

Detection and Structural Characterization of Natural Ara h 7, the Third Peanut Allergen of the 2S Albumin Family

Hendrik Schmidt,^{||,†} Susanne Krause,^{||,‡} Christoph Gelhaus,[§] Arnd Petersen,^{*,‡}
Ottmar Janssen,^{||,†} and Wolf-Meinhard Becker^{||,‡}

Molecular Immunology, Institute for Immunology, Christian-Albrechts-University of Kiel, Building 17, Arnold-Heller-Strasse 3, D-24105 Kiel, Germany, Clinical and Molecular Allergology, Research Center Borstel, Parkallee 22, D-23845 Borstel, Germany, Department of Zoophysiology, Zoological Institute, Christian-Albrechts-University of Kiel, Olshausenstrasse 40, D-24098 Kiel, Germany

Received March 16, 2010

In recent years, several novel relevant peanut allergens have been identified. Among those, a new member of the conglutinin family was cloned by a phage display approach and initially annotated as Ara h 7.0101. Later, however, recloning of Ara h 7 revealed an alternate isoform, termed Ara h 7.0201. Because the natural Ara h 7 counterpart had not been found at the protein level in peanut extracts, the aim of the present study was to search for authentic natural Ara h 7 protein(s). To this end, enriched low molecular mass proteins (<20 kDa) from peanut extracts were separated by 2D electrophoresis and subjected to mass spectrometric analyses. Fifty of 65 analyzed spots were identified. Interestingly, Ara h 7.0101 was not identified, but Ara h 7.0201 and Ara h 7.0202, a different Ara h 7 isoallergen containing an additional pro-peptide cleavage site, were. In accordance with the conserved cysteine pattern of conglutinins, Ara h 7.0201 possesses eight cysteine residues, in contrast to the six cysteines present in the previously cloned Ara h 7.0101. Moreover, a putative cleavage site in the Ara h 7.0202 isoform points to the characteristic biological function of conglutinins as amylase/trypsin inhibitors.

Keywords: Peanut allergens • Ara h 7 • *Arachis hypogaea* • phage display • 2D electrophoresis • N-terminal sequencing • mass spectrometry

Introduction

Several approaches have been used to identify and characterize allergens. Normally, allergens are identified by patients' IgE reactivities and characterized by Western blotting, N-terminal sequencing, and additional methods including the use of antibodies as allergen-specific probes. Some of these tools can be used to screen a cDNA expression library of the respective allergen source to clone the allergen of interest. In the case of homologous or similar allergens, degenerated or specific primers can be deduced from conserved sequence regions to fish homologous DNA sequences from genomic DNA or cDNA. The cloning of Ara h 8 and Ara h 9 represents two examples for this approach.^{1,2} Another effective and straightforward cloning strategy is the phage display technology, which has been applied for the cloning of several peanut allergens.³ With this method, four novel peanut allergens, Ara h 4, Ara h 5, Ara h 6, and Ara h 7, were identified and characterized at that time. One advantage of the phage display technology is

the cloning of allergens that are represented with low DNA copy numbers in the allergenic source.⁴

Still, if an allergen is identified at the RNA level, it remains elusive whether the translated protein is properly represented in the natural allergenic material. Although recombinant allergens proved to be suitable for allergy diagnosis or therapy, there is still a scientific need to identify and isolate the respective natural and therefore authentic counterpart. Moreover, the often time-consuming verification of equal molecular and immunological properties of natural allergens and their recombinant counterparts is a prerequisite before the introduction of recombinant material as "standards" for diagnosis and therapy. Nearly three years elapsed from the description of recombinant Ara h 8 (rAra h 8) before the natural Ara h 8 (nAra h 8) was biochemically identified in peanut extracts.⁵ Even more puzzling, eight years after the description and cloning of rAra h 7, the natural Ara h 7 has not been identified. One hindrance to the detection of nAra h 7 is that several peanut allergens with similar molecular masses accumulate in the range where natural Ara h 7 is expected (Table 1), and no Ara h 7-specific sera or monoclonal antibody are available to detect the natural Ara h 7 exclusively.

A reason for that might be the potential cross-reactivity between the three 2S peanut allergens Ara h 2, Ara h 6, and Ara h 7. One way to overcome this problem is the use of multiantigenic peptides (MAP),⁶ representing specific peptides

* Corresponding author (telephone +49-4537188497; fax +49-4537188686; e-mail apetersen@fz-borstel.de).

[†] Molecular Immunology, Institute for Immunology, Christian-Albrechts-University of Kiel.

[‡] Clinical and Molecular Allergology, Research Center Borstel.

[§] Department of Zoophysiology, Zoological Institute, Christian-Albrechts-University of Kiel.

^{||} Equally contributing first or last authors.

Table 1. Overview of Peanut Allergens <20 kDa

allergen	NCBI accession no.	MW (Da)	pI	% identity to Ara h 2
Ara h 2.0201	gil26245446	17993	5.51	
Ara h 2.0101	gil224747150	16637	5.50	100
Ara h 6	gil57118278	14719	5.24	53
Ara h 7.0101	gil5931948	16322	5.56	42
Ara h 7.0201	gil158121995	17374	7.49	44
Ara h 7.0201 fragment	aa14-aa144	15847	6.19	
Ara h 7.0202 fragment	aa14-aa144	15839	6.19	
2S protein 2	gil52001231	16491	6.14	
2S protein 2 fragment	aa14-aa138	14929	5.47	
Ara h 8.0201	gil145904609	16281	5.07	
Ara h-oleosin	gil47156059	18448	9.57	
Ara h 10.0101	gil122246173	17753	9.61	
Ara h 10.0201	gil75114091	15527	9.36	
Ara h 11.0101	gil122218541	14308	10.08	

of Ara h 7, for immunization, or the purification of an antiserum against rAra h 7 by affinity chromatography using immobilized specific Ara h 7 peptides. Such a peptide could be derived from the N terminus of the mature Ara h 7 covering the first 16 amino acids (aa) (TRWDPDRGRSGSRWDA).

Of note, in contrast to the characteristic cysteine pattern of conglutins with at least eight cysteines (8 in Ara h 2 and even 10 in Ara h 6),^{7,8} the cloned rAra h 7 contains only six cysteines. This structural divergence was ascribed to the C-terminal part of Ara h 7.

In the present study, we therefore recloned the peanut allergen Ara h 7, and we report the identification of two separate isoforms of the allergen. To identify the authentic natural Ara h 7 counterpart in peanut extracts, we enriched proteins below 20 kDa by size exclusion chromatography (SEC) from peanut extracts and subsequently separated this fraction by high-resolution 2D electrophoresis. Candidate protein spots were isolated and characterized by mass spectrometry. With Ara h 7.0201 and Ara h 7.0202, we characterized two allergen isoforms that fit into the conserved cysteine scheme of conglutins. Interestingly, in peanut extracts we were not able to detect a nAra h 7 corresponding to the cloned rAra h 7.0101.

Materials and Methods

Characterization of the Ara h 7 Gene Sequence from mRNA. Total RNA was isolated from freshly prepared meal of frozen germinable, that is, vital, peanuts. Thereafter, 30 mg of peanut meal was lysed using the SV Total RNA Isolation System (Promega, Mannheim, Germany), and RNA was prepared following the manufacturer's instructions. For the generation of cDNA, a First Strand cDNA Synthesis Kit (GE Healthcare, Munich, Germany) was used with a gene-specific 5' oligonucleotide Ara h 7.f (5'-AAGCTTATGATGGTCAAGCTCAGCATC-3'), binding to the start region of Ara h 7.0101, and the *NotI* d(T)₁₈-Primer (GE Healthcare) to identify the 3' gene sequence.

Expression of Ara h 7. The signal peptide of Ara h 7 spanning 20 amino acids was estimated by SignalP 3.0 software (www.cbs.dtu.dk/services/SignalP; provided by Center for Biological Sequence Analysis, Technical University of Denmark, Lyngby, Denmark).⁹ The sequence comprising the mature protein was optimized for expression in *Escherichia coli* and ligated as a His-tagged fusion protein into the *EcoRI*/*EcoRV* restriction sites of pET32b vector (Novagen, Madison, WI). Gene synthesis was performed at Geneart (Regensburg, Germany). The transformed BL21(DE3) cells (Stratagene, La Jolla, CA) were grown at 37 °C

in LB medium supplemented with 50 µg/mL ampicillin to a cell density of OD₆₀₀ = 0.5–0.7. Isopropyl-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. The incubation was continued for 4 h at 37 °C; afterward, the cells were harvested by centrifugation and stored at –70 °C.

Purification of Recombinant Ara h 7. To purify rAra h 7, the cell pellet of a 500 mL culture was resuspended in 30 mL of binding buffer (50 mM NaH₂PO₄, 300 mM sodium chloride, pH 8.0) supplemented with 3 mL of BugBuster (Novagen) and incubated under the addition of benzonase for 30 min on a roller at room temperature. Insoluble material was removed by centrifugation at 10000g for 30 min. The supernatant was bound batchwise to 5 mL Talon Metal Affinity resin (Clontech-Takara, Saint-Germain-en-Laye, France). Afterward, a column was packed and equilibrated with 20 mL of binding buffer and 40 mL of washing buffer (50 mM NaH₂PO₄, 300 mM sodium chloride, 50 mM imidazole, pH 8.0). The His-tagged protein was eluted with 10 mL of elution buffer (50 mM NaH₂PO₄, 300 mM sodium chloride, 1 M imidazole, pH 8.0). The different fractions were analyzed by 12% SDS-PAGE using Coomassie staining, and respective eluates were pooled. The protein concentration was determined after extensive dialysis against ddH₂O according to the Bradford method using a Coomassie protein reagent (Pierce, Rockford, IL) with bovine serum albumin as standard.¹⁰ The purified protein was lyophilized and stored at –20 °C.

Peanut Extract. Peanut extract was prepared as described elsewhere.¹¹ Briefly, 30 g of peanut flour was extracted overnight in 250 mL of 0.1 M NH₄HCO₃, pH 8.0. The extract was dialyzed against ddH₂O overnight and lyophilized.

Size Exclusion Chromatography (SEC). Peanut extract was separated by SEC using a calibrated Superdex 200 column, which was run on a HPLC system (ÄKTApurifier, Amersham Biosciences, Germany). Ten milligrams of protein per run was applied onto a prepacked 10/300 GL column (length = 34 cm, width = 1.5 cm) in a buffer containing 50 mM K₂HPO₄ and 0.5 M NaCl, pH 7.5. Fractions with proteins under 20 kDa were dialyzed against ddH₂O and lyophilized.

Sera and Antibodies. Patients allergic to peanuts were selected according to a convincing case history, skin test reactivity to peanut extract, and/or positive tests by CAP-FEIA. Sera were collected at the Research Center Borstel (Borstel, Germany), after the patients gave informed consent. The patients' sera were specified elsewhere.^{3,4,12}

Rabbit anti-Ara h 7 was raised against rAra h 7.0101 and affinity-purified by immobilized N-terminal peptide TRWDPDRGRSGSRWDA (kindly provided by S. Bade, Research Center Borstel, Borstel, Germany).

Electrophoresis and Immunoblotting. Discontinuous SDS-PAGE¹³ was performed for separation of low molecular weight proteins (<25 kDa) using NuPAGE Bis-Tris gels (Invitrogen, Groningen, The Netherlands) according to the manufacturer's instructions. After electrophoresis, proteins were either stained with Coomassie blue or, for Western blot analysis, transferred to polyvinylidene difluoride (PVDF) membranes (Millipore, Eschborn, Germany) by semidry blotting¹⁴ at 0.8 mA/cm² for 30 min. Blocking of PVDF membranes was achieved by shaking in 5% nonfat dry milk/TBS (Tris-buffered saline; 100 mM Tris/HCl, 100 mM NaCl, 2.5 mM MgCl₂, pH 7.4) under gentle agitation for 2 h. Subsequently, membranes were incubated in an appropriate dilution of primary antibody in TBST (TBS + Tween-20, 0.05%), human sera, 1:10, or specific rabbit serum, 1:1000. Specific binding of human IgE or rabbit antibodies was

Table 2. Specifications of Performed 2D Experiments^a

expt	material	amount (μg)	1D	2D (%)
1	<20 kDa EEx	75	3–11NL	16
2	<20 kDa EEx	100	3–11NL	16
3	EEx	80	3–11NL	12.5

^aEEx, peanut extract; separation conditions pH 3–11 in the first dimension and 16 or 12.5% polyacrylamide gels in the second dimension.

detected using either a mouse anti-human-IgE antibody (1:2000; Allergopharma, Reinbek, Germany) or a goat anti-rabbit-IgG antibody (1:5000; Dianova, Hamburg, Germany) conjugated with alkaline phosphatase (AP). Bound antibodies were visualized after three washes with TBST by incubation with NBT/BCIP as substrate in AP buffer (0.1 M TBS, 0.1 M NaCl, 5 mM MgCl₂, pH 9.5).¹⁵

2D Electrophoresis for N-Terminal Sequencing. For the first dimension, proteins were separated by isoelectric focusing (IEF) on a precast polyacrylamide gel strip with an immobilized pH gradient (IPG) of 3–10 (Invitrogen). The IPG strip was rehydrated in sample buffer [8 M urea, 2% 3-[3-cholamidopropyl]-dimethylammonio-1-propanesulfonate (CHAPS), 0.5% ampholytes (pH 3–10), 0.002% bromphenol blue, and 20 mM DTT] containing 100 μg of protein sample. The IEF conditions for the Zoom IPG Runner system (Invitrogen) were set as follows: 200 V for 20 min, 450 V for 15 min, 750 V for 15 min, and 1 h at 2000 V and 1 mA. The determination of the pI and the molecular masses of the proteins was done by using an IEF marker pH 3–10 (Serva, Heidelberg, Germany) and Multi Mark prestained standard (Invitrogen). After 15 min of equilibration in LDS sample buffer containing 0.5 M DTT or 125 mM iodoacetamide, the strip was transferred to the well. For the second dimension 4–20% Tris–Glycine–Zoom gels (NuPage) were used with 2-(N-morpholino)ethanesulfonic acid-containing running buffer (MES) (0.05 M MES, 0.05 M Tris base, 3.47 mM SDS, 20.5 mM EDTA, pH 7.3). The gel run was performed according to the manufacturer's instruction for the Zoom IPG Runner System at 50–100 V.

Proteins separated by 2D gels were transferred onto methanol-activated PVDF membranes (Schleicher & Schuell, Dassel, Germany) by semidry electroblotting with transfer buffer (25 mM bicine, 25 mM Bis-Tris, 1.025 mM EDTA, 0.05 mM chlorobutanol, and 20% methanol) according to the manufacturer's instructions (Invitrogen). They were visualized by Coomassie staining for 10 min (0.06 mM Coomassie R-250 in 50% methanol) and destaining for 7 min in destaining solution (50% methanol).¹⁶ Protein spots or bands were cut from the PVDF membranes and sequences on a Procise sequencer (Applied Biosystems, Foster City, CA).

High-Resolution 2D Electrophoresis. Lyophilized total peanut kernel protein or SEC-fractionated extracts were dissolved in 2D-lysis buffer [7 M urea, 2 M thiourea, 4% CHAPS, 1% (v/v) IPG buffer pH 3–11NL and 1% (w/v) DTT] on ice. The experimental design for different technical/biological replicates in three independent experiments, including different loading amounts and protein prefractionation, is detailed in Table 2. Between 75 and 100 μg of protein was applied by cup loading on 18 cm IPG strips (pH 3–11NL) rehydrated at room temperature for a minimum of 10 h in 340 μL of DeStreak solution supplemented with 0.5% (v/v) IPG buffer (GE Healthcare). Proteins were focused using an IPGphor 3 electrophoresis unit (GE Healthcare) at 20 °C to reach 53.3 kVh (150 V for 3 h, 300 V for 3 h, gradient to 1000 V over 6 h, gradient to 10000 V over

3 h, and finally 10000 V until total kVh was reached). Reduction and alkylation in equilibration buffer were performed as described elsewhere.¹⁷ For the second dimension, protein separation was performed on 12.5 or 16% polyacrylamide gels using a Protean II cell (Bio-Rad, Munich, Germany). For spot picking, the gels were stained with FlamingoPink (Bio-Rad) following the manufacturer's instructions. Image Master 6.0 software (GE Healthcare) was used for spot selection, and selected spots were picked automatically on an Ettan Spot-picker (GE Healthcare) with a 2 mm picking head. Picked gels were rescanned to verify the locations of the cut spots.

Proteolytic Cleavage and Mass Spectrometry. Proteolytic cleavage was performed as described elsewhere.¹⁸ Briefly, picked gel pieces were washed twice with HPLC grade water (Roth, Karlsruhe, Germany), dehydrated with 25 mM ammonium bicarbonate (ABC) in 50% acetonitrile (ACN), and finally shrunk in pure ACN. After rehydration for 20 min, with 100 ng of porcine sequencing-grade trypsin (Serva) in 10 μL of 5 mM ABC, another 10 μL of 5 mM ABC was added. Peptides were extracted after an incubation period of 14 h at 37 °C. Twenty microliters of 0.3% trifluoroacetic acid (TFA) in ACN was added, and the samples were sonicated in a water bath (Sonorex Super RK100; Bandelin, Berlin, Germany) for 15 min. Liquid phases were collected, lyophilized, redissolved in 0.8 μL of MALDI matrix solution [3.2 mg/mL α-cyanohydroxycinnamic acid (LaserBio Laboratories, Sophia-Antipolis Cedex, France) in 65% ACN/0.1% TFA], spotted onto a 192-well stainless steel MALDI plate, and air-dried.

Mass spectrometry was performed on a 4700 Proteomics analyzer MALDI-TOF/TOF instrument (Applied Biosystems, Framingham, MA) using 4000 Series Explorer Software V3.6. Protein identification was performed by combining MS and MS/MS spectral data for database searches. MS spectra were acquired in positive reflector mode by accumulation of 1000–2000 laser shots. For peak picking, a minimum signal-to-noise ratio of 30 was chosen. In addition, precursor ions were fragmented by collision-induced dissociation. Precursor ions were selected automatically according to the following criteria: a maximum of five masses per spot and a signal-to-noise ratio of >35. Peak intensity was used to select the most intense peaks excluding peaks from trypsin autolysis products with a 0.2 m/z precursor mass tolerance (see below). For MS and MS/MS measurements, the instrument default calibration was updated before the run. Internal calibration was applied in MS mode, and the porcine trypsin autolysis products were used at 25 ppm mass tolerance. In MS/MS mode, air was used as collision gas, and the collision energy was set to 1 kV. 8000 laser shots were accumulated per MS/MS spectrum. For database searches expected tryptic autolysis peaks were excluded from the selection with 50 ppm precursor mass tolerance.¹⁹ A signal-to-noise ratio of >10 was selected for peak picking. Combined MS and MS/MS spectral data were processed by GPS Explorer 3.6 software (Applied Biosystems) and submitted to the MASCOT server (version 2.0, Matrix Science, London, U.K.) for search against an in-house database (NCBI nr, date 26/07/2009) including all proteins from *Arachis hypogaea* found in the NCBI database (with 875 protein entries) and the DNA-based peptide predictions described here. MASCOT was used with the following parameters: Trypsin's enzymatic specificity was considered, autolysis products of trypsin were excluded from the searches, modifications on cysteine residues by carboxyamidomethylation were set as obligate, potential

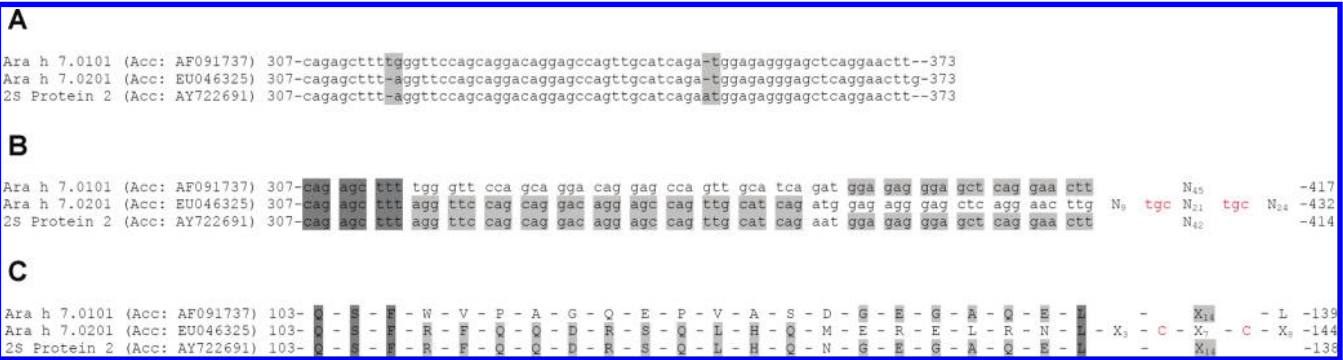


Figure 1. Alignment of the different Ara h 7 isoforms at the nucleotide (A) and protein level (B, C). Panel A demonstrates the high accordance between these allergens in the 3' region at the nucleotide sequence and shows the insertion of a single nucleotide in two isoforms, causing a reading frame shift, which is responsible for the extreme differences in the C termini of the primary structures of the expressed proteins. Panels B and C demonstrate the translation of the codons (B); the aa number corresponds to the N termini (T₁RW..) of the mature proteins.

modification of methionine was by oxidation, and a maximum of one missed tryptic cleavage site in a fragment was allowed.

