCC	ATC	TCC	TGG	AAG	ACG	ACG	GTG	GAC	ACG	TGG	AGC	GAC	GAG	AAG	AAG	AGC	TAC	576	
193	H	Y	G	S	N	T	C	D	Q	G	K	M	C	G	H	Y	T	A	V	V		
	CAC	TAC	GGC	TCC	AAC	ACC	TGC	GAC	CAA	GGC	AAG	ATG	TGC	GGC	CAC	TAC	ACC	GCC	GTC	GTG	636	
213	W	K	D	T	T	S	V	G	C	G	R	V	L	C	D	D	K	K	D	T		
	TGG	AAG	GAC	ACC	ACC	AGC	GTC	GGA	TGC	GGA	CGC	GTC	CTC	TGC	GAC	GAC	AAG	AAG	GAC	ACC	696	
233	M	I	M	C	S	Y	W	P	P	G	N	Y	E	N	Q	K	P	Y				
	ATG	ATC	ATG	TGC	AGC	TAC	TGG	CCG	CCG	GGC	AAC	TAT	GAA	AAC	CAG	AAG	CCC	TAC			750	

B

Cyn d 24	-----STQLPSDEPLNGLNDKAIQDILNEHNMFRAKEHVPPLTWNTTLA	44
Hordeum	-----MQTPKLVILLALAMSAAMVNLQAQNSP..YVSP..AA..AVG.GAVS.S.K.Q	54
Triticum	-----MQTPKLAILLALAMSAAMANLSQAQNSP..Y.SP..AA..AVG.GAV..S.K.Q	54
Zea	-----MAPRLACLLALAMAAIVVAPCTAQNSP..YVDP..AA..DVG.G.VS.D..V	53
Nicotiana	MGFVLFSQLPSFLLVSTLL.FLVISHSCRAQNSQ..Y.DA..TA..DVG.E....DDQV.	60
	* * *	
Cyn d 24	KFSQDYAESKLKDKCMVHSDSPYGENLMFGSGAISWKT--VDTWSDEKKSYPHYGSNTC	102
Hordeum	A.A.N..N-QRIN...LQ..GG.....IFW..AGAD..ASDA.NS.VS...D.D.....	113
Triticum	G.A.S..N-QRIN...LQ..GG.....IFW..AGAD..AADA.NA.VG...D.D.....	113
Zea	AYA.S..A-QRQG...LI..GG.....FW..AGAD.SASDA.GS.VS..QY.DHDT.S.	112
Nicotiana	AYA.N..S-Q.AA..NL...HGQ.....AE..-GDFMTAAKA.EM.V...QY.DHD....	118
	* * * *	
Cyn d 24	DQGKMCBGHYTAVVWKDTTSVGCGRVLCDDKKDDTMIMCSYWPPGNYENQKPY	153
Hordeum	AA..V.....Q...RAS..I..A..V.NNNGRVF.T.N.E.R..IVG....	164
Triticum	AA..V.....Q...RAS..I..A..V.NNNGRVF.T.N.E.R..IIG....	164
Zea	AE.QV.....Q...R.S.AI..A..V..NNAGRVF.I...N...VVGES..	163
Nicotiana	S..QV.....Q...RNSVR...A..Q.NNG-GYVVS.N.D.....RGES..	168

Fig. 3. cDNA sequence and sequence alignment of Cyn d 24. (A) Nucleotide and deduced amino acid sequences of Cyn d 24. The numbers on the right of the figure indicate the positions of the nucleotide sequence. The numbers on the left of the figure indicate the positions of the deduced amino acid sequence. N-terminal segment determined by protein sequencing is underlined. (B) Comparison of the amino acid sequence of Cyn d 24 with those of various PR-1. The accession numbers of PR-1 in the protein database are Hordeum (SwissProt: P35793), Triticum (SwissProt: Q94F73), Zea (SwissProt: O82086) and Nicotiana (SwissProt: Q40557). The numbering system is based on Cyn d 24 sequence. Dashes are introduced for optimal alignment and to give maximal homology between all compared sequences. Identical amino acids are shown as dots. The highly conserved and consensus amino acid residues involved in six Cys residues are indicated by asterisks. The glycosylation site is boxed.

purified protein from the pollen (18 411 Da). This difference suggested the existence of post-translational modification of the putative N-glycosylation site. Alignment of the Cyn d 24 amino acid sequence with those of proteins contained in the Swiss-Prot/EMBL database revealed similarity with various PR-1 s. The sequence similarities of four PR-1 s from barley (*Hordium vulgare*), wheat (*Triticum aestivum*), maize (*Zea mays*) and tobacco (*Nicotiana tabacum*) were 49.6, 48.9, 45.2 and 48.5% identity, respectively, compared with Cyn d 24 (Fig. 3B). The mature protein sequence contains six cysteines and two highly conserved domains (109–119 and 136–147).

