



Primary structure, recombinant expression, and molecular characterization of Phl p 4, a major allergen of timothy grass (*Phleum pratense*)[☆]

Andreas Nandy^{a,*}, Arnd Petersen^b, Martin Wald^a, Roland Suck^a, Helga Kahlert^a, Bernhard Weber^a, Wolf-Meinhard Becker^b, Oliver Cromwell^a, Helmut Fiebig^a

^a Allergopharma Joachim Ganzer KG, Research and Development, Reinbek, Germany

^b Research Center Borstel, Molecular and Clinical Allergology, Borstel, Germany

Received 6 September 2005

Abstract

Grass pollen allergy is one of the most important allergic diseases world-wide. Several meadow grasses, like timothy grass and rye grass, contribute to allergic sensitizations, but also allergens from extensively cultivated cereals, especially rye, make a profound contribution. The group 4 allergens are well known as important major allergens of grasses. We have cloned for the first time group 4 sequences from *Phleum pratense*, *Lolium perenne*, *Secale cereale*, *Triticum aestivum*, and *Hordeum vulgare*, and investigated the IgE-reactivity of recombinant Phl p 4 as a candidate for allergy diagnostic and therapeutic applications.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Type I allergy; Grass group 4 allergen; Phl p 4; *Phleum pratense*; Flavoprotein; Recombinant allergen

IgE-mediated hypersensitivity (type I allergy) is an important disease the prevalence of which has been increasing such that it affects about one-quarter of the population of developed countries. Hay fever is caused by grass pollen of several grass species like timothy grass (*Phleum pratense*), and rye grass (*Lolium perenne*), but extensively cultivated cereals, especially rye (*Secale cereale*), also play an important role.

The first group 4 allergen of grasses has been described as a high molecular weight allergen from *L. perenne* in 1983 [1]. Phl p 4, from *P. pratense*, was first described in 1985 as a basic glycoprotein with an apparent molecular weight of

54 kDa and an isoelectric point of 9.45 [2]. Phl p 4, Phl p 1, and Phl p 5 allergens show the highest IgE-frequencies, and Phl p 4 is recognized by 70–88% of grass pollen allergic subjects [3–7]. High cross-reactivity of the specific IgE of allergic subjects and monoclonal antibodies with group 4 allergens of different grass species has been reported [2,6–14].

Whereas the cDNA sequences of several grass pollen allergens are known, the coding sequence for Phl p 4 and other group 4 allergens defied elucidation. Four short internal peptide sequences were reported from *Dactylus glomerata* Dac g 4 and a homology to berberine bridge enzyme (BBE) was proposed [10]. Eleven peptides were identified for the group 4 allergen from *Cynodon dactylon* BG-60 [15] and two peptides of *L. perenne* Lol p 4 [16].

Here we report the molecular cloning of cDNA and genomic DNA sequences of timothy grass Phl p 4, the expression of the gene in *Escherichia coli* and in *Pichia pastoris*, and the biochemical and immunological characterization of the purified recombinant molecules. This is

[☆] Nucleotide sequence data reported are available in the EMBL database under Accession Nos. AJ512487, AJ512488, AJ862835, AJ862830, AJ862831, AJ862832, AJ862833, and AJ862834. Genetic variants have been named provisionally according to the rules of the WHO/IUIS allergen subcommittee nomenclature (www.allergen.org).

* Corresponding author. Fax: +49 40 727 65 252.

E-mail address: andreas.nandy@allergopharma.de (A. Nandy).

the first report of the DNA sequence and deduced amino acid sequence of a grass group 4 allergen. With the knowledge of the Phl p 4 sequence we have identified the primary structures of further group 4 allergens, namely from rye grass (Lol p 4), rye (Sec c 4), wheat (Tri a 4), and barley (Hor v 4), and could align the sequences with BBE, a member of a flavoprotein oxidoreductase superfamily [17].

Materials and methods

N-terminal and internal sequence analysis of natural Phl p 4. Natural Phl p 4 has been purified from *P. pratense* pollen (Allergon) according to the method described earlier [18]. The protein was subjected to enzymatic digestion with endopeptidase Glu-C (Promega) or with endopeptidase Lys-C (Roche). Chemical protein cleavage has been achieved by CNBr treatment. The resulting fragments were purified and subjected to N-terminal protein sequencing by Edman degradation (ABI 476A protein sequencer). Hydroxyprolines have been identified using a PTH-Hyp standard (Pierce).

Isolation of genomic DNA, RNA, and cDNA synthesis. Wild isolates of *P. pratense* inflorescences collected from environmental meadows and leaves of *L. perenne*, *S. cereale*, *Hordeum vulgare*, and *Triticum aestivum* from cultured plants were used to prepare genomic DNA using the DNeasy Plant Maxi Kit (Qiagen). Total RNA was isolated from *P. pratense* pollen using the GlassMAX RNA Microisolation Kit (Invitrogen). Oligo(dT) primed first strand cDNA was generated using a first strand cDNA-Synthesis Kit (Roche) and the Oligo(dT) adapter primer GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTT (Gibco).

