Crystal structure and immunologic characterization of the major grass pollen allergen Phl p 4

Domen Zafred, MSc,[a] Andreas Nandy, PhD,[b] Linda Pump,[b] Helga Kahlert, PhD,[b] and Walter Keller, PhD[a] *Graz, Austria, and Reinbeck, Germany*

Background: Phl p 4 is a major pollen allergen but exhibits lower allergenicity than other major allergens. The natural protein is glycosylated and shows cross-reactivity with related and structurally unrelated allergens.
Objective: We sought to determine the high-resolution crystal structure of Phl p 4 and to evaluate the immunologic properties of the recombinant allergen in comparison with natural Phl p 4.
Methods: Different isoallergens of Phl p 4 were expressed, and the nonglycosylated mutant was crystallized. The specific role of protein and carbohydrate epitopes for allergenicity was studied by using IgE inhibition and basophil release assays.
Results: The 3-dimensional structure was determined by using x-ray crystallography at a resolution of 1.9 Å. The allergen is a glucose dehydrogenase with a bicovalently attached flavin adenine dinucleotide. Glycosylated and nonglycosylated recombinant Phl p 4 showed identical inhibition of IgE binding, but compared with natural Phl p 4, all recombinant isoforms displayed a reduced IgE-binding inhibition. However, the recombinant protein exhibited an approximately 10-fold higher potency in basophil release assays than the natural protein.
Conclusion: The crystal structure reveals the compact globular nature of the protein, and the observed binding pocket implies the size of the natural substrate. Plant-derived cross-reactive carbohydrate determinants (CCDs) appear to reduce the allergenicity of the natural allergen, whereas the *Pichia pastoris*–derived glycosylation does not. Our results imply yet undescribed mechanism of how CCDs dampen the immune response, leading to a novel understanding of the role of CCDs.
(J Allergy Clin Immunol 2013;132:696-703.)

Key words: *Grass pollen allergy, recombinant allergen, Phl p 4, Sec c 4, cross-reactive carbohydrate determinant, basophil release assay, dehydrogenase, bicovalently attached flavin adenine dinucleotide*

Abbreviations used
BBE: Reticuline oxidase (informally known as the berberine bridge enzyme)
CCD: Cross-reactive carbohydrate determinant
CD: Circular dichroism
FAD: Flavin adenine dinucleotide
FSAS: Fractional solvent-accessible surfaces
GPA: Grass pollen allergen
RMSD: Root-mean-square deviation

Nearly 30 years ago a high-molecular-weight grass pollen allergen (GPA) was found in perennial ryegrass *(Lolium perenne)*.[1] Later, a basic glycoprotein from timothy grass *(Phleum pratense)* was described as a group 4 GPA, which had an apparent molecular weight of 54 kDa and an isoelectric point of 9.45.[2,3] Two back-to-back studies were published in 2005 and 2006 describing the cloning, expression, and biochemical characterization of recombinant Phl p 4 (rPhl p 4). When expressed in *Escherichia coli*, the flavin adenine dinucleotide (FAD)–dependent enzyme was not properly folded, was missing the cofactor, and did not bind IgE antibodies from allergic patients' sera.[4,5] Natural Phl p 4 (nPhl p 4), along with Phl p 1, Phl p 2, Phl p 5, and Phl p 6, shows the highest IgE frequency in patients with GPA allergy and is recognized by 70% to 88% of adult allergic subjects.[6-9] One important feature of the allergen is its high cross-reactivity with group 4 allergens of different grass species, as well as structurally unrelated allergens.[8] Surprisingly, nPhl p 4 evokes only weak immune response in the majority of allergic patients[10] and a part of these individuals develop IgE antibodies, which are directed towards cross-reactive carbohydrate determinants (CCDs).[9,11] The low observed allergenicity is a unique behavior of a major pollen allergen that has not yet been explained on a structural or molecular level.

Natural extracts have been used for allergy therapy for more than a century, but recently, most of the major allergens have been expressed as recombinant proteins. They can be produced with high purity and homogeneity and on a large scale, which makes them invaluable in research. First recombinant allergens have successfully passed small-scale clinical trials and some have already entered phase III clinical trials.[12-14] Engineering of hypoallergenic derivatives, which evoke considerably less side effects compared with natural or natural-like recombinant allergens in vaccination trials,[12,13] will be the challenge in future vaccine developments.[15] Structural data of allergens are therefore of key importance in structure-based epitope predictions.[16,17]

Here we report an extensive immunologic characterization of the natural timothy grass pollen allergen Phl p 4, as well as 5 recombinant allergens: rPhl p 4.0101, rPhl p 4.0201, rPhl p 4.0202, the rye pollen allergen rSec c 4.0101, and the nonglycosylated rPhl p 4.0201, which was crystallized for a thorough study of the structural and surface features.