Oligosaccharide analyses

In an attempt to isolate glycopeptides, Cyn d 24 was digested with Lysyl-C endoproteinase and subjected to HPLC; the fraction containing Con A-binding activity was subjected to MALDI-TOF mass spectrometry analysis, which gave two clusters of peaks (Fig. 4A). The first cluster contained a major signal at m/z 1816.8 and the second was dominated by two major signals at m/z 2646.4 and 2668.4, the mass interval of which indicated a protonated and sodiated molecular ion, respectively. Other signals at a higher mass level also contained similar heterogeneity, i.e. pairs of signals separated by 22 μm . Taking the more abundant $(M + \text{Na})^+$ peak, the signals at m/z 2778.4 and 2800.4 corresponded, respectively, to a pentose increment from the peak at m/z 2646.4 and 2668.4. As shown in Fig. 4B, all these peaks disappeared after glycoamidase-A digestion, confirming that they represented glycopeptides, and a new prominent signal was detected at m/z 1608.7, which corresponded to the loss of $\text{dHex}_1\text{Hex}_3\text{HexNAc}_2$ from the putative protonated glycopeptide at m/z 2646.4 concomitant with a mass increment of 1 μm due to conversion of the asparagine (N) residue to aspartic acid (D) as a consequence of de-N-glycosylation by glycoamidase A. This was corroborated by post source decay (PSD) analysis (Fig. 4C) of the protonated peptide at m/z 1608.7, which led to the determination of the peptide sequence as EHVPLITWDTTILIAK, where DTT corresponds to the original N-glycosylation site, NTT. As isoleucine (I) and leucine (L) have the same mass, PSD analysis could not differentiate these two amino acids in the sequence. However, Edman sequencing of this peptide fraction not only confirmed the derived sequence, but also showed that the amino acids at positions 6 and 13 were L.

To define the oligosaccharides present, N-glycans released sequentially by trypsin digestion and glyco-

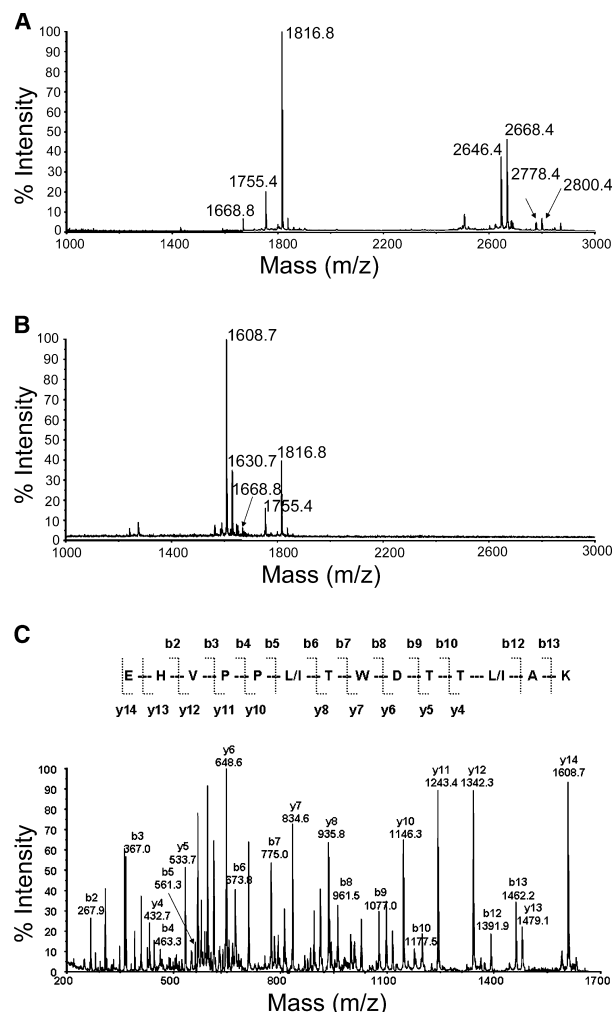


Fig. 4. MALDI-TOF and PSD analyses of the glycopeptide isolated from Cyn d 24. (A) Four signals at m/z 2646.4, 2668.4, 2778.4 and 2800.4 were obtained for the glycopeptide before glycoamidase-A digestion. (B) One signal at m/z 1608.7 appeared for the glycopeptide and four signals at m/z 2646.4, 2668.4, 2778.4 and 2800.4 disappeared after glycoamidase-A digestion. The peaks at m/z 1816.8, 1668.8 and 1755.4, were inferred to be nonglycosylated peptides, since they did not show any change after glycopeptidase-A digestion. (C) Amino acid sequence for the glycopeptide with a signal at m/z 1608.7 obtained by PSD analysis.