PCR amplification of the Phl p 4 sequence from genomic DNA and cDNA. PCRs were performed with a Biometra T-cycler (Biometra). Ethidium bromide stained fragments of expected length were excised from agarose gels, cloned into the pGEM T-easy vector (Promega), and subsequently sequenced. The 147 bp internal Phl p 4 gene fragment from genomic DNA was amplified by degenerate primer pools #30 and #37 with PCR conditions: 5 min at 98 °C, adding Triple Master polymerase (Eppendorf) at 85 °C followed by 45 cycles (15 s 95 °C, 20 s 50 °C, and 1 min 72 °C), and a final elongation for 10 min at 72 °C. Oligo(dT) adapter primer primed first-strand cDNA was used to amplify the 3'-end of the Phl p 4 sequence using a 3'-RACE approach with Triple Master polymerase, primer #82, and AUAP primer GGCCACGCGTCGACTAGTAC (Gibco) in the first amplification round (5 min 98 °C, 35 cycles [15 s 94 °C, 20 s 60 °C, and 1 min 72 °C], 10 min 72 °C). In a second amplification round 0.5 µl of the first PCR was used as template for primer #82 and the UAP primer CUACUACUACUAGGCCACGCGTCGACTAGTAC (Gibco): 5 min 98 °C, 35 cycles (15 s 94 °C, 80 s 72 °C), 10 min 72 °C. A "sense side mixed oligonucleotide, antisense side nested unique primer PCR" approach was used to amplify a 1320 bp fragment of Phl p 4. Here

genomic *P. pratense* DNA served as a template for the degenerate sense primer pool #29 and the specific primer #86 in the first PCR: 5 min 98 °C, addition of *Taq*-polymerase (Eppendorf) at 85 °C, 45 cycles (15 s 95 °C, 20 s 58.3 °C, 60 s 72 °C), 10 min at 72 °C, using HiFi-buffer (Eppendorf) supplemented with DMSO (5% v/v final concentration); and with primer #29 and #85 in the second (nested) PCR round after purifying the PCR with the High-Pure PCR purification kit (Roche). Additional genomic and cDNA clones were amplified using primer combinations #88 and #86, #206 and #189, #206 and #215, and #206 and #217, Triple Master polymerase and HiFi-Buffer with DMSO (5% v/v): 5 min 95 °C, 45 cycles (15 s 95 °C, 20 s 60 °C, and 180 s 72 °C) followed by 10 min at 72 °C.

Expression and purification of recombinant Phl p 4. The gene for mature Phl p 4 was cloned into standard *E. coli* expression vectors and the resulting protein was expressed in inclusion bodies. After cell disruption, the isolated inclusion bodies were finally purified by size exclusion chromatography (Superdex 200, GE) in 6 M guanidinium hydrochloride. For expression in the methylotrophic yeast *P. pastoris* the mature form of Phl p 4, extended by a C-terminal His-Tag, was fused to the *Saccharomyces cerevisiae* α -factor signal sequence of the AOX1-promoter containing vector pPICZ α (Invitrogen). Recombinant Phl p 4 was expressed in KM71H cells using standard protocols and recovered in the culture media. The heterologous protein was purified using immobilized metal ion affinity chromatography (HisTrap HP, GE) followed by size exclusion chromatography step (Superdex 75, GE).

Characterization of recombinant Phl p 4. SDS-PAGE was performed using 10% Bis-Tris gels and SDS-Mops running buffer (Invitrogen) and gels were subsequently stained with Coomassie brilliant blue R-350 (GE). For Western blot analysis monoclonal mouse anti-Phl p 4 antibodies 3C4 [9] and 5H1 [11] were used. IgE inhibition assays used purified natural Phl p 4-coated ELISA plates as solid phase. Sera of allergic donors were preincubated with a dilution series of recombinant or natural Phl p 4 and subsequently applied to the precoated plates. Bound IgE was detected by peroxidase coupled anti-human IgE secondary antibodies. The inhibitory capacity of recombinant and natural Phl p 4 was expressed as percent inhibition [19]. The ability of natural and recombinant Phl p 4 to activate basophils of allergic donors was determined by flow cytometric analysis of CD203c expression [20].

Results and discussion

Amino acid sequencing of natural Phl p 4

N-terminal sequencing of natural Phl p 4 (nPhl p 4) resulted in six sequences, Pl a-f (Table 1), starting with the amino acid sequence YFP. Two of the first 7 prolines were identified as hydroxyprolines. The occurrence of

Table 1
N-terminal and internal peptide sequences of Phl p 4

Peptide	Sequence	Position	Fragment
P 1 a	YFPP'P'AAKEDFLGXLVKEIPRLLYAKSSPAYP	1–33	N-terminal
P 1 b	YFPP'P'AAKEDFLGXLVKE-PPRLLYAKSSP	1–30	N-terminal
P 1 c	YFPXXAAKEDFLGXL	1–15	N-terminal
P 1 d	YFPXXAKKEDFLGXL	1–15	N-terminal
P 1 e	YFPXXAAKDDFLGXL	1–15	N-terminal
P 1 f	YFPXXLANEDF	1–11	N-terminal
P 2	SATPFXHRKGVLFNIQYV	384–401	Glu-C
P 3	GLXYRXLXPE	90–99	Glu-C
P 4	KXMGDDHFXAVR	194–205	Lys-C
P 5	APEGAVDII	238–246	Lys-C
P 6	MEPYVSINPVQAYANY	425–439	CNBr

Positions of the fragments are related to the mature Phl p 4 sequence (Fig. 1). The different N-terminal sequences presumably represent different variants of mature Phl p 4. Amino acid residues that were not confirmed on the DNA level are shown in bold. The prolines marked P' have been confirmed as hydroxyprolines.

hydroxyprolines in plant proteins is quite common [21] and they are often O-glycosylated with arabinosides as described for Phl p 1 [22,23] and Phl p 5 [24]. Five internal peptide fragments P 2 to P6 were obtained from enzymatically or chemically cleaved nPhl p 4 (Table 1) and served as the basis for degenerate Phl p 4 primer design.