From [a]the Institute for Molecular Biosciences (IMB), University of Graz, and [b]the Research and Development Division, Allergopharma Joachim Ganzer KG, Reinbek.
Supported by the Austrian National Science Fund (FWF) through the SFB projects F1805 and F4604.
Disclosure of potential conflict of interest: D. Zafred and W. Keller have received grants from the Austrian Science Foundation (FDF). The rest of the authors declare that they have no relevant conflicts of interest.
Received for publication April 12, 2012; revised March 9, 2013; accepted for publication March 19, 2013.
Available online May 14, 2013.
Corresponding author: Walter Keller, PhD, Institute for Molecular Biosciences (IMB), University of Graz, Humboldtstraße 50/III, 8010 Graz, Austria. E-mail: walter.keller@uni-graz.at.
0091-6749/$36.00

http://dx.doi.org/10.1016/j.jaci.2013.03.021

METHODS

Cloning, expression, and purification

Genes for the constructs were cloned into the *P pastoris* expression vector pPICZα (Life Technologies, Carlsbad, Calif), as described by Nandy et al.[4] Standard PCR techniques were used to generate the nonglycosylated variant rPhl p 4.0201 N61Q N330Q. Recombinant allergens were expressed in the methylotrophic yeast *P pastoris* and purified, as previously described.[4] nPhl p 4 was purified from grass pollen extract, as described by Suck et al.[18]

Purity, concentration, and fold determination

SDS-PAGE and Western blotting were performed as previously described.[4] The murine anti–Phl p 4 mAbs 3C4 and 5H1[19,20] were used for detection of group 4 allergens in combination with a goat anti-mouse IgG–alkaline phosphatase conjugate (Sigma-Aldrich, Taufkirchen, Germany). Gel filtration was performed on a Superdex 200 HR 10/30 column (GE Healthcare Life Sciences, Uppsala, Sweden) in 20 mmol/L sodium phosphate (pH 7.4) and 150 mmol/L NaCl at 0.5 mL/min. Before each set of experiments, a fresh aliquot of purified protein was thawed and centrifuged, followed by a protein concentration measurement by means of UV absorbance at 280 nm. Circular dichroism (CD) spectra were collected on a Jasco J715 spectropolarimeter (JASCO, Tokyo, Japan). Glycosylation was detected with the DIG Glycan detection kit (Roche, Mannheim, Germany).

IgE inhibition and basophil release assays

Sera from individuals with GPA allergy were collected from blood donors.[21] Immunoblotting and IgE inhibition assays were performed, as previously described.[21] Basophil activation assays were performed as described by Kahlert et al,[22] with 8 ten-fold serial dilutions of the protein in the range of 10 nmol/L to 0.001 pmol/L.

Biochemical and enzymatic characterization

Enzymatic activity assays were performed in 80 mmol/L KH_2PO_4/Na_2HPO_4 buffer (pH 6.4), 200 mmol/L NaCl, and 0.05 mg/mL BSA with 4 nmol/L nonglycosylated allergen. The reaction was monitored with the Spectronic 20 Genesys spectrophotometer (Thermo Fisher Scientific, Uppsala, Sweden) at 600 nm or with the Biotrak II plate reader (Amersham Biosciences, Uppsala, Sweden) at 595 nm. The redox state of FAD was measured at 450 nm on the Lambda 35 UV/VIS Spectrophotometer (PerkinElmer, Waltham, Mass).

Crystallization

Batch crystallization setup with the Index screen (Hampton Research, Aliso Viejo, Calif) was prepared with the Oryx8 automatic system (Douglas Instruments, Berkshire, United Kingdom).[23] Optimized crystals were obtained with a sitting drop vapor diffusion setup with 50 μL of reservoir solution in 2 different conditions. Solution I was 1.1 mol/L sodium malonate (pH 7.0), 0.1 mol/L HEPES (pH 7.0), and 0.5% vol/vol Jeffamine ED-2001 (pH 7.0) mixed 1:1 (vol/vol) with the protein. Crystals were soaked in 1× crystallization buffer with 10% vol/vol glycerol before flash cooling in liquid nitrogen to avoid ice formation. Solution II was 0.2 mol/L lithium sulfate monohydrate, 0.1 mol/L HEPES (pH 7.5), and 25% wt/vol polyethylene glycol 3350 mixed 1:1 with the protein and with glucose added to a final concentration of 10 mmol/L. Crystals were soaked in 1× crystallization buffer with 50 mmol/L glucose. A protein stock of 4.5 mg/mL of the nonglycosylated variant in 10 mmol/L Tris (pH 7.4) and 50 mmol/L NaCl was used in all crystallization trials.