Data Interpretation Using ProteinProspector. In addition to MASCOT, MS data interpretation was performed for selected protein spots using MS-Bridge in ProteinProspector 5.3.2 (<http://prospector.ucsf.edu>).²⁰ Parameters were set as follows: cysteine modification by carboxyamidomethylation was set as constant, and oxidation on methionine and changes of glutamine on the peptides' N termini to pyroglutamic acid were set as variable. Masses of matrix clusters and methylated porcine trypsin autolysis fragments were excluded (see Table S1 of the Supporting Information).¹⁹ The mass range was set from 500 to 4000 with a mass tolerance of 50 ppm and a minimum peptide length of 5 amino acids (data set S2 in the Supporting Information).

Results

Identification of a New Ara h 7 Isoform. The published Ara h 7 (rAra h 7.0101, gi|5931948) displays a structural divergence to the conserved cysteine pattern of conglutins, which is situated in the C terminus. To explore whether it is possible to get a structure that is in full accordance with conglutin characteristics, the 3' end of the annotated Ara h 7 gene sequence was analyzed at the mRNA level by a PCR employing a gene-specific forward primer (Ara h 7.f) and an oligonucleotide binding to the poly(A) tail. The analysis of the obtained sequence revealed 16 nucleotide substitutions and one deletion in the Ara h 7 gene in comparison to the published sequence³ (Figure 1A). The consequence of this deletion is a shift in the reading frame, which generates a protein with a different C terminus starting at aa117 (in the numbering system used throughout this paper the signal peptide is not counted). Importantly, the deduced primary structure of the new Ara h 7 isoform (Ara h 7.0201) is in accordance with the cysteine pattern of conglutins (Figure 1B).

The IgE reactivity of the expressed and purified protein was tested with patients' sera (Figure 2). Five of the six patients' sera, which were preselected by Ara h 7.0101 reactivity, showed positive IgE binding with the novel Ara h 7. Consequently, the new isoform was classified by the I.U.I.S. allergen nomenclature subcommittee as Ara h 7.0201. In comparison with Ara h 7.0101, the new isoform possesses a higher molecular weight and a more alkaline pI [with reference to the mature proteins (without signal peptides): Ara h 7.0101, 16.3 kDa, pI, 5.56; Ara h 7.0201, 17.3 kDa, pI, 7.5].

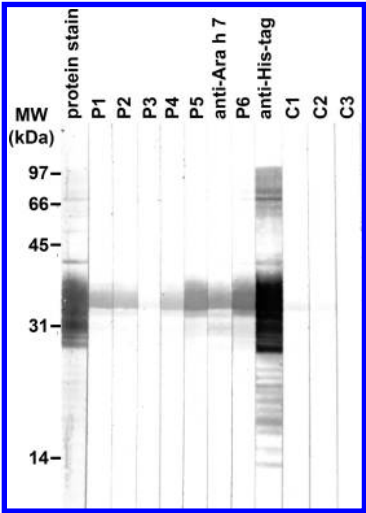


Figure 2. Western blot analysis of rAra h 7.0201 (pET32b-Ara h 7.02) to demonstrate its IgE reactivity. Equal amounts of rAra h 7.0201 were transferred onto a membrane and incubated either with patients' sera (P1–P6) or with control sera (C1–C3). Additionally, recombinant protein was detected by an N terminus specific rabbit antiserum and anti-His-tag antibodies. Protein stain was performed using India Ink.

Identification of Ara h 7 Isoforms in Peanut Extracts. In a next step, we tried to identify natural Ara h 7 protein in peanut extracts. Therefore, on the basis of the calculated molecular weight, proteins below 20 kDa from peanut extracts were enriched by SEC. Seventy-five micrograms of protein of this fractionation was separated by 2D electrophoresis (Table 2, expt 1). FlamingoPink stained the separated proteins, and individual spots were picked and analyzed by mass spectrometry (Figure 3). Fifty of 65 detected spots were identified on the basis of respective database analyses (Table 3). Notably, several of those spots represented different isoforms of the Ara h 7 related conglutins Ara h 2 or Ara h 6. Referring to the published (Ara h 7.0101) and the new Ara h 7.0201 sequences in an in-house database, Ara h 7.0201 but not Ara h 7.0101 was identified in four different spots (17–20) at an estimated molecular weight of 17 kDa and a pI ranging from 5.8 to 7.0. Besides peptide mass fingerprinting, protein fragments representing the peptides NLPQNCGFR and ANLRPCEEHIR were identified by MS/MS analysis with two prominent monoisotopic masses of 1105.5 and 1394.7, respectively. The assignment of additional masses

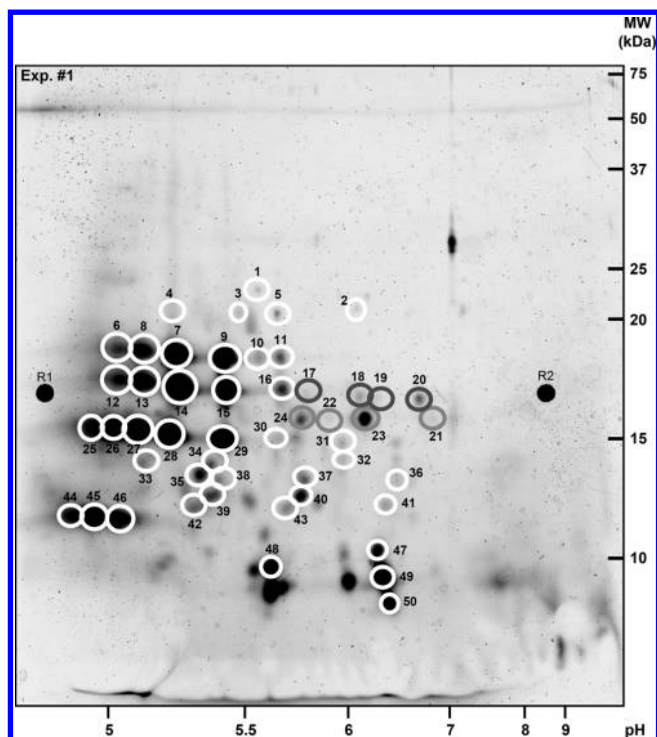


Figure 3. Proteome map of prefractionated peanut extracts below 20 kDa. Proteins were first separated on pH 3–11 NL IPG strips, followed by second-dimension separation on 16% Bis-Tris Laemmli gels. Identified proteins and individual protein isoforms are circled and annotated with a spot number corresponding to numbers given in Table 3. Dark and light gray circles mark isoforms of allergens Ara h 7.0201 and Ara h 7.0202, respectively. R1 and R2 are reference markers for spot picking. One representative experiment of three is shown. For replicate experiments see the Supporting Information, Figure S1.

led to a sequence coverage of up to 60% (86/144 aa, without signal sequence) of mature Ara h 7.0201 in spot 18 (Table 4).

Moreover, spots 22–24 were identified as 2S protein 2 of *A. hypogaea* without an apparent classification as peanut allergen, that is, Ara h 7.0101. The mature 2S protein 2, however, displays a identity of 87% to the mature Ara h 7.0101 and of 84% to Ara h 7.0201 and starts with the N terminus TRW, which is also characteristic for Ara h 7 (Supporting Information Figure S1). Further analyses revealed no sequence coverage in the C terminus of 2S protein 2 starting at aa117. In addition, assignment of the peptides to the sequence of Ara h 7.0201 resulted in a reduced coverage in the N-terminal part of the protein. The sequence coverage of spot 23 was 64% for 2S protein 2 and 61% for Ara h 7.0201 on the basis of the identified peptide masses (see Table 4 and the Supporting Information Figure S2). However, by reconstructing a hypothetical protein isoform, termed Ara h 7.0202, consisting of the N-terminal part of 2S protein 2 (TRW-aa116) and the C-terminal part of Ara h 7.0201 (aa117–144), this construct showed the highest sequence coverage (83% coverage of 23). Because peptides within the N-terminal as well as the C-terminal part of this hypothetical protein were identified by MS/MS, we concluded that Ara h 7.0202 was composed as indicated (Figure 4C; Figures S4B and S5 of the Supporting Information). Similarly, increased sequence coverage for Ara h 7.0202 was

also achieved by theoretical comparison for spots 21–24 (Table 4). It should be stressed that the main differences in the N termini were found in peptides VEKEQEQQDEY-PYIQR (aa44–60 for Ara h 7.0201, only spot 360 of expt 3; Figure S3 of the Supporting Information) and VEQEQQDEY-PYSQR (for 2S protein 2 and Ara h 7.0202) and GQRPGES-DEDQEQR (aa64–77 for Ara h 7.0201) and GRRPGES-EDQEQR (for the 2S protein 2 and Ara h 7.0202), respectively. The C termini of Ara h 7.0201 and Ara h 7.0202 are characterized by the peptide NLPQNCGFR (aa123–131).

As detailed in Supporting Information data set S2, the individual masses were assigned to peptides of the hypothetical Ara h 7.0202 using ProteinProspector 5.3.2.²⁰ For a better overview, the identified peptides of spots 17–24 were aligned to the sequences of Ara h 7.0101 (gil5931948), Ara h 7.0201 (gil158121995), 2S protein 2 (gil52001231), and the aforementioned hypothetical Ara h 7.0202. Additional experiments verified our results: We analyzed 100 μ g of SEC-isolated extract proteins below 20 kDa, focusing on an area between Ara h 2 (16.6 kDa) and Ara h 6 (14.7 kDa) in the pI range of 6–8. Here, spot 160 (equivalent to spot 23 of expt 1, Figure S5 of the Supporting Information) was annotated as Ara h 7.0202 with a sequence coverage of 49% also including the aforementioned isoform-specific peptides. In contrast, spot 166 (corresponding to spot 20 in expt 1) was identified as Ara h 7.0201 (Figure S3 of the Supporting Information).

Knowing the exact position of the novel Ara h 7 isoforms, we succeeded in the first identification of natural Ara h 7 in whole peanut extracts (Supporting Information, Figure S1, expt 3). Again, one spot (365) was identified as Ara h 7.0202 with a coverage of 52% (and thus corresponded to spot 23). In addition, spot 360, which represents Ara h 7.0201, corresponded to spot 20 (Figures S1 and S3 of the Supporting Information).

As far as the first isoform Ara h 7.0201 is concerned, we identified peptides covering aa112–122 and/or aa123–131 in spots 17–20 (Table 3). The peptides proved the significant differences in the C termini between structures of Ara h 7.0101 and Ara h 7.0201 and the 2S protein 2. In essence, Ara h 7.0201 contains two additional cysteines at its C terminus (by mass spectrometry the location of one of these cysteines was confirmed).

Identification of the N Terminus. The MALDI-based MS/MS analyses did not reveal the N termini of Ara h 7.0201 or Ara h 7.0202. Identified peptides covered fragments from aa20 downstream of the predicted start. To identify the individual N termini, Edman sequencing was performed after 2D electrophoresis of low molecular mass (<20 kDa) peanut proteins on minigels and transferred to PVDF membrane (Figure S6 of the Supporting Information). Putative Ara h 7 containing spots were chosen according to their two-dimensional location, excised, and subjected to N-terminal sequencing. The N terminus of nAra h 7 was determined as WDApS (with “p” indicating a hydroxyproline residue), which is 14 aa downstream of the theoretical start of the (mature) Ara h 7, as predicted by the SignalP 3.0 software for the cloned Ara h 7.0201.

Discussion

Phage display technology proved to be a powerful tool to identify and clone the “allergome” of an allergenic source in combination with patients’ IgE even if the cDNA or the IgE-reactive entity is rare. However, although Ara h 7.0101 was

Table 3. Identified Proteins of <20 kDa Fraction of Peanut Extract

spot	protein	NCBI accession no.	isoallergen	MW	pI	identified peptides	MASCOT score	sequence coverage ^a (%)	matched/unmatched queries
1	glycinin	gil47933675	Ara h 3	58368	5.4	2	88	4	3/30
2	arachin Ahy-1	gil52001219	Ara h 3/4	61803	5.5	6	113	15	9/33
3	allergen II	gil15418705	Ara h 2	18612	5.7	4	133	30	7/28
4	seed storage protein SSP2	gil46560482	Ara h 2	20907	6.0	7	144	36	12/37
5	glycinin	gil47933675	Ara h 3/4	58368	5.4	5	183	14	7/38
6	allergen II	gil15418705	Ara h 2	18612	5.7	9	250	54	14/34
7	allergen II (Ara h 2.01)	gil15418705	Ara h 2	18612	5.7	6	112	37	10/37
8	allergen II	gil15418705	Ara h 2	18612	5.7	11	329	62	17/45
9	allergen II (Ara h 2.01)	gil15418705	Ara h 2	18612	5.7	11	371	60	17/43
10	seed storage protein SSP2	gil46560482	Ara h 2	20907	6.0	6	101	30	8/34
11	seed storage protein SSP2	gil46560482	Ara h 2	20907	6.0	8	198	37	11/34
12	seed storage protein	gil28912768	Ara h 2	19765	5.7	10	284	57	16/43
13	seed storage protein	gil28912768	Ara h 2	19765	5.7	10	329	57	15/42
14	seed storage protein	gil28912768	Ara h 2	19765	5.7	11	290	57	17/54
15	seed storage protein	gil28912768	Ara h 2	19765	5.7	5	183	35	10/59
16	seed storage protein	gil28912768	Ara h 2	19765	5.7	10	210	57	15/46
17	Ara h 7 allergen Ara h 7.0201	gil158121995	Ara h 7	19783	8.2	5	161	31	7/21
18	Ara h 7 allergen Ara h 7.0201	gil158121995	Ara h 7	19783	8.2	12	243	52	14/41
19	Ara h 7 allergen Ara h 7.0201	gil158121995	Ara h 7	19783	8.2	4	77	24	6/26
20	Ara h 7 allergen Ara h 7.0201	gil158121995	Ara h 7	19783	8.2	11	235	42	14/42
21	Ara h 7 allergen (Ara h 7.0202)	gil158121995	Ara h 7	19783	8.2	2	48	12	4/27
22	2S protein 2 (Ara h 7.0202)	gil52001231	Ara h 7	18786	6.7	7	128	46	9/28
23	2S protein 2 (Ara h 7.0202)	gil52001231	Ara h 7	18786	6.7	11	326	50	14/65
24	2S protein 2 (Ara h 7.0202)	gil52001231	Ara h 7	18786	6.7	9	309	46	12/40
25	conglutin	gil17225991	Ara h 6	17349	5.7	11	208	68	22/65
26	conglutin	gil17225991	Ara h 6	17349	5.7	12	238	68	18/58
27	conglutin	gil17225991	Ara h 6	17349	5.7	12	294	80	21/57
28	conglutin	gil17225991	Ara h 6	17349	5.7	8	406	55	19/62
29	conglutin	gil75114094	Ara h 6	17479	6.1	11	350	73	23/44
30	conglutin	gil17225991	Ara h 6	17349	5.7	6	100	38	9/38
31	conarachin-fragment	gil46560478	Ara h 1	66934	6.7	4	52	6	5/27
32	conarachin-fragment	gil46560478	Ara h 1	66934	6.7	4	45	5	5/32
33	seed storage protein SSP2	gil46560482	Ara h 2	20907	6.0	4	43	19	6/25
34	seed storage protein SSP2	gil46560482	Ara h 2	20907	6.0	6	113	36	9/33
35	seed storage protein SSP2	gil46560482	Ara h 2	20907	6.0	5	188	25	8/28
36	seed storage protein SSP2	gil46560482	Ara h 2	20907	6.0	4	72	19	7/35
37	seed storage protein SSP2	gil46560482	Ara h 2	20907	6.0	5	86	25	8/36
38	seed storage protein SSP2	gil46560482	Ara h 2	20907	6.0	4	61	19	6/22
39	seed storage protein SSP2	gil46560482	Ara h 2	20907	6.0	5	134	25	8/25
40	seed storage protein SSP2	gil46560482	Ara h 2	20907	6.0	6	202	25	9/31
41	seed storage protein SSP2	gil46560482	Ara h 2	20907	6.0	4	58	19	7/39
42	seed storage protein SSP2	gil46560482	Ara h 2	20907	6.0	4	45	19	6/23
43	seed storage protein SSP2	gil46560482	Ara h 2	20907	6.0	4	49	19	6/29
44	conglutin	gil17225991	Ara h 6	17349	5.7	5	61	30	8/19
45	conglutin	gil17225991	Ara h 6	17349	5.7	8	129	37	15/47
46	conglutin	gil75114094	Ara h 6	17479	6.1	6	152	44	14/29
47	conglutin	gil17225991	Ara h 6	17349	5.7	4	86	29	6/38
48	conglutin	gil17225991	Ara h 6	17349	5.7	4	104	19	9/31
49	conglutin	gil17225991	Ara h 6	17349	5.7	4	89	19	9/29
50	allergen II	gil15418705	Ara h 2	18612	5.7	4	56	19	6/40

^a The coverage refers to the sequences composed of signal and mature protein sequence.

identified as a putative peanut allergen by the phage display technology, up until now its natural counterpart has not been identified in peanut extracts. This is not surprising in view of results obtained in the course of the present work. As an example, if one compares the spot intensities of natural Ara h

7 (e.g., spot 17–24 in Figure 3) to those of the closely related conglutins, Ara h 2 and Ara h 6 are manifold more abundant. Moreover, the identification of natural Ara h 7 in extracts was hindered by an accumulation of peanut allergens with similar molecular weights, the potential cross-reactivity between the

Table 4. Sequence Coverage of Identified Peptides to Selected Sequences

spot	sequence coverage ^a (in %) to			
	Ara h 7.0101	Ara h 7.0201	2S protein 2	Ara h 7.0202
17	9	35	17	26
18	9	60	33	48
19	4	28	12	18
20	12	51	24	40
21	0	14	18	24
22	9	34	53	57
23	17	61	64	83
24	24	52	59	74

^a Highest coverage is displayed in bold type.

three 2S conglutins, and the absence of specific antibodies as analytical tools.

In addition, it was noted that the primary sequence of cloned Ara h 7 did not fully match the conserved conglutin pattern. In the case of Ara h 7.0101, the last two of eight typically conserved cysteine residues were missing in this protein. This discrepancy prompted us to search for a putative Ara h 7 cDNA that would fulfill the conglutin consensus criteria. Because the N terminus of Ara h 7.0101 presumably matched the described conglutin pattern, mRNA amplification was performed using a 5' forward oligonucleotide in accordance with the putative start region of the Ara h 7.0101 gene sequence. To identify alternative C termini that would fit the conserved conglutin pattern, for the initial reverse transcription we used a reverse oligonucleotide pairing with the poly(A) tail of the mRNA. As a result, a new Ara h 7 isoform was identified that we named Ara h 7.0201. As expected, this isoform contained additional cysteines and formed the basis for the identification of the natural counterpart by high-resolution 2D electrophoresis and mass spectrometry.

To minimize background and to bias for proteins in the molecular mass range of Ara h 7, we enriched proteins below 20 kDa by SEC of basic peanut extracts. This method has proven to be especially helpful for the enrichment of low molecular mass proteins in crude extracts because many of the storage proteins including Ara h 1 and Ara h 3/4 may form large protein aggregates of 200 kDa and above.⁸ SEC-enriched fractions were separated by conventional 2D electrophoresis and visualized with sensitive fluorescent dyes. Spots were picked, digested with trypsin, and analyzed by peptide mass fingerprinting and peptide sequencing using a MALDI-tof/tof approach. Not surprisingly, most of the picked and identified spots represented isoforms of Ara h 2 and Ara h 6.

We got especially interested in spot 23 due to its prominent location between Ara h 2 (short form) and Ara h 6 with respect to its apparent molecular mass and due to its characteristic isoelectric point, which appears 1 unit shifted to basic site relative to those of Ara h 2 and Ara h 6. Tentatively this is in accordance with the theoretical (calculated) molecular weight and pI of Ara h 7 (Table 1). However, spots 22–24 were initially identified as (*A. hypogaea*) 2S protein 2 isoforms by the MASCOT software. Thereby, the published Ara h 7.01 was not significantly identified, whereas the new Ara h 7 isoform was ranked as second hit in the database search. Nonetheless, when we studied the primary structure of the “2S protein 2” in more detail, the N terminus of the mature protein in fact pointed to a clear relation to Ara h 7 (Figure 4A,B). Knowing about the limitation of PMF analyses and their dependency on known

structures in databases employed by interpretation software such as MASCOT, we decided to control the results individually. We therefore compared the identified fragments of spots 21–24 to the “2S protein 2” and “Ara h 7.0201” sequence data. Apparently, the specific fragments VEKEQEQQDEYPYIQR and GQRPGEDEDQEQR in the N terminus of Ara h 7.0201 were missing, whereas the C-terminal part starting at aa117 matched Ara h 7.0201. Hence, we concluded that spots 21–24 represent an Ara h 7.02 isoform which is composed of an N-terminal portion of the “2S protein 2” and a C-terminal part of Ara h 7.0201 (Figure 4C). This assumption is further supported by MS/MS fragmentation data of the key peptides NLPQNCGFR and ANLRPCEEHIR and an increased theoretical sequence coverage after reconstruction (Table 4).

This finding was substantiated by the PMF analysis of spots 17–20 with their prominent spot 20. Here, the identified peptide masses of each spot were matched with the primary structure of the allergens Ara h 7.0101, Ara h 7.0201, 2S protein 2, and a hypothetical Ara h 7.0202 that was composed as described above. With regard to the sequence coverage, this spot accumulation represents the natural Ara h 7.0201.

Until now it was not possible to identify the mature nAra h 7 by N-terminal sequencing, because all attempts led to a truncated N terminus starting with aa14. This is in line with the lower molecular weight of spot 23 (spot 160 in expt 2 or spot 365 in expt 3; Figure S1 of the Supporting Information) and its higher abundance in peanut extracts as compared to spot 20 (spot 166 in expt 2 and spot 360 in expt 3). Interestingly, the N terminus of the mature Ara h 7 depicts 23 amino acids to the first conserved cysteine residue in contrast to the mature Ara h 2, which shows 11 amino acids, or mature Ara h 6, which presents 13 amino acids to the first conserved cysteine residue (Figure 4D). Therefore, a truncated N terminus of Ara h 7 with 10 amino acids to the first conserved cysteine would be in line with the other conglutins.