Cloning strategy and gene sequence of Phl p 4

The peptide sequences of Phl p 4 (Table 1) were used to design a collection of degenerate Phl p 4 oligonucleotide primer pools of which only the successful ones are shown in Table 2 (primers #29, #30, and #37). With primer pools #30 and #37 consisting of 192 and 16 different oligonucleotides, respectively, one of the resulting fragments obtained with genomic *P. pratense* DNA as the template proved to be a 149 bp fragment of the Phl p 4 gene. This fragment connects the Glu-C fragment P2 and the CNBr-fragment P6 (Table 1) and served to create the first specific Phl p 4 primer #82. This sense primer was used to amplify the 3'-end of the Phl p 4 cDNA in a 3'-RACE PCR. Seven clones have been sequenced and assigned to 2 groups of Phl p 4 sequences with 489 and 557 bp, respectively, that show apparent differences downstream of the stop codons. With the knowledge of the 3'-end of the Phl p 4 transcript new specific antisense PCR primers #85 and #86 were designed. A first successful amplification of the main part of the Phl p 4 gene was achieved using the degenerate primer pool #29, consisting of 256 different oligonucleotides, and the two specific primers #85 and #86 in a half-sided nested PCR using genomic DNA as template. This PCR approach may be called a “sense side mixed oligonucleotide,

antisense side nested unique primer PCR” and led to two long fragments of the Phl p 4 gene among a large number of unspecific amplifications. To improve the PCR specificity, a new gene specific primer #88 adjacent to the degenerate primer #29 has been designed. The primer pair #88 and #86 was used to amplify a Phl p 4 DNA fragment of 1336 bp from genomic *P. pratense* DNA as well as from cDNA. This fragment together with the N-terminal peptide P1 enabled us to present for the first time the complete sequence of mature Phl p 4. Since a 5'-RACE approach to get the full N-terminal sequence was not successful, we used primers based on the 3'-untranslated region (#215, #216, and #217) together with primer #206 based on the nucleotide sequence coding for the putative signal peptide of Tri a 4 from wheat, a homologue of Phl p 4 that has been detected later (see Cloning strategy and gene sequence of Lol p 4, Sec c 4, Tri a 4 and Hor v 4). Some additional 30 genomic and 10 cDNA clones have been sequenced. Due to the large number of sequences from independent PCRs, we were able to exclude PCR errors. The analysis revealed two main isoforms of Phl p 4 as well as some minor variants (see Phl p 4 variants and characterization of the deduced amino acid sequence). A comparison of the genomic and cDNA sequences revealed that the Phl p 4 gene has no intron sequences. The Phl p 4 sequence presented here is the first published sequence of a grass group 4 allergen. Various factors have probably contributed to the fact that this important allergen has defied being cloned for so long, including the relatively high GC-content (~65%) and the assumed strong secondary structures of the DNA/RNA that required the development of the special cloning strategy described here.

Table 2
Oligonucleotides

Name	Sequence ^a	Description
#029	YTNTAYGCNAARWSNWSNCCNGCNTAYCC	Degenerate Phl p 4 sense primer
#030	CAYMGNAARGGNGTNYTNTTYAAYATMC	Degenerate Phl p 4 sense primer
#037	TARTTNGCRTANGCYTGNACNGGRTT	Degenerate Phl p 4 antisense primer
#082	ACTACTGGTTCGCCCCGGGAGCC	Phl p 4 sense 3'-RACE primer
#083	GGCTCCCGGGGCGAACCAGTAG	Phl p 4 antisense primer
#085	TGAAGTATTTCTGGCCCCACACCAAAACC	Phl p 4 nested antisense primer
#086	CCCTTGGTGATGGCGAGCCTCTGG	Phl p 4 antisense primer
#087	ACCAACGCCTCCACATCCAGCT	Phl p 4 sense primer
#088	CTCAGTCCTGGGGCAGACCATCC	Phl p 4 sense primer
#189	GATAAGCTTCTCGAGTGATTAGTACTTTTGATCAGCGGCGGGATGCTC	Phl p 4 antisense primer
#194	GCCTTGCTCTGCCACCACGCCGCCGCCACC	Hor v 4 signal peptide sense primer
#195	GCTCTCGATCGGCTACAATGGCG	Hor v 4 signal peptide sense primer
#198	CACGCACTACAAATCTCCATGCAAG	Hor v 4 3'-UTR antisense primer
#199	CACGCACTAAATCTCCATGCAAG	Tri a 4 3'-UTR antisense primer
#202	CATGCTTGATCCTTATTCTACTAGTTGGGC	Hor v 4 3'-UTR antisense primer
#203	TACGCACGATCCTTATTCTACTAGTTGGGC	Tri a 4 3'-UTR antisense primer
#206	GGTGCTCCTCTCTGCGCCTTGTC	Tri a 4 signal peptide sense primer
#215	GGCCATTGAAACACCAACCATGC	Phl p 4 B 3'-UTR antisense primer
#216	GGAAAAAGAAAGTACACGCCTAACCGC	Phl p 4 B 3'-UTR antisense primer
#217	AGCACCTAACTAATTCTCAAAACCTGTCTC	Phl p 4 A 3'-UTR antisense primer

^a Y = C or T, W = A or T, S = G or C, M = A or C, R = A or G, N = inosine.

Phl p 4 variants and characterization of the deduced amino acid sequence

Two main isoforms of the Phl p 4 gene have been identified and provisionally designated Phl p 4.0101 and Phl p 4.0201 (Fig. 1). The two isoforms show 92.8% DNA sequence identity in the coding regions for mature Phl p 4 and 91% identity on the amino acid level. Due to different 3'-untranslated regions the antisense primer #217 specifically amplifies the Phl p 4.0101 gene, whereas Phl p 4.0201 could be specifically amplified using primer #215 or #216. Three variants of Phl p 4.0101 differing at single amino acid positions (P141T, S289G or E351R) could be detected. Three variants of Phl p 4.0201 with amino acid differences at one position (S76T), two positions (A186P and P227L), or three positions (S76T, A186P, and P227L) were identified. Additionally some silent mutations have been found on the DNA level. The N-terminal sequence P1a and the fragment P2 determined by amino acid sequencing (Table 1) correspond to Phl p 4.0201; P3 and P4 match both Phl p 4.0101 and Phl p 4.0201. Despite being found at the DNA level the existence of the Phl p 4.0101 form could not be verified by amino acid sequencing. However, additional mutations that were not found in the DNA but in the analysed peptides is indicative of the existence of additional variants at positions 6 (A6L), 7 (A7K), 8 (K8N), 239 (S239P), 431 (K431I), and 434 (R434V), and a deletion of position 19 (Δ I19) (Table 1). This may be due to different isoform distributions in the *P. pratense* material used for peptide sequencing and PCR analysis.