amidase-F and glycoamidase-A were permethylated and subjected to MALDI MS and MS/MS analyses, taking advantage of the enzyme's specificity. Only very small amounts of N-glycans were released by glycoamidase-F digestion and were determined by MALDI-MS to be $\text{Hex}_3\text{HexNAc}_2$ and $\text{Hex}_6\text{HexNAc}_2$ (data not shown). In contrast, strong signals were produced by the glycoamidase-A in the released fractions (Fig. 5), the most abundant of which could be assigned to be $\text{Hex}_3\text{HexNAc}_2\text{Fuc}$ (m/z 1345.6). Definitive structural

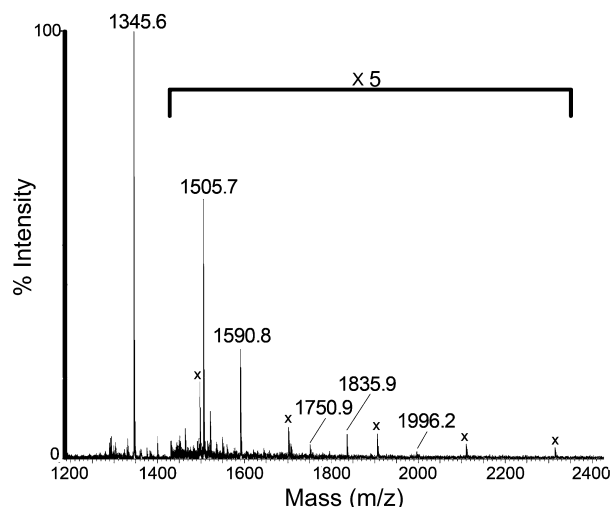


Fig. 5. MALDI mass spectrum of permethylated Cyn d 24 oligosaccharides. Six major molecular ion signals were detected as indicated. The spectrum was magnified five-fold from m/z 1450 to show more clearly the less abundant peaks.

characterization was not attempted, but subsequent MS/MS analysis localized the Fuc at the reducing end, HexNAc, and the results are consistent with a 'pauci' mannose structure with core α -(1,3)-fucosylation, making it resistant to glycoamidase-F digestion. Other components present included Hex₃-HexNAc₂Fuc-Pent (m/z 1505.7), HexNAc-Hex₃HexNAc₂Fuc (m/z 1590.6), HexNAc-Hex₃HexNAc₂FucPent (m/z 1750.9), HexNAc₂-Hex₃HexNAc₂Fuc (m/z 1835.9), and HexNAc₂-Hex₃HexNAc₂FucPent (m/z 1996.2). The Pent is probably core xylosylation, as commonly seen for plant glycoproteins. To summarize, the major structures detected were paucimannose-type (Man₃GlcNAc₂-based structure), most of which were core α -(1,3)-fucosylated, and a small portion also carries core xylosylation and/or additional GlcNAc. This pattern of heterogeneity was demonstrated by MS analysis of both the glycopeptides and the released glycans. It should, however, be noted that MS cannot distinguish between isomeric structures. Based on the molecular masses of Cyn d 24 (18 411 Da) and the major oligosaccharide (Man₃GlcNAc₂Fuc) (1057 Da), the carbohydrate content was approximately 5.7%.

Discussion

Using 2D electrophoresis and immunoblotting techniques, our recent studies have indicated that BGP contains about 230 proteins including isoforms/degraded protein products and around 65 of these are IgE-binding proteins [19]. Only a few of these proteins have been

purified and characterized so far. In the present study, we isolated a 21 kDa protein which bound serum IgE from BGP-allergic patients using a combination of CM-TSK and RP-HPLC. The final material was homogenous, as shown by the presence of a single band on SDS/PAGE, a single sharp peak on RP-HPLC, and a single band on immunoblotting with human antibodies. The percentage of serum samples from BGP allergic patients that contained IgE reactive with Cyn d 24 was about 65% in this study, which is higher than previously reported [5]; this is probably because the sera used in this study were from patients who gave a high prick test response (value of > 3) to BGP crude extract. To examine the role of the carbohydrate moiety of Cyn d 24 in antibody binding, enzyme cleavage of the carbohydrate moiety of Cyn d 24 was performed using glycoamidase-A. The result showed that the carbohydrate moiety of Cyn d 24 is involved in the serum IgE binding (data not shown). The importance of carbohydrate moieties to IgE binding has been demonstrated by various research groups [12,24–31].

