Hemagglutinating activity and corresponding putative sequence identity from *Curcuma aromatica* rhizome

Ponpimol Tiptara,[1] Amorn Petsom,[1,2] Sophon Roengsumran[1] and Polkit Sangvanich[1]*

[1]*Research Center for Bioorganic Chemistry, Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand*
[2]*Institute of Biotechnology and Genetic Engineering, Chulalongkorn University, Bangkok, 10330, Thailand*

Abstract

BACKGROUND: *Curcuma aromatica* is a medicinal plant belonging to the Zingiberaceae family with an incomplete genome sequence. It has been reported that extract from the rhizome of this plant contains haemagglutinating activity. In this study the profile of fractions containing hemagglutinating activity is described.

RESULTS: Following extraction with saline buffer, the protein solution was fractionated by ammonium sulfate precipitation. Ion-exchange chromatography was completed on fast-flow SP-Sepharose, as well as gel filtration chromatography on Superdex 75. The active fractions were then separated by one-dimensional sodium dodecyl sulfate–polyacrylamide gel electrophoresis and labeled proteins were digested with trypsin. The digest bands were analyzed by reversed-phase liquid chromatography – tandem mass spectrometry. Inferred peptide sequences were used in Mascot searching and mass spectrometry-driven BLAST (MS-BLAST) homology searches allowed the recognition of related proteins in other species of Viridiplantae. Six putative proteins from nine bands showed similarity with lectin sequences.

CONCLUSION: This study reports the identification of six lectins from the *Curcuma aromatica* rhizome achieved by mass spectrometry using MS-BLAST algorithms to search for homology between *de novo* determined peptide sequences and protein sequences available in sequence databases.

Keywords: *Curcuma aromatica*; lectin; LC-MS/MS; *de novo* sequencing; MS BLAST

INTRODUCTION

Turmeric is an important plant, widely used as a food flavoring agent (e.g., in curry) as well as in traditional medicinal applications.[1–4] In addition, some species are commonly used in industrial application.[5] Several common herbal drugs are derived from the *Curcuma* species and its essential oil, mainly extracted from the rhizome, has been reported to possess a variety of biological properties.[6–9] From previous studies, hemagglutinating (lectin) activity was found in crude protein extract from various *Curcuma* species.[10] Lectins or agglutinins are proteins that recognize and bind to sugar complexes attached to proteins and lipids. The specific agglutination properties of lectins are based on their defined recognition of and binding to sugar.[11] They are widespread throughout the plant kingdom, occurring in a number of species.[12] Their functions in plant tissue are as a defense against phytopathogenic microorganisms, phytophagous insects and plant-eating animals.[13] The defensive role of lectins has stimulated research on possible application of lectins in crop protection. In addition, from the specificity of protein–sugar interaction, lectins could be used as carrier molecules to target drugs specifically to different cells and tissue.[14]

Crude protein extracted from *Curcuma aromatica* exhibits a strong hemagglutinating activity (5.9×10^{-5} mg mL).[10] However, information on the protein constituents extracted from the rhizome of this plant is scant and has not been investigated before. This led us to investigate the protein components and their hemagglutinating activity, which suggested the presence of lectin-type proteins. This was used to guide the separation step. Identification of gel-separated protein by the combination of mass spectrometry

* Correspondence to: Polkit Sangvanich, Research Center for Bioorganic Chemistry, Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand
E-mail: polkit@gmail.com
Contract/grant sponsor: Thailand Research Foundation (TRF)
Contract/grant sponsor: Chulalongkorn University Graduate Scholarship
(*Received 27 March 2007; revised version received 7 July 2007; accepted 23 October 2007*)
Published online 29 January 2008; DOI: 10.1002/jsfa.3184

 J Sci Food Agric 0022–5142/2008/$30.00

and database searching has become a powerful tool in modern protein analysis.[15,16] In addition, mass spectrometric sequence analysis from tandem mass spectra generates highly specific information in the fragmentation pattern, being useful for organisms of which genomic information is limited. This paper describes the identification of proteins from the rhizome of *Curcuma aromatica* achieved by the powerful tool of tandem mass spectrometry combined with database searching.

EXPERIMENTAL

Plant material

Fresh rhizomes of *Curcuma aromatica* were purchased from the local market (Bangkok, Thailand). They were stored at room temperature until use.