Mature Phl p 4 consists of 500 amino acids with calculated masses of 55,895 Da (Phl p 4.0101) and 55,624 Da (Phl p 4.0201) and calculated pI values of 8.8 and 9.2, respectively. There are 4 or 2 potential sites of N-glycosylation (pattern: Asn-Xaa-Ser/Thr) at N61, N187, N298, and N330 (Phl p 4.0101), or N61 and N330 (Phl p 4.0201). Phl p 4 has 5 cysteines at positions 14, 71, 150, 281, and 302 with the potential to form 2 internal disulphide bridges.

Cloning strategy and gene sequence of Lol p 4, Sec c 4, Tri a 4, and Hor v 4

The PCR conditions to amplify the Triticeae and *Lolium* group 4 DNAs were essentially the same as described for Phl p 4 (see PCR amplification of the Phl p 4 sequence from genomic DNA and cDNA). Based on the discovered Phl p 4 DNA five Lol p 4 genomic DNA fragments could be amplified from *L. perenne* genomic DNA using the Phl p 4 specific primers #87 and #189. Five independent genomic DNA fragments could be amplified from *S. cereale* genomic DNA using the Phl p 4 specific primer pair #83 and #87. A databank search using this partial Sec c 4 sequence found single, but not overlapping, EST-fragments in *H. vulgare* and *T. aestivum* EST-databases. Four primers were designed on the basis of the *H. vulgare* EST-fragments GenBank BM442329 (#194 and #195), and GenBank AL510095 (#198 and #202). Four primers were designed

on the basis of the *T. aestivum* EST-fragments GenBank BJ254228 (#199 and #203), GenBank CA594522 (#204), and GenBank BJ243609 (#206). The primer combinations #195 and #189, #195 and #202, #195 and #203, and #195 + #198 were used to amplify 8 independent Sec c 4 genomic sequences. Two Sec c 4 variants Sec c 4.0101 and Sec c 4.0201 have been detected. From *H. vulgare* genomic DNA 17 independent Hor v 4 DNA sequences could be amplified using the primer combinations #195 and #198, #195 and #202, #194 and #198, and #194 and #202. The resulting sequence has been provisionally designated Hor v 4.0101. From *T. aestivum* genomic DNA 13 independent sequences have been amplified using the primer combinations #204 and #203, #204 and #199, #206 and #203, and #206 and #199. Two Sec c 4 sequences have been provisionally designated Sec c 4.0101 and Sec c 4.0201. A comparison with the Phl p 4 cDNA sequence revealed that the *Lolium* and Triticeae sequences, as Phl p 4, do not contain any intron sequences.

Comparison of grass group 4 sequences and homology to the berberine bridge enzyme-like protein family

The alignment of the group 4 allergen sequences from *Phleum*, *Lolium*, and Triticeae species shows a high degree of homology (Fig. 1). They represent a family of basic proteins with molecular weights of about 55 kDa and calculated pI values far above 8 (Table 3). In rye, wheat, and *P. pratense* distinct isoforms with amino acid identities of 88–94% could be detected. The inter-species homology ranges from 83% (Phl p 4 to Triticeae species) to 95% (Sec c 4 to Tri a 4). The occurrence of cross-reacting isoforms in distinct species with amino acid homologies that are comparable to those of different group 4 molecules across the species border is remarkable. Since recombinant group 4 allergens may be important for a future recombinant allergen based specific immunotherapy, efforts should be made to evaluate the cross-reactive therapeutic potential of the different group 4 allergens and their isoforms. A sequence database homology search revealed similarities of grass group allergens to berberine bridge enzyme-like oxidoreductases. More than 20 homologous sequences can be found in the *Arabidopsis thaliana* database. The basic oxidoreductase from *A. thaliana* (pI 9.7) shown in Fig. 1 shares 43% identical residues with Phl p 4, whereas Phl p 4 and the acidic BBE from *E. californica* [25] (pI 5.2) share 34% amino acids only.

Expression of recombinant Phl p 4 in *E. coli* and *P. pastoris*

Since the Phl p 4 gene is quite GC-rich with a GC-content of 65% and because it contains several codons that are rarely found in genes of bacterial or yeast expression hosts we used PCR techniques with overlapping primers to generate a completely artificial codon optimized Phl p 4.0201 gene with a final GC-content of 44%. Following expression in *E. coli* the Phl p 4 protein was found in insoluble