The cleavage site resulting in the truncated N terminus might also represent a target sequence for different proteases, which makes the N terminus of Ara h 7 fragile in whole peanut extracts. If Ara h 7 loses its N-terminal peptide, representing a putative Ara h 7-specific “tag”, specific identification by antibodies or sera could be impaired. This might explain why the rabbit anti-Ara h 7 antiserum, affinity purified by the immobilized “tag”, fails to detect nAra h 7 in peanut extracts.

The apparent low abundance of Ara h 7.0201 and Ara h 7.0202 in peanut extracts raises the question of its biological function. Because of its low abundance in peanut extracts, its function as a storage protein seems rather unlikely. However, general features of conglutins include trypsin or amylase inhibitor function. Therefore, we hypothesize that Ara h 7.0201 or Ara h 7.0202 might in fact represent a novel form of trypsin or amylase inhibitors. Thus, the mature Ara h 7.0202 may act as a pro-enzyme, which depicts a predicted cleavage site between aa66 and aa67 (<http://www.cbs.dtu.dk/services/ProP>).²¹

Conclusion

The identification of Ara h 7.0201 and Ara h 7.0202 but not the previously annotated Ara h 7.0101 as authentic allergens in peanuts justified the recloning of Ara h 7 on the basis of its aberration from the cysteine pattern of conglutins. Moreover, the cloning of a novel recombinant Ara h 7 isoform, Ara h 7.0201, yielded a protein matching the conglutin characteristics. Given this information, we could identify the natural Ara h 7

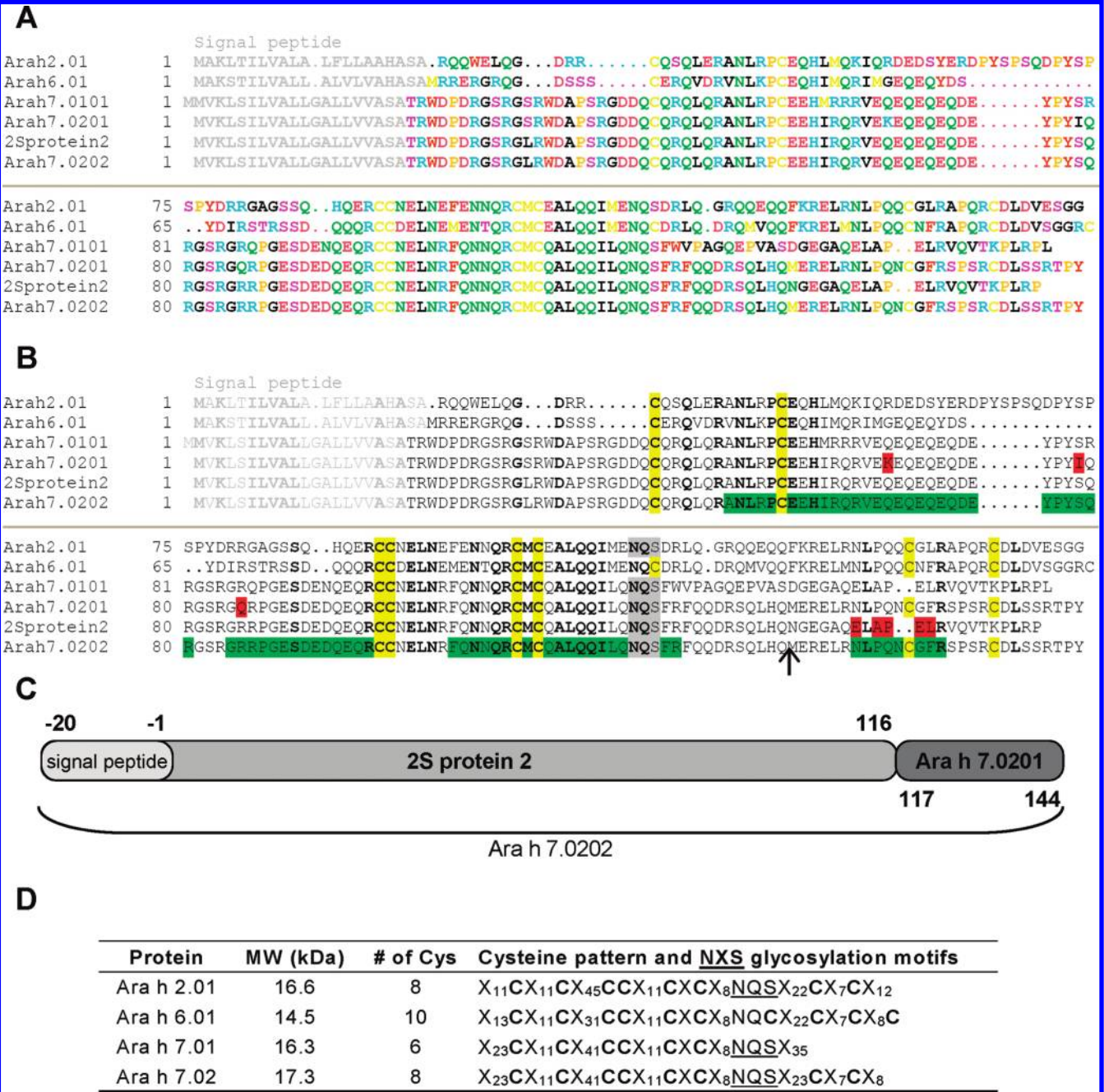


Figure 4. Sequence alignment of 2S peanut allergens. Ara h 2 (gi|15418705), Ara h 6 (gi|75114094), Ara h 7.0101 (gi|5931948), Ara h 7.0201 (gi|158121995), and 2S protein 2 (gi|52001231) were aligned with a hybrid sequence of the theoretical Ara h 7.0202 using Strap alignment software (www.charite.de/bioinf/strap). The chemical similarities between the sequences of the single proteins are displayed in panel A. Additionally, the biochemically similar amino acids between the allergens of the 2S protein family are in bold in panel B. Peptides of Ara h 7.0202 sequenced by MS/MS are highlighted in green. Differences within these sequenced peptides from Ara h 7.0201 and 2S protein 2 are highlighted in red. Cysteine patterns are marked in yellow and glycosylation motifs in gray. (C) Model of Ara h 7.0202 in which the N terminus is homologous to 2S protein S and the C terminus to Ara h 7.0201. The switch between the two proteins at aa116/117 is caused by a frame shift at the DNA level (marked by an arrow) (B). A summary of the number of cysteines and NXS-glycosylation motifs is displayed in panel D.

in peanut extracts. Here, automatic data interpretation by MASCOT was helpful, but only manual assignment of peptide masses led to the identification of the natural Ara h 7.0201 and its new isoform Ara h 7.0202 as authentic allergens in peanut extracts.

Acknowledgment. We gratefully acknowledge the expert technical assistance of Melanie Nebendahl. This study was supported by the Medical Faculty of the Christian-

Albrechts University of Kiel, the Cluster of Excellence Inflammation@Interfaces and the Innovationsfond Schleswig-Holstein. Additionally, this work was supported by 'Freundeskreis Borstel und Sparkassenstiftung Südholstein' (Borstel, Germany) and we would like to thank Stefanie Fox for excellent technical assistance.

Supporting Information Available: Gel pictures of additional experiments, sequence alignment of identified pep-

tides using ProteinProspector, list of masses measured in selected protein spots, homology calculations of respective allergens, and complete mass spectrometry data for each of the 50 identified spots. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Mittag, D.; Akkerdaas, J.; Ballmer-Weber, B. K.; Vogel, L.; Wensing, M.; Becker, W. M.; Koppelman, S. J.; Knulst, A. C.; Helbling, A.; Hefle, S. L.; Van, R. R.; Vieths, S. Ara h 8, a Bet v 1-homologous allergen from peanut, is a major allergen in patients with combined birch pollen and peanut allergy. *J. Allergy Clin. Immunol.* **2004**, *114* (6), 1410–1417.
- (2) Krause, S.; Reese, G.; Randow, S.; Zennaro, D.; Quarantino, D.; Palazzo, P.; Ciardiello, M. A.; Petersen, A.; Becker, W. M.; Mari, A. Lipid transfer protein (Ara h 9) as a new peanut allergen relevant for a Mediterranean allergic population. *J. Allergy Clin. Immunol.* **2009**, *124* (4), 771–778.
- (3) Kleber-Janke, T.; Crameri, R.; Appenzeller, U.; Schlaak, M.; Becker, W. M. Selective cloning of peanut allergens, including profilin and 2S albumins, by phage display technology. *Int. Arch. Allergy Immunol.* **1999**, *119* (4), 265–274.
- (4) Kleber-Janke, T.; Crameri, R.; Scheurer, S.; Vieths, S.; Becker, W. M. Patient-tailored cloning of allergens by phage display: peanut (*Arachis hypogaea*) profilin, a food allergen derived from a rare mRNA. *J. Chromatogr., B: Biomed. Sci. Appl.* **2001**, *756* (1–2), 295–305.
- (5) Riecken, S.; Lindner, B.; Petersen, A.; Jappe, U.; Becker, W. M. Purification and characterization of natural Ara h 8, the Bet v 1 homologous allergen from peanut, provides a novel isoform. *Biol. Chem.* **2008**, *389* (4), 415–423.
- (6) Gruber, P.; Suhr, M.; Frey, A.; Becker, W. M.; Hofmann, T. Development of an epitope-specific analytical tool for the major peanut allergen Ara h 2 using a high-density multiple-antigenic peptide strategy. *Mol. Nutr. Food Res.* **2004**, *48* (6), 449–458.
- (7) Kreis, M.; Forde, B. G.; Rahman, S.; Mifflin, B. J.; Shewry, P. R. Molecular evolution of the seed storage proteins of barley, rye and wheat. *J. Mol. Biol.* **1985**, *183* (3), 499–502.
- (8) Shewry, P. R.; Jenkins, J.; Beaudoin, F.; Mills, E. N. C. Plant protein families as a basis for predicting the allergenicity of food proteins. In *Allergy Matters: New Approaches to Allergy Prevention and Management*; Glissen, L. J. W. J., Wichers, H. J., Savelkoul, H. F. J., Bogers, R. J., Eds.; Springer: Dordrecht, The Netherlands, 2006; pp 39–44.
- (9) Bendtsen, J. D.; Nielsen, H.; von Heijne, G.; Brunak, S. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **2004**, *340* (4), 783–795.
- (10) Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Anal. Biochem.* **1976**, *72*, 248–254.
- (11) Suhr, M.; Wicklein, D.; Lepp, U.; Becker, W. M. Isolation and characterization of natural Ara h 6: evidence for a further peanut allergen with putative clinical relevance based on resistance to pepsin digestion and heat. *Mol. Nutr. Food Res.* **2004**, *48* (5), 390–399.
- (12) Boldt, A.; Fortunato, D.; Conti, A.; Petersen, A.; Ballmer-Weber, B.; Lepp, U.; Reese, G.; Becker, W. M. Analysis of the composition of an immunoglobulin E reactive high molecular weight protein complex of peanut extract containing Ara h 1 and Ara h 3/4. *Proteomics* **2005**, *5* (3), 675–686.
- (13) Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227* (5259), 680–685.
- (14) Kyhse-Andersen, J. Electrophoretic blotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. *J. Biochem. Biophys. Methods* **1984**, *10* (3–4), 203–209.
- (15) Leary, J. J.; Brigati, D. J.; Ward, D. C. Rapid and sensitive colorimetric method for visualizing biotin-labeled DNA probes hybridized to DNA or RNA immobilized on nitrocellulose: Bio-blot. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80* (13), 4045–4049.
- (16) Edman, P. Sequence determination. *Mol. Biol. Biochem. Biophys.* **1970**, *8*, 211–255.
- (17) Schmidt, H.; Gelhaus, C.; Nebendahl, M.; Lettau, M.; Watzl, C.; Kabelitz, D.; Leippe, M.; Janssen, O. 2-D DIGE analyses of enriched secretory lysosomes reveal heterogeneous profiles of functionally relevant proteins in leukemic and activated human NK cells. *Proteomics* **2008**, *8* (14), 2911–2925.
- (18) Schmidt, H.; Gelhaus, C.; Latendorf, T.; Nebendahl, M.; Petersen, A.; Krause, S.; Leippe, M.; Becker, W. M.; Janssen, O. 2-D DIGE analysis of the proteome of extracts from peanut variants reveals striking differences in major allergen contents. *Proteomics* **2009**, *9* (13), 3507–3521.
- (19) Harris, W. A.; Janneck, D. J.; Reilly, J. P. Use of matrix clusters and trypsin autolysis fragments as mass calibrants in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **2002**, *16* (18), 1714–1722.
- (20) Clauser, K. R.; Baker, P.; Burlingame, A. L. Role of accurate mass measurement (± 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Anal. Chem.* **1999**, *71* (14), 2871–2882.
- (21) Duckert, P.; Brunak, S.; Blom, N. Prediction of proprotein convertase cleavage sites. *Protein Eng. Des. Sel.* **2004**, *17* (1), 107–112.

PR1002406

Supplementary material

Detection and structural characterization of natural Ara h 7, the third peanut allergen of the 2S albumin family.

Hendrik Schmidt*, Susanne Krause*, Christoph Gelhaus, Arnd Petersen, Ottmar Janssen* and Wolf-Meinhard Becker*

*equally contributing authors

Figure S1

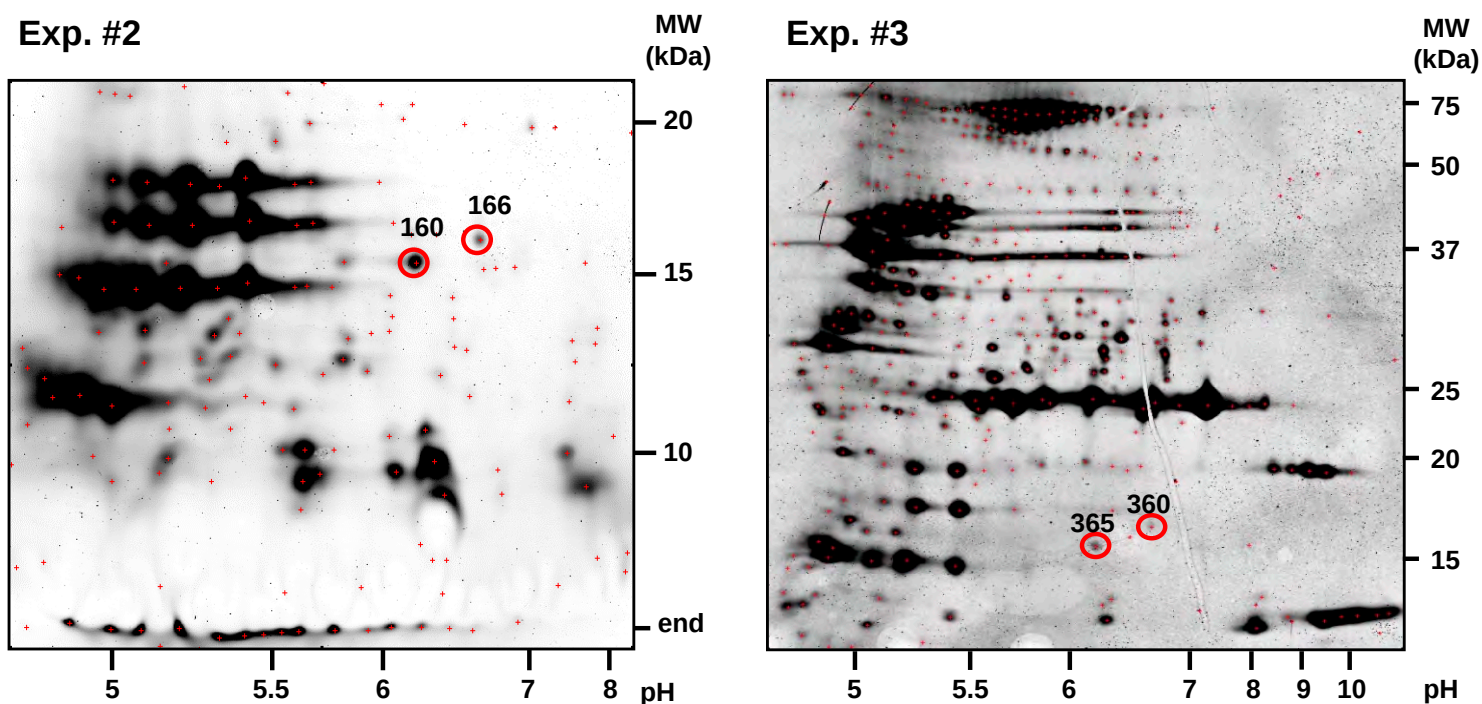


Figure S1. Prefractionated (#2) and whole (#3) peanut extracts were separated on two dimensional gels as indicated in table 2. Circled spots were picked and analyzed by combined PMF and MS/MS sequencing. Spots 160 (#2) and 365 (#3) correspond to spot 23 (#1) and spots 166 (#2) and 360 (#3) to spot 20 (#1) in the annotated master gel (figure 3).

Figure S2 A

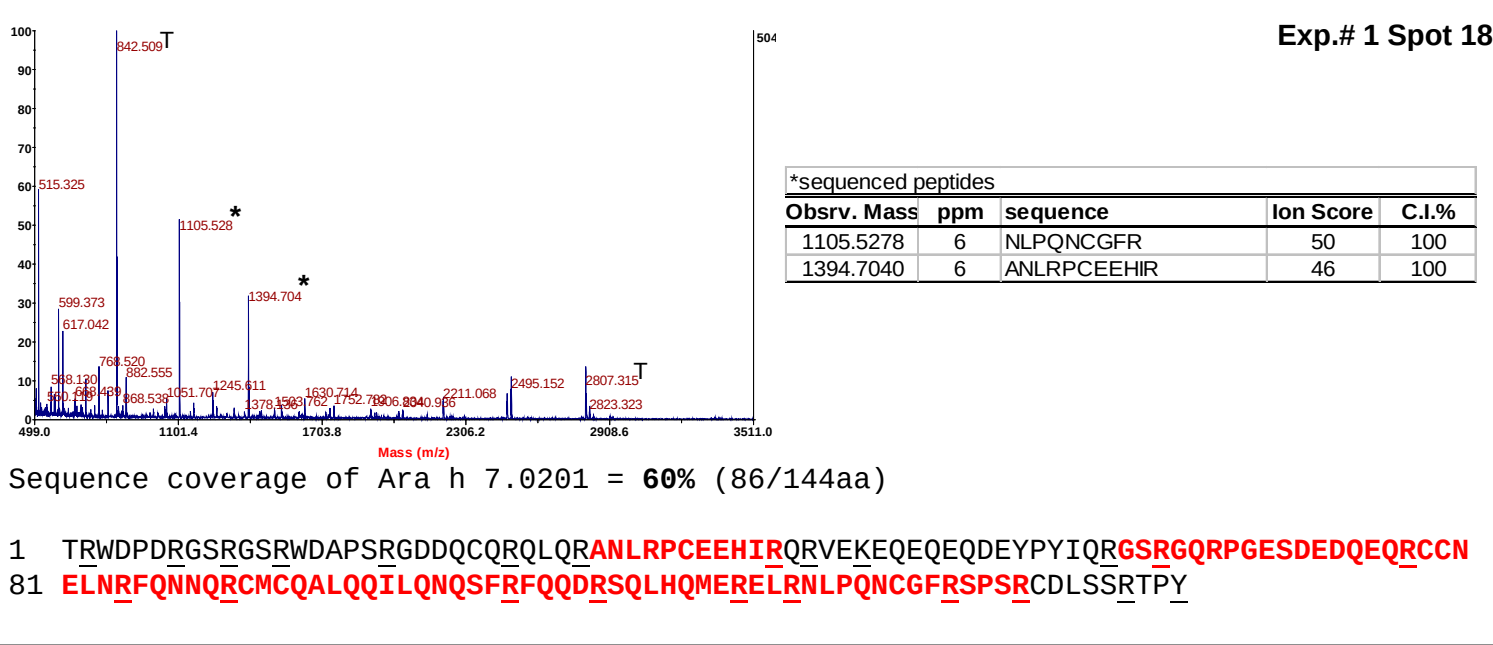
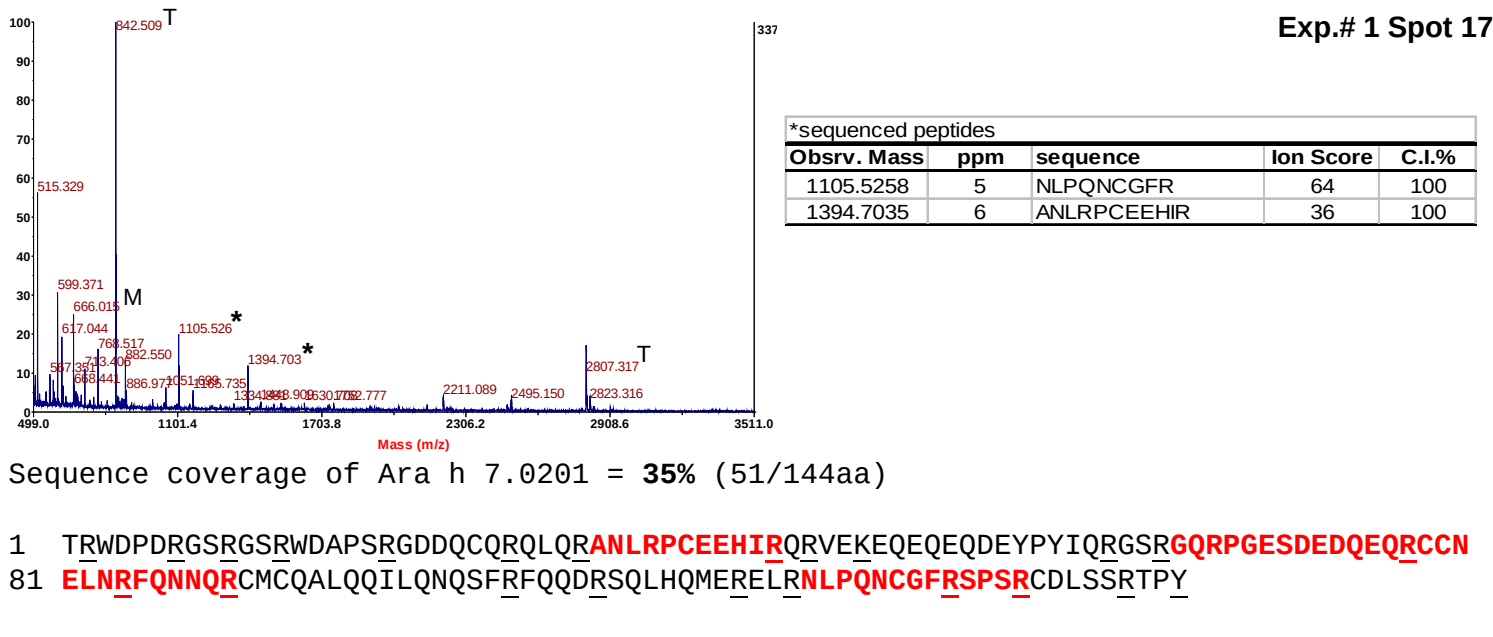