A feature shared by all the oligosaccharides of Cyn d 24 is the Man₃GlcNAc₂Fuc structure. This Man₃GlcNAc₂Fuc, which makes up about 5.7% of the total weight of the glycoprotein, has an L-Fuc α -(1,3)-linked to an Asn-linked GlcNAc, which does not have a Xyl β -(1,2)-linked to the branching Man. This structure was previously reported by us as a major oligosaccharide (68.3% of the total carbohydrate weight) of BG60 from BGP [13]. Thus, this structure is probably a unique feature of, and the predominant component of, the oligosaccharides of BGP glycoproteins, whereas it is reported to be only a minor constituent in soybean peroxidase and horseradish peroxidase [32,33]. Formation of this type of oligosaccharide in BG60 has been suggested to be the result of degradative reactions, rather than imperfect biosynthesis [13].

Of the plant allergens listed in the official allergen database of the IUIS, about 25% belong to various pathogenesis-related protein groups and these have been categorized into nine of the 14 groups. In this study, structure analysis of the Cyn d 24 amino acid sequence revealed that Cyn d 24 contains two highly conserved sequences at the C-terminus (109–119, 136–147) [34] and six highly conserved cysteine residues, which are characteristics of the cysteine-rich secretory protein (CRISP). Some relevant members of the CRISP [35] family are plant PR-1 [36], rodent sperm-coating glycoprotein (SCP) [37], mammalian testis-specific protein Tpx-1 [38], Venom allergen 5 (Ag5) from vespid wasps [39], proteins Sc7 and Sc14 from *Schizophyllum commune* [40], and mammalian glioma pathogenesis-related protein (GliPR) [41]. These family proteins also possess similar

conserved domains. However, Cyn d 24 showed higher sequence similarity to the PR-1 s from barley, wheat, maize, rice, and tobacco where the identity ranged from 45 to 50%, but there was lower sequence similarity to other related CRISP family members such as SCP, Tpx-1, Ag5, and GliPR, where identity ranged from 32% to 43%. The plant PR-1 family contains six highly conserved and consensus Cys residues, but other family proteins contain different numbers of cysteines from three (SCP) [37] to 17 (CRISP) [35]. In addition, the sequence GHYTQVW is a significantly conserved region in the plant PR-1 s. It has been suggested that this domain may play an important functional role in the plant defense-related activity [22]. Their weak similarity to the group allergens five from insect venom could link pollen allergy and hypersensitivity to insect sting in some patients. Taken together, these results suggest that Cyn d 24 is most likely to be a PR-1 protein. In conclusion, we have purified, cloned and characterized Cyn d 24 as a novel pathogenesis-related protein from BGP. Additionally, the identification of Cyn d 24 has identified the involvement of a novel class of PR proteins in pollen allergy. This finding may have a significant impact in diagnosis and therapeutic applications.

Experimental procedures

Bermuda grass pollen was purchased from International Biologicals (Piedmont, OK, USA). Biotinylated lectins and avidinylated horseradish peroxidase were from Sigma (St Louis, MO, USA). Horseradish peroxidase-conjugated goat antihuman IgG and antimouse IgG antibodies were from Jackson ImmunoResearch (West Grove, PN, USA). Glycoamidase-A and glycoamidase-F were from Calbiochem (San Diego, CA, USA).

Purification of nCyn d 24

All purifications were carried out at 4 °C. The AS1 fraction was obtained as described previously [11]; briefly, the crude extract was brought to 90% saturation with solid ammonium sulfate, stirred slowly for 40 min at 4 °C. After centrifugation, the supernatant was applied to a Sephadex G-25 column with the break-through material forming the G1 fraction. Solid ammonium sulfate was added to 70% saturation and the mixture stirred at 4 °C for 40 min, then the precipitate was collected by centrifugation, dissolved in 20 mM phosphate buffer (pH 6.0). The supernatant was applied to a CM-TSK column (33 × 1.6 cm, Tosoh Co., Tokyo, Japan) pre-equilibrated with the starting buffer (20 mM phosphate buffer, pH 6.0, 0.02% sodium azide). The column was washed with starting buffer and eluted with a linear gradient of 0–0.5 M NaCl in starting buffer at

a flow rate of 60 mL·h⁻¹. Fractions were examined by SDS/PAGE and tested for the binding of human IgE on immunoblots. Fractions positive for IgE binding were pooled, lyophilized and dissolved in a small aliquot of 0.1% trifluoroacetic acid. Sample was further applied to a semipreparative RP-HPLC column (C₄-300 Å, 200 × 10 mm, Vydac, Hesperia, CA, USA) equilibrated with 0.1% trifluoroacetic acid and the column eluted for 100 min at a flow rate of 1 mL·min⁻¹ using a linear gradient of 0–100% acetonitrile.