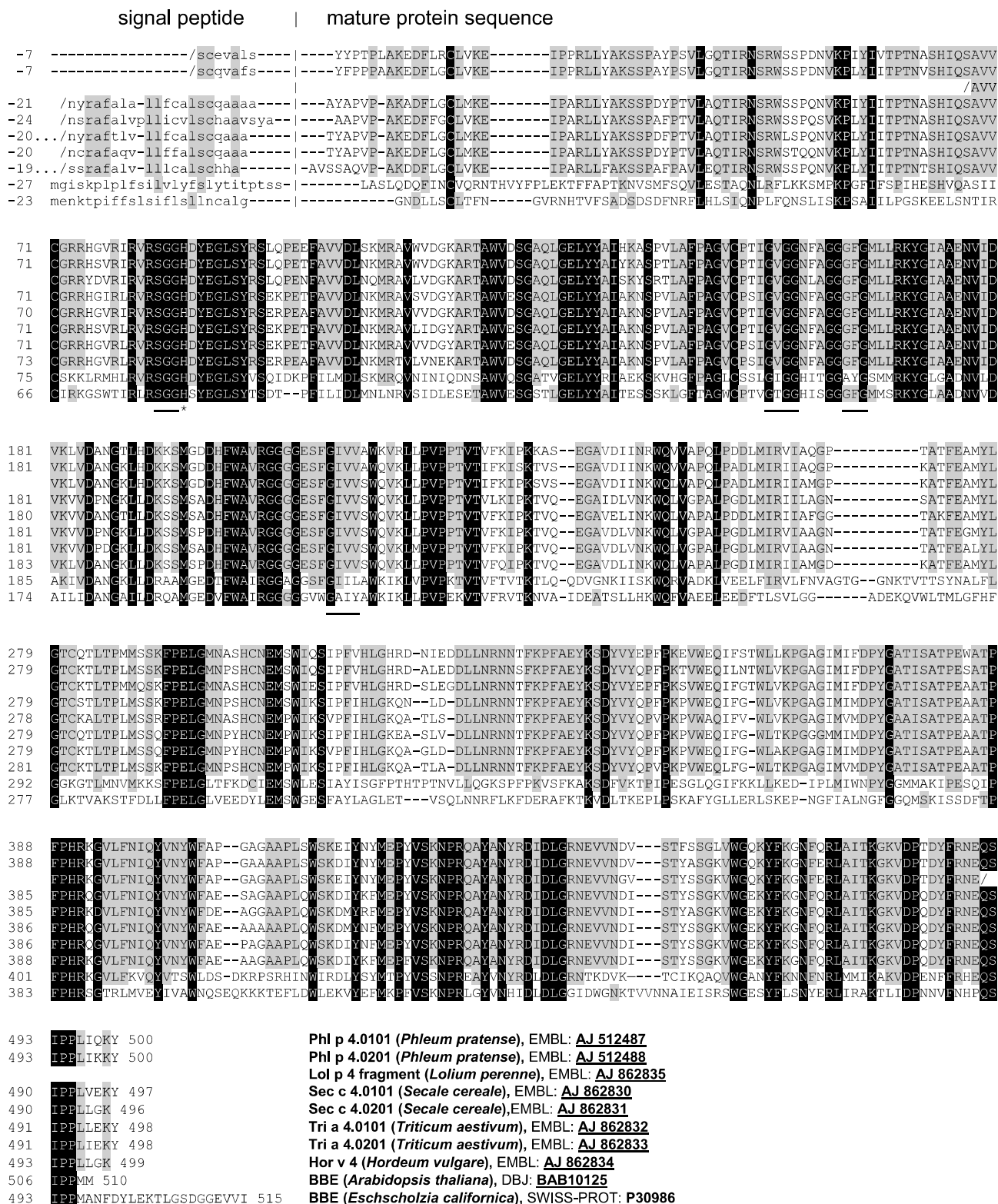


Fig. 1. Amino acid sequence alignment of Phl p 4 variants Phl p 4.0101 and Phl p 4.0201, their homologues from ryegrass, rye, wheat, and barley, and two members of the berberine bridge enzyme oxidoreductase family, BBE from *A. thaliana* and *E. californica*. The numbering corresponds to that of the mature proteins. The start of the mature sequence of BBE and Phl p 4.0201 is known from N-terminal sequencing. The signal peptide cleavage sites of the remaining proteins have been predicted using the TargetP program at www.cbs.dtu.dk/services/TargetP [28]. Truncated sequences are marked with a slash (/). Conserved residues that are identical in all sequences are shown in black, residues that are conserved in the grass spp. are shaded in gray. The oxidoreductase family flavin-binding consensus sequences [17] are marked with black lines, the conserved histidine to which the FAD is covalently bound in BBE is marked with an asterisk.

Table 3
Grass group 4 sequence analysis

Allergen	Sequence length (amino acids)	Molecular weight (Da)	Isoelectric point (pI)
Phl p 4.0101	500	55,895	8.8
Phl p 4.0201	500	55,624	9.2
Lol p 4 ^a	423 ^a	— ^a	8.8 ^a
Sec c 4.0101	497	55,002	9.1
Sec c 4.0201	496	54,670	9.3
Tri a 4.0101	498	55,339	8.8
Tri a 4.0201	498	55,252	8.8
Hor v 4	499	55,074	9.3

The sequence length, isoelectric point, and molecular weight calculations were made on the basis of the mature proteins (Fig. 1).

^a The Lol p 4 sequence is only partial and contains about 85% of the mature Lol p 4 sequence.

inclusion bodies. Recombinant Phl p 4 purified from inclusion bodies shows an apparent molecular weight of about 56 kDa (Fig. 2A) that is in agreement with the mass calculated from the amino acid sequence. The mass of natural Phl p 4 is somewhat higher due to posttranslational glycosylation (Fig. 2A) and has been published as 61,300 Da [6]. In contrast to the situation in *E. coli*, the *P. pastoris* derived His-tagged rPhl p 4.0201 is highly soluble and has been

purified to homogeneity. The protein shows as a single band in SDS-PAGE (Fig. 2A) and eluted as a single peak on SEC corresponding to a monomer. Due to post-translational glycosylation in *P. pastoris* the apparent molecular weight is slightly higher (62,066 Da by mass spectrometry, unpublished) than that of the unglycosylated protein from *E. coli*, but comparable to that of glycosylated nPhl p 4 purified from grass pollen (Fig. 2A). The shift in the SDS-PAGE of reduced and non-reduced nPhl p 4 and rPhl p 4 from *P. pastoris* indicates the existence of internal disulfide bridges and is indicative for a correct folding, in contrast to the *E. coli* derived protein (Fig. 2A).

