doi:10.1016/S0022-2836(03)00403-0 J. Mol. Biol. (2003) 329, 721–730

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Plant-based Heterologous Expression of Mal d 2, a Thaumatin-like Protein and Allergen of Apple (*Malus domestica*), and its Characterization as an Antifungal Protein

Monika Krebitz[1], Birgit Wagner[1], Fatima Ferreira[2], Clemens Peterbauer[3] Nuria Campillo[4], Michael Witty[4], Daniel Kolarich[5], Herta Steinkellner[5] Otto Scheiner[1] and Heimo Breiteneder[1]*

[1]*Department of Pathophysiology, University of Vienna, AKH-EBO-3Q Waehringer Guertel 18-20 Vienna 1090, Austria*

[2]*Institute of Genetics University of Salzburg Salzburg, Austria*

[3]*Institute of Food Technology University of Agricultural Sciences, Vienna, Austria*

[4]*Department of Biochemistry University of Cambridge Cambridge, UK*

[5]*Institute of Chemistry and Center of Applied Genetics University of Agricultural Sciences, Vienna, Austria*

**Corresponding author*

Mal d 2 is a thaumatin-like protein and important allergen of apple fruits that is associated with IgE-mediated symptoms in apple allergic individuals. We obtained a full-length cDNA clone of Mal d 2 from RNA isolated from ripe apple (*Malus domestica* cv. Golden Delicious). The cDNA's open reading frame encodes a protein of 246 amino acid residues including a signal peptide of 24 residues and two putative glycosylation sites. The deduced amino acid sequence of the mature Mal d 2 protein results in a predicted molecular mass of 23,210.9 Da and a calculated pI of 4.55. Sequence comparisons and molecular modeling place Mal d 2 among those pathogenesis-related thaumatin-like proteins that contain a conserved acidic cleft. In order to ensure the correct formation of the protein's eight conserved disulfide bridges we expressed Mal d 2 in *Nicotiana benthamiana* plants by the use of a tobacco mosaic viral vector. Transfected *N. benthamiana* plants accumulated Mal d 2 to levels of at least 2% of total soluble protein. MALDI-TOF mass spectrometric analyses of the recombinant Mal d 2 and its proteolytic fragments showed that the apple-specific leader peptide was correctly cleaved off by the host plant and that the mature recombinant protein was intact and not glycosylated. Purified recombinant Mal d 2 displayed the ability to bind IgE from apple-allergic individuals equivalent to natural Mal d 2. In addition, the recombinant thaumatin-like Mal d 2 exhibited antifungal activity against *Fusarium oxysporum* and *Penicillium expansum*, implying a function in plant defense against fungal pathogens.

© 2003 Elsevier Science Ltd. All rights reserved

Keywords: recombinant apple allergen; transient expression; tobacco mosaic virus; antifungal pathogenesis-related protein; structural modeling

Introduction

Along with changes in life style and eating habits and with the increase of fruit and vegetable consumption for health reasons, the prevalence of allergic reactions to plant food has increased over the last ten years. Approximately 2% of the adult population suffer from food-induced allergic disorders.[1] According to some studies it was estimated that as much as 10%[2] of the population is affected by IgE-mediated food allergy with symptoms ranging from oral allergy syndrome, asthma, skin manifestations, symptoms in the gastro-intestinal-tract to anaphylactic shock. Patients suffering from type I allergy caused by birch pollen frequently demonstrate an intolerance to fruits and/or vegetables.[3] The most frequently observed adverse reactions to fruits are those induced by the consumption of apples (*Malus domestica*).[4] Interestingly, apple also accounts for nearly 60% of the world fruit tree production.[5] Mal d 1, an allergen homologous to the major birch pollen allergen Bet

Abbreviations used: n, natural; PR, pathogenesis-related; r, recombinant; TLP, thaumatin-like protein; TMV, tobacco mosaic virus.

E-mail address of the corresponding author: heimo.breiteneder@akh-wien.ac.at

 © 2003 Elsevier Science Ltd. All rights reserved

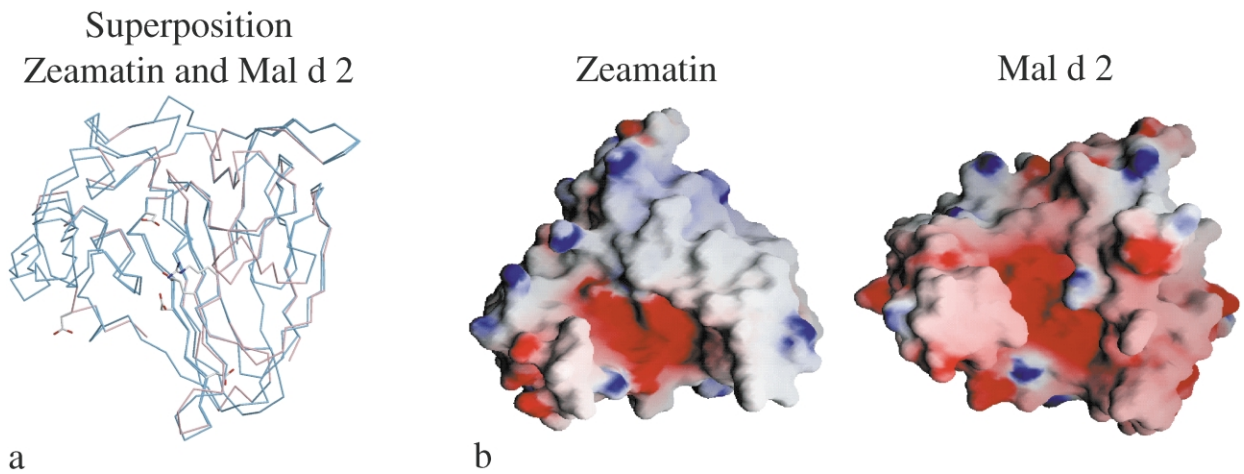

Figure 1. (a) A view of the α-carbon traces of zeamatin (blue) and of the model of Mal d 2 (pink) superimposed. Residues forming the acidic cleft are shown as ball-and-stick models. This Figure was generated with MOLSCRIPT[44] and RASTER3D.[45] (b) View of the surface topology of zeamatin and the model of Mal d 2 showing the surface electrostatic potential ranging from blue (most positive) to white (neutral) to red (most negative). The Mal d 2 model was generated using GRASP.[46]

v 1, was the first apple allergen to be characterized on the molecular level.[6]

Hsieh *et al.* reported a 31 kDa apple protein that appeared to be involved in plant disease resistance reactions.[7] Interestingly, in this study 75% of the apple allergic patients' sera reacted with this 31 kDa protein, indicating a major allergen of at least the same importance as Mal d 1. N-terminal sequencing of the 31 kDa protein revealed approximately 50% identity to the superfamily of thaumatin-like proteins (TLP)[8] that belong to the group 5 of pathogenesis-related (PR) proteins.[9] TLPs contain 16 conserved cysteine residues that form eight disulfide bonds essential for the overall fold, the antifungal and possibly the allergenic activity of the protein.[10,11]

Here we describe the use of a tobacco mosaic virus (TMV)-based expression system for the production of a recombinant plant protein whose folding is dependent on the correct formation of several disulfide bonds. Transient plant expression systems have been used for the production of a variety of heterologous proteins since they became popular in 1995.[12] Recently, we applied such a system for the successful expression of the major birch pollen allergen Bet v 1, a protein that is neither glycosylated nor harbors any disulfide bridges.[13] Here we report for the first time the TMV-directed expression of recombinant (r) Mal d 2, a TLP and important allergen from apple, in *N. benthamiana* plants. Furthermore, we investigated the antifungal activity of rMal d 2.