SDS/PAGE, immunoblotting and periodic acid Schiff stain

The protein peaks of the different purification steps were separated by 12.5% SDS/PAGE and electroblotted onto poly(vinylidene difluoride) membranes; immunoblotting for IgE-binding proteins were detected as described previously [5]. The patients' sera used for immunoblotting were diluted five-fold. Carbohydrate staining was performed using a Glycoprotein staining kit (Pierce, Rock, IL, USA) according to the manufacturer's instruction.

Lectin-binding assays

The ability of Cyn d 24 to bind various lectins was examined using seven biotinylated lectins from *Canavalia ensiformis* (Con A), *Dolichos biflorus* (horse gram), *Glycine max* (soybean), *Helix pomatia* (edible snail), *Maclura pomifera* (osage orange), *Tetragonolobus purpureas* (winged pea), and *Triticum vulgaris* (wheat germ) (Sigma). Briefly, 96-well plates (Costar, Cambridge, MA, USA), coated and blocked as for ELISA, were incubated sequentially for 3 h at 37 °C with 50 µL of biotinylated lectin (1 µg·mL⁻¹), followed by 50 µL of avidinylated horseradish peroxidase (1 µg·mL⁻¹), bound horseradish peroxidase activity being measured as described for ELISA.

Enzyme-linked immunosorbant assay (ELISA)

ELISA was carried out essentially as described previously [5]. Wells in ELISA plates (Costar) were coated overnight at 4 °C with 50 µL of the antigen (5 µg·mL⁻¹), diluted in carbonate buffer (15 mmol·L⁻¹, pH 9.6), and blocked for 30 min at 37 °C with blocking solution (1% normal goat serum in NaCl/P_i containing 0.1% Tween 20). For IgE binding assays, serum samples from allergic patients were diluted 10-fold with blocking solution and incubated overnight at 4 °C with the immobilized antigen. After washes, alkaline phosphatase-conjugated mouse antihuman IgE antibody (1000-fold diluted) was added for 3 h at 37 °C, the alkaline phosphatase activity was measured using disodium *p*-nitrophenyl phosphate as substrate by measuring the absorbance at 405 nm. All assays were in triplicate.

Proteolytic treatment and amino acid sequence analysis

Cyn d 24 was first denatured and reduced with 2-mercaptoethanol and the reduced protein was digested overnight at 37 °C with endoproteinase Lysyl-C (Lys-C), at an enzyme/substrate ratio of 1 : 50. The Lys-C digestions were carried out in 0.1 M pyridine/acetate/collidine (pH 8.2). The resulting peptides were fractionated by reversed-phase HPLC on a Beckman ODS column (Beckman, Fullerton, CA, USA) using a linear gradient of 5–60% acetonitrile in 0.06% trifluoroacetic acid and a flow rate of 1 mL·min⁻¹. Peptide elution was monitored at 220 nm and all fractions were collected and analyzed by N-terminal sequencing in a Procise ABI 494 protein sequencer (Applied Biosystems, Foster City, CA, USA). MALDI mass-spectroscopic analysis was performed on a Voyager DE-STR mass spectrometer (PerSeptive Biosystems, Framingham, MA, USA).

PCR-based cloning strategy of Cyn d 24 cDNA

Total RNA was extracted from the pollen of *C. dactylon* with the TRIzol reagent kit (Life Technologies, Eggenstein, Germany) according to the manufacturer's instructions. Poly(A)⁺ RNA were purified by oligo(dT) cellulose chromatography. The rapid amplification of cDNA ends (RACE) method was used to produce cDNA fragments coding for Cyn d 24 using a Marathon cDNA amplification kit (Clontech Laboratories, Palo Alto, CA, USA). Two degenerate primers based on sequences near the N-terminal and internal to the gene were synthesized. The sense primer used was 5'-AAYGAYAARGCSATYCARGA-3', encoding seven amino acids, NDKAIQD, near the N terminal, while the antisense primer was 5'-CCRTARTGRTAGSWYTT-3', encoding the internal sequence KSYHYG. As a result of the PCR a fragment of 240 bp was amplified. To obtain the 5' and 3' portions of the Cyn d 24 cDNA, the RACE PCR protocol was used as described previously [42]. The 5'-end was amplified by 5'-RACE using the anchor primer AP2, 5'-GCCGCTCGAGCCCTATAGTGAGT-3', and gene-specific primer 5'-CATGTTGTGCTCGTTGAGGATGTC-3', corresponding to the partial sequence of 240 bp fragment. The 3'-end was amplified by 3'-RACE using the same anchor primer, AP2, and a gene-specific primer 5'-ATGTTCCGGCTCCGGCGCATCTC-3', corresponding to the partial sequence of 240 bp fragment. The amplified PCR product was analyzed by electrophoresis, subcloned into the pGEM-T vector, and then transformed into *Escherichia coli* strain JM109. After transformation, plasmids from positive clones were subjected to sequence analysis using an ABI 377 sequencer (Applied Biosystems, Foster City, CA, USA). Similarity searches were performed using the BLAST program, and multiple amino acid sequence alignments were performed using the GENEDOC program.