Phl p 4 contains a covalently bound flavin

Leduc-Brodard et al. [10] proposed that Phl p 4 may have a homology to BBE, an oxido-reductase that contains a covalently bound flavin as co-factor and is involved in alkaloid biosynthesis in plants [26]. For natural BG-60, the group 4 homologue from Bermuda grass (*Cynodon dactylon*) the occurrence of a flavin co-factor was proposed [15]. The sequence of Phl p 4 shows the conserved flavin binding motif (Fig. 1) of the oxido-reductase superfamily described by Fraije et al. [17]. The absorption spectrum of recombinant Phl p 4 shows a typical flavin spectrum (Fig. 3). Since the autofluorescence of Phl p 4 is retained in SDS-denatured nPhl p 4 and *P. pastoris* derived rPhl p 4 on polyacrylamide gel electrophoresis (Fig. 2B), a covalently bound flavin is very likely. The recombinant Phl p 4 produced in *E. coli* did not show any autofluorescence indicating the absence of the co-factor. The candidate amino acid for the covalent attachment of the flavin is His-86, that aligns with His-81, the attachment site in BBE [26]. However, the enzymatic function of Phl p 4 is not known since the BBE substrate (*S*)-reticuline is not a substrate for recombinant Phl p 4 (data not shown).

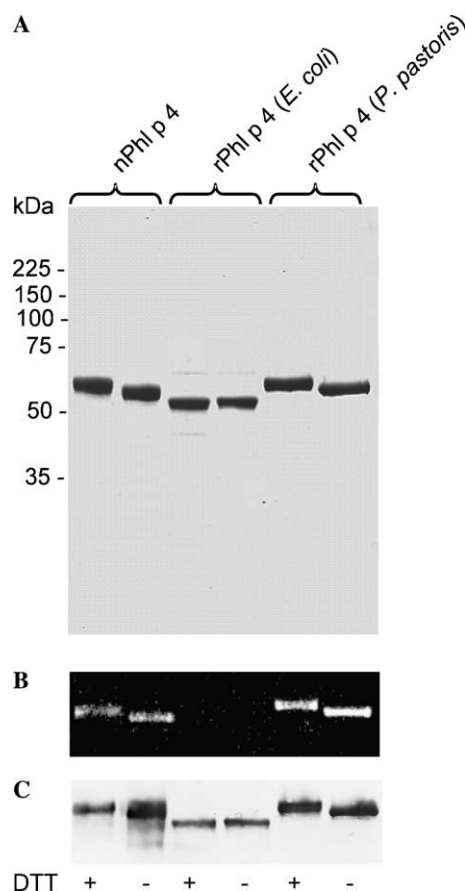


Fig. 2. SDS-PAGE comparison of natural Phl p 4, and recombinant Phl p 4 expressed in *E. coli*, or *P. pastoris*. (A) Stained with Coomassie brilliant blue. (B) Autofluorescence visualized by trans-illumination at 314 nm. (C) Western blot developed with anti-Phl p 4 antibody 3C4.

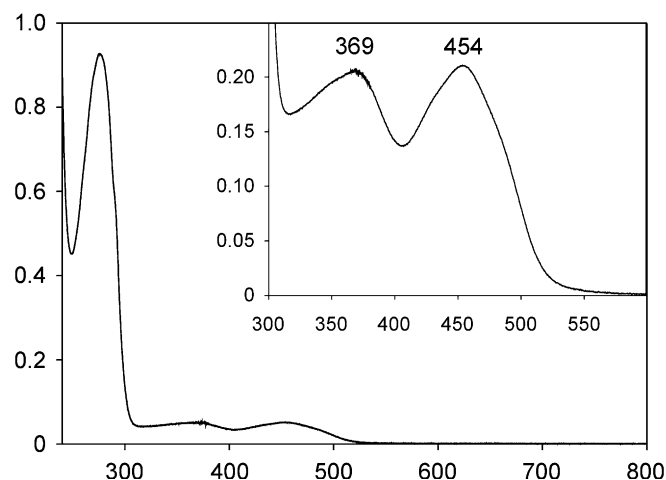


Fig. 3. Absorption spectrum of ~8 and ~30 μM rPhl p 4, respectively. rPhl p 4 shows a typical flavoprotein spectrum with absorption maxima at 369 and 454 nm.

Immunological characterization of Phl p 4–IgE and monoclonal antibody reactivity

Recombinant Phl p 4 and nPhl p 4 show similar reactions with monoclonal anti-Phl p 4 antibodies in Western blot experiments, the reactions with mAb 3C4 are shown in Fig. 2C. Immunoblot analysis using a panel of IgE containing allergic patients' sera showed a good correlation of the reactivity of nPhl p 4 and rPhl p 4. Some sera showed reduced or no reactivity to rPhl p 4. These differences may be explained by IgE directed to carbohydrate determinants and the different nature of the carbohydrate moieties of nPhl p 4 and *P. pastoris* derived rPhl p 4. The reduction of IgE binding to nPhl p 4 carbohydrates in the range of 20–87% after periodate treatment has been described [6]. However, these cross-reactive carbohydrate determinants are believed to be of less importance for the allergic disease

[27]. Therefore, rPhl p 4 will be an important tool to detect true Phl p 4 IgE reactivity. In human IgE-inhibition assays, a good correlation of the reactivity of soluble recombinant and natural Phl p 4 is seen (Fig. 4A). If the lack of epitopes that is seen in some patients (Fig. 4A, right) may be due to isoform differences or due to anti-CCD IgE needs to be elaborated in the future. Basophil activation assays show that rPhl p 4 is able to stimulate basophils of allergic subjects (Fig. 4B), being slightly more potent than natural Phl p 4 in cross-linking IgE on effector cells.