Figure S2 B

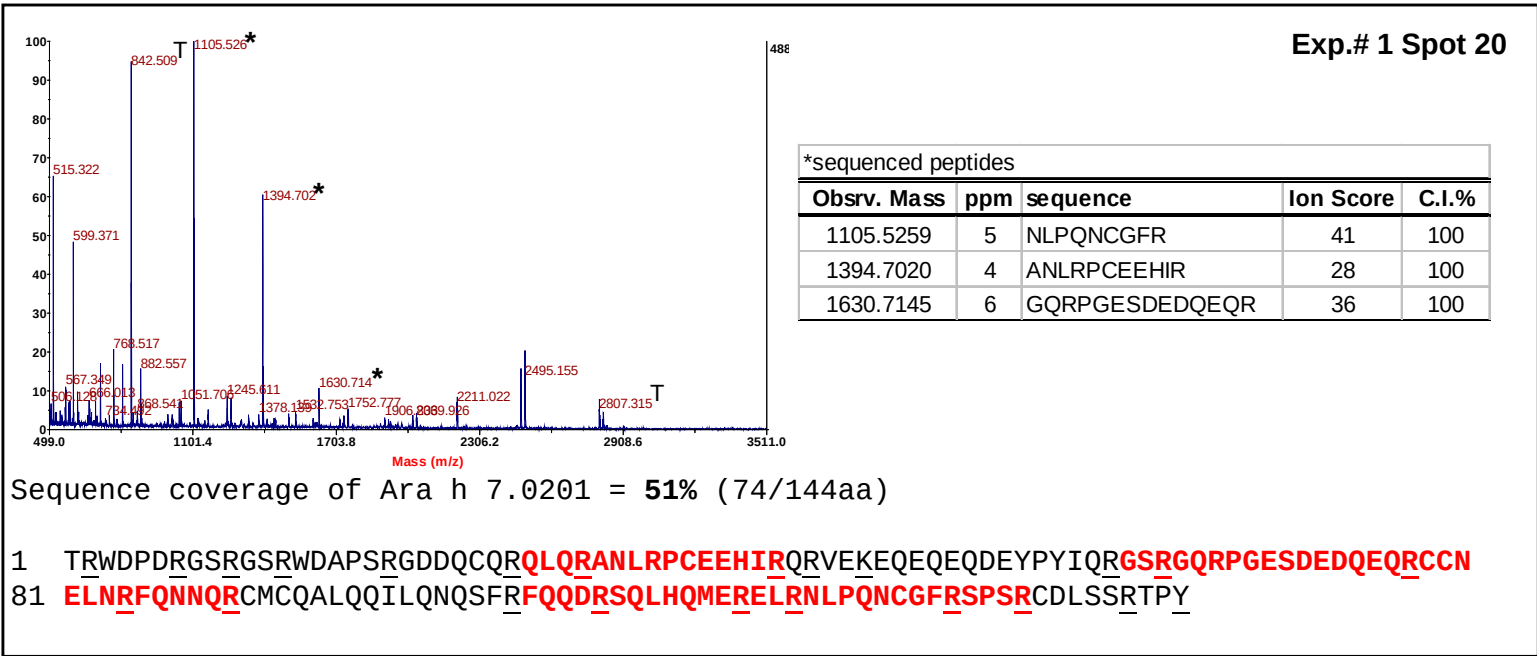
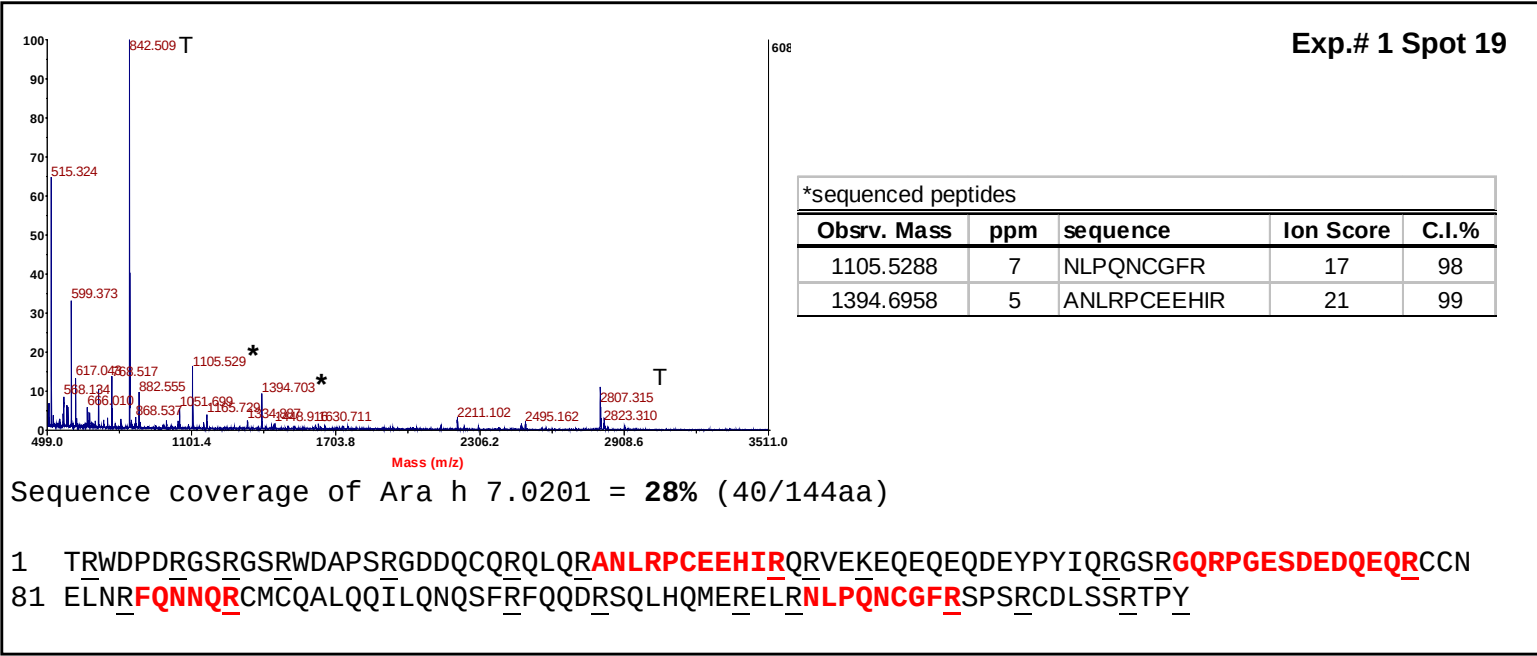


Figure S2 A,B. MALDI data interpretation of Ara h 7.0201 spots 17, 18 **(A)** and 19, 20 **(B)**. Mass spectra were annotated according to known masses of methylated porcine trypsin autolysis fragments (T) and matrix clusters (M). Sequenced peptides in MS/MS mode are marked (*) and listed with the observed mass, the variation to the calculated mass, the ion score and the confidence intervall. Other protein specific peptide masses are listed in the supplemental MASCOT data and table S1. Sequence coverage of Ara h 7.0201 excluding the signal sequence (aa 1-20) is displayed in red letters.

Figure S3

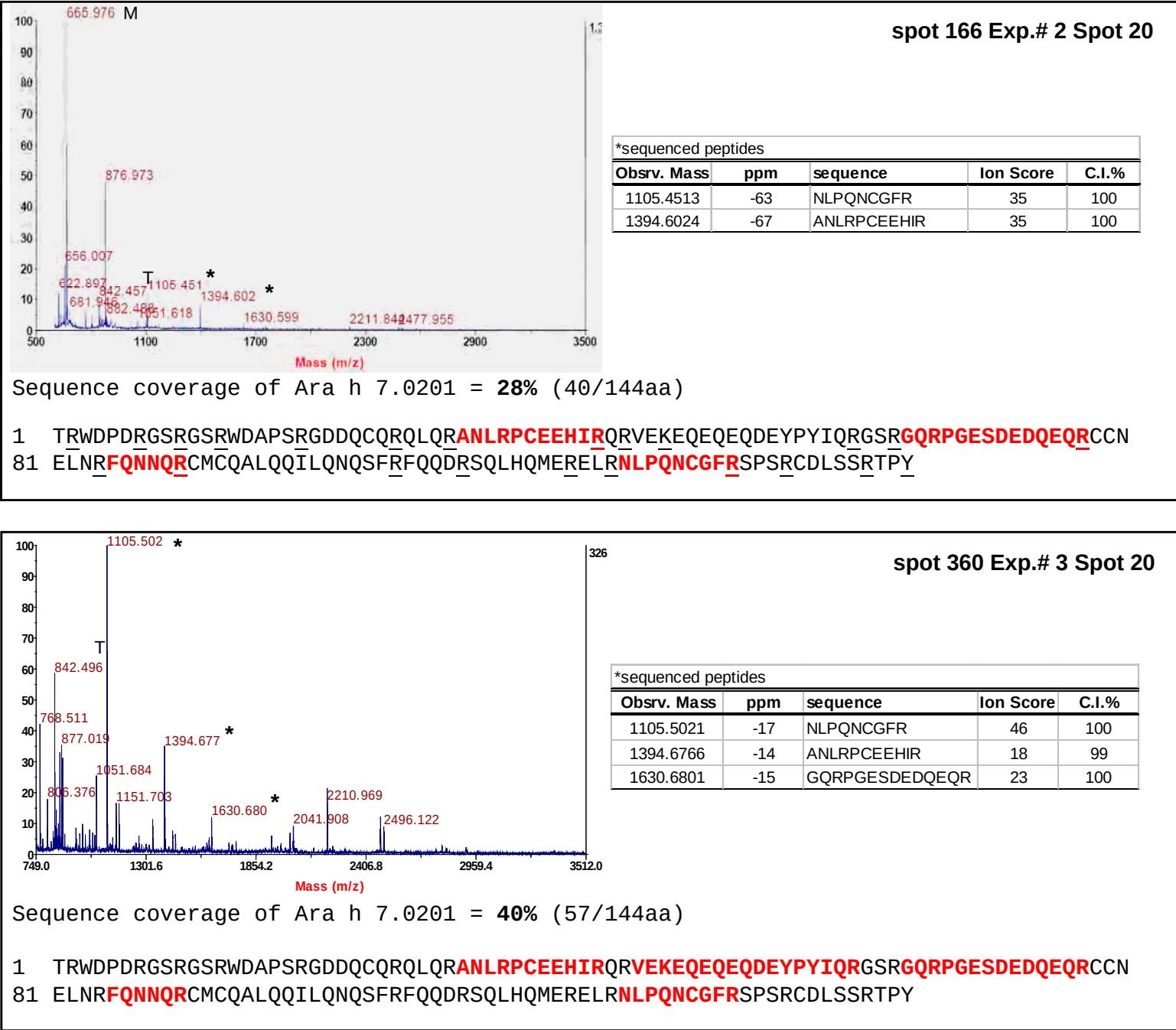


Figure S3. MALDI data interpretation of spot 20 (Ara h 7.0201) in two additional experiments (figure S1). Mass spectra were annotated according to known masses of methylated porcine trypsin autolysis fragments (T) and matrix clusters (M). Sequenced peptides in MS/MS mode are marked (*) and listed with the observed mass, the variation to the calculated mass (in ppm), the ion score and the confidence intervall. Sequence coverage of Ara h 7.0201 excluding the signal sequence (aa 1-20) is displayed in red letters.

Figure S4 A

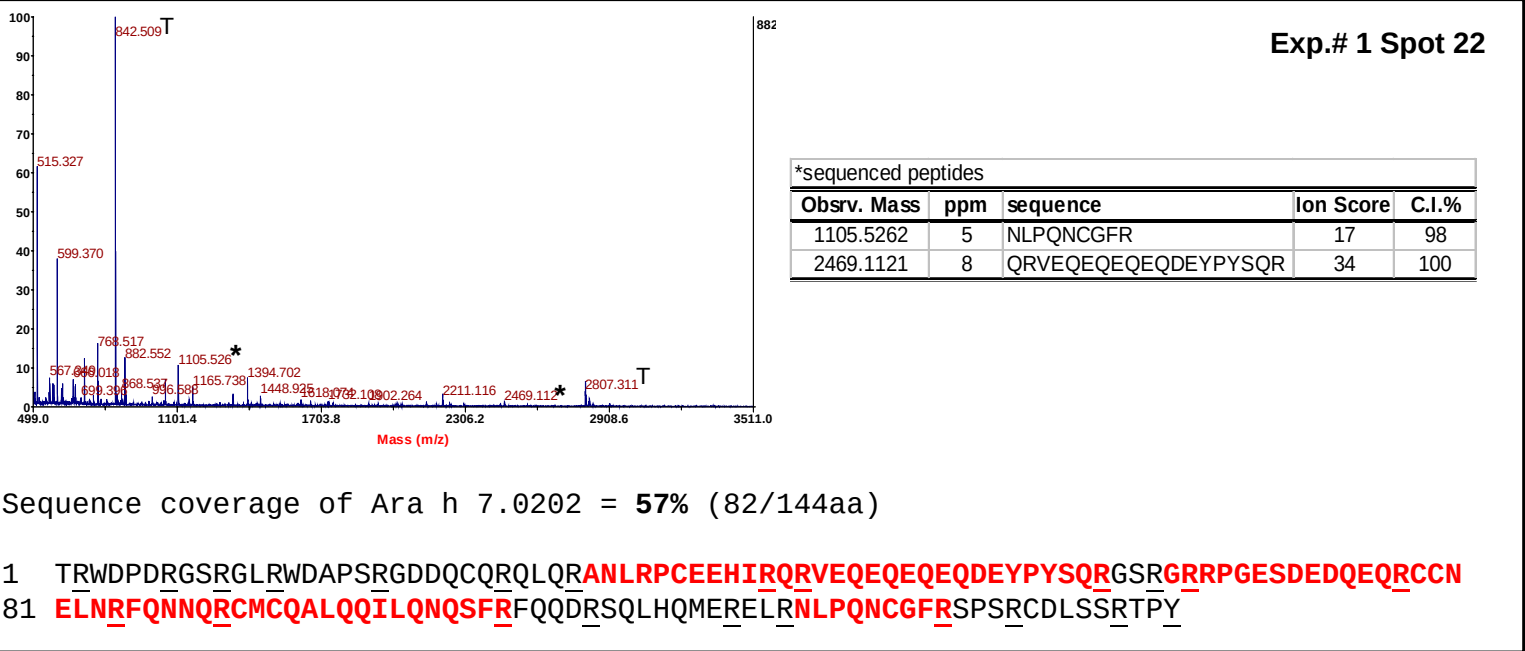
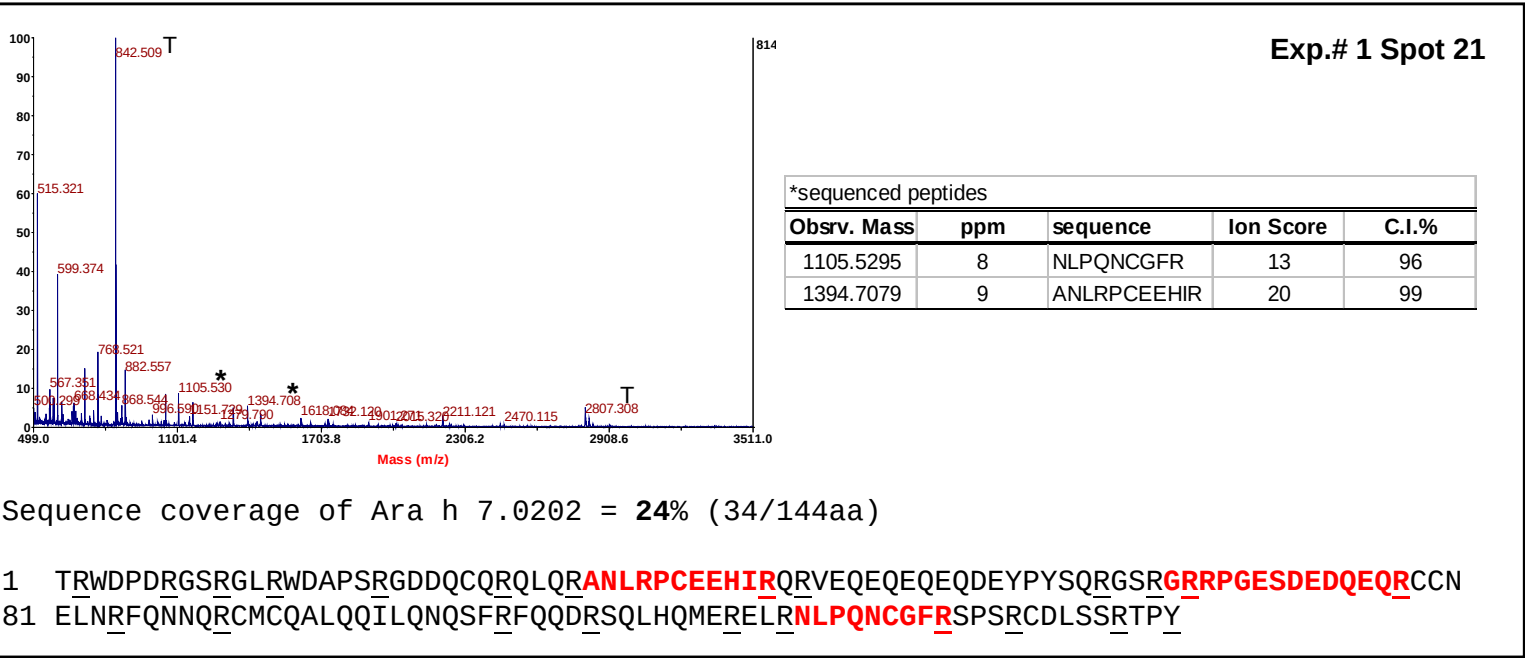