Glycopeptide identification and analysis

For MALDI-TOF mass spectrometry analysis of glycopeptides, Cyn d 24 was digested with above condition. Peptide fractions were collected and analyzed by Con A binding assay as described previously. A single fraction with Con A-binding activity was identified and subjected to MALDI-TOF mass spectrometry analysis using the appropriate matrix (α -cyano-4-hydrocinnamic acid). The derivatized glycopeptide of interest was isolated using timed ion selection and analyzed by PSD MS/MS.

For release of N-glycans, Cyn d 24 was treated with trypsin in 50 mmol·L⁻¹ ammonium bicarbonate (pH 8.5), then *N-glycans* were cleaved with glycoamidase A. de-N-glycosylated peptide separated using a C₁₈ Sepharose – pak cartridge (Waters, Milford, MA, USA) and eluted using 5% aqueous acetic acid. The N-glycan samples were permethylated using the NaOH/dimethyl sulfoxide slurry method as described previously [43]. For MALDI-TOF mass spectrometry glycan profiling, the permethyl derivatives (in acetonitrile) were mixed 1 : 1 with a 2,5-dihydroxybenzoic acid matrix (10 mg·mL⁻¹ in acetonitrile) and spotted onto the target plate. Data acquisition was performed manually on a benchtop MALDI LR system (Micromass, Manchester, UK) operated in reflectron mode. Glycan mass profiling was also performed on a Q-TOF Ultima MALDI instrument (Micromass, Manchester, UK), in which case the permethylated sample in acetonitrile was mixed 1 : 1 with α -cyano-4-hydrocinnamic acid matrix (in acetonitrile: 0.1% trifluoroacetic acid, 99 : 1, v/v) for spotting onto the target plate. Mass spectrometry survey data were acquired manually and the decision to switch over to CID MS/MS acquisition mode for a particular parent ion was made on-the-fly on examination of the summed spectra.

Acknowledgements

We thank Ms. J. M. Chen (Institute of Biochemistry, Academia Sinica) for performing the oligosaccharide analysis. This work was supported in part by Promotion of Research-oriented university program from the Ministry of Education and Grant NSC 92-2320-B-002-174 from the National Science Council, Taiwan (LPC) and VGH-93-302 from the Taipei Veterans General Hospital, Taiwan (SNS).

References

- 1 Trinca JC (1962) Pollens used in the treatment of hay fever and asthma in eastern Australia. *Med J Aust* **94**, 969–972.
- 2 Chang H, Lin F & Chang Y (1984) Correlation of skin test to RAST and PRIST in asthma patients. *Clin Med J* **34**, 376–383.