Chemicals

All reagents were of analytical grade. Solvents used in preparation and analysis by high-performance liquid chromatography (HPLC) were purchased from Pharmco (Brookfield, CT, USA). C18 chromatography medium was purchased from Waters (Milford, MA, USA). SP-Sepharose and Superdex 75 were purchased from GE Healthcare (Uppsala, Sweden). All chemicals, protein and peptide standards were obtained from Sigma (St Louis, MO, USA). Formic acid (for mass spectrometry) was from Fluka (Seelze, Germany), Trifluoroacetic acid (TFA) (spectrophotometric grade) was from Aldrich (Singapore). HPLC-grade water, acetonitrile (ACN) and methanol were purchased from Sigma-Aldrich.

Instrumentation

Peptide identification by liquid chromatography–tandem mass spectrometry (LC-MS/MS) was performed using a quadrupole time-of-flight (Q-ToF I; Micromass, Manchester, UK) equipped with a Z-spray ion source. The mass spectrometer was operated in the positive ion mode and the sample cone voltage was fixed at 49 V. For MS mode, spectra were acquired in continuum mode with a scan time of 1 s and an inter-scan delay of 0.1 s over a 400–2000 *m/z* range. For MS/MS analysis, the collision energy was set to change according to precursor *m/z*, but ranged from 25 (2+, *m/z* 500) to 40 (3+, *m/z* 1500). Argon was used as the collision gas. Data were acquired over a 50–2000 *m/z* range. During LC-MS analysis, peaks were selected for collision-induced dissociation (CID) according to their signal intensity and charge state using MassLynx data-dependent acquisition functions. Peaks with an intensity of >10 counts s^{-1} and a charge state of 2+ and 3+ triggered switching from MS to MS/MS mode. Deconvolution of raw mass spectra was performed using the ProteinLynx 1.0 module of Masslynx (version 3.4). Chromatography was performed using an Ultimate/Switchos/Famos nanoflow LC system (Dionex, Camberley, UK) coupled to the QToF. The spray tip voltage was set at 2500 V.

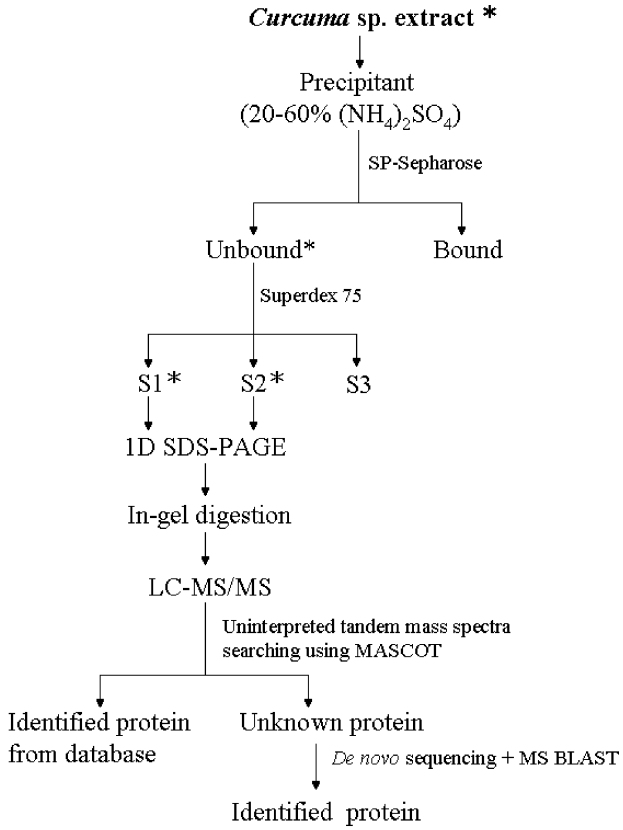

Figure 1. Scheme of fractionation and identification of proteins from *Curcuma aromatica* rhizome.