X-ray data collection, processing, and refinement

Data sets were collected at X06DA and X06SA beam lines of the Paul Scherrer Institute (Villigen, Switzerland). Data sets were processed by using XDS[24] and scaled with SCALA in the CCP4i program package.[25,26] Molecular replacement was done with Phaser,[27] applying a poly-Ala model based on the structure of reticuline oxidase (berberine bridge enzyme [BBE], PDB ID 3FW9).[28] Cofactor and malonate molecules were built with ARP/wARP ligand.[29,30] Structures were refined iteratively by using Refmac5[31] and manual model building in Coot,[32] with 5% of the data set aside for R_{FREE} calculation (see the footnote for Table II for definition).[33] The cif library file for the reduced form of FAD (for which the 3 letter PDB code FDA is used in the PDB) was calculated by using JLigand, the new CCP4 GUI for LIBCHECK.[34] Structures were analyzed and validated with the MolProbity server.[35]

TABLE I. Sequence alignment of the recombinant isoallergens

Full protein name	UniProtKB/TrEMBL	Sequence identity with Phl p 4.0201 (%)
Phl p 4.0201	Q5ZQK4	100
Phl p 4.0201 N61Q N330Q	N/A	100*
Phl p 4.0202	B2ZWE9	100†
Phl p 4.0101	Q5ZQK5	91
Sec c 4.0101	Q5TIW8	85

N/A, Not applicable.
*Two amino acids are different.
†One amino acid is different (S76 in Phl p 4.0201 corresponds to T76 in Phl p 4.0202).

TABLE II. Basic data collection and refinement statistics

Data collection and refinement		
Space group	$P6_122$	$C222_1$
RCSB/PDB ID	3TSH	3TSJ
Average R_{SYM}	0.059	0.061
Maximal resolution (Å)	1.9	2.0
R_{CRYST}/R_{FREE}	0.16/0.19	0.16/0.21

$R_{CRYST} = \Sigma||F_o|-|F_c||/\Sigma|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes. Calculated for 95% of the data set used for refinement.
$R_{FREE} = R_{CRYST}$ but with 5% randomly chosen measurements omitted from the refinement.
$R_{SYM} = \Sigma|I-\langle I\rangle|/\Sigma I$, where I is the intensity of a reflection and $\langle I\rangle$ its mean value.

Bioinformatic tools

Multiple sequence alignments were made with the ClustalW server.[36] Biochemical properties of proteins were calculated by using the ProtParam tool.[37] Solvent-accessible surfaces[38] were calculated with MSMS,[39] and the fractional solvent-accessible surfaces (FSASs) were calculated by using the program SPADE.[17] The 3-dimensional protein search was conducted with the DALI server,[40] and the 3-dimensional protein alignments were made with the PDBeFold server.[41] Structure images were created with the PyMOL Molecular Graphics System (Schrödinger, Pasadena, Calif).[42]

RESULTS

Expression, purification, and detection

nPhl p 4 was purified from natural extract, yielding pure protein as proved by using SDS-PAGE and specific anti–nPhl p 4 and anti–nPhl p 13 mAbs (see Fig E1, *A*, in this article's Online Repository at www.jacionline.org). Recombinant proteins (Table I) were purified to a high extent, as shown on SDS-PAGE (see Fig E2 in this article's Online Repository at www.jacionline.org). Gel filtration was performed as the final step of purification. Elution diagrams showed that the proteins formed no dimeric or higher oligomeric species but eluted as a single monomeric peak (see Fig E3 in this article's Online Repository at www.jacionline.org). Proteins on Western blots were detected with the murine anti–Phl p 4 IgG mAbs 3C4 and 5H1. The latter bound to all recombinant isoallergens, whereas 3C4 did not bind to rSec c 4.0101 (see Fig E2). Natural and recombinant proteins were completely folded and exhibited very similar CD spectra, which is indicative of a preserved secondary structure between the different isoforms (see Fig E4 in this article's Online Repository at www.jacionline.org). Both