Results

Isolation and characterization of a Mal d 2 cDNA clone

The analysis of six independent cDNA clones identical in their coding regions and 5′ and 3′ untranslated regions yielded the reported sequence, available from the EMBL database under the accession number AJ243427. The Mal d 2 cDNA comprised 975 bp and had an open reading frame of 738 bp, which encoded a polypeptide of 246 amino acid residues. The deduced Mal d 2 amino acid sequence includes a signal peptide of 24 residues. The removal of this leader sequence results in a mature protein with a predicted molecular mass of 23,210.9 Da and a calculated pI of 4.55. There are two possible glycosylation sites in the mature Mal d 2 protein (N^{174}NTP and N^{205}STF, respectively). Furthermore, a thaumatin family signature corresponding to the consensus pattern, GX(G/F)XCXT(G/A)DC$X_{1,2}$G$X_{2,3}$C (position 63–79 in the mature protein), was detected. Besides the three cysteine residues included in the signature sequence a further 13 cysteine residues, all of them involved in disulfide bonds characteristic for TLPs, were identified in their conserved positions.

The deduced amino acid sequence of Mal d 2 was compared to sequences in the EMBL and SWISS-PROT databases. The sequence of the mature Mal d 2 protein isolated from Golden Delicious apples was identical to a previously reported deduced amino acid sequence of the cDNA clone Mdtl1 isolated from Fuji apples.[5] Interestingly, the leader peptides of Mal d 2 (24 residues) and Mdtl1 (23 residues) encoded by these cDNA clones differed by 37%. Mal d 2 showed significant sequence similarity to members of family 5 of the PR proteins, also referred to as TLPs.[9] Mal d 2 shares 72% sequence identity with Pru av 2, a major allergen and TLP from cherry (*Prunus avium*),[14,15] and 50% identity with Jun a 3, an allergen from pollen of mountain cedar (*Juniperus ashei*).[16] Sequence identities to other plant TLPs with determined crystal structures

were around 40% (PR-5d from *Nicotiana tabacum*,[10] zeamatin from *Zea mays*,[17] thaumatin from *Thaumatococcus daniellii*).[18] Like its homologous proteins thaumatin, zeamatin and PR-5d from tobacco, Mal d 2 consists of three domains. Together with domain I, domain II forms a large cleft. As previously reported, this cleft is highly acidic in PR-5d and zeamatin, whereas thaumatin mainly displays a basic surface in the cleft region.[10] The acidic residues involved in the formation of the acidic cleft of the mature tobacco PR-5d are Arg45, Glu85, Asp98, Asp103, and Asp186, and these residues are also found in mature Mal d 2 (Arg50, Glu92, Asp105, Asp110, and Asp203; Figure 1).

Protein modeling

Four models of Mal d 2 were built based on templates of known crystal structures of homologous TLPs. Model 1 was based on the tobacco TLP PR-5d (PDB file 1aun) as it showed the highest sequence identity with the target sequence. Model 2 was based on thaumatin (PDB file 1thw) as its structure had been solved to the highest resolution. Model 3 was built on both structures (PDB files 1aun and 1thw), and for model 4 the template of zeamatin (PDB file 1du5) was employed. The models were analyzed by the Ramachandran plot[19] and based on values of the root mean square deviations (RMSD). Model 3 showed the lowest value of RMSD with all used template structures, indicating a high degree of structural similarity within the TLP family, as could be expected from the high degree of sequence similarity. Figure 1(a) shows the superposition of the model of Mal d 2 with zeamatin (PDB 1du5), the only other plant food TLP whose crystal structure is known to date.[17] The acidic cleft conserved in zeamatin,[17] PR-5d[10] and in the model of Jun a 3[11] is also conserved in Mal d 2, as shown by the surface electrostatic potential of zeamatin and Mal d 2 (Figure 1(b)).

Expression of rMal d 2 in *N. benthamiana* plants and its purification

The coding sequence of Mal d 2 was placed under the control of the TMV-U1 coat protein subgenomic promoter by ligating the cDNA sequence encoding Mal d 2, including its own signal peptide sequence, and a sequence for a hexahistidyl tag into the 30B based TMV-expression vector. A recombinant clone containing the correct insert was used to produce infectious viral RNA by *in vitro* transcription with T7-RNA polymerase. Mechanical inoculation of *N. benthamiana* plants with *in vitro* transcripts resulted in symptomatic infection after ten days, visible as mild leaf deformations, some variable leaf mottling and growth retardation. RT-PCR was performed with RNA isolated from newly formed upper leaves two weeks after inoculation. Sequencing of the RT-PCR products confirmed the correct sequence of Mal d 2 mRNA transcribed from the viral subgenomic promoter in infected leaves.

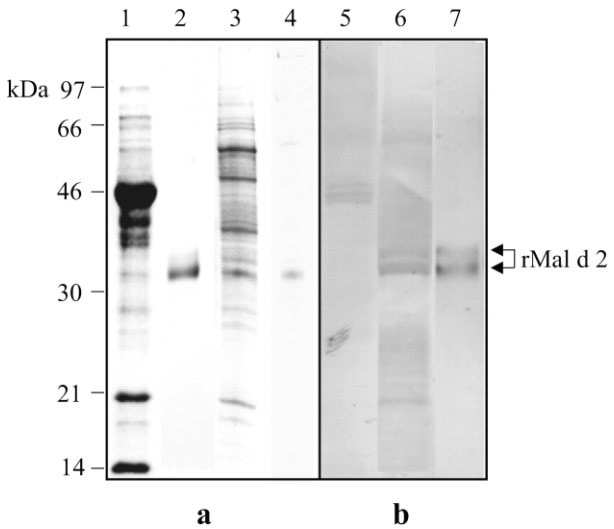

Figure 2. (a) Coomassie-stained SDS-PAGE gel of total protein extract from infected *N. benthamiana* leaves (lane 1), rMal d 2 from infected *N. benthamiana* plants purified by Ni-NTA chromatography (lane 2), total apple extract (lane 3), and nMal d 2 purified by reversed phase HPLC (lane 4). (b) Western blot detection of rMal d 2 by anti-$(His)_5$ antibody: *N. benthamiana* leaf extract after infection with the viral construct not harboring the Mal d 2 coding sequence (lane 5), rMal d 2 present in crude protein extracts of *N. benthamiana* leaves (lane 6), purified rMal d 2 (lane 7). Arrows indicate the positions of rMal d 2 conformers.