- 3 Orren A & Dowdle EB (1977) Studies on Bermuda grass pollen allergens. *S Afr Med J* **51**, 586–591.
- 4 Ford SA & Baldo BA (1987) Identification of Bermuda grass pollen allergens by electroblotting. *J Allergy Clin Immunol* **79**, 711–720.
- 5 Shen HD, Wang S, Tang RB, Chang ZN, Su SN & Han SH (1988) Identification of allergens and antigens of Bermuda grass pollen by immunoblot analysis. *Clin Allergy* **18**, 401–409.
- 6 Matthiesen F, Schumacher MJ & Lowenstein H (1991) Characterization of the major allergen of *Cynodon dactylon* (Bermuda grass) pollen Cyn d I. *J Allergy Clin Immunol* **88**, 763–774.
- 7 Chang ZN, Liu C, Tam MH, Peng HJ, Tsai JJ & Han SH (1995) Characterization of the isoforms of the group I allergen of *Cynodon dactylon*. *J Allergy Clin Immunol* **95**, 1206–1214.
- 8 Smith PM, Sulphilglu C, Griffith IJ, Theriault K, Knox RB & Singh MB (1996) Cloning and expression in yeast *Pichia pastoris* of a biologically active form of Cyn d I, the major allergen of Bermuda grass pollen. *J Allergy Clin Immunol* **98**, 331–343.
- 9 Su SN, Peng HJ, Yang SY, Tsai LC, Chow LP & Huang SW (2003) Purification and characterization of a novel isoallergen of a major Bermuda grass pollen allergen, Cyn d I. *J Biomed Sci* **10**, 111–119.
- 10 Su SN, Lau GX, Yang SY, Shen HD, Tsai JJ & Han SH (1990) Isolation and partial characterization of Bermuda grass pollen allergen: BG-60b. *Proc Natl Sci Counc* **14**, 85–90.
- 11 Su SN, Lau GX, Lee MJ, Tsai JJ, Yang SY, Shen HD & Han SH (1991) Isolation and partial characterization of Bermuda grass pollen allergen, BG-60a. *Clin Exp Allergy* **21**, 449–455.
- 12 Su SN, Shu P, Lau GX, Yang SY & Lee YC (1996) Immunologic and physicochemical studies of Bermuda grass pollen allergen, BG60. *J Allergy Clin Immunol* **98**, 486–494.
- 13 Ohsuga H, Su SN, Takahashi N, Yang SY, Nakagawa H, Shimada I, Arata Y & Lee YC (1996) The carbohydrate moiety of the Bermuda grass antigen BG60, new oligosaccharides of plant origin. *J Biol Chem* **271**, 26653–26658.
- 14 Liaw SH, Lee DY, Chow LP, Lau GX & Su SN (2001) Structural characterization of the 60-kDa Bermuda grass pollen isoallergens, a covalent flavoprotein. *BBRC* **280**, 738–743.
- 15 Liaw SH, Lee DY, Yang SY & Su SN (1999) Crystallization and preliminary diffraction data of 60-kDa glycosylated pollen isoallergens from Bermuda grass. *J Str Biol* **127**, 83–87.
- 16 Suphioglu C, Ferreira F & Knox RB (1997) Molecular cloning and immunological characterization of Cyn d 7, a novel calcium-binding allergen from Bermuda grass pollen. *FEBS Lett* **402**, 167–172.
- 17 Focke M, Hemmer W, Hayek B, Gotz M & Jarisch R (1998) Identification of allergens in oilseed rape (*Brassica napus*) pollen. *Int Arch Allergy Immunol* **117**, 105–112.
- 18 Astruias JA, Arilla MC, Gomez-Bayon N, Martinez J, Martinez A & Palacios R (1997) Cloning and high level expression of *Cynodon dactylon* (Bermuda grass) pollen profilin (Cyn d 12) in *Escherichia coli*: purification and characterization of the allergen. *Clin Exp Allergy* **27**, 1307–1313.
- 19 Wu W, Tam MH, Peng HJ, Tsai LC, Chi CW & Chang ZN (2001) Isolation and partial characterization of a 46-kD allergen of Bermuda grass pollen. *J Biomed Sci* **8**, 342–348.
- 20 Kao SH, Su SN, Tsai JJ, Huang SW & Chow LP (2005) Sub-proteome analysis of novel IgE-binding proteins of Bermuda grass pollen. *Proteomics* **5**, 3805–3813.
- 21 Stintzi A, Heitz T, Prasadm V, Wiedemann-Merdinoglu S, Kauffmann S, Geoffroy P, Legrand M & Fritig B (1993) Plant 'pathogenesis-related' proteins and their role in defense against pathogens. *Biochimie* **75**, 687–706.
- 22 Hoffmann-Sommergruber K (2002) Pathogenesis-related (PR)-proteins identified as allergens. *Biochem Soc Transactions* **30**, 930–935.
- 23 King TP, Hoffman D, Lowenstein H, Marsh DG, Platts-Mills TAE & Thomas W (1995) Allergen nomenclature. *Allergy* **50**, 765–774.
- 24 Okano M, Nishizaki K, Satoskar AR, Yoshino T, Masuda Y & Harn DA Jr (1999) Involvement of carbohydrate on phospholipase A2, a bee-venom allergen, *in vivo* antigen-specific IgE synthesis in mice. *Allergy* **54**, 811–818.
- 25 Okano M, Kino K, Takishita T, Hattori H, Ogawa T, Yoshino T, Yokoyama M & Nishizaki K (2001) Roles of carbohydrate on Cry j 1, the major allergen of Japanese cedar pollen, in specific T-cell responses *J Allergy Clin Immunol* **108**, 101–108.
- 26 Aalberse RC, Koshte V & Clemens JG (1981) Immunoglobulin E antibodies that cross-react with vegetable foods, pollen, and Hymenoptera venom. *J Allergy Clin Immunol* **68**, 356–364.
- 27 Hiemori M, Bando N, Ogawa T, Shimada H, Tsuji H, Yamanishi R & Terao J (2000) Occurrence of IgE antibody-recognizing N-linked glycan moiety of a soybean allergen, Gly m Bd 28K. *Int Arch Allergy Immunol* **122**, 238–245.
- 28 Iacovacci P, Pini C, Afferni C, Barletta B, Tinghino R, Schinina E, Federico R, Mari A & Di Felice G (2001) A monoclonal antibody specific for a carbohydrate epitope recognizes an IgE-binding determinant shared by taxonomically unrelated allergenic pollens. *Clin Exp Allergy* **31**, 458–465.
- 29 van Ree R, Cabanes-Macheteau M, Akkerdaas J, Milazzo JP, LoutelieR-Bourhis C, Villalba M,