Figure S4 B

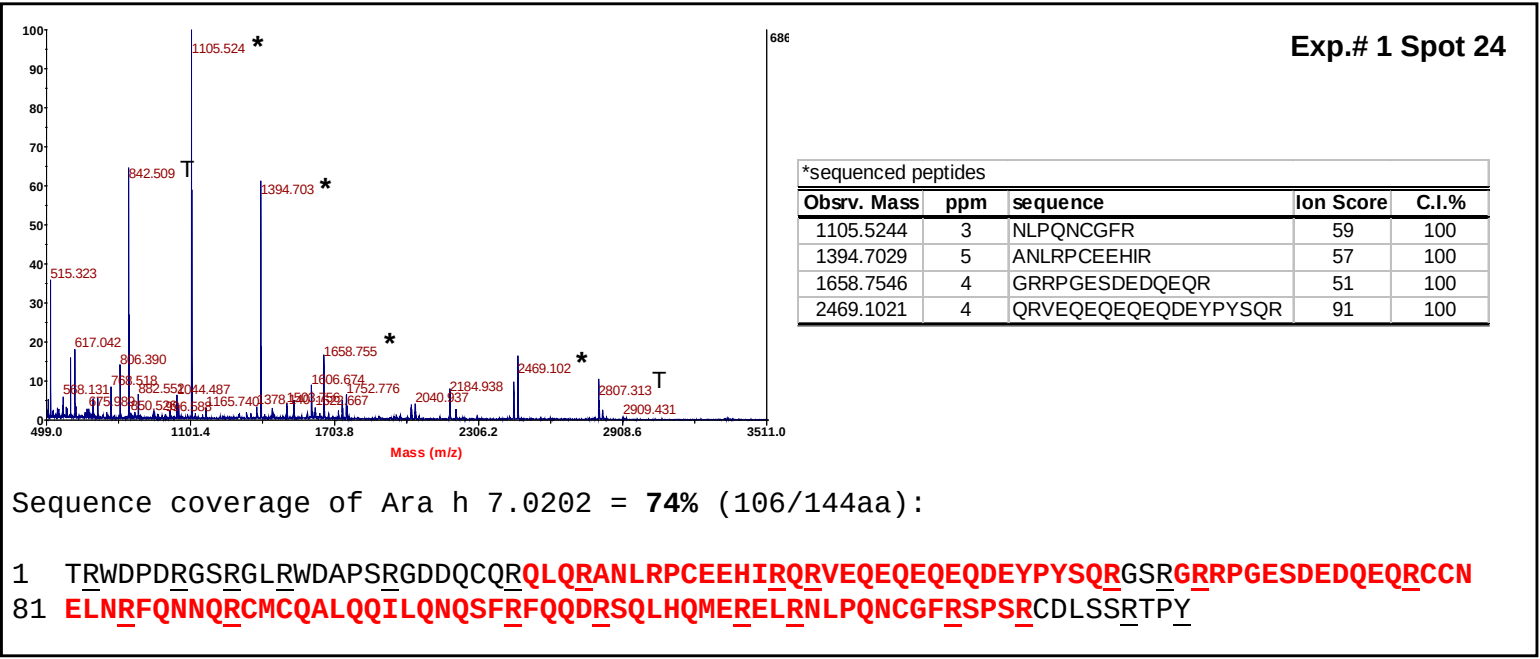
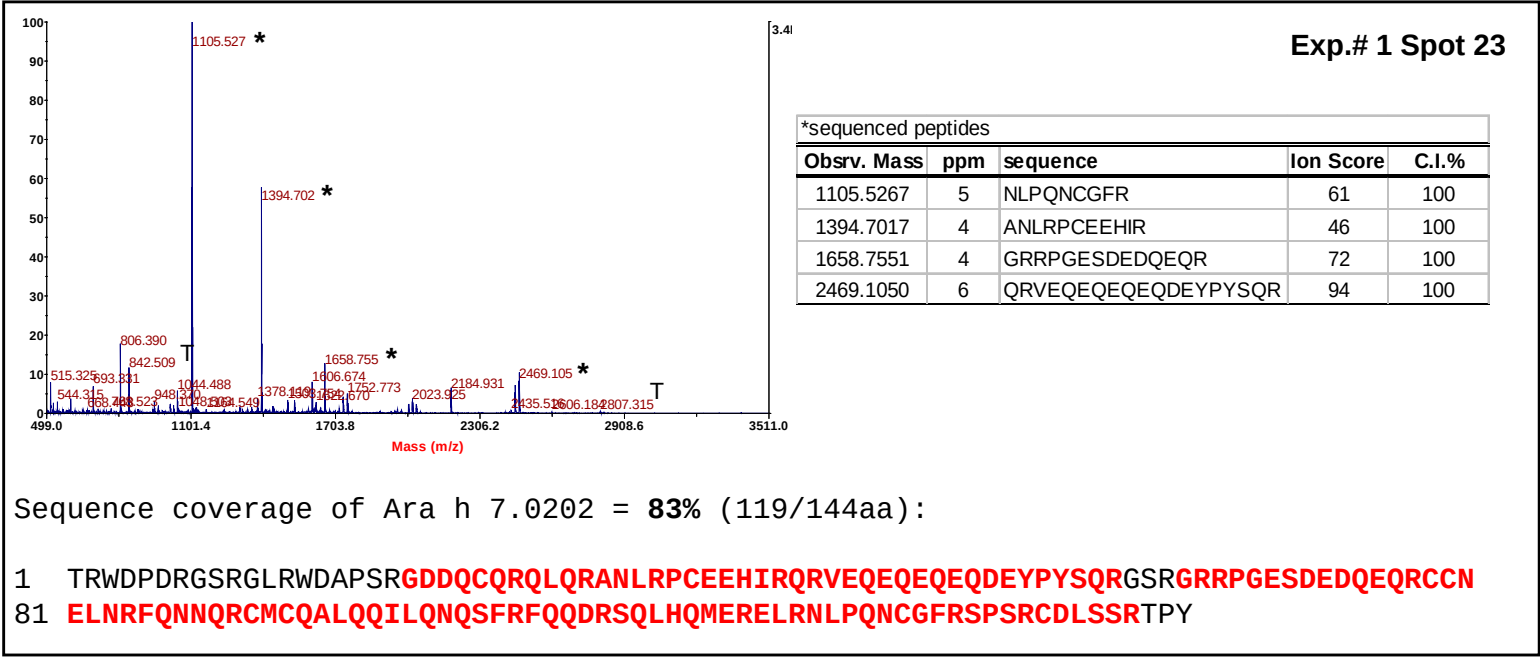


Figure S4 A,B. MALDI data interpretation of Ara h 7.0202 spots 21, 22 **(A)** and 23, 24 **(B)**. Mass spectra were annotated according to known masses of methylated porcine trypsin autolysis fragments (T) and matrix clusters (M). Sequenced peptides in MS/MS mode are marked (*) and listed with the observed mass, the variation to the calculated mass, the ion score and the confidence intervall. Other protein specific peptide masses were searched against Ara h 7.0202 sequence (Figure 4B and data set S2) using the software MS-Bridge in ProteinProspector 5.3.0 (UCSF, USA). The search results are listed additional to the supplemental MASCOT data. Sequence coverage of Ara h 7.0202 is displayed in red letters

Figure S5

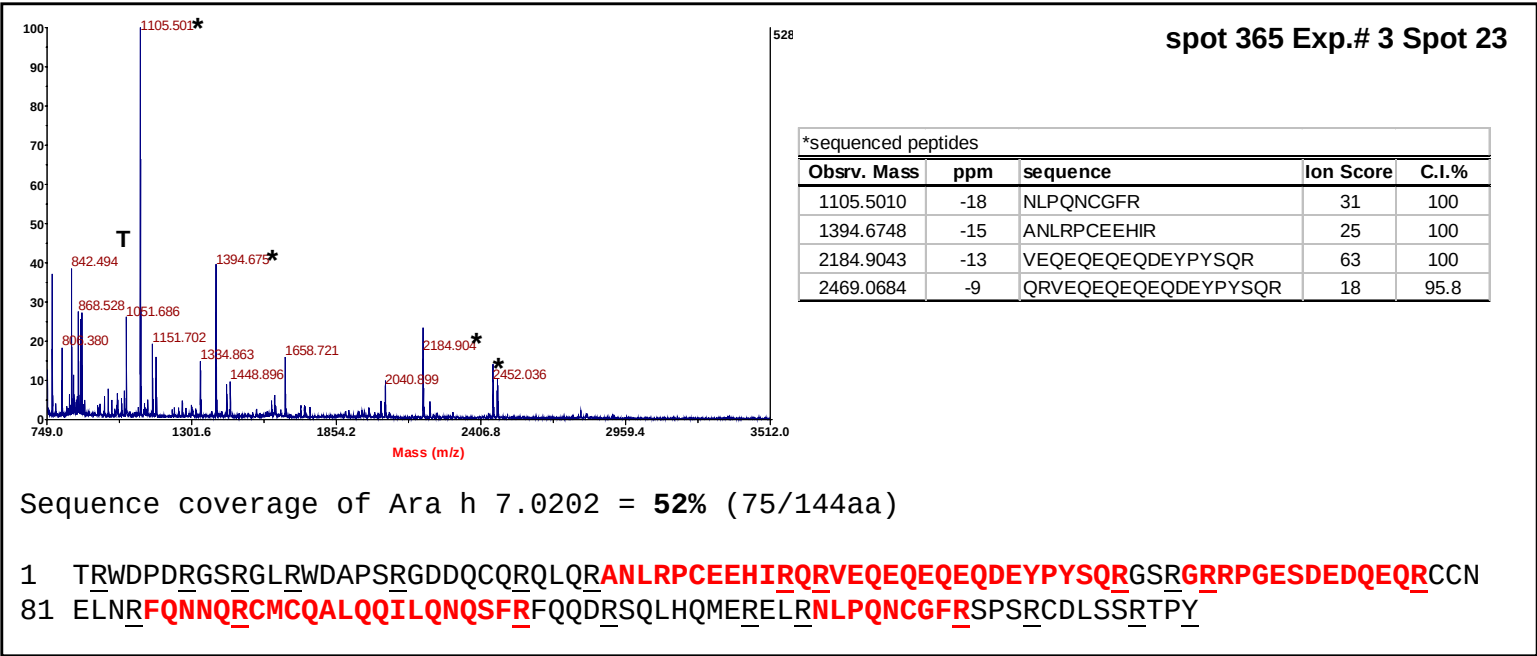
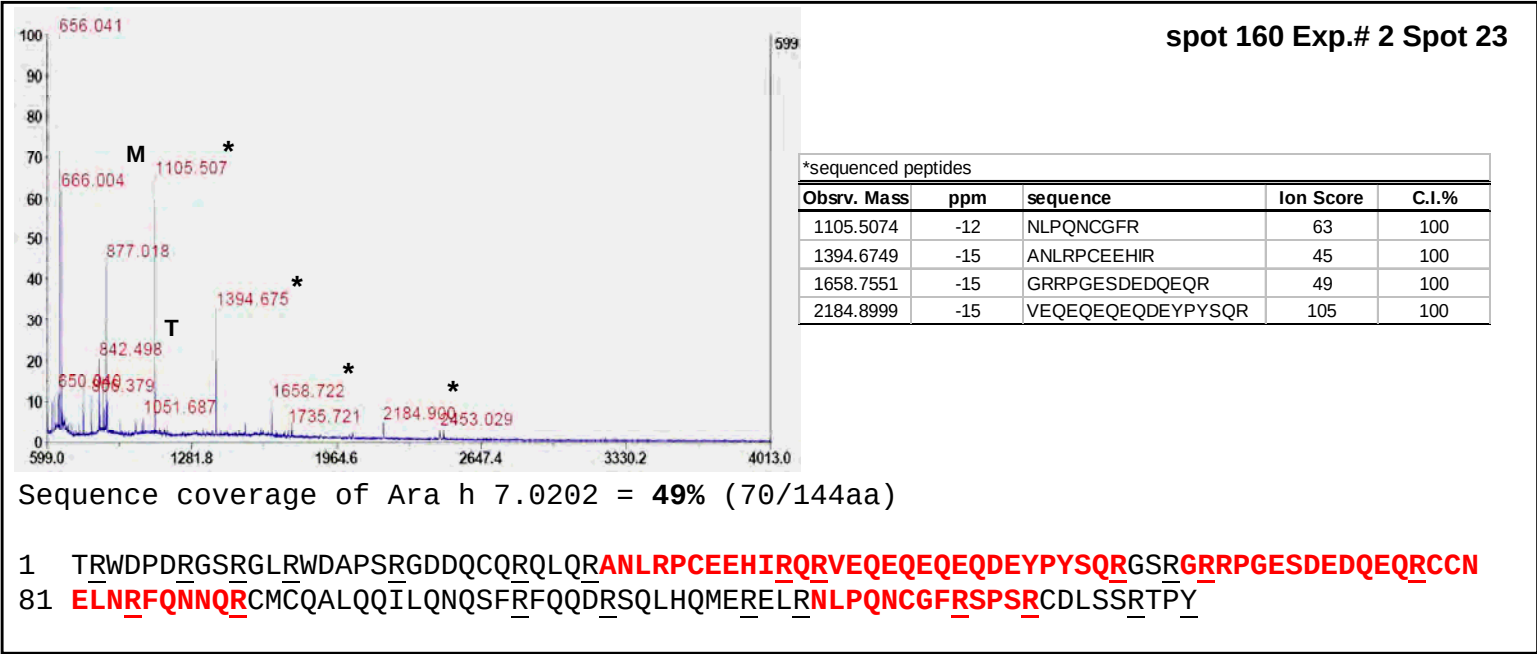


Figure S5. MALDI data interpretation of spot 23 (Ara h 7.0202) in two additional experiments (figure S1). Mass spectra were annotated according to known masses of methylated porcine trypsin autolysis fragments (T) and matrix clusters (M). Sequenced peptides in MS/MS mode are marked (*) and listed with the observed mass, the variation to the calculated mass (in ppm), the ion score and the confidence intervall. Measured masses were searched against tryptic peptides of Ara h 7.0202 (Figure 4B and data set S2) using software MS-Bridge in ProteinProspector 5.3.0 (UCSF, USA). The search results are listed additional to the supplemental MASCOT data. Sequence coverage of Ara h 7.0202 is displayed in red letters.

Figure S6

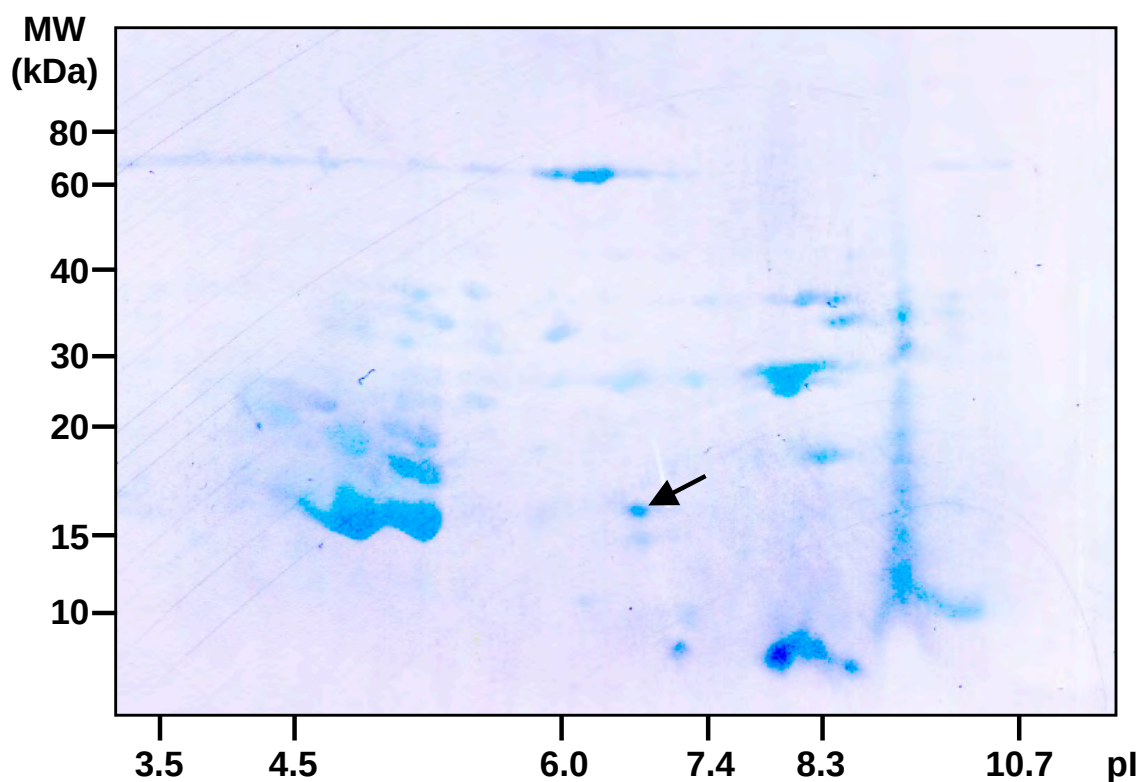


Figure S6. 2D-PAGE for N-terminal sequencing. Protein extracts of <20 kDa fraction were separated on a 2D-PAGE, transferred to PVDF-membrane and stained with Coomassie blue. The arrow marks the spot which was excised for N-terminal sequencing.

Table S1. List of measured masses of selected spots (#17-24).

spot #								contaminants*	
17	18	19	20	21	22	23	24	matrix clusters	autolysis fragments
500.30063	500.30313	500.29974	500.3028	500.29855	500.30234	500.28812	500.30112	644.0473	842.5099
506.12479	506.12891	506.12543	506.1275	506.12317	506.12817	506.12585	506.12064	666.0293	870.5412
515.32947	515.32501	515.32373	515.32172	515.32104	515.32727	515.32477	515.3233	680.9954	1126.5645
565.35394	524.1416	565.3446	524.14471	551.32861	565.34143	524.14142	527.28979	833.0899	1424.7648
567.35132	550.11932	567.35126	527.29669	565.34247	567.3493	527.28796	544.31659	855.0719	1623.8856
568.13367	565.35553	568.13391	544.31274	567.35107	568.13281	544.31458	550.11365	871.0458	1940.9351
581.36969	567.354	581.37061	565.34094	568.13153	581.36914	565.3454	567.35376	907.9938	1987.0784
585.36682	568.12988	585.35883	567.34882	581.36542	585.35681	567.3537	568.13086	1060.0884	2211.1045
599.37128	568.36469	599.37335	568.12939	585.36188	599.36987	568.13379	581.36987	1082.0703	2225.1201
617.04437	581.37201	617.0426	581.36829	599.37396	617.04059	581.37097	585.35693	1098.0442	2239.1357
622.04181	585.36517	622.03253	585.36072	617.03827	622.03406	585.35498	599.37256		2283.1807
622.99902	599.37317	666.00995	599.37109	622.03497	660.02252	595.37433	617.04163		2297.1963
666.01465	617.04224	668.43628	617.03931	660.00354	666.01752	599.37799	622.03467		2564.1812
668.44104	640.34308	675.97052	622.03265	666.01276	668.43665	617.04529	660.01959		2707.4165
675.9787	666.01105	699.39233	660.00635	668.4339	675.98706	622.04279	668.44116		2720.3635
681.987	668.43854	713.40619	666.01251	675.97607	697.98322	650.34442	675.98944		2748.3948
697.97644	675.98157	734.48877	668.43658	681.99341	699.39563	666.01782	693.32788		2807.3142
713.40637	693.32813	750.51965	675.97247	699.39417	713.40411	668.44305	713.40594		2914.5059
750.51953	699.39105	768.51672	693.33521	713.40576	734.49036	675.97894	750.51373		3346.6816
768.51685	713.40863	806.39282	699.38538	734.48511	750.51263	693.33051	768.51758		
842.50928	734.49017	842.50903	713.4043	736.49609	768.51685	706.30847	806.39026		
850.52765	750.51477	850.52429	734.49182	750.51691	782.4696	713.41016	842.50885		
882.54987	768.52002	868.53687	750.51575	768.52075	806.39093	720.30298	850.52649		
886.97687	782.47516	882.5545	768.51746	782.47064	842.50885	737.32727	868.53664		
996.58716	806.39423	886.98981	782.47107	842.50879	850.53101	747.33057	882.55182		
1051.69934	842.50897	996.59033	806.39258	850.52881	868.53668	768.52313	948.37292		
1105.52576	850.53802	1045.55115	842.50897	864.53735	882.5517	806.39044	965.40417		
1165.73486	864.54944	1051.6991	850.52795	868.54425	886.99261	842.50897	996.58844		
1334.88123	868.53827	1105.52881	868.54065	882.55725	982.57025	850.53552	1014.52606		
1394.70349	882.55493	1151.72998	882.55701	886.99835	996.58777	878.36639	1044.48706		
1448.90881	982.57135	1165.72949	886.98914	951.61688	1045.56213	882.56262	1051.70618		
1532.75037	996.59247	1334.88696	996.59125	982.57861	1051.70239	942.43726	1105.52441		
1630.70837	1044.49023	1394.70325	1014.5271	996.59045	1105.52625	948.37012	1121.51245		
1752.77698	1051.70691	1448.91565	1044.48938	1033.69055	1151.72754	965.39423	1165.73962		
2144.04834	1105.52783	1619.06738	1051.70593	1045.56091	1165.73828	980.50452	1303.70068		
2211.08936	1151.72742	1630.71082	1105.52588	1051.70959	1334.89026	1014.52118	1334.88831		
2478.13037	1165.74414	1733.10486	1121.5188	1105.52954	1394.7019	1016.52069	1353.724		
2495.14966	1245.61108	2143.04321	1124.50818	1151.72876	1448.92456	1028.5	1378.14038		
2807.31738	1261.60596	2211.10205	1165.73792	1165.74609	1618.07361	1044.48816	1394.70288		
2823.31567	1334.89038	2240.14502	1245.61145	1279.78992	1658.76062	1048.50256	1442.72449		
2840.32104	1378.13599	2478.1272	1261.6084	1334.90015	1732.10791	1088.52673	1503.75623		
2909.43823	1394.70398	2495.16187	1334.88867	1394.70789	1735.75513	1088.94995	1532.74878		
	1442.72998	2807.31543	1378.13904	1434.91284	1752.77173	1105.52673	1606.67419		
	1448.92334	2813.3103	1394.70203	1448.93433	1902.26355	1117.52722	1618.68762		
	1503.76196	2823.31006	1413.68848	1618.08423	2143.0415	1119.52014	1622.66711		
	1532.7489	2841.32349	1442.73047	1658.76978	2184.94702	1121.51416	1658.75464		
	1606.67432	2909.45654	1448.927	1732.12024	2211.11646	1124.50891	1678.8634		
	1630.71448		1503.75574	1901.27087	2469.11206	1127.51074	1718.8114		
	1718.81677		1532.75293	2015.31958	2807.31079	1164.54871	1735.75012		
	1735.75317		1606.67468	2144.05225	2812.31763	1235.55652	1752.77588		
	1752.78223		1630.71448	2211.12061	2823.30396	1303.69446	2023.9303		
	1906.83447		1718.81396	2470.11475	2908.44043	1305.67444	2039.91882		
	1922.82239		1735.75	2807.30835		1314.7041	2040.93701		
	1930.87646		1752.77734	2823.28931		1337.68042	2056.94019		
	2023.92664		1906.8363	2909.45435		1353.7019	2184.93823		
	2039.92639		1922.83472			1377.68408	2211.11328		
	2040.93604		1930.87195			1378.11853	2452.07788		
	2144.04858		2023.92395			1394.70166	2469.10205		
	2211.06763		2039.92603			1410.69531	2807.31348		
	2300.18066		2211.02246			1413.68896	2823.30737		
	2478.12524		2478.13159			1426.72302	2909.4314		
	2495.15234		2495.15503			1442.71875			
	2807.31494		2807.31494			1445.64648			
	2823.323		2823.31982			1503.75354			
	2841.31226		2839.3064			1532.74487			
	2909.45874		2909.47266			1589.65735			
						1590.10876			
						1606.67432			
						1618.68518			
						1622.66956			
						1641.21704			
						1642.20178			
						1658.75513			
						1678.86279			
						1718.80786			
						1735.74707			
						1752.77295			
						1771.75879			
						1950.93774			
						1959.91785			
						1961.52881			
						1978.55249			
						2007.92285			
						2023.92468			
						2039.91992			
						2040.93237			
						2057.96094			
						2184.93115			
						2435.51563			
						2452.08105			
						2469.10498			
						2606.18359			
						2807.31494			
						2823.31079			

* according to Harris et al.