Three weeks after inoculation, upper leaves of *N. benthamiana* plants were harvested and proteins extracted. Samples of the supernatant of infected leaves were subsequently separated by SDS-PAGE, stained by Coomassie brilliant blue (Figure 2(a), lane 1) or transferred to nitrocellulose membranes and analyzed by Western blotting (Figure 2(b), lane 6). The expression in TMV-infected *N. benthamiana* plants yielded a diffuse double band at approximately 31 kDa, both bands differing by 1–2 kDa in molecular mass as detected by immunoblotting (Figure 2(b), lane 6). The recombinant protein was purified under native conditions by Ni-NTA affinity chromatography and eluted at 50 mM imidazole. The eluted fraction containing the two rMal d 2 bands was visualized by SDS-PAGE (Figure 2(a), lane 2). Both bands were recognized by an alkaline phosphatase-conjugated anti-His-tag monoclonal antibody (Figure 2(b), lane 7).

The final yield of the purified recombinant protein was calculated to be 240 μg/g fresh leaf material (0.024%). Total soluble protein extracted from leaves was determined by the method described by Bradford to be 10 mg/g fresh leaf. Therefore, rMal d 2 was at least accumulated to

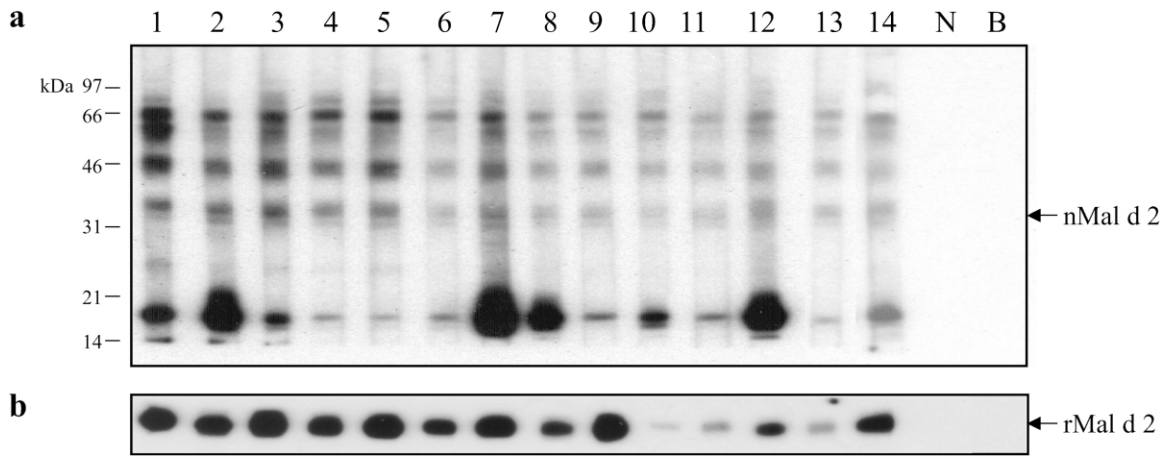

Figure 3. IgE-immunoblot analysis of apple extract (a) and rMal d 2 (b). Nitrocellulose-blotted apple extract and rMal d 2 were incubated with individual sera of patients allergic to apple (lanes 1–14). Controls were a serum pool from ten atopic individuals with high IgE titers against house dust mite but no detectable IgE against apple proteins (lane N), and buffer instead of serum (lane B). IgE binding to natural and recombinant Mal d 2 are indicated.

2.4% of the soluble protein in leaves as measured after purification.

N-terminal sequencing of rMal d 2 and mass spectroscopy analysis

To confirm the identity of the recombinant protein the first 12 amino acid residues of the two bands were determined by protein microsequencing, revealing the same peptide sequence, AKITFNNCPNTV. These experimental data are in accordance with the predicted N terminus of the mature protein corroborating the presence of a 24 amino acid residue leader sequence. Mass analysis of the purified rMal d 2-$(His)_6$ revealed a predominant peak on MALDI-TOF at a molecular mass of 23,878.1 Da compared to the calculated mass of 24,033.7 Da (including the hexahistidyl tag) of the deduced amino acid sequence of Mal d 2 (data not shown). Six additional minor peaks were detected with mass shifts of 140, 219, 1012, 1162, 1312 and 1518 (data not shown). These values were obtained with different laser energies and different matrix preparations.

Mass analysis of the reduced and alkylated rMal d 2 revealed only one sharp peak at a molecular mass of 24,800.41 Da (data not shown). These data indicate that some of the cysteine residues were modified due to covalently bound molecules forming minor products when the protein was not completely reduced, resulting in indistinct protein bands in SDS-PAGE.[20] The N-terminal sequencing of proteolytic fragments treated with trypsin or endoproteinase Glu-C revealed that the difference of 155.59 Da between measured and theoretical mass was due to the loss of one His residue. The protein was not glycosylated. Peptides of mature rMal d 2 containing putative glycosylation sites (N^{174}NTP or N^{205}STF, respectively) had the predicted masses and no peaks with mass shifts were found (data not shown).

Immunological reactivity of rMal d 2 and immunoblot inhibition experiments

Immunoblot analysis of apple extract and of rMal d 2 was performed with sera from individuals allergic to apple. IgE immunoblots with serum samples containing IgE against natural (n) Mal d 2 are shown in Figure 3(a). rMal d 2 was able to bind IgE from all sera tested (Figure 3(b)). IgE binding to the rMal d 2 was visualized as one single band (Figure 3(b)), corresponding to the predominant and lower of the two Coomassie-stained bands of rMal d 2 (Figure 2). No IgE-binding could be detected with the pool of sera from atopic individuals without allergy to apple (Figure 3(a) and (b), lane N).