In conclusion we report for the first time the cloning of grass group 4 allergens, the expression of mature Phl p 4 in a soluble immunologically reactive form and its further characterization. The ability to produce recombinant Phl p 4 may represent a key step for the development of future diagnostic and immunotherapeutic preparations and may be of special importance for those allergic persons that

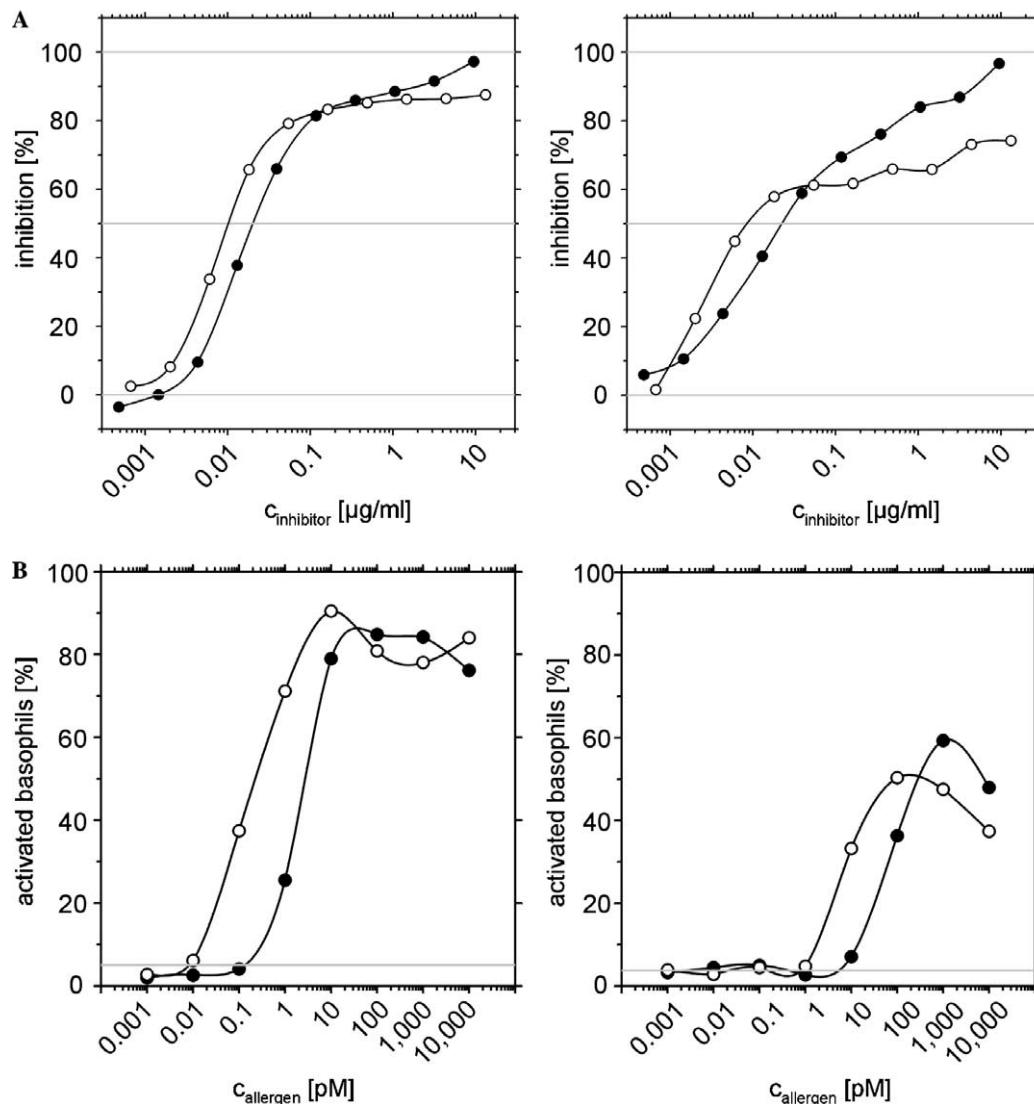


Fig. 4. (A) IgE inhibition assay with sera of two grass pollen allergic blood donors and purified natural Phl p 4 as solid phase. The inhibitory capacity of natural Phl p 4 (●) and that of recombinant Phl p 4-His expressed in *P. pastoris* (○) are compared. (B) Basophil activation test with basophils of two grass pollen allergic blood donors. The baseline is indicated by the grey line. The activation of basophils by natural Phl p 4 (●) and recombinant Phl p 4-His expressed in *P. pastoris* (○) is comparable.

show a dominant IgE response to Phl p 4. Further studies on the cross reactivity of group 4 allergens (e.g., rye) will be important for the correct diagnosis and efficient therapy of grass and rye pollen allergy. Recombinant group 4 allergens expressed in *P. pastoris* are promising candidates for these applications.