Isolation of protein from *Curcuma aromatica*

The method for protein separation and identification is outlined in Fig. 1. Fresh rhizomes (1.5 kg) were washed, cut into small pieces and homogenized in a blender with 50 mmol L^{-1} phosphate buffer, pH 7.0, containing 0.1 mol L^{-1} NaCl and 1.0 mmol L^{-1} EDTA at 1:3 (g mL^{-1}). The homogenate was centrifuged at $15\,000 \times g$ at 4 °C for 30 min, and $(NH_4)_2SO_4$ was added to the resulting supernatant to 20–60% saturation.[17] The resulting suspension was centrifuged at $13\,000 \times g$ at 4 °C for 30 min, the supernatant discarded, and the precipitate was resuspended in 50 mmol L^{-1} phosphate buffer, pH 7.0. The solution was then dialyzed against the same buffer for 24 h (using dialysis tubing with molecular weight cut off 3500 Da). Insoluble proteins were removed by centrifugation at $13\,000 \times g$ at 4 °C for 30 min. The solution (≈300 mL) was applied to a column (1.6 × 15 cm) of fast-flow SP-Sepharose installed in an AKTA prime instrument (GE Healthcare, Uppsala, Sweden) equilibrated with buffer A (50 mmol L^{-1} phosphate buffer, pH 7.0). Unbound protein fraction was collected and bound

DOI: 10.1002/jsfa

fraction was eluted with 0–100% linear gradient of buffer B (50 mmol L^{-1} phosphate buffer, pH 7.0, containing 0.35 mol L^{-1} NaCl) at a flow rate of 1.0 mL min^{-1} over 15 column volumes. The unbound fraction was lyophilized, dissolved in water and dialyzed against 0.1 mol L^{-1} NH_4HCO_3, pH 7.8. Insoluble proteins were removed by centrifugation at 13 000 × *g* at 4 °C for 30 min and the supernatant applied to a column (1.6 × 56 cm) of Superdex 75 (GE Healthcare) equilibrated in the same buffer at a flow rate of 0.5 mL min^{-1}. Protein fractions were lyophilized for further characterization by gel electrophoresis. Protein concentrations were determined by the method of Bradford[18] using bovine serum albumin (Sigma) as a standard and Coomassie Plus – the better Bradford™ assay reagent (Pierce, Rockford, IL, USA).

Hemagglutinating activity test

Fractions were tested for hemagglutinating activity by the standard method.[19] Briefly, serial twofold dilutions of the protein in microtiter U-plates (50 μL) were mixed with 50 μL of a 2% (v/v) suspension of rabbit erythrocytes in phosphate-buffered saline, pH 7.2, at room temperature. Precipitation of erythrocytes was assessed visually after 1 h when the blank had fully sedimented.

Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE)

One-dimensional SDS-PAGE was performed using standard methods on the Bio-Rad Mini-Protean II system. It is a discontinuous system with 12% separating gel (pH 8.8) and 4% stacking gel (pH 6.8) of size 7 cm × 10 cm × 1 cm. Prior to electrophoresis, protein samples were redissolved in Laemmli buffer[20] and boiled in the presence of dithiothreitol (DTT) for 5 min at 100 °C. For the molecular weight markers, Sigma Marker™ wide molecular weight range (Sigma) was used. The electrophoresis was performed at 10 mA per gel. The gels were stained with Brilliant Blue R concentrate (Sigma-Aldrich) for 30 min and were destained in 50% (v/v) methanol, 5% acetic acid (v/v) for 30 min or until bands appeared.

In-gel digestion

In-gel digestion was performed using features from standard methods.[21] Each gel band was excised, cut into small pieces (*ca* 1 mm^3) and washed with 100 μL deionized water. The gel pieces were destained by adding 200 μL acetonitrile/25 mmol L^{-1} NH_4HCO_3 (2:1) for 15 min. This step was performed several times until the gel pieces were completely destained. The supernatant was removed and gels were dehydrated for 15 min by adding 200 μL acetonitrile prior to drying in a vacuum centrifuge. 50 μL of 10 mmol L^{-1} DTT in 100 mmol L^{-1} NH_4HCO_3 were added, and the proteins were reduced for 1 h at 56 °C. After cooling to room temperature, the DTT solution was replaced with the same volume of 55 mmol L^{-1} iodoacetamide in 100 mmol L^{-1} NH_4HCO_3 and gels were incubated for 45 min at room temperature in the dark. After alkylation, the solution was removed and then gel pieces were dehydrated with acetonitrile and dried prior the addition of 10 μL trypsin solution (proteomics Grade, Sigma) (10 ng μ L^{-1} in 50 mmol L^{-1} NH_4HCO_3). After allowing the gel plug to swell for 15 min at 4 °C, 30 μL of 50 mmol L^{-1} NH_4HCO_3 was added. The digestion was performed at 37 °C overnight. The supernatant was collected following centrifugation at 10 000 × *g* for 1 min. Peptides remaining in the gel were extracted with a solution of 50% (v/v) acetronitrile containing 5% (v/v) formic acid for 10 min with shaking, and subsequently pooled with the supernatant and taken to dryness.