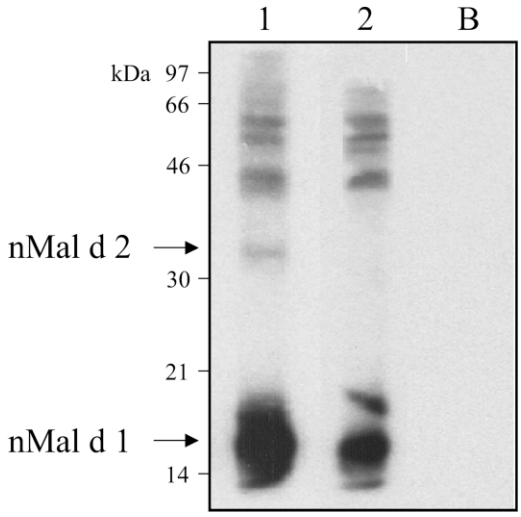

Figure 4. Inhibition of patients' IgE-binding to nMal d 2 in apple extract by rMal d 2. The serum pool of patients allergic to apple was pre-incubated with BSA (lane 1) or with 30 μg purified rMal d 2 (lane 2) and then incubated with apple extract blotted onto nitrocellulose strips. Lane B, buffer control without addition of serum. IgE binding to nMal d 1 and nMal d 2 is indicated.

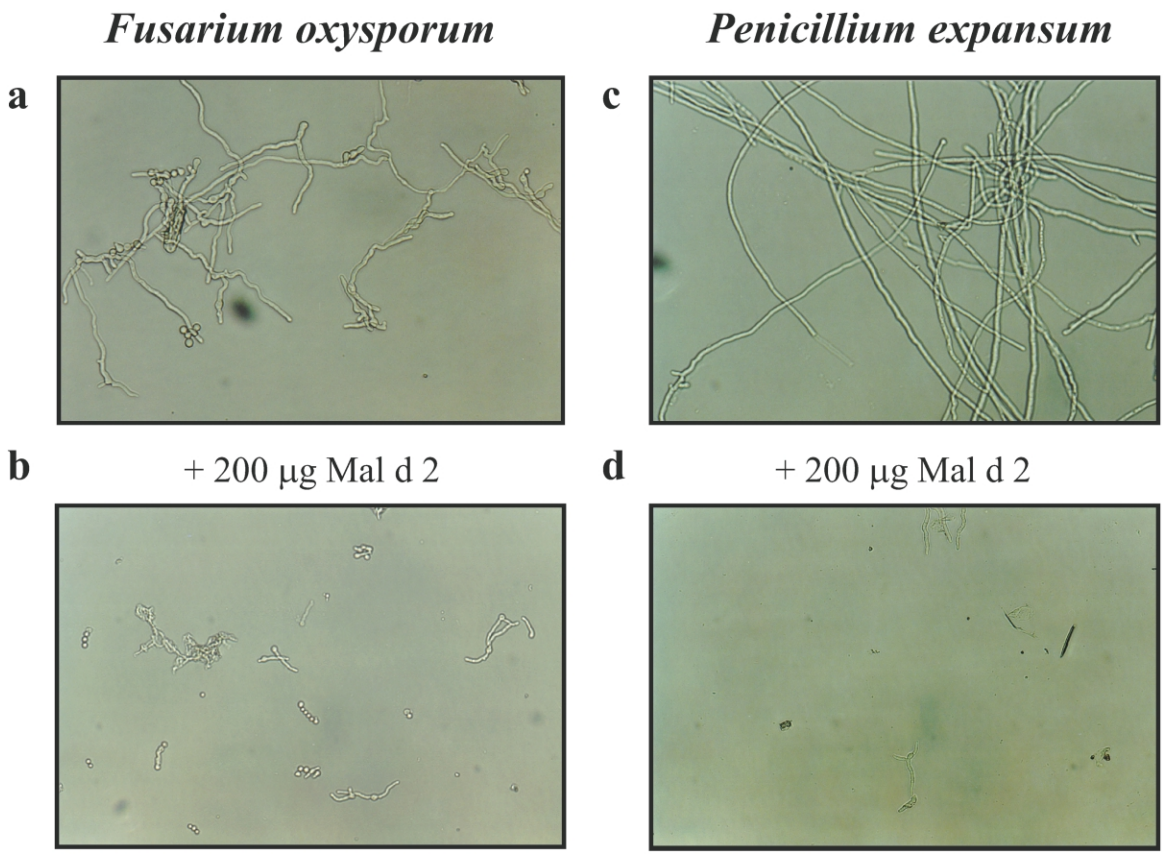

Figure 5. Antifungal activity assay of rMal d 2. The germination rate and the length of the germ tubes of *F. oxysporum* (a) were reduced after an incubation period of 24 hours by the addition of 200 µg/ml rMal d 2 (b). *P. expansum* (c) showed a greatly reduced germination rate after addition of 200 µg/ml rMal d 2 as well as reduced lengths of germ tubes (d).

An inhibition assay was performed using the serum pool of apple allergic patients (Figure 4). IgE-binding to nMal d 2 was not inhibited after pre-incubating the serum pool with BSA (Figure 4, lane1), whereas total inhibition was observed when the serum pool was pre-incubated with 30 µg rMal d 2 (Figure 4, lane 2).

Fungal growth inhibition assay

The germination rate of *Fusarium oxysporum* determined in triplicates was reduced from a mean value of 89% without the addition of rMal d 2 (Figure 5(a)) to a mean of 53% at 200 µg/ml rMal d 2 (Figure 5(b)) after a 24 hour incubation period (mean inhibition of 40%). Additionally, the lengths of the germ tubes were reduced by approximately 90% (Figure 5(b)). *Penicillium expansum* spores showed a more prominent reduced germination rate from 72% without addition of rMal d 2 (Figure 5(c)) to 51% at 100 µg/ml rMal d 2 to 11% at 200 µg/ml rMal d 2 (mean inhibition levels of 29% and 85%, respectively; Figure 5(d)), as well as reduced lengths of germ tubes (Figure 5(d)). Due to the observed necrotic hyphae the effect of the addition of rMal d 2 was considered as fungicidal.[21] No effect on germination or growth of *Alternaria alternata* was observed (data not shown).

Discussion

The present study reports for the first time the production of an allergenic TLP in a transient plant virus-based expression system. We were able to isolate a cDNA clone of the 31 kDa apple allergen Mal d 2 from Golden Delicious apple, a very popular cultivar in Europe. The open reading frame of the Mal d 2 cDNA encodes a polypeptide of 246 amino acid residues with a predicted molecular mass of 23,210.9 Da and a calculated pI of 4.55. The first 24 amino acid residues were absent in the mature protein as determined by microsequencing and were therefore taken to constitute a signal peptide. The presence of this leader indicates that Mal d 2 is not a cytoplasmic protein. Interestingly, the amino acid sequence of a fruit-specific TLP deduced from a cDNA clone isolated from *M. domestica* cv. Fuji shows complete identity with the mature Mal d 2, but differs in the length and the sequence of the leader peptide.[5]

Generally, TLPs are divided into two subgroups, a basic form found in vacuoles and an acidic form that is apoplastic.[22] The presence of a signal

peptide, the absence of a C-terminal extension as found in vacuolar forms, and the acidic character of Mal d 2 suggest that this protein accumulates in the apoplast.