- Koppelman S, Aalberse R, Rodriguez R *et al.* (2000) Beta(1,2)-xylose and alpha(1,3)-fucose residues have a strong contribution in IgE binding to plant glycoallergens. *J Biol Chem* **275**, 11451–11458.
- 30 Kurosaka A, Yano A, Itoh N, Kuroda Y, Nakagawa T & Kawasaki T (1991) The structure of a neural specific carbohydrate epitope of horseradish peroxidase recognized by anti-horseradish peroxidase antiserum. *J Biol Chem* **266**, 4168–4172.
- 31 Bublin M, Radauer C, Wilson IB, Kraft D, Scheiner O, Breiteneder H & Hoffmann-Sommergruber K (2003) Cross-reactive *N-glycans* of Api g 5, a high molecular weight glycoprotein allergen from celery, are required for immunoglobulin E binding and activation of effector cells from allergic patients. *FASEB* **17**, 1697–1699.
- 32 Gray JS, Yang BY, Hull SR, Venzke DP & Montgomery R (1996) The glycans of soybean peroxidase. *Glycobiology* **6**, 23–32.
- 33 Yang BY, Gray JS & Montgomery R (1996) The glycans of horseradish peroxidase. *Carbohydrate Res* **287**, 203–212.
- 34 Falquet L, Pagni M, Bucher P, Hulo N, Sigrist CJ, Hofmann K & Bairoch A (2002) The PROSITE database, its status in 2002. *Nucleic Acids Res* **30**, 235–238.
- 35 Foster JA & Gerton GL (1996) Autoantigen 1 of the guinea pig sperm acrosome is the homologue of mouse Tpx-1 and human TPX1 and is a member of the cysteine-rich secretory protein (CRISP) family. *Mol Reprod Dev* **44**, 221–229.
- 36 Muradov A, Petrasovits L, Davidson A & Scott KJ (1993) A cDNA clone for a pathogenesis-related protein 1 from barley. *Plant Mol Biol* **23**, 439–442.
- 37 Mizuki N & Kasahara M (1992) Mouse submandibular glands express an androgen-regulated transcript encoding an acidic epididymal glycoprotein-like molecule. *Mol Cell Endocrinol* **89**, 25–32.
- 38 Kasahara M, Gutknecht J, Brew K, Spurr N & Goodfellow PN (1989) Cloning and mapping of a testis-specific gene with sequence similarity to a sperm-coating glycoprotein gene. *Genomics* **5**, 527–534.
- 39 Lu G, Villalba M, Coscia MR, Hoffman DR & King TP (1993) Sequence analysis and antigenic cross-reactivity of a venom allergen, antigen 5, from hornets, wasps, and yellow jackets. *J Immunol* **150**, 2823–2830.
- 40 Schuren FHJ, Asgeirsdottir SA, Kothe EM, Scheer JMJ & Wessels JG (1993) The Sc7/Sc14 gene family of *Schizophyllum commune* codes for extracellular proteins specifically expressed during fruit-body formation. *J Gen Microbiol* **139**, 2083–2090.
- 41 Murphy EV, Zhang Y, Zhu W & Biggs J (1995) The human glioma pathogenesis-related protein is structurally related to plant pathogenesis-related proteins and its gene is expressed specifically in brain tumors. *Gene* **159**, 131–135.
- 42 Chow LP, Chou MH, Ho CY, Chuang CC, Pan FM, Wu SH & Lin JY (1999) Purification, characterization and molecular cloning of trichoanguin, a novel type-I ribosome-inactivating protein from the seeds of *Trichosanthes anguina*. *Biochem J* **338**, 211–219.
- 43 Dell A, Reason AJ, Khoo KH, Panico M, McDowell RA & Morris HR (1994) Mass spectrometry of carbohydrate-containing biopolymers. *Methods Enzymol* **230**, 108–132.