LC-MS/MS

Peptides were desalted online through a 300 mm × 5 mm C_{18} trapping cartridge (LC Packings, San Francisco, CA, USA) at 100% mobile phase A (30 μL min^{-1}) for 3 min. The peptides were transferred and separated at a flow rate of 180 nL min^{-1} on a 15 cm × 75 μm i.d. Pep-Map™ C_{18} reversed-phased column (Dionex), with bound peptides being eluted over a 35 min gradient to 60% mobile phase B (mobile phase A = 2% (v/v) acetonitrile, 0.06% formic acid; mobile phase B = 95% acetronitrile, 0.05% formic acid).

Data interpretation and MS-BLAST homology searching

As an initial step of protein identification all tandem mass spectra were used for database searches with Mascot software[22] against all sequences in the National Center for Biotechnology Information (NCBI) at the Matrix Science Ltd server (http://www.matrixscience.com/). The following parameters were set for Mascot: trypsin with up to one missed cleavage allowed; fixed modifications: carbamidomethyl (C); variable modifications: oxidation (M); peptide tolerance ±200 ppm; MS/MS tolerance ±200 mmu; peptide charge +2, +3 (monoisotopic); instrument type: ESI-QUAD-ToF. If separated proteins did not match to any protein in the database, *de novo* sequencing was performed manually using the peptide-sequencing tool implemented in MassLynx (Micromass) and the resulting peptides were searched against the non-redundant database with the modified mass spectrometry-driven BLAST searching protocol (MS-BLAST)[23] using the European Molecular Biology Laboratory (EMBL) web interface (http://dove.emblheidelberg.de/Blast2/msblast.html) and default settings.

RESULTS AND DISCUSSION

Isolation of protein

Ion-exchange chromatography was chosen to be the first step in purification because large volumes of sample can be processed with it. However, most of the protein components were not bound with

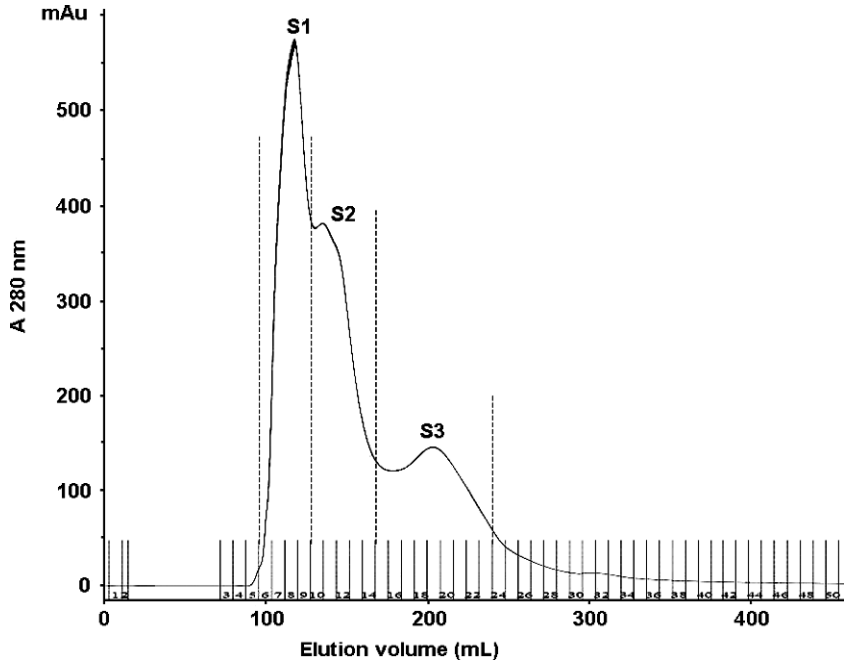

Figure 2. Gel filtration chromatography of unbound fraction from ion-exchange column on the Superdex 75 column (1.6 × 56 cm) in 0.1 mol L^{-1} NH_4HCO_3, pH 7.8.