Most of the PR-5 type proteins consist of three domains containing 16 conserved cysteine residues involved in the formation of eight disulfide bridges. We used the crystal structure of three PR-5 type proteins (PR-5d from tobacco,[10] zeamatin from *Z. mays*,[17] thaumatin from *T. daniellii*[18]), all around 40% identical to Mal d 2, as templates to build homology models for the allergen. Comparison of the three-dimensional molecular structures revealed that many of those with antifungal activity had a common motif, a negatively charged surface cleft.[10] Secondary and tertiary structure analyses showed that the amino acid residues that comprise the acidic cleft are conserved[10] and also found in the TLPs identified as allergens, namely Pru av 2,[14] Jun a 3,[16] and Mal d 2 (Figure 1). Although the contribution of the acidic cleft to PR-5 antifungal activity is not clear, it may be the site of interaction between PR-5 members and their receptors on the plasma membrane of target microorganisms and may be directly involved in cell permeabilization.[17]

When expressing recombinant thaumatin or rTLPs in bacteria, yeast, insect cells, and fungi several problems have been reported such as incorrect processing, incorrect three-dimensional folding, insolubility, and insufficient yields of the pure recombinant protein.[23–26] A major drawback of non-plant-based expression is the need for additional purification procedures to produce a recombinant protein free of toxins and fit for clinical applications. Therefore, a transient plant expression system that was recently evaluated for the expression of plant allergens[13,27] was applied here.

Purified rMal d 2 appeared as a double band at 31–32 kDa on SDS-PAGE (Figure 2, lane 2). Interestingly, both bands had identical N-terminal sequences. To throw light on this phenomenon purified rMal d 2 was subjected to MALDI-TOF before and after a treatment with reducing agents. Mass analysis of the reduced and alkylated rMal d 2 showed only one sharp peak at a molecular mass of 24,800.41 Da. These data support the presumption that in a small amount of the Mal d 2 sample some of the eight disulfide bridges were reduced to SH-groups during the purification process, resulting in different migration of the conformers on SDS-PAGE and decreased IgE binding capacity. This supports the view that the correct three-dimensional structure of this protein may be crucial for the presentation of conformational epitopes and, hence, for the binding of IgE.

Carbohydrate moieties and their relevance in allergic diseases are subjects of ongoing discussions.[28] There are two possible glycosylation sites in the mature Mal d 2 protein, one at position 205 (N^{205}STF) and one at position 174 (N^{174}NTP). As there is a proline C-terminal to the threonine, the probability of glycosylation of the second site would be only 50%.[29] However, MALDI-TOF mass-spectroscopy of the mature protein as well as of proteolytic fragments showed that rMal d 2 was not glycosylated.

Mal d 2 represents the second important apple allergen that has been identified. The recombinant apple TLP was recognized by IgE from the apple allergic patients included in this study (Figure 3(b)). In addition, rMal d 2 could inhibit IgE binding to the 31 kDa protein band in apple extract (Figure 4, lane 2), indicating that the recombinant allergen possesses all IgE epitopes present in nMal d 2 and is therefore immunologically equivalent to its natural counterpart.

Soman and co-workers were the first to identify continuous epitopes of a thaumatin-like allergen.[11] These epitopes were found clustered in one area of the surface of Jun a 3, easily accessible for interaction with IgE.[11] The corresponding homologous area in Mal d 2 may also represent IgE epitopes, as the overall structural models of the two proteins are highly similar.

Inhibition of the growth and branching of the hyphae of several fungal species was observed as one of the proposed defense roles of TLPs in plants.[30] The demonstration of fungicidal activity of rMal d 2 against *F. oxysporum* and *P. expansum* concurs with the protein's expected biological function of a TLP (Figure 5). The protein's biological activity indicates a correct folding of the recombinant protein in addition to its already demonstrated IgE-reactivity. The observed inhibition values of rMal d 2 are roughly comparable with the data obtained by Lorito and co-workers using an identical growth inhibition protocol.[21] The glucan 1,3-β-glucosidase from *Trichoderma harzianum* P1 (now *T. atroviride*) showed 55% and 75% inhibition on spores of *Botrytis cinerea* at 100 μg/ml and 200 μg/ml, respectively.[21] In a different assay based on microspectrophotometry Hu and Reddy found a 50% growth inhibition of *F. oxysporum* at a concentration of zeamatin or osmotin around 30 μg/ml.[23] This assay, however, does not count the germinating spores microscopically but infers growth inhibition from optical densities independent of the length of the growing hyphae.

There is an ongoing discussion about the allergenicity of transgenic plants.[31] TLPs have already been used to enhance the resistance of transgenic plants.[32,33] As TLPs of cherry,[14] cedar pollen,[16] bell pepper,[34] as well as the TLP of apple described here, are now characterized as allergens, the risk of creating foods with a higher content of proteins with a known allergenic potential should be considered when introducing TLP-encoding sequences into transgenic crop plants.

Materials and Methods

Isolation of apple RNA and natural Mal d 2

Fruits of the apple *Malus domestica* cv. Golden Delicious,

purchased at a local market, were used for this study. RNA isolation and protein extraction were performed from fruits of the same harvest. Total RNA was prepared from 100 g apple fruit (mesocarp and skin), according to the acidic guanidinium–HCl method including a CsCl-purification.[35] Proteins from Golden Delicious apples were extracted as described by Vanek-Krebitz and co-workers.[6] Natural Mal d 2 was purified from the apple protein extract by reversed-phase HPLC (RP-HPLC) employing a linear gradient of 2-propanol at room temperature (C8 Hypersil WP 300, 10 μm, 8 mm × 250 mm; solvent A: 0.1% trifluoroacetic acid in water; solvent B: 90% 2-propanol, 0.1% trifluoroacetic acid; gradient: of 0–80% of solvent B). The identity of the 30 kDa protein fraction was confirmed by SDS-PAGE and N-terminal sequencing using a Procise 491 sequencer (Applied Biosystems, Foster City, California).

cDNA synthesis and isolation of a cDNA encoding Mal d 2

Poly(A)-enriched RNA was prepared from total RNA by biotinylated oligo-(dT) and streptavidin coupled to paramagnetic particles provided by the PolyATract mRNA Isolation Sytstem (Promega, Madison, WI, USA). For cDNA synthesis 1 μg of poly(A)-enriched RNA was reverse transcribed (GeneAmpRNAPCR kit, Perkin Elmer Cetus, Norwalk, CT, USA). The cDNA synthesis was primed with a degenerated oligodeoxynucleotide of 36 bp length (5′-GGA GAA GGA T_{25} (A/G/C) (A/G/C/T)-3′). The product of the first-strand cDNA synthesis was directly subjected to DNA amplification using AmpliTaq DNA Polymerase (Perkin Elmer). From the known N-terminal peptide sequence [7] a degenerate sense oligonucleotide was designed for use in PCR: 5′-AA(C/T) ACN GTN TGG CCN GGN ATC-3′. Amplification products of 850 bp length were cloned into the pGEM®-T Easy vector (Promega) and sequenced using a DNA sequencer model 4000L (LICOR, NE, USA). To complete the cDNA, the fragment was extended by 5′-rapid amplification of cDNA ends using the Ampli-Finder™ kit (Clontech, Palo Alto, CA, USA).