Data set S1.

Sequence alignments of 2S protein 2 with Ara h 7.0101 and Ara h 7.0201 using SIM - Alignment Tool for protein sequences (<http://www.expasy.ch/tools/sim-prot.html>)

Parameters for alignment:

Comparison matrix: BLOSUM62
Number of alignments computed: 1
Gap open penalty: 3
Gap extension penalty: 1

Sequence 1: Ara h 7.0101, (139 residues)

Sequence 2: 2Sprotein2, (138 residues)

87.0% identity in 138 residues overlap; Score: 655.0; Gap frequency: 0.0%

```
Arah7.0101      1 TRWDPDRGSRGSRWDAPSRGDDQCQRQLQRANLRPCEEHMRRRVEQEQQEQDEYPYSRR
2Sprotein2      1 TRWDPDRGSRGLRWDAPSRGDDQCQRQLQRANLRPCEEHIRQRVEQEQQEQDEYPYSQR
                *****
Arah7.0101     61 GSRGRQPGESDENQEQRCCNELNRFQNNQRCMCQALQQILQNQSFVWPAGQEPVASDGEG
2Sprotein2     61 GSRGRRPGESDEDQEQRCCNELNRFQNNQRCMCQALQQILQNQSFQFQDRSQLHQNGEG
                *****
Arah7.0101    121 AQELAPELRVQVTKPLRP
2Sprotein2    121 AQELAPELRVQVTKPLRP
                *****
```

Sequence 1: Ara h 7.0201, (144 residues)

Sequence 2: 2Sprotein2, (138 residues)

83.8% identity in 142 residues overlap; Score: 629.0; Gap frequency: 8.5%

```
Arah7.0201      1 TRWDPDRGSRGSRWDAPSRGDDQCQRQLQRANLRPCEEHIRQRVEKEQEQQEQDEYPYIQR
2Sprotein2      1 TRWDPDRGSRGLRWDAPSRGDDQCQRQLQRANLRPCEEHIRQRVEQEQQEQDEYPYSQR
                *****
Arah7.0201     61 GSRGQRPGESDEDQEQRCCNELNRFQNNQRCMCQALQQILQNQSFQFQDRSQLHQM-E-
2Sprotein2     61 GSRGRRPGESDEDQEQRCCNELNRFQNNQRCMCQALQQILQNQSFQFQDRSQLHQNGEG
                *****
Arah7.0201    119 -RELRLNPQNCGFR---SPSR
2Sprotein2    121 AQELA--PE---LRVQVTKPLR
                **  *  *  *
```

Data set S2: This data set contains the ProteinProspector search results of the measured peptide masses (excluding masses of matrix clusters and trypsin autolysis fragments, see table S1) against four different sequences of putative Ara h 7 related isoforms namely Ara h 7.0101, Ara h 7.0102, 2S protein 2 and Ara h 7.0202. Masses were searched with depicted parameters (see below). Results display theoretical pI and MW as well as the sequence coverage of the respective sequence.

Parameters for MS-Bridge in ProteinProspector 5.3.2

Database: **User Protein**

Considered modifications: | **Gln->pyro-Glu (N-term Q)** | **Oxidation (M)** |

Digest Used: **Trypsin**

Max. # Missed Cleavages: **1**

Constant Modification: **Carbamidomethyl (C)**

Minimum Digest Fragment Mass: **500**

Maximum Digest Fragment Mass: **4000**

Minimum Digest Fragment Length: **5**

Mass tolerance: 50 ppm

Spot 17

Ara h 7.0101

Index Number: 17

pI of Protein: 5.6

Protein MW: 16323

Amino Acid Composition: A7 C6 D8 E15 F2 G9 H1 I1 K1 L9 M2 N7 P10 Q21 R20 S8 T2 V5 W3 Y2

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHM RRRVEQEQQ EQDEYPYSRR GSRGRQPGES DENQEQRCCN

81 ELNRFQNNQR CMCQALQQIL QNQSFVWPAG QEPVASDGEG AQELAPELRV QVTKPLRPL

The matched peptides cover 9% (13/139AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
1752.7770	1752.7653	100.0	6.67		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)

Ara h 7.0201

Index Number: 17

pI of Protein: 7.5

Protein MW: 17375

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I3 K1 L9 M2 N8 P8 Q24 R23 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQVEKEQQ EQDEYPYIQR GSRGQRPGES DEDQEQRCCN

81 ELNRFQNNQR CMCQALQQIL QNQSFQFQD RSQLHQMERL LNLPQNCGF RSPSRCDLSS RTPY

The matched peptides cover 35% (51/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
1105.5258	1105.5207	100.0	4.54		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.7035	1394.6957	100.0	5.55		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1532.7504	1532.7387	100.0	7.63		123	135	1	(R)NLPQNC(Carbamidomethyl)GFRSPSR(C)
1630.7084	1630.7052	100.0	1.96		64	77	0	(R)GQRPGEDEDQEQR(C)
1752.7770	1752.7653	100.0	6.67		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)

2S protein 2

Index Number: 17

pI of Protein: 6.1

Protein MW: 16492

Amino Acid Composition: A5 C6 D9 E14 F3 G8 H2 I2 K1 L10 M1 N7 P8 Q25 R21 S7 T2 V3 W2 Y2

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR **ANLRPCEEHI** RQRVEQEQQEQ EQDEYPYSQR GSRGRRPGES DEDQEQR**CCN**

81 **ELNRFQNNQR** CMCQALQQIL QNQSF~~R~~FQQD RSQLHQNGEG AQELAPELRV QVTKPLRP

The matched peptides cover 17% (24/138AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
1394.7035	1394.6957	100.0	5.55		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1752.7770	1752.7653	100.0	6.67		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)

Ara h 7.0202

Index Number: 17

pI of Protein: 7.5

Protein MW: 17403

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I2 L10 M2 N8 P8 Q24 R24 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR **ANLRPCEEHI** RQRVEQEQQEQ EQDEYPYSQR GSRGRRPGES DEDQEQR**CCN**

81 **ELNRFQNNQR** CMCQALQQIL QNQSF~~R~~FQQD RSQLHQMERE LR**NLPQNC**GF **RSPSR**CDLSS RTPY

The matched peptides cover 26% (37/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
1105.5258	1105.5207	100.0	4.54		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.7035	1394.6957	100.0	5.55		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1532.7504	1532.7387	100.0	7.63		123	135	1	(R)NLPQNC(Carbamidomethyl)GFRSPSR(C)
1752.7770	1752.7653	100.0	6.67		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)

Spot 18

Ara h 7.0101

Index Number: 18

pI of Protein: 5.6

Protein MW: 16323

Amino Acid Composition: A7 C6 D8 E15 F2 G9 H1 I1 K1 L9 M2 N7 P10 Q21 R20 S8 T2 V5 W3 Y2

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHM RRRVEQEQQ EQDEYPYSRR GSRGRQPGES DENQEQRCCN

81 ELNRFQNNQR CMCQALQQIL QNQSFVWPAG QEPVASDGEG AQELAPELRV QVTKPLRPL

The matched peptides cover 9% (13/139AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
806.3942	806.3904	100.0	4.78		85	90	0	(R)FQNNQR(C)
1752.7822	1752.7653	100.0	9.66		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)

Ara h 7.0201

Index Number: 18

pI of Protein: 7.5

Protein MW: 17375

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I3 K1 L9 M2 N8 P8 Q24 R23 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQRVEKEQQ EQDEYPYIQR GSRGQRPGES DEDQEQRCCN

81 ELNRFQNNQR CMCQALQQIL QNQSFREFQD RSQLHQMERE LRNLQPNCGF RSPSRCDLSS RTPY

The matched peptides cover 60% (86/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
693.3281	693.3315	100.0	-4.81		107	111	0	(R)FQQDR(S)
806.3942	806.3904	100.0	4.78		85	90	0	(R)FQNNQR(C)
1044.4902	1044.4891	100.0	1.08	1Oxidation	112	119	0	(R)SQLHQMER(E)
1105.5278	1105.5207	100.0	6.41		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.7040	1394.6957	100.0	5.91		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)

1442.7300	1442.7169	100.0	9.09	1Oxidation	112	122	1	(R)SQLHQMERELR(N)
1503.7620	1503.7485	100.0	8.95		120	131	1	(R)ELRNLPQNC(Carbamidomethyl)GFR(S)
1532.7489	1532.7387	100.0	6.67		123	135	1	(R)NLPQNC(Carbamidomethyl)GFRSPSR(C)
1630.7145	1630.7052	100.0	5.70		64	77	0	(R)GQRPGESDEDQEQR(C)
1718.8168	1718.8027	100.0	8.17	1Oxidation	107	119	1	(R)FQQDRSQLHQMER(E)
1752.7822	1752.7653	100.0	9.66		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)
1930.8765	1930.8598	100.0	8.64		61	77	1	(R)GSRGQRPGESDEDQEQR(C)
2040.9360	2040.9412	100.0	-2.53	1Oxidation	91	106	0	(R)C(Carbamidomethyl)MC(Carbamidomethyl)QALQQILQNQSFR(F)

2S protein 2

Index Number: 18

pl of Protein: 6.1

Protein MW: 16492

Amino Acid Composition: A5 C6 D9 E14 F3 G8 H2 I2 K1 L10 M1 N7 P8 Q25 R21 S7 T2 V3 W2 Y2

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR **ANLRPCEEHI** **RQRVEQEQQE** EQDEYPYSQR GSRGRRPGES DEDQEQR**CCN**

81 **ELNRFQNNQR** **CMCQALQQIL** **QNQSFRFQQD** **R**SQLHQNGEG AQELAPELRV QVTKPLRP

The matched peptides cover 33% (45/138AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
693.3281	693.3315	100.0	-4.81		107	111	0	(R)FQQDR(S)
806.3942	806.3904	100.0	4.78		85	90	0	(R)FQNNQR(C)
1394.7040	1394.6957	100.0	5.91		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1752.7822	1752.7653	100.0	9.66		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)
2040.9360	2040.9412	100.0	-2.53	1Oxidation	91	106	0	(R)C(Carbamidomethyl)MC(Carbamidomethyl)QALQQILQNQSFR(F)

Ara h 7.0202
Index Number: 18
pI of Protein: 7.5
Protein MW: 17403
Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I2 L10 M2 N8 P8 Q24 R24 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR **ANLRPCEEHI** RQRVEQEQQEQ EQDEYPYSQR GSRGRRPGES DEDQEQR**CCN**
81 **ELNRFQNNQR** **CMCQALQQIL** **QNQSFRFQQD** **RSQLHQMERE** **LRNLPQNCGF** **RSPSR**CDLSS RTPY

The matched peptides cover 48% (69/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
693.3281	693.3315	100.0	-4.81		107	111	0	(R)FQQDR(S)
806.3942	806.3904	100.0	4.78		85	90	0	(R)FQNNQR(C)
1044.4902	1044.4891	100.0	1.08	1Oxidation	112	119	0	(R)SQLHQMER(E)
1105.5278	1105.5207	100.0	6.41		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.7040	1394.6957	100.0	5.91		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1442.7300	1442.7169	100.0	9.09	1Oxidation	112	122	1	(R)SQLHQMERELR(N)
1503.7620	1503.7485	100.0	8.95		120	131	1	(R)ELRNLPQNC(Carbamidomethyl)GFR(S)
1532.7489	1532.7387	100.0	6.67		123	135	1	(R)NLPQNC(Carbamidomethyl)GFRSPSR(C)
1718.8168	1718.8027	100.0	8.17	1Oxidation	107	119	1	(R)FQQDRSQLHQMER(E)
1752.7822	1752.7653	100.0	9.66		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)
2040.9360	2040.9412	100.0	-2.53	1Oxidation	91	106	0	(R)C(Carbamidomethyl)MC(Carbamidomethyl)QALQQILQNQSFR(F)

Spot 19

Ara h 7.0101

Index Number: 19

pI of Protein: 5.6

Protein MW: 16323

Amino Acid Composition: A7 C6 D8 E15 F2 G9 H1 I1 K1 L9 M2 N7 P10 Q21 R20 S8 T2 V5 W3 Y2

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHM RRRVEQEQQ EQDEYPYSRR GSRGRQPGES DENQEQRCCN

81 ELNR**FQNNQR** CMCQALQQIL QNQSFVWPAG QEPVASDGEG AQELAPELRV QVTKPLRPL

The matched peptides cover 4% (6/139AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
806.3928	806.3904	100.0	3.03		85	90	0	(R)FQNNQR(C)

Ara h 7.0201

Index Number: 19

pI of Protein: 7.5

Protein MW: 17375

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I3 K1 L9 M2 N8 P8 Q24 R23 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR **ANLRPCEEHI** RQRVEKEQQ EQDEYPYIQR GSR**GQRPGES** **DEDQEQR**CCN

81 ELNR**FQNNQR** CMCQALQQIL QNQSF**RFQOD** **RSQ**LHQMER**E** LR**NLPQNCGF** **R**SPSRCDLSS **R**TPY

The matched peptides cover 28% (40/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
806.3928	806.3904	100.0	3.03		85	90	0	(R)FQNNQR(C)
1105.5288	1105.5207	100.0	7.30		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.7033	1394.6957	100.0	5.38		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1630.7108	1630.7052	100.0	3.46		64	77	0	(R)GQRPGESEDEDQEQR(C)

2S protein 2

Index Number: 19

pI of Protein: 6.1

Protein MW: 16492

Amino Acid Composition: A5 C6 D9 E14 F3 G8 H2 I2 K1 L10 M1 N7 P8 Q25 R21 S7 T2 V3 W2 Y2

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR **ANLRPCEEHI** RQRVEQEQQEQ EQDEYPYSQR GSRGRRPGES DEDQEQRCCN

81 ELNR**FQNNQR** CMCQALQQIL QNQSFRFQQD RSQLHQNGEG AQELAPELRV QVTKPLRP

The matched peptides cover **12%** (17/138AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
806.3928	806.3904	100.0	3.03		85	90	0	(R)FQNNQR(C)
1394.7033	1394.6957	100.0	5.38		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)

Ara h 7.0202

Index Number: 19

pI of Protein: 7.5

Protein MW: 17403

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I2 L10 M2 N8 P8 Q24 R24 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR **ANLRPCEEHI** RQRVEQEQQEQ EQDEYPYSQR GSRGRRPGES DEDQEQRCCN

81 ELNR**FQNNQR** CMCQALQQIL QNQSFRFQQD RSQLHQMERE LR**NLPQNCGF** RSPSRCDLSS RTPY

The matched peptides cover **18%** (26/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
806.3928	806.3904	100.0	3.03		85	90	0	(R)FQNNQR(C)
1105.5288	1105.5207	100.0	7.30		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.7033	1394.6957	100.0	5.38		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)

Spot 20

Ara h 7.0101

Index Number: 20

pI of Protein: 5.6

Protein MW: 16323

Amino Acid Composition: A7 C6 D8 E15 F2 G9 H1 I1 K1 L9 M2 N7 P10 Q21 R20 S8 T2 V5 W3 Y2

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHM RRRVEQEQQ EQDEYPYSRR GSRGRQPGES DENQEQRCCN

81 ELNRFQNNQR CMCQALQQIL QNQSFVWPAG QEPVASDGEG AQELAPELRV QVTKPLRPL

The matched peptides cover 12% (17/139AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
527.2967	527.2936	100.0	5.82	1Gln->pyro-Glu	27	30	0	(R)QLQR(A)
544.3127	544.3202	100.0	-13.7		27	30	0	(R)QLQR(A)
806.3926	806.3904	100.0	2.73		85	90	0	(R)FQNNQR(C)
1752.7773	1752.7653	100.0	6.87		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)

Ara h 7.0201

Index Number: 20

pI of Protein: 7.5

Protein MW: 17375

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I3 K1 L9 M2 N8 P8 Q24 R23 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQVEKEQQ EQDEYPYIQR GSRGQRPGES DEDQEQRCCN

81 ELNRFQNNQR CMCQALQQIL QNQSFQFD RSQHQMERE LRNLQNCGF RSPSRCDLSS RTPY

The matched peptides cover 51% (74/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
527.2967	527.2936	100.0	5.82	1Gln->pyro-Glu	27	30	0	(R)QLQR(A)
544.3127	544.3202	100.0	-13.7		27	30	0	(R)QLQR(A)
693.3352	693.3315	100.0	5.40		107	111	0	(R)FQQDR(S)

806.3926	806.3904	100.0	2.73		85	90	0	(R)FQNNQR(C)
1044.4894	1044.4891	100.0	0.264	1Oxidation	112	119	0	(R)SQLHQMER(E)
1105.5259	1105.5207	100.0	4.65		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.7020	1394.6957	100.0	4.51		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1442.7305	1442.7169	100.0	9.43	1Oxidation	112	122	1	(R)SQLHQMERELR(N)
1503.7557	1503.7485	100.0	4.81		120	131	1	(R)ELRNLPQNC(Carbamidomethyl)GFR(S)
1532.7529	1532.7387	100.0	9.30		123	135	1	(R)NLPQNC(Carbamidomethyl)GFRSPSR(C)
1630.7145	1630.7052	100.0	5.70		64	77	0	(R)GQRPGESDEDQEQR(C)
1718.8140	1718.8027	100.0	6.54	1Oxidation	107	119	1	(R)FQQDRSQLHQMER(E)
1752.7773	1752.7653	100.0	6.87		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)
1930.8719	1930.8598	100.0	6.30		61	77	1	(R)GSRGQRPGESDEDQEQR(C)

2S protein 2

Index Number: 20

pl of Protein: 6.1

Protein MW: 16492

Amino Acid Composition: A5 C6 D9 E14 F3 G8 H2 I2 K1 L10 M1 N7 P8 Q25 R21 S7 T2 V3 W2 Y2

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQRVEQEQQ EQDEYPYSQR GSRGRRPGES DEDQEQRCCN

81 ELNRFQNNQR CMCQALQQIL QNQSFRFQQD RSQLHQNGEG AQELAPELRV QVTKPLRP

The matched peptides cover 24% (33/138AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
527.2967	527.2936	100.0	5.82	1Gln->pyro-Glu	27	30	0	(R)QLQR(A)
544.3127	544.3202	100.0	-13.7		27	30	0	(R)QLQR(A)
693.3352	693.3315	100.0	5.40		107	111	0	(R)FQQDR(S)
806.3926	806.3904	100.0	2.73		85	90	0	(R)FQNNQR(C)
1394.7020	1394.6957	100.0	4.51		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1752.7773	1752.7653	100.0	6.87		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)

Ara h 7.0202
Index Number: 20
pI of Protein: 7.5
Protein MW: 17403
Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I2 L10 M2 N8 P8 Q24 R24 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQVEQEQQEQ EQDEYPYSQR GSRGRRPGES DEDQEQRCCN
81 ELNRFQNNQR CMCQALQQIL QNQSFRRQQD RSQLHQMERE LRNLQPQNCGF RSPSRCDLSS RTPY

The matched peptides cover 40% (57/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
527.2967	527.2936	100.0	5.82	1Gln->pyro-Glu	27	30	0	(R)QLQR(A)
544.3127	544.3202	100.0	-13.7		27	30	0	(R)QLQR(A)
693.3352	693.3315	100.0	5.40		107	111	0	(R)FQQDR(S)
806.3926	806.3904	100.0	2.73		85	90	0	(R)FQNNQR(C)
1044.4894	1044.4891	100.0	0.264	1Oxidation	112	119	0	(R)SQLHQMER(E)
1105.5259	1105.5207	100.0	4.65		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.7020	1394.6957	100.0	4.51		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1442.7305	1442.7169	100.0	9.43	1Oxidation	112	122	1	(R)SQLHQMERELR(N)
1503.7557	1503.7485	100.0	4.81		120	131	1	(R)ELRNLPQNC(Carbamidomethyl)GFR(S)
1532.7529	1532.7387	100.0	9.30		123	135	1	(R)NLPQNC(Carbamidomethyl)GFRSPSR(C)
1718.8140	1718.8027	100.0	6.54	1Oxidation	107	119	1	(R)FQQDRSQLHQMER(E)
1752.7773	1752.7653	100.0	6.87		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)

Spot 21

Ara h 7.0101

Index Number: 21

pI of Protein: 5.6

Protein MW: 16323

Amino Acid Composition: A7 C6 D8 E15 F2 G9 H1 I1 K1 L9 M2 N7 P10 Q21 R20 S8 T2 V5 W3 Y2

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHM RRRVEQEQQE EQDEYPYSRR GSRGRQPGES DENQEQRCCN
81 ELNRFQNNQR CMCQALQQIL QNQSFVWPAG QEPVASDGEG AQELAPELRV QVTKPLRPL

The matched peptides cover 0% (0/139AA's) of the protein.
No hits.