the SP-Sepharose column (data not shown). This unbound fraction containing hemagglutinating activity toward rabbit erythrocytes at 0.55 mg mL^{-1} was then lyophilized and further purified by gel filtration chromatography on a Superdex 75 column. The isolated proteins yielded three peaks called S1, S2 and S3 (Fig. 2). Hemagglutinating activity was detected in S1 (tubes 6–9) and S2 (tubes 10–14) peaks at 0.19 mg mL^{-1} and 0.14 mg mL^{-1}, respectively. The proteins from both peaks were separated by SDS-PAGE and visualized with Coomassie Blue because of its compatibility with MS analysis. S1 and S2 protein fractions resolved in SDS-PAGE are shown in Fig. 3. By gel electrophoresis, only one distinct region was apparent in the S1 lane and three distinct regions in the S2 lane. The dominant bands marked in Fig. 3 were excised and comprised four bands from the S1 lane (called a, b, c, d) and five bands from the S2 lane (called a, b, c, d, e). Excised gel bands were subjected to in-gel digestion by trypsin, and the tryptic peptides hydrolyzed at the carboxyl side of lysine and arginine residues by trypsin were dried under vacuum prior to analysis by LC-MS/MS.

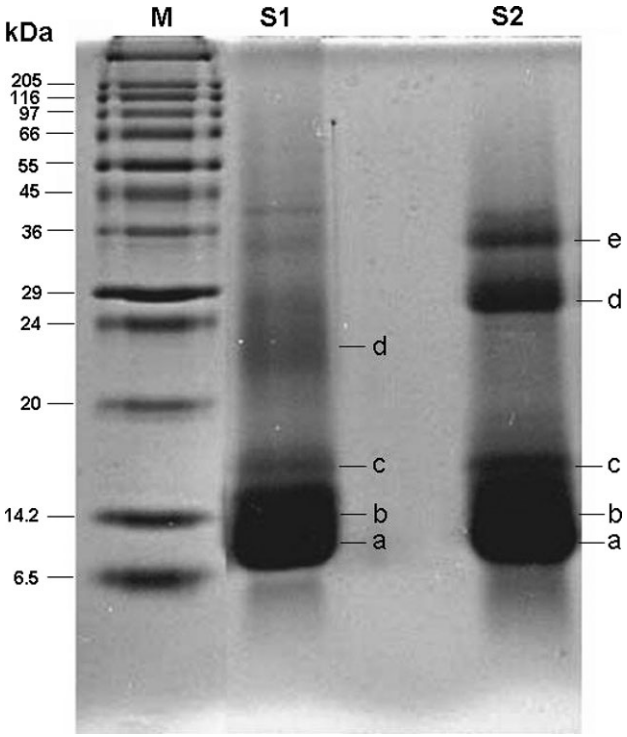

Figure 3. SDS-PAGE profile for S1 and S2 fractions from *Curcuma aromatica* rhizome. M, wide molecular weight marker (Sigma).

Protein identification

Protein bands, which can be detected by Coomassie Blue, are generally present in sufficient quantity to permit identification by MS/MS. This technique generally requires at least 1–2 pmol of protein and it is useful for identifying peptides from organisms with incomplete genome sequence through the partial or complete determination of amino acid sequences by manual or automated *de novo* sequencing. An amino acid sequence can be interpreted directly from spectra without comparison to database. The candidate sequences are then submitted for similarity searching against a protein database. In this study nine proteins bands from two fractions containing hemagglutinating activity, S1 and S2, from *Curcuma aromatica* were identified. Only bands S2/d and S2/e were identified by Mascot as cysteine proteinase GP-I (P82473) (Table 1) found previously in ginger (*Zingiber officinale*). The peptide sequence CGIAISPSYPIK of doubly charged precursor ion with *m*/*z* of 653.31 (Fig. 4B)

Table 1. Characteristics of S2/d and S2/e proteins isolated from *Curcuma aromatica* and identified as cysteine proteinase by Mascot

Band	Accession number	Org.	Mass (Da)	$(M + 2H)^{2+}$	Sequence	M_r (expt.)	M_r (calc.)	ΔM (Da)	Ion score
S2/d	P82473	ZO	24 628	653.3176	CGIAISPSYPIK	1304.6406	1304.6798	−0.0392	49
S2/e	P82473	ZO	24 628	653.3455	CGIAISPSYPIK	1304.6764	1304.6798	−0.0034	59