Molecular modeling

The amino acid sequence alignments were performed using the program FUGUE,[36] by searching position-specific environmental substitution tables generated from the HOMSTRAD database of homologous structures.[37] The resulting alignment of the target sequence Mal d 2 to known structures was visually inspected using the program JOY[38] and manually edited using the SEAVIEW program.[39] A recursive approach was adopted comprising the sequence alignment and model building.[40] From the best alignment of template structures to target sequences, 15 3D models containing all non-hydrogen atoms were obtained automatically using the method implemented in MODELLER (Version 4.0).[41] Minimizations of the models were performed automatically by the program. The model corresponding to the lowest value of the objective function was used for further analysis. The cycle of realignment, modeling and structure validation was repeated until no further improvement of the structure was observed.

Construction of the 30B-Mal d 2 expression plasmid and inoculation of *N. benthamiana* plants with recombinant viral RNA

*Pac*I and *Avr*II restriction sites and an additional sequence coding for a C-terminal hexahistidyl affinity tag were added to the Mal d 2 coding sequence by PCR. The digested amplification product was ligated to respective sites of the 30B-based TMV vector[42] pBSG735A66 to create 30B-Mal d 2. pBSG735A66 was kindly provided by Large Scale Biology Corporation (Vacaville, California). The ligation product was used to transform competent *Escherichia coli* XL1-Blue cells (Clontech, Palo Alto, CA, USA).

N. benthamiana plants were inoculated with *in vitro* transcripts of the *Sse*8387 I-linearized plasmid 30B-Mal d 2 as described.[13] Two weeks after inoculation total RNA was isolated by grinding newly formed upper leaves from visibly infected plants and subjected to RT-PCR using the GeneAmp RNA PCR kit (Perkin Elmer Cetus) and TMV specific primers. Sequence analysis of the RT-PCR product was performed as described above.

Protein extraction of *N. benthamiana* expressing rMal d 2 and its purification

Upper leaves of infected *N. benthamiana* plants were harvested 21 days after inoculation, frozen in liquid nitrogen and ground in a fourfold excess of 10 mM potassium phosphate-buffer (pH 8.0) containing 10 mM diethyldithiocarbamate, 2% (w/v) polyvinylpolypyrrolidone, and 1 mM PMSF. The supernatants obtained by centrifugation for one hour at 40,000*g* and 4 °C were filtered and dialyzed against 10 mM potassium phosphate-buffer (pH 8.0). Protein concentrations were determined by the Bradford method according to the Bio-Rad-Microassay Procedure (Bio-Rad, Hercules, CA, USA).

For purification, the His-tagged rMal d 2 present in the protein extract of TMV-infected *N. benthamiana* plants was subjected to nickel chelate chromatography (nickel-nitrilo-tri-acetic (Ni-NTA) agarose, Qiagen). Elution of the Ni-NTA agarose-bound His-tagged recombinant protein was performed under native conditions by competition with increasing concentrations of imidazole following the vendor's instructions. The course of the purification of rMal d 2 from the *N. benthamiana* extract was monitored by SDS-PAGE on 12% (w/v) gels.

Molecular characterization of rMal d 2

Purified recombinant Mal d 2, migrating as a double band at approximately 31 kDa in 7% SDS gels, was electroblotted onto a polyvinylidene fluoride membrane (Millipore, Bedford, MA, USA), Coomassie-stained, excised and sequenced using a Procise 491 sequencer (Applied Biosystems, Foster City, California).

For the reduction of disulfide bonds and alkylation of thiol groups of rMal d 2, 80 μg (20 nmol) rMal d 2 containing 320 nmol cysteine residues were incubated with 7.5 μmol DTT followed by incubation with 15 μmol iodoacetamide and then 7.5 μmol DTT again as previously described.[43]

To produce tryptic peptides, solutions of purified rMal d 2 (30 μg in 50 μl of double distilled water) were incubated at 37 °C with 0.3 μg trypsin (sequencing grade, Boehringer Mannheim) after heat treatment at 95 °C. Digestion of Mal d 2 with endoproteinase Glu-C

(sequencing grade, Boehringer Mannheim) was performed for four hours at 37 °C using 0.6 μg of the protease.

MALDI-TOF mass spectrometry was performed on a Compact Maldi IV instrument (Kratos Analytical, Manchester, UK). Protein samples were dissolved in the matrix mixture (saturated solution of 3,5-dimethoxy-4-hydroxycinnamic acid in a 2:3 mixture of acetonitrile/0.1% aqueous trifluoroacetic acid) at a concentration of 1 μg/μl. An aliquot of 0.5 μl of the protein solution was applied to the sample target and dried in an air stream. Masses were determined in the linear flight mode with time-delayed extraction. Trypsin from bovine pancreas (Sigma) was used as internal standard. For the mass measurements of proteolytic fragments of Mal d 2, 0.5 μl of peptide solutions were mixed with 0.5 μl of matrix (10 mg/ml gentisic acid in 40% (v/v) acetonitrile/0.1% aqueous trifluoroacetic acid), applied to the sample target, and dried in an air stream. Masses were determined in the reflectron flight mode with time-delayed extraction.

SDS-PAGE and immunoblotting experiments

Proteins extracted from infected plants or from apple fruits were separated by 12% SDS-PAGE and visualized by staining with Coomassie brilliant blue R-250 (BioRad) or transferred to Protran nitrocellulose membranes (Schleicher & Schuell, Germany).

For IgE immunoblots with patients' sera, a total of 14 individual serum samples from patients with positive case histories, positive skin prick test, positive radioallergosorbent tests (RAST classes higher than 4) and characteristic type I allergic reactions to apple were used in this study. A serum pool from ten healthy atopic individuals with high IgE levels for house dust mite, but negative skin prick tests, and negative RAST results to apple was used as negative control. The IgE-binding ability of individual proteins was detected applying sera (diluted 1:5 in blocking buffer) from apple allergic patients. Bound IgE was detected using ^{125}I-labeled rabbit anti-human IgE antibodies (Pharmacia, Uppsala, Sweden) diluted 1:10 and visualized on Biomax™ MS film (Kodak, NY, USA). The negative control serum pool was tested in parallel. For inhibition studies with rMal d 2, a serum pool of the 14 apple-allergic patients was pre-incubated overnight at 4 °C with purified rMal d 2 (1 μg, 10 μg, 30 μg). The incubated serum pool was then used to probe nitrocellulose strips containing natural protein extract from apples. As control, the serum pool was either untreated or pre-incubated with BSA (20 μg/μl). Subsequently, the IgE immunoblot was continued as described above.