References

- [1] A.K. Ekramoddoullah, F.T. Kisil, A.H. Schon, Immunochemical characterization of a high molecular weight basic allergen (HMBA) of rye grass (*Lolium perenne*) pollen, *Mol. Immunol.* 20 (1983) 465–473.
- [2] S. Haavik, B.S. Paulsen, J.K. Wold, Glycoprotein allergens in pollen of Timothy. II. Isolation and characterization of a basic glycoprotein allergen, *Int. Arch. Allergy Appl. Immunol.* 78 (1985) 260–268.
- [3] R. Valenta, S. Vrtala, C. Ebner, D. Kraft, O. Scheiner, Diagnosis of grass pollen allergy with recombinant timothy grass (*Phleum pratense*) pollen allergens, *Int. Arch. Allergy Immunol.* 97 (1992) 287–294.
- [4] S. Fischer, M. Grote, B. Fahlbusch, W.D. Müller, D. Kraft, R. Valenta, Characterization of Phl p 4, a major timothy grass (*Phleum pratense*) pollen allergen, *J. Allergy Clin. Immunol.* 98 (1996) 189–198.
- [5] R.E. Rossi, G. Monasterolo, S. Monasterolo, Measurement of IgE antibodies against purified grass-pollen allergens (Phl p 1, 2, 3, 4, 5, 6, 7, 11, and 12) in sera of patients allergic to grass pollen, *Allergy* 56 (2001) 1180–1185.
- [6] S. Stumvoll, J. Lidholm, R. Thunberg, A. Marknell DeWitt, P. Eibensteiner, I. Swoboda, A. Bugajska-Schretter, S. Spitzauer, L. Vangelista, L. Kazemi-Shirazi, W.R. Sperr, P. Valent, D. Kraft, R. Valenta, Purification, structural and immunological characterization of a timothy grass (*Phleum pratense*) pollen allergen, Phl p 4, with cross-reactive potential, *Biol. Chem.* 383 (2002) 1383–1396.
- [7] A. Mari, Skin test with a timothy grass (*Phleum pratense*) pollen extract vs. IgE to a timothy extract vs. IgE to rPhl p 1, rPhl p 2, nPhl p 4, rPhl p 5, rPhl p 6, rPhl p 7, rPhl p 11, and rPhl p 12: epidemiological and diagnostic data, *Clin. Exp. Allergy* 33 (2003) 43–51.
- [8] S.N. Su, G.X. Lau, J.J. Tsai, S.Y. Yang, H.D. Shen, S.H. Han, Isolation and partial characterization of Bermuda grass pollen allergen, BG-60a, *Clin. Exp. Allergy* 21 (1991) 449–455.
- [9] B. Fahlbusch, W.D. Müller, C. Diener, L. Jäger, Detection of crossreactive determinants in grass pollen extracts using monoclonal antibodies against group IV and group V allergens, *Clin. Exp. Allergy* 23 (1993) 51–60.
- [10] V. Leduc-Brodard, F. Inacio, M. Jaquinod, E. Forest, B. David, G. Peltre, Characterization of Dac g 4, a major basic allergen from *Dactylis glomerata* pollen, *J. Allergy Clin. Immunol.* 98 (1996) 1065–1072.
- [11] B. Fahlbusch, W.D. Müller, O. Rudeschko, L. Jäger, O. Cromwell, H. Fiebig, Detection and quantification of group 4 allergens in grass pollen extracts using monoclonal antibodies, *Clin. Exp. Allergy* 28 (1998) 799–807.
- [12] M. Gavrović-Jankulović, T. Cirković, M. Bukilica, B. Fahlbusch, S. Petrović, R.M. Jankov, Isolation and partial characterization of Fes p 4 allergen, *Invest. Allergol. Clin. Immunol.* 10 (2000) 361–367.
- [13] M. Grote, S. Stumvoll, R. Reichelt, J. Lidholm, R. Valenta, Identification of an allergen related to Phl p 4, a major timothy grass pollen allergen, in pollens, vegetables, and fruits by immunogold electron microscopy, *Biol. Chem.* 383 (2002) 1441–1445.
- [14] K. Andersson, J. Lidholm, Characteristics and immunobiology of grass pollen allergens, *Int. Arch. Allergy Immunol.* 130 (2003) 87–107.
- [15] S.H. Liaw, D.Y. Lee, L.P. Chow, G.X. Lau, S.N. Su, Structural characterization of the 60-kDa Bermuda grass pollen isoallergens, a covalent flavoprotein, *Biochem. Biophys. Res. Commun.* 280 (2001) 738–743.
- [16] K.S. Jaggi, A.K. Ekramoddoullah, F.T. Kisil, Allergenic fragments of ryegrass (*Lolium perenne*) pollen allergen Lol p IV, *Int. Arch. Allergy Appl. Immunol.* 89 (1989) 342–348.
- [17] M.W. Fraaije, J.H. van Berkel, J.A.E. Benen, J. Visser, A. Mattevi, A novel oxidoreductase family sharing a conserved FAD-binding domain, *TIBS* 23 (1998) 206–207.
- [18] R. Suck, S. Hagen, O. Cromwell, H. Fiebig, The high molecular mass allergen fraction of timothy grass pollen (*Phleum pratense*) between 50 and 60 kDa is comprised of two major allergens: Phl p 4 and Phl p 13, *Clin. Exp. Allergy* 30 (2000) 1395–1402.
- [19] O. Cromwell, R. Suck, H. Kahlert, A. Nandy, B. Weber, H. Fiebig, Transition of recombinant allergens from bench to clinical application, *Methods* 32 (2004) 300–312.
- [20] H. Kahlert, O. Cromwell, H. Fiebig, Measurement of basophil-activating capacity of grass pollen allergens, allergoids and hypoallergenic recombinant derivatives by flow cytometry using anti-CD203c, *Clin. Exp. Allergy* 33 (2003) 1266–1272.
- [21] D.T.A. Lampion, D.H. Miller, Hydroxyproline arabinosides in the plant kingdom, *Plant Physiol.* 48 (1971) 454–456.
- [22] A. Petersen, W.M. Becker, M. Schlaak, Characterization of grass group I allergens in timothy grass pollen, *J. Allergy Clin. Immunol.* 92 (1993) 789–796.
- [23] D. Wicklein, B. Lindner, H. Moll, D. Kolarich, F. Altmann, W.M. Becker, A. Petersen, Carbohydrate moieties can induce mediator release: a detailed characterization of two major timothy grass pollen allergens, *Biol. Chem.* 385 (2004) 397–407.
- [24] A. Petersen, W.M. Becker, M. Schlaak, Characterisation of isoforms of the major allergen Phl p V by two-dimensional immunoblotting and microsequencing, *Int. Arch. Allergy Immunol.* 98 (1992) 105–109.
- [25] H. Dittrich, T.M. Kutchan, Molecular cloning, expression, and induction of berberine bridge enzyme, an enzyme essential to the formation of benzophenanthridine alkaloids in the response of plants to pathogenic attack, *Proc. Natl. Acad. Sci. USA* 88 (1991) 9969–9973.
- [26] T.M. Kutchan, H. Dittrich, Characterization and Mechanism of the berberine bridge enzyme, a covalently flavinylated oxidase of benzophenanthridine alkaloid biosynthesis in plants, *J. Biol. Chem.* 41 (1995) 24475–24481.
- [27] R. van Ree, R.C. Aalberse, Specific IgE without clinical allergy, *J. Allergy Clin. Immunol.* 103 (1999) 1000–1001.
- [28] H. Nielsen, J. Engelbrecht, S. Brunak, G. von Heijne, Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites, *Protein Eng.* 10 (1997) 1–6.