Ara h 7.0201

Index Number: 21

pI of Protein: 7.5

Protein MW: 17375

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I3 K1 L9 M2 N8 P8 Q24 R23 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQRVEKEQQE EQDEYPYIQR GSRGRQPGES DEDQEQRCCN
81 ELNRFQNNQR CMCQALQQIL QNQSFVFQOD RSQ LHQMERE LRNL PQNCGF RSPSRCDLSS RTPY

The matched peptides cover 14% (20/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
1105.5295	1105.5207	100.0	7.96		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.7079	1394.6957	100.0	8.71		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)

2S protein 2

Index Number: 21

pI of Protein: 6.1

Protein MW: 16492

Amino Acid Composition: A5 C6 D9 E14 F3 G8 H2 I2 K1 L10 M1 N7 P8 Q25 R21 S7 T2 V3 W2 Y2

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR **ANLRPCEEHI** RQRVEQEQQEQ EQDEYPYSQR GSR**GRRPGES** **DEDQEQR**CCN

81 ELNRFQNNQR CMCQALQQIL QNQSF~~R~~FQQD RSQLHQNGEG AQELAPELRV QVTKPLRP

The matched peptides cover 18% (25/138AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
1394.7079	1394.6957	100.0	8.71		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1658.7698	1658.7477	100.0	13.3		64	77	1	(R)GRRPGESDEDQEQR(C)

Ara h 7.0202

Index Number: 21

pI of Protein: 7.5

Protein MW: 17403

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I2 L10 M2 N8 P8 Q24 R24 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR **ANLRPCEEHI** RQRVEQEQQEQ EQDEYPYSQR GSR**GRRPGES** **DEDQEQR**CCN

81 ELNRFQNNQR CMCQALQQIL QNQSF~~R~~FQQD RSQLHQMERE LR**NLPQNC**GF RSPSRCDLSS RTPY

The matched peptides cover 24% (34/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
1105.5295	1105.5207	100.0	7.96		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.7079	1394.6957	100.0	8.71		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1658.7698	1658.7477	100.0	13.3		64	77	1	(R)GRRPGESDEDQEQR(C)

Spot 22

Ara h 7.0101

Index Number: 22

pI of Protein: 5.6

Protein MW: 16323

Amino Acid Composition: A7 C6 D8 E15 F2 G9 H1 I1 K1 L9 M2 N7 P10 Q21 R20 S8 T2 V5 W3 Y2

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHM RRRVEQEQQ EQDEYPYSRR GSRGRQPGES DENQEQRCCN
81 ELNRFQNNQR CMCQALQQIL QNQSFVWPAG QEPVASDGEG AQELAPELRV QVTKPLRPL

The matched peptides cover 9% (13/139AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
806.3909	806.3904	100.0	0.687		85	90	0	(R)FQNNQR(C)
1752.7717	1752.7653	100.0	3.67		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)

Ara h 7.0201

Index Number: 22

pI of Protein: 7.5

Protein MW: 17375

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I3 K1 L9 M2 N8 P8 Q24 R23 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQRVEKEQQ EQDEYPYIQR GSRGRQPGES DEDQEQRCCN
81 ELNRFQNNQR CMCQALQQIL QNQSFRFQQD RSQLHQMERE LRNLQPNCGF RSPSRCDLSS RTPY

The matched peptides cover 34% (49/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
806.3909	806.3904	100.0	0.687		85	90	0	(R)FQNNQR(C)
1105.5262	1105.5207	100.0	4.99		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.7019	1394.6957	100.0	4.41		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1752.7717	1752.7653	100.0	3.67		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)
2812.3176	2812.3188	100.0	-0.427		85	106	1	(R)FQNNQRC(Carbamidomethyl)MC(Carbamidomethyl)QALQQILQNQSFR(F)

2S protein 2

Index Number: 22

pI of Protein: 6.1

Protein MW: 16492

Amino Acid Composition: A5 C6 D9 E14 F3 G8 H2 I2 K1 L10 M1 N7 P8 Q25 R21 S7 T2 V3 W2 Y2

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQVEQEQQEQ EQDEYPYSQR GSRGRRPGES DEDQEQRCCN

81 ELNRFQNNQR CMCQALQQIL QNQSFRFQQD RSQLHQNGEG AQELAPELRV QVTKPLRP

The matched peptides cover 53% (73/138AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
806.3909	806.3904	100.0	0.687		85	90	0	(R)FQNNQR(C)
1394.7019	1394.6957	100.0	4.41		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1658.7606	1658.7477	100.0	7.78		64	77	1	(R)GRRPGESDEDQEQR(C)
1752.7717	1752.7653	100.0	3.67		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)
2184.9470	2184.9316	100.0	7.05		44	60	0	(R)VEQEQQEQEQDEYPYSQR(G)
2469.1121	2469.0913	100.0	8.41		42	60	1	(R)QVEQEQQEQEQDEYPYSQR(G)
2812.3176	2812.3188	100.0	-0.427		85	106	1	(R)FQNNQRC(Carbamidomethyl)MC(Carbamidomethyl)QALQQILQNQSF(R)

Ara h 7.0202

Index Number: 22

pI of Protein: 7.5

Protein MW: 17403

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I2 L10 M2 N8 P8 Q24 R24 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQVEQEQQEQ EQDEYPYSQR GSRGRRPGES DEDQEQRCCN

81 ELNRFQNNQR CMCQALQQIL QNQSFRFQQD RSQLHQMERE LRNLQPQNCGF RSPSRCDLSS RTPY

The matched peptides cover 57% (82/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
806.3909	806.3904	100.0	0.687		85	90	0	(R)FQNNQR(C)
1105.5262	1105.5207	100.0	4.99		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)

1394.7019	1394.6957	100.0	4.41	31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1658.7606	1658.7477	100.0	7.78	64	77	1	(R)GRRPGESDEDQEQR(C)
1752.7717	1752.7653	100.0	3.67	78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)
2184.9470	2184.9316	100.0	7.05	44	60	0	(R)VEQEQQEQDEYPYSQR(G)
2469.1121	2469.0913	100.0	8.41	42	60	1	(R)QRVEQEQQEQDEYPYSQR(G)
2812.3176	2812.3188	100.0	-0.427	85	106	1	(R)FQNNQRC(Carbamidomethyl)MC(Carbamidomethyl)QALQQILQN QSFR(F)

Spot 23

Ara h 7.0101

Index Number: 23

pI of Protein: 5.6

Protein MW: 16323

Amino Acid Composition: A7 C6 D8 E15 F2 G9 H1 I1 K1 L9 M2 N7 P10 Q21 R20 S8 T2 V5 W3 Y2

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHM RRRVEQEQQ EQDEYPYSRR GSRGRQPGES DENQEQRCCN
81 ELNRFQNNQR CMCQALQQIL QNQSFVWPAG QEPVASDGEG AQELAPELRV QVTKPLRPL

The matched peptides cover 17% (24/139AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
527.2880	527.2936	100.0	-10.7	1Gln->pyro-Glu	27	30	0	(R)QLQR(A)
544.3146	544.3202	100.0	-10.3		27	30	0	(R)QLQR(A)
806.3904	806.3904	100.0	0.0794		85	90	0	(R)FQNNQR(C)
878.3664	878.3421	100.0	27.6		20	26	0	(R)GDDQC(Carbamidomethyl)QR(Q)
965.3942	965.3928	100.0	1.52		78	84	0	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNR(F)
1752.7730	1752.7653	100.0	4.37		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)

Ara h 7.0201
Index Number: 23
pI of Protein: 7.5
Protein MW: 17375
Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I3 K1 L9 M2 N8 P8 Q24 R23 S11 T2 V1 W2 Y3
1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQRVEKEQE EQDEYPYIQR GSRGQRPGES DEDQEQRCCN
81 ELNRFQNNQR CMCQALQQIL QNQSFRFQQD RSQLHQMERE LRNLQPQNCGF RSPSRCDLSS RTPY

The matched peptides cover 61% (88/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
527.2880	527.2936	100.0	-10.7	1Gln->pyro-Glu	27	30	0	(R)QLQR(A)
544.3146	544.3202	100.0	-10.3		27	30	0	(R)QLQR(A)
693.3305	693.3315	100.0	-1.38		107	111	0	(R)FQQDR(S)
737.3273	737.3247	100.0	3.53		136	141	0	(R)C(Carbamidomethyl)DLSSR(T)
806.3904	806.3904	100.0	0.0794		85	90	0	(R)FQNNQR(C)
878.3664	878.3421	100.0	27.6		20	26	0	(R)GDDQC(Carbamidomethyl)QR(Q)
965.3942	965.3928	100.0	1.52		78	84	0	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNR(F)
1028.5000	1028.4942	100.0	5.65		112	119	0	(R)SQLHQMER(E)
1044.4882	1044.4891	100.0	-0.904	1Oxidation	112	119	0	(R)SQLHQMER(E)
1105.5267	1105.5207	100.0	5.42		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1164.5487	1164.5426	100.0	5.25		132	141	1	(R)SPSRC(Carbamidomethyl)DLSSR(T)
1394.7017	1394.6957	100.0	4.24		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1426.7230	1426.7220	100.0	0.745		112	122	1	(R)SQLHQMERELR(N)
1442.7188	1442.7169	100.0	1.30	1Oxidation	112	122	1	(R)SQLHQMERELR(N)
1503.7535	1503.7485	100.0	3.35		120	131	1	(R)ELRNLPQNC(Carbamidomethyl)GFR(S)
1532.7449	1532.7387	100.0	4.05		123	135	1	(R)NLPQNC(Carbamidomethyl)GFRSPSR(C)
1678.8628	1678.8554	100.0	4.38		31	43	1	(R)ANLRPC(Carbamidomethyl)EEHIRQR(V)
1718.8079	1718.8027	100.0	2.99	1Oxidation	107	119	1	(R)FQQDRSQLHQMER(E)
1752.7730	1752.7653	100.0	4.37		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)
2040.9324	2040.9412	100.0	-4.33	1Oxidation	91	106	0	(R)C(Carbamidomethyl)MC(Carbamidomethyl)QALQQILQNQSFR(F)

2S protein 2

Index Number: 23

pI of Protein: 6.1

Protein MW: 16492

Amino Acid Composition: A5 C6 D9 E14 F3 G8 H2 I2 K1 L10 M1 N7 P8 Q25 R21 S7 T2 V3 W2 Y2

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQRVEQEQQE EQDEYPYSQR GSRGRRPGES DEDQEQRCCN

81 ELNRFQNNQR CMCQALQQIL QNQSFRFQQD RSQLHQNGEG AQELAPELRV QVTKPLRP

The matched peptides cover 64% (89/138AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
527.2880	527.2936	100.0	-10.7	1Gln->pyro-Glu	27	30	0	(R)QLQR(A)
544.3146	544.3202	100.0	-10.3		27	30	0	(R)QLQR(A)
693.3305	693.3315	100.0	-1.38		107	111	0	(R)FQQDR(S)
806.3904	806.3904	100.0	0.0794		85	90	0	(R)FQNNQR(C)
878.3664	878.3421	100.0	27.6		20	26	0	(R)GDDQC(Carbamidomethyl)QR(Q)
965.3942	965.3928	100.0	1.52		78	84	0	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNR(F)
1394.7017	1394.6957	100.0	4.24		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1445.6465	1445.6251	100.0	14.8		66	77	0	(R)RPGESDEDQEQR(C)
1658.7551	1658.7477	100.0	4.47		64	77	1	(R)GRRPGESDEDQEQR(C)
1678.8628	1678.8554	100.0	4.38		31	43	1	(R)ANLRPC(Carbamidomethyl)EEHIRQR(V)
1752.7730	1752.7653	100.0	4.37		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)
2040.9324	2040.9412	100.0	-4.33	1Oxidation	91	106	0	(R)C(Carbamidomethyl)MC(Carbamidomethyl)QALQQILQNQSFR(F)
2184.9311	2184.9316	100.0	-0.211		44	60	0	(R)VEQEQQEQQEYPYSQR(G)
2452.0810	2452.0648	100.0	6.65	1Gln->pyro-Glu	42	60	1	(R)QRVEQEQQEQQEYPYSQR(G)
2469.1050	2469.0913	100.0	5.54		42	60	1	(R)QRVEQEQQEQQEYPYSQR(G)

Ara h 7.0202

Index Number: 23

pI of Protein: 7.5

Protein MW: 17403

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I2 L10 M2 N8 P8 Q24 R24 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQRVEQEQQ EQDEYPYSQR GSRGRRPGES DEDQEQRCCN

81 ELNRFQNNQR CMCQALQQIL QNQSFRFQQD RSQLHQMERE LRNLQPQNCGF RSPSRCDLSS RTPY

The matched peptides cover **83%** (119/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
527.2880	527.2936	100.0	-10.7	1Gln->pyro-Glu	27	30	0	(R)QLQR(A)
544.3146	544.3202	100.0	-10.3		27	30	0	(R)QLQR(A)
693.3305	693.3315	100.0	-1.38		107	111	0	(R)FQQDR(S)
737.3273	737.3247	100.0	3.53		136	141	0	(R)C(Carbamidomethyl)DLSSR(T)
806.3904	806.3904	100.0	0.0794		85	90	0	(R)FQNNQR(C)
878.3664	878.3421	100.0	27.6		20	26	0	(R)GDDQC(Carbamidomethyl)QR(Q)
965.3942	965.3928	100.0	1.52		78	84	0	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNR(F)
1028.5000	1028.4942	100.0	5.65		112	119	0	(R)SQLHQMER(E)
1044.4882	1044.4891	100.0	-0.904	1Oxidation	112	119	0	(R)SQLHQMER(E)
1105.5267	1105.5207	100.0	5.42		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1164.5487	1164.5426	100.0	5.25		132	141	1	(R)SPSRC(Carbamidomethyl)DLSSR(T)
1394.7017	1394.6957	100.0	4.24		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1426.7230	1426.7220	100.0	0.745		112	122	1	(R)SQLHQMERELR(N)
1442.7188	1442.7169	100.0	1.30	1Oxidation	112	122	1	(R)SQLHQMERELR(N)
1445.6465	1445.6251	100.0	14.8		66	77	0	(R)RPGESDEDQEQR(C)
1503.7535	1503.7485	100.0	3.35		120	131	1	(R)ELRNLPQNC(Carbamidomethyl)GFR(S)
1532.7449	1532.7387	100.0	4.05		123	135	1	(R)NLPQNC(Carbamidomethyl)GFRSPSR(C)
1658.7551	1658.7477	100.0	4.47		64	77	1	(R)GRRPGESDEDQEQR(C)
1678.8628	1678.8554	100.0	4.38		31	43	1	(R)ANLRPC(Carbamidomethyl)EEHIRQR(V)
1718.8079	1718.8027	100.0	2.99	1Oxidation	107	119	1	(R)FQQDRSQLHQMER(E)
1752.7730	1752.7653	100.0	4.37		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)

2040.9324	2040.9412	100.0	-4.33	1Oxidation	91	106	0	(R)C(Carbamidomethyl)MC(Carbamidomethyl)QALQQILQNQSFR(F)
2184.9311	2184.9316	100.0	-0.211		44	60	0	(R)VEQEQQEQDEYPYSQR(G)
2452.0810	2452.0648	100.0	6.65	1Gln->pyro-Glu	42	60	1	(R)QRVEQEQQEQDEYPYSQR(G)
2469.1050	2469.0913	100.0	5.54		42	60	1	(R)QRVEQEQQEQDEYPYSQR(G)

Spot 24

Ara h 7.0101

Index Number: 24

pI of Protein: 5.6

Protein MW: 16323

Amino Acid Composition: A7 C6 D8 E15 F2 G9 H1 I1 K1 L9 M2 N7 P10 Q21 R20 S8 T2 V5 W3 Y2

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHM RRRVEQEQQEQ EQDEYPYSRR GSRGRQPGES DENQEQRCCN
81 ELNRFQNNQR CMCQALQQIL QNQSFVWPAG QEPVASDGEG AQELAPELRV QVTKPLRP

The matched peptides cover 24% (33/139AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
527.2898	527.2936	100.0	-7.27	1Gln->pyro-Glu	27	30	0	(R)QLQR(A)
544.3166	544.3202	100.0	-6.58		27	30	0	(R)QLQR(A)
806.3903	806.3904	100.0	-0.144		85	90	0	(R)FQNNQR(C)
965.4042	965.3928	100.0	11.8		78	84	0	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNR(F)
1752.7759	1752.7653	100.0	6.04		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)
2056.9402	2056.8730	100.0	32.6		44	59	0	(R)VEQEQQEQDEYPYSR(R)

Ara h 7.0201
Index Number: 24
pI of Protein: 7.5
Protein MW: 17375
Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I3 K1 L9 M2 N8 P8 Q24 R23 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQRVEKEQEQ EQDEYPYIQR GSRGQRPGES DEDQEQRCCN
81 ELNRFQNNQR CMCQALQQIL QNQSFRFQQD RSQLHQMERE LRNLQPQNCGF RSPSRCDLSS RTPY

The matched peptides cover 52% (75/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
527.2898	527.2936	100.0	-7.27	1Gln->pyro-Glu	27	30	0	(R)QLQR(A)
544.3166	544.3202	100.0	-6.58		27	30	0	(R)QLQR(A)
693.3279	693.3315	100.0	-5.17		107	111	0	(R)FQQDR(S)
806.3903	806.3904	100.0	-0.144		85	90	0	(R)FQNNQR(C)
965.4042	965.3928	100.0	11.8		78	84	0	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNR(F)
1044.4871	1044.4891	100.0	-1.96	1Oxidation	112	119	0	(R)SQLHQMER(E)
1105.5244	1105.5207	100.0	3.32		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.7029	1394.6957	100.0	5.12		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1442.7245	1442.7169	100.0	5.28	1Oxidation	112	122	1	(R)SQLHQMERELR(N)
1503.7562	1503.7485	100.0	5.14		120	131	1	(R)ELRNLPQNC(Carbamidomethyl)GFR(S)
1532.7488	1532.7387	100.0	6.60		123	135	1	(R)NLPQNC(Carbamidomethyl)GFRSPSR(C)
1678.8634	1678.8554	100.0	4.75		31	43	1	(R)ANLRPC(Carbamidomethyl)EEHIRQR(V)
1718.8114	1718.8027	100.0	5.05	1Oxidation	107	119	1	(R)FQQDRSQLHQMER(E)
1752.7759	1752.7653	100.0	6.04		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)
2040.9370	2040.9412	100.0	-2.06	1Oxidation	91	106	0	(R)C(Carbamidomethyl)MC(Carbamidomethyl)QALQQILQNQSFR(F))

2S protein 2
Index Number: 24
pI of Protein: 6.1
Protein MW: 16492
Amino Acid Composition: A5 C6 D9 E14 F3 G8 H2 I2 K1 L10 M1 N7 P8 Q25 R21 S7 T2 V3 W2 Y2

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQRVEQEQQE EQDEYPYSQR GSRGRRPGES DEDQEQRCCN
81 ELNRFQNNQR CMCQALQQIL QNQSFRFQQD RSQLHQNGEG AQELAPELRV QVTKPLRP

The matched peptides cover 59% (82/138AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
527.2898	527.2936	100.0	-7.27	1Gln->pyro-Glu	27	30	0	(R)QLQR(A)
544.3166	544.3202	100.0	-6.58		27	30	0	(R)QLQR(A)
693.3279	693.3315	100.0	-5.17		107	111	0	(R)FQQDR(S)
806.3903	806.3904	100.0	-0.144		85	90	0	(R)FQNNQR(C)
965.4042	965.3928	100.0	11.8		78	84	0	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNR(F)
1394.7029	1394.6957	100.0	5.12		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1658.7546	1658.7477	100.0	4.18		64	77	1	(R)GRRPGESDEDQEQR(C)
1678.8634	1678.8554	100.0	4.75		31	43	1	(R)ANLRPC(Carbamidomethyl)EEHIRQR(V)
1752.7759	1752.7653	100.0	6.04		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)
2040.9370	2040.9412	100.0	-2.06	1Oxidation	91	106	0	(R)C(Carbamidomethyl)MC(Carbamidomethyl)QALQQILQNQSFR(F))
2184.9382	2184.9316	100.0	3.03		44	60	0	(R)VEQEQQEQDEYPYSQR(G)
2452.0779	2452.0648	100.0	5.35	1Gln->pyro-Glu	42	60	1	(R)QRVEQEQQEQDEYPYSQR(G)
2469.1020	2469.0913	100.0	4.35		42	60	1	(R)QRVEQEQQEQDEYPYSQR(G)

Ara h 7.0202

Index Number: 24

pI of Protein: 7.5

Protein MW: 17403

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I2 L10 M2 N8 P8 Q24 R24 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQRVEQEQQE EQDEYPYSQR GSRGRRPGES DEDQEQRCCN

81 ELNRFQNNQR CMCQALQQIL QNQSFRFQQD RSQLHQMERE LRNLQPNCGF RSPSRCDLSS RTPY

The matched peptides cover 74% (106/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
527.2898	527.2936	100.0	-7.27	1Gln->pyro-Glu	27	30	0	(R)QLQR(A)
544.3166	544.3202	100.0	-6.58		27	30	0	(R)QLQR(A)
693.3279	693.3315	100.0	-5.17		107	111	0	(R)FQQDR(S)
806.3903	806.3904	100.0	-0.144		85	90	0	(R)FQNNQR(C)
965.4042	965.3928	100.0	11.8		78	84	0	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNR(F)
1044.4871	1044.4891	100.0	-1.96	1Oxidation	112	119	0	(R)SQLHQMER(E)
1105.5244	1105.5207	100.0	3.32		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.7029	1394.6957	100.0	5.12		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1442.7245	1442.7169	100.0	5.28	1Oxidation	112	122	1	(R)SQLHQMERELR(N)
1503.7562	1503.7485	100.0	5.14		120	131	1	(R)ELRNLPQNC(Carbamidomethyl)GFR(S)
1532.7488	1532.7387	100.0	6.60		123	135	1	(R)NLPQNC(Carbamidomethyl)GFRSPSR(C)
1658.7546	1658.7477	100.0	4.18		64	77	1	(R)GRRPGESDEDQEQR(C)
1678.8634	1678.8554	100.0	4.75		31	43	1	(R)ANLRPC(Carbamidomethyl)EEHIRQR(V)
1718.8114	1718.8027	100.0	5.05	1Oxidation	107	119	1	(R)FQQDRSQLHQMER(E)
1752.7759	1752.7653	100.0	6.04		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)
2040.9370	2040.9412	100.0	-2.06	1Oxidation	91	106	0	(R)C(Carbamidomethyl)MC(Carbamidomethyl)QALQQILQNQSFR(F))
2184.9382	2184.9316	100.0	3.03		44	60	0	(R)VEQEQQEQQEYYPYSQR(G)
2452.0779	2452.0648	100.0	5.35	1Gln->pyro-Glu	42	60	1	(R)QRVEQEQQEQQEYYPYSQR(G)
2469.1020	2469.0913	100.0	4.35		42	60	1	(R)QRVEQEQQEQQEYYPYSQR(G)