Immunodetections of the recombinant apple allergen were also performed with a His-tag specific monoclonal antibody in a 1:2000 dilution (Qiagen, Hilden, Germany). For the detection alkaline phosphatase-conjugated rabbit anti-mouse IgG (JacksonImmuno Lab. Inc., West Grove, PA, USA) diluted 1:1000 was used and developed with 5-bromo-4-chloro-3-indolylphosphate and nitroblue-tetrazolium chloride (Boehringer Mannheim, Germany).

Fungal growth inhibition assay

ATCC strains *A. alternata* 34509, *P. expansum* 359745, and *F. oxysporum* 4254 were used in this study. *In vitro* bioassays were done as follows: 20 μl of 3 × MPD (malt extract/peptone/dextrose, 3:1:1% final concentration) medium, 20 μl spore suspension of the test fungus (at approx. 10^6 spores/ml) and 20 μl of protein solution or elution buffer to give final concentrations of 0 μg/ml, 100 μg/ml, and 200 μg/ml rMal d 2 were mixed and incubated for 24 hours at 28 °C. On microscope slides the germination and growth of the first 100 spores randomly appearing were evaluated. Samples were prepared in triplicates, and experiments were repeated on different days. Given values are means of these single determinations.

Sequence data

The complete cDNA sequence and its deduced amino acid sequence is available from the EMBL database under the accession number AJ243427.

Acknowledgements

This work was supported by the Austrian Science Fund grants P12838-GEN and SFB F018-02. We gratefully thank Large Scale Biology Corporation (Vacaville, California) for kindly providing the TMV-based expression vector pBSG735A66.

References

1. Sampson, H. A. (1999). Food allergy. Part I: immunopathogenesis and clinical disorders. *J. Allergy Clin. Immunol.* **103**, 717–728.
2. Altman, D. R. & Chiaramonte, L. T. (1996). Public perception of food allergy. *J. Allergy Clin. Immunol.* **97**, 1247–1251.
3. Ebner, C., Hirschwehr, R., Bauer, L., Breiteneder, H., Valenta, R., Ebner, H. *et al.* (1995). Identification of allergens in fruits and vegetables: IgE cross-reactivities with the important birch pollen allergens Bet v 1 and Bet v 2 (birch profilin). *J. Allergy Clin. Immunol.* **95**, 962–969.
4. Fritsch, R., Bohle, B., Vollmann, U., Wiedermann, U., Jahn-Schmid, B., Krebitz, M. *et al.* (1998). Bet v 1, the major birch pollen allergen, and Mal d 1, the major apple allergen, cross-react at the level of allergen-specific T helper cells. *J. Allergy Clin. Immunol.* **102**, 679–686.
5. Oh, D. H., Song, K. J., Shin, Y. U. & Chung, W. I. (2000). Isolation of a cDNA encoding a 31-kDa, pathogenesis-related 5/thaumatin-like (PR5/TL) protein abundantly expressed in apple fruit (*Malus domestica* cv. Fuji). *Biosci. Biotechnol. Biochem.* **64**, 355–362.
6. Vanek-Krebitz, M., Hoffmann-Sommergruber, K., Laimer da Camara Machado, M., Susani, M., Ebner, C., Kraft, D. *et al.* (1995). Cloning and sequencing of Mal d 1, the major allergen from apple (*Malus domestica*), and its immunological relationship to Bet v 1, the major birch pollen allergen. *Biochem. Biophys. Res. Commun.* **214**, 538–551.
7. Hsieh, L. S., Moos, M., Jr & Lin, Y. (1995). Characterization of apple 18 and 31 kD allergens by microsequencing and evaluation of their content during storage and ripening. *J. Allergy Clin. Immunol.* **96**, 960–970.