Spot 166, exp #2 is equivalent to spot 20

Ara h 7.0201

Index Number: 1

pI of Protein: 7.5

Protein MW: 17375

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I3 K1 L9 M2 N8 P8 Q24 R23 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQRVEKEQEQ EQDEYPYIQR GSRGQRPGES DEDQEQRCCN
81 ELNRFQNNQR CMCQALQQIL QNQSFRRFQQD RSQLHQMERE LRNLPQNCGF RSPSRCDLSS RTPY

The matched peptides cover 28% (40/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta* ppm	Modificatio ns	Start	End	Missed Cleavages	Sequence
806.3426	806.3904	100.0	-59.3		85	90	0	(R)FQNNQR(C)
1105.4513	1105.5207	100.0	-62.8		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.6024	1394.6957	100.0	-66.9		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1630.5985	1630.7052	100.0	-65.4		64	77	0	(R)GQRPGESDEDQEQR(C)

*exception: mass tolerance was set to 100 ppm

Spot 360, exp #3 is equivalent to spot 20

Ara h 7.0201

Index Number: 360

pI of Protein: 7.5

Protein MW: 17375

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I3 K1 L9 M2 N8 P8 Q24 R23 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GSRWDAPSRG DDQCQRQLQR **ANLRPCEEHI** **RQRVEKEQEQ** **EQDEYPYIQR** GSR**GQRPGES** **DEDQEQR**CCN
81 ELNR**FQNNQR** CMCQALQQIL QNQSFRRFQQD RSQLHQMERE LR**NLPQNCGF** **RSPSRCDLSS** RTPY

The matched peptides cover 40% (57/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
806.3759	806.3904	100.0	-18.0		85	90	0	(R)FQNNQR(C)
1105.5021	1105.5207	100.0	-16.9		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.6766	1394.6957	100.0	-13.7		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1630.6800	1630.7052	100.0	-15.4		64	77	0	(R)GQRPGESDEDQEQR(C)
2210.9690	2211.0200	100.0	-23.1		44	60	1	(R)VEKEQEQQDEYPYIQR(G)

Spot 160, exp. #2 is equivalent to spot 23

Ara h 7.0202

Index Number: 160

pI of Protein: 7.5

Protein MW: 17403

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I2 L10 M2 N8 P8 Q24 R24 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR **ANLRPCEEHI** **RQRVEQEQQE** **EQDEYPYSQR** GSR**GRRPGES** **DEDQEQRCCN**

81 **ELNRFQNNQR** CMCQALQQIL QNQSFRRFQQD RSQLHQMERE LR**NLPQNC**GF **RSPSR**CDLSS RTPY

The matched peptides cover 49% (70/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
806.3791	806.3904	100.0	-14.0		85	90	0	(R)FQNNQR(C)
1105.5075	1105.5207	100.0	-12.0		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.6749	1394.6957	100.0	-14.9		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1532.7154	1532.7387	100.0	-15.1		123	135	1	(R)NLPQNC(Carbamidomethyl)GFRSPSR(C)
1658.7224	1658.7477	100.0	-15.3		64	77	1	(R)GRRPGESDEDQEQR(C)
1752.7419	1752.7653	100.0	-13.3		78	90	1	(R)C(Carbamidomethyl)C(Carbamidomethyl)NELNRFQNNQR(C)
2184.8999	2184.9316	100.0	-14.5		44	60	0	(R)VEQEQQEQQEYPYSQR(G)
2469.0689	2469.0913	100.0	-9.09		42	60	1	(R)QRVEQEQQEQQEYPYSQR(G)

Spot 365, exp. #3 is equivalent to spot 23

Ara h 7.0202

Index Number: 365

pI of Protein: 7.5

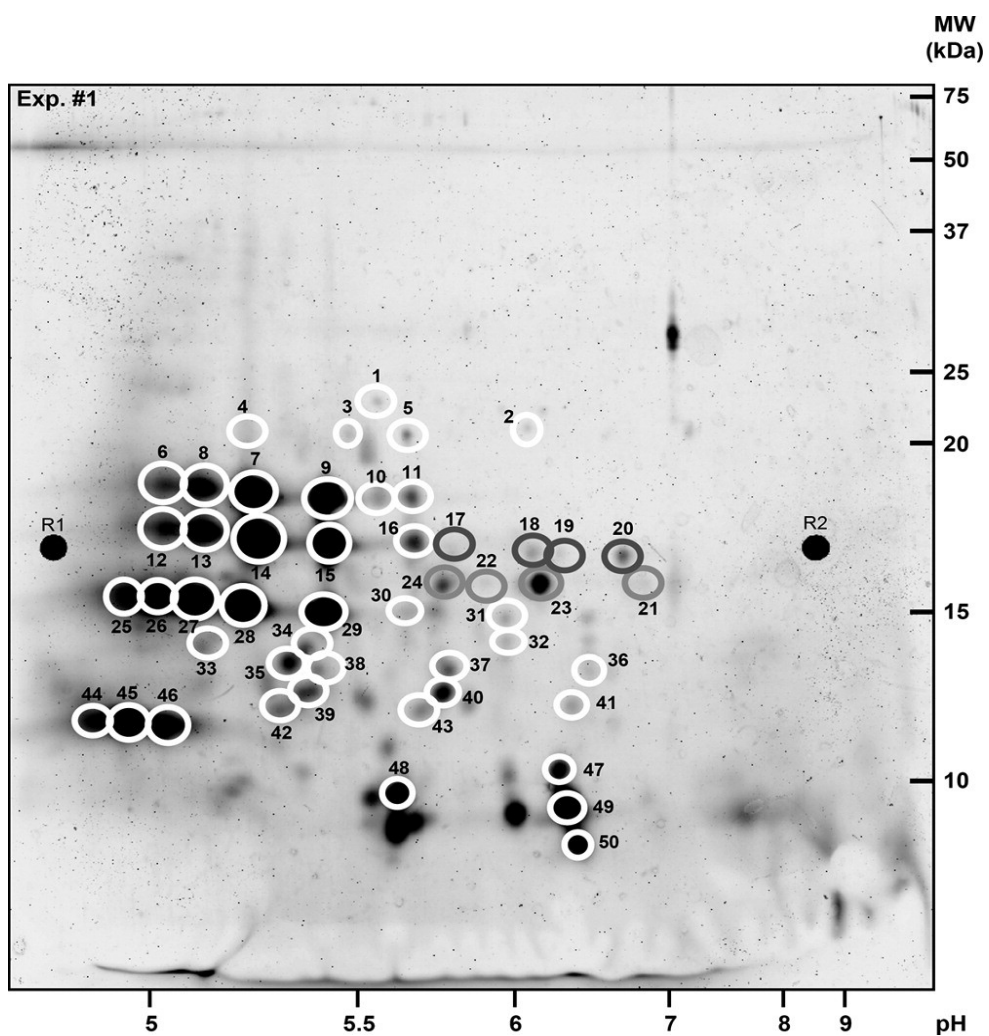
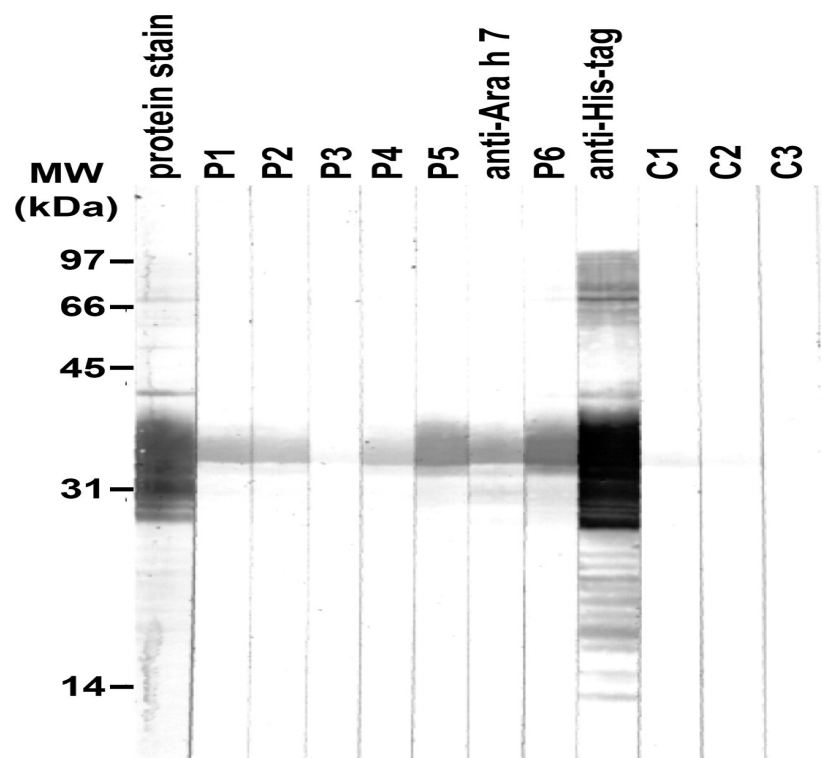
Protein MW: 17403

Amino Acid Composition: A3 C8 D10 E13 F4 G7 H2 I2 L10 M2 N8 P8 Q24 R24 S11 T2 V1 W2 Y3

1 TRWDPDRGSR GLRWDAPSRG DDQCQRQLQR ANLRPCEEHI RQRVEQEQQE EQDEYPYSQR GSRGRRPGES DEDQEQRCCN
81 ELNR FQNNQR CMCQALQQIL QNQSFR FQQD RSQLHQMERE LRNLQPQNCGF RSPSRCDLSS RTPY

The matched peptides cover 52% (75/144AA's) of the protein.

m/z Submitted	MH ⁺ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
806.3796	806.3904	100.0	-13.4		85	90	0	(R)FQNNQR(C)
1105.5010	1105.5207	100.0	-17.9		123	131	0	(R)NLPQNC(Carbamidomethyl)GFR(S)
1394.6748	1394.6957	100.0	-15.0		31	41	0	(R)ANLRPC(Carbamidomethyl)EEHIR(Q)
1658.7214	1658.7477	100.0	-15.8		64	77	1	(R)GRRPGESDEDQEQR(C)
2040.8994	2040.9412	100.0	-20.5	1Oxidation	91	106	0	(R)C(Carbamidomethyl)MC(Carbamidomethyl)QALQQILQNQSFR(F)
2184.9043	2184.9316	100.0	-12.5		44	60	0	(R)VEQEQQEQQEYPYSQR(G)
2452.0356	2452.0648	100.0	-11.9	1Gln->pyro-Glu	42	60	1	(R)QRVEQEQQEQQEYPYSQR(G)
2469.0684	2469.0913	100.0	-9.29		42	60	1	(R)QRVEQEQQEQQEYPYSQR(G)



A

Signal peptide

Arah2.01 1 MAKLITILVALA.LFLLAHASA.RQQWELQG...DRR.....CQSQLERANLRPCQEHLMQKIQRDEDSYERDPYSPSQDPYSP

Arah6.01 1 MAKSTILVALL.ALVLVAHASA.MRRERGRQG...DSSS.....CERQVDRVNLKPCQEHIMQRIHQEQEQYDS.....

Arah7.0101 1 MMVKLSILVALLGALLVVASA.TRWDPDRGSRGSRWDAPSRGDDQCQRQLQRANLRPCQEEHMRRRVEQEQQEQDE.....YPYSR

Arah7.0201 1 MVKLSILVALLGALLVVASA.TRWDPDRGSRGSRWDAPSRGDDQCQRQLQRANLRPCQEEHMRRRVEQEQQEQDE.....YPYIQ

2Sprotein2 1 MVKLSILVALLGALLVVASA.TRWDPDRGSRGLRWDAPSRGDDQCQRQLQRANLRPCQEEHMRRRVEQEQQEQDE.....YPYSQ

Arah7.0202 1 MVKLSILVALLGALLVVASA.TRWDPDRGSRGLRWDAPSRGDDQCQRQLQRANLRPCQEEHMRRRVEQEQQEQDE.....YPYSQ

Arah2.01 75 SPYDRRGAGSSQ..HQERCCNELNEFENNQRCCMCALQQIMENQSDRIQ.GRQQEQQFKRELRLNLPQQCGIRAPQRCDDLDES

Arah6.01 65 ..YDIRSTRSSD..QQQRCCDELNEMENTQRCCMCALQQIMENQCDRIQ.DRQMVQQFKRELRLNLPQQCNFRAPQRCDDLDES

Arah7.0101 81 RGSRGQPGESDENQEQRCNELNRFQNNQRCCMCALQQILQNSFWVPAGQEPVADGEGAGQELAP..ELRVQVTKPLRPL

Arah7.0201 80 RGSRGQPGESDENQEQRCNELNRFQNNQRCCMCALQQILQNSFWVPAGQEPVADGEGAGQELAP..ELRVQVTKPLRPL

2Sprotein2 80 RGSRGRRPGESDENQEQRCNELNRFQNNQRCCMCALQQILQNSFWVPAGQEPVADGEGAGQELAP..ELRVQVTKPLRPL

Arah7.0202 80 RGSRGRRPGESDENQEQRCNELNRFQNNQRCCMCALQQILQNSFWVPAGQEPVADGEGAGQELAP..ELRVQVTKPLRPL

B

Signal peptide

Arah2.01 1 MAKLITILVALA.LFLLAHASA.RQQWELQG...DRR.....CQSQLERANLRPCQEHLMQKIQRDEDSYERDPYSPSQDPYSP

Arah6.01 1 MAKSTILVALL.ALVLVAHASA.MRRERGRQG...DSSS.....CERQVDRVNLKPCQEHIMQRIHQEQEQYDS.....

Arah7.0101 1 MMVKLSILVALLGALLVVASA.TRWDPDRGSRGSRWDAPSRGDDQCQRQLQRANLRPCQEEHMRRRVEQEQQEQDE.....YPYSR

Arah7.0201 1 MVKLSILVALLGALLVVASA.TRWDPDRGSRGSRWDAPSRGDDQCQRQLQRANLRPCQEEHMRRRVEQEQQEQDE.....YPYIQ

2Sprotein2 1 MVKLSILVALLGALLVVASA.TRWDPDRGSRGLRWDAPSRGDDQCQRQLQRANLRPCQEEHMRRRVEQEQQEQDE.....YPYSQ

Arah7.0202 1 MVKLSILVALLGALLVVASA.TRWDPDRGSRGLRWDAPSRGDDQCQRQLQRANLRPCQEEHMRRRVEQEQQEQDE.....YPYSQ

Arah2.01 75 SPYDRRGAGSSQ..HQERCCNELNEFENNQRCCMCALQQIMENQSDRIQ.GRQQEQQFKRELRLNLPQQCGIRAPQRCDDLDES

Arah6.01 65 ..YDIRSTRSSD..QQQRCCDELNEMENTQRCCMCALQQIMENQCDRIQ.DRQMVQQFKRELRLNLPQQCNFRAPQRCDDLDES

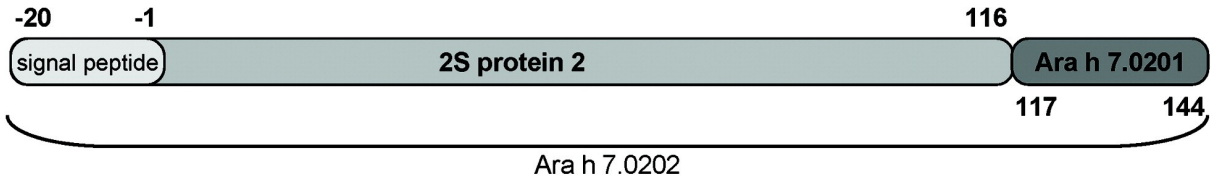
Arah7.0101 81 RGSRGQPGESDENQEQRCNELNRFQNNQRCCMCALQQILQNSFWVPAGQEPVADGEGAGQELAP..ELRVQVTKPLRPL

Arah7.0201 80 RGSRGRRPGESDENQEQRCNELNRFQNNQRCCMCALQQILQNSFWVPAGQEPVADGEGAGQELAP..ELRVQVTKPLRPL

2Sprotein2 80 RGSRGRRPGESDENQEQRCNELNRFQNNQRCCMCALQQILQNSFWVPAGQEPVADGEGAGQELAP..ELRVQVTKPLRPL

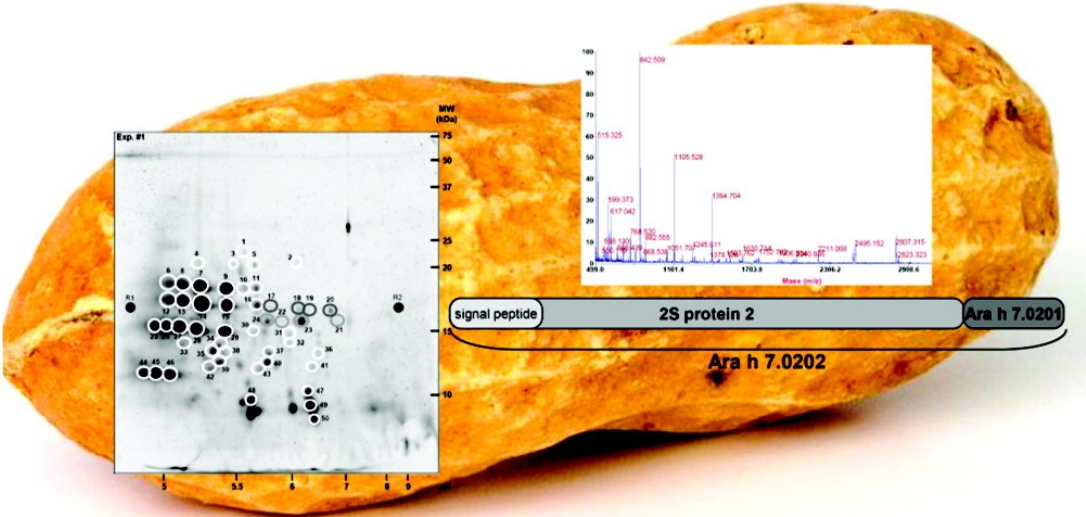
Arah7.0202 80 RGSRGRRPGESDENQEQRCNELNRFQNNQRCCMCALQQILQNSFWVPAGQEPVADGEGAGQELAP..ELRVQVTKPLRPL

C



D

Protein	MW (kDa)	# of Cys	Cysteine pattern and <u>NXS</u> glycosylation motifs
Ara h 2.01	16.6	8	X ₁₁ CX ₁₁ CX ₄₅ CCX ₁₁ CXCX ₈ NQSX ₂₂ CX ₇ CX ₁₂
Ara h 6.01	14.5	10	X ₁₃ CX ₁₁ CX ₃₁ CCX ₁₁ CXCX ₈ NQCX ₂₂ CX ₇ CX ₈ C
Ara h 7.01	16.3	6	X ₂₃ CX ₁₁ CX ₄₁ CCX ₁₁ CXCX ₈ NQSX ₃₅
Ara h 7.02	17.3	8	X ₂₃ CX ₁₁ CX ₄₁ CCX ₁₁ CXCX ₈ NQSX ₂₃ CX ₇ CX ₈



A

Ara h 7.0101 (Acc: AF091737) 307-cagagcttttgggttcagcaggacaggagccagttgcatcaga-tggagagggagctcaggaactt--373
Ara h 7.0201 (Acc: EU046325) 307-cagagcttt-aggttcagcaggacaggagccagttgcatcaga-tggagagggagctcaggaacttg-373
2S Protein 2 (Acc: AY722691) 307-cagagcttt-aggttcagcaggacaggagccagttgcatcagaatggagagggagctcaggaactt--373

B

Ara h 7.0101 (Acc: AF091737) 307-cag agc ttt tgg gtt cca gca gga cag gag cca gtt gca tca gat gga gag gga gct cag gaa ctt
Ara h 7.0201 (Acc: EU046325) 307-cag agc ttt agg ttc cag cag gac agg agc cag ttg cat cag atg gag agg gag ctc agg aac ttg N₅ tgc N₄₅ -417
2S Protein 2 (Acc: AY722691) 307-cag agc ttt agg ttc cag cag gac agg agc cag ttg cat cag aat gga gag gga gct cag gaa ctt N₄₂ tgc N₂₁ tgc N₂₄ -432
-414

C

Ara h 7.0101 (Acc: AF091737) 103-Q - S - R - W - V - P - A - G - Q - E - P - V - A - S - D - G - E - G - A - Q - E - L - - X₃ - C - X₇ - L -139
Ara h 7.0201 (Acc: EU046325) 103-Q - S - R - R - F - Q - Q - D - R - S - Q - L - H - Q - M - E - R - E - L - R - N - L - - X₃ - C - X₇ - C - X₅ -144
2S Protein 2 (Acc: AY722691) 103-Q - S - R - R - F - Q - Q - D - R - S - Q - L - H - Q - N - G - E - G - A - Q - E - L - - X₁₄ -138