8. Vigers, A. J., Wiedemann, S., Roberts, W. K., Legrand, M., Selitrennikoff, C. P. & Fritig, B. (1992). Thaumatin-like pathogenesis-related proteins are antifungal. *Plant Sci.* **83**, 155–161.
9. Van Loon, L. C. & van Strein, E. A. (1999). The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiol. Mol. Plant Pathol.* **55**, 85–97.
10. Koiwa, H., Kato, H., Nakatsu, T., Oda, J., Yamada, Y. & Sato, F. (1999). Crystal structure of tobacco PR-5d protein at 1.8 Å resolution reveals a conserved acidic cleft structure in antifungal thaumatin-like proteins. *J. Mol. Biol.* **286**, 1137–1145.
11. Soman, K. V., Midoro-Horiuti, T., Ferreon, J. C., Goldblum, R. M., Brooks, E. G., Kurosky, A. *et al.* (2000). Homology modeling and characterization of IgE binding epitopes of mountain cedar allergen Jun a 3. *Biophys. J.* **79**, 1601–1609.
12. Turpen, T. H., Reinl, S. J., Charoenvit, Y., Hoffman, S. L., Fallarme, V. & Grill, L. K. (1995). Malarial epitopes expressed on the surface of recombinant tobacco mosaic virus. *Biotechnology*, **13**, 53–57.
13. Krebitz, M., Wiedermann, U., Essl, D., Steinkellner, H., Wagner, B., Turpen, T. H. *et al.* (2000). Rapid production of the major birch pollen allergen Bet v 1 in *Nicotiana benthamiana* plants and its immunological *in vitro* and *in vivo* characterization. *FASEB J.* **14**, 1279–1288.
14. Inschlag, C., Hoffmann-Sommergruber, K., O'Riordain, G., Ahorn, H., Ebner, C., Scheiner, O. & Breiteneder, H. (1998). Biochemical characterization of Pru a 2, a 23-kD thaumatin-like protein representing a potential major allergen in cherry (*Prunus avium*). *Int. Arch. Allergy Immunol.* **116**, 22–28.
15. Fils-Lycaon, B. R., Wiersma, P. A., Eastwell, K. C. & Sautiere, P. (1996). A cherry protein and its gene, abundantly expressed in ripening fruit, have been identified as thaumatin-like. *Plant Physiol.* **111**, 269–273.
16. Midoro-Horiuti, T., Goldblum, R. M., Kurosky, A., Wood, T. G. & Brooks, E. G. (2000). Variable expression of pathogenesis-related protein allergen in mountain cedar (*Juniperus ashei*) pollen. *J. Immunol.* **164**, 2188–2192.
17. Batalia, M. A., Monzingo, A. F., Ernst, S., Roberts, W. & Robertus, J. D. (1996). The crystal structure of the antifungal protein zeamatin, a member of the thaumatin-like, PR-5 protein family. *Nature Struct. Biol.* **3**, 19–23.
18. Ogata, C. M., Gordon, P. F., de Vos, A. M. & Kim, S. H. (1992). Crystal structure of a sweet tasting protein thaumatin I, at 1.65 Å resolution. *J. Mol. Biol.* **228**, 893–908.
19. Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). Main-chain bond lengths and bond angles in protein structures. *J. Mol. Biol.* **231**, 1049–1067.
20. Keough, T. W., Sun, W., Barnett, B. L., Lacey, M. P., Bauer, M. D., Wang, E. S. & Erwin, C. R. (1996). Rapid analysis of single-cysteine variants of recombinant proteins. In *Methods in Molecular Biology: Protein and Peptide Analysis by Mass Spectrometry* (Chapman, J. R., ed.), vol. 61, pp. 171–183, Humana Press Inc, Totowa, NJ.
21. Lorito, M., Hayes, C. K., Di Pietro, A., Woo, S. L. & Harman, G. E. (1994). Purification, characterization, and synergistic activity of a glucan 1,3-β-glucosidase and an *N*-acetyl-β-glucosaminidase from *Trichoderma harzianum*. *Phytopathology*, **84**, 398–405.
22. Stintzi, A., Heitz, T., Prasad, V., Wiedemann-Merdinoglu, S., Kauffmann, S., Geoffroy, P. *et al.* (1993). Plant 'pathogenesis-related' proteins and their role in defense against pathogens. *Biochimie*, **75**, 687–706.
23. Hu, X. & Reddy, A. S. (1997). Cloning and expression of a PR5-like protein from *Arabidopsis*: inhibition of fungal growth by bacterially expressed protein. *Plant Mol. Biol.* **34**, 949–959.
24. Faus, I. (2000). Recent developments in the characterization and biotechnological production of sweet-tasting proteins. *Appl. Microbiol. Biotechnol.* **53**, 145–151.
25. Moralejo, F. J., Cardoza, R. E., Gutierrez, S., Sisniega, H., Faus, I. & Martin, J. F. (2000). Overexpression and lack of degradation of thaumatin in an aspergillopepsin A-defective mutant of *Aspergillus awamori* containing an insertion in the pepA gene. *Appl. Microbiol. Biotechnol.* **54**, 772–777.
26. Velazhahan, R., Zen, K. C. & Muthukrishnan, S. (2001). Expression of a rice thaumatin-like protein (PR-5) gene in a baculovirus expression system. *Acta Phytopathol. Entomol. Hung.* **36**, 311–316.
27. Breiteneder, H., Krebitz, M., Wiedermann, U., Wagner, B., Essl, D., Steinkellner, H. *et al.* (2001). Rapid production of recombinant allergens in *Nicotiana benthamiana* and their impact on diagnosis and therapy. *Int. Arch. Allergy Immunol.* **124**, 48–50.
28. van Ree, R. (2002). Carbohydrate epitopes and their relevance for the diagnosis and treatment of allergic diseases. *Int. Arch. Allergy Immunol.* **129**, 189–197.
29. Gavel, Y. & von Heijne, G. (1990). Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. *Protein Eng.* **3**, 433–442.
30. Vigers, A. J., Roberts, W. K. & Selitrennikoff, C. P. (1991). A new family of plant antifungal proteins. *Mol. Plant Microbe Interact.* **4**, 315–323.
31. Lack, G., Chapman, M., Kalsheker, N., King, V., Robinson, C. & Venables, K. (2002). Report on the potential allergenicity of genetically modified organisms and their products. *Clin. Exp. Allergy*, **32**, 1131–1143.
32. Liu, D., Raghothama, K. G., Hasegawa, P. M. & Bressan, R. A. (1994). Osmotin overexpression in potato delays development of disease symptoms. *Proc. Natl Acad. Sci. USA*, **91**, 1888–1892.
33. Chen, W. P. & Punja, Z. K. (2002). Transgenic herbicide- and disease-tolerant carrot (*Daucus carota L.*) plants obtained through Agrobacterium-mediated transformation. *Plant Cell. Rep.* **20**, 929–935.
34. Fuchs, H. C., Hoffmann-Sommergruber, K., Wagner, B., Krebitz, M., Scheiner, O. & Breiteneder, H. (2002). Heterologous expression in *Nicotiana benthamiana* of Cap a 1, a thaumatin-like protein and major allergen from bell pepper (*Capsicum annuum*). *J. Allergy Clin. Immunol.* **109**, S134.
35. Kingston, R. (1995). Method for plant RNA preparation. In *Current Protocols in Molecular Biology* (Chanda, V. B., ed.), Wiley, New York.
36. Shi, J., Blundell, T. L. & Mizuguchi, K. (2001). FUGUE: sequence–structure homology recognition using environment-specific substitution tables and structure-dependent gap. *J. Mol. Biol.* **310**, 243–257.
37. Mizuguchi, K., Deane, C. M., Blundell, T. L. & Overington, J. P. (1998). HOMSTRAD: a database of protein structure alignments for homologous families. *Protein Sci.* **7**, 2469–2471.
38. Mizuguchi, K., Deane, C. M., Blundell, T. L., Johnson, M. S. & Overington, J. P. (1998). JOY: protein

sequence–structure representation and analysis. *Bioinformatics*, **14**, 617–623.
39. Galtier, N., Gouy, M. & Gautier, C. (1996). SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci.* **12**, 543–548.
40. Burke, D. F., Deane, C. M., Nagarajaram, H. A., Campillo, N., Martin-Martinez, M., Mendes, J. *et al.* (1999). An iterative structure-assisted approach to sequence alignment and comparative modeling. *Proteins: Struct. Funct. Genet.* **37**, 55–60.
41. Sali, A. & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815.
42. Shivprasad, S., Pogue, G. P., Lewandowski, D. J., Hidalgo, J., Donson, J., Grill, L. K. & Dawson, W. O. (1999). Heterologous sequences greatly affect foreign gene expression in tobacco mosaic virus-based vectors. *Virology*, **255**, 312–323.
43. Himly, M., Jahn-Schmid, B., Dedic, A., Kelemen, P., Wopfner, N., Altmann, F. *et al.* (2003). Art v 1, the major allergen of mugwort pollen, is a modular glycoprotein with a defensin-like and a hydroxyproline-rich domain. *FASEB J.* **17**, 106–108.
44. Esnouf, R. M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph. Model.* **15**, 132–134. see also 112–133..
45. Merritt, E. A. & Murphy, M. E. P. (1994). Raster3D version 2.0—a program for photorealistic molecular graphics. *Acta Crystallog sect. D, Biol Crystallog sect. D*, **50**, 869–873.
46. Nicholls, A. & Honig, B. (1991). A rapid finite difference algorithm, utilizing successive over relaxation to solve the Poisson–Boltzman equation. *J. Comput. Chem.* **12**, 435–445.

Edited by N.-H. Chua

(Received 29 October 2002; received in revised form 20 March 2003; accepted 20 March 2003)