ZO, *Zingiber officinale*.
All identifications met statistical confidence criteria according to Mascot with $p < 0.05$.[22]

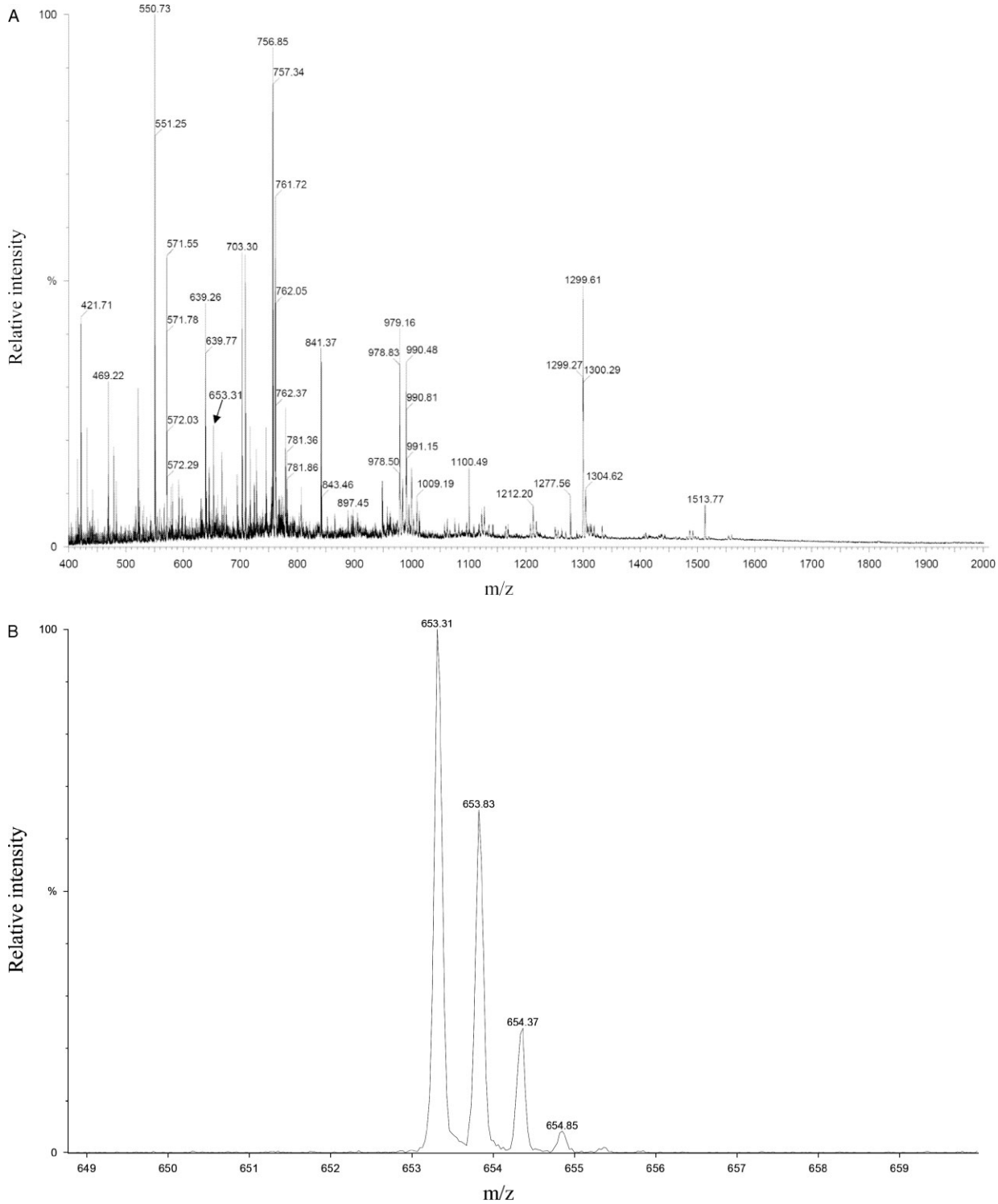

Figure 4. Identification of S2/d protein from *Curcuma aromatica* by mass spectrometry. (A) Mass spectrum survey scan from 400 to 2000 Da. (B) Doubly charged precursor ion *m/z* of 635.31. (C) Tandem mass spectrum of precursor ion *m/z* of 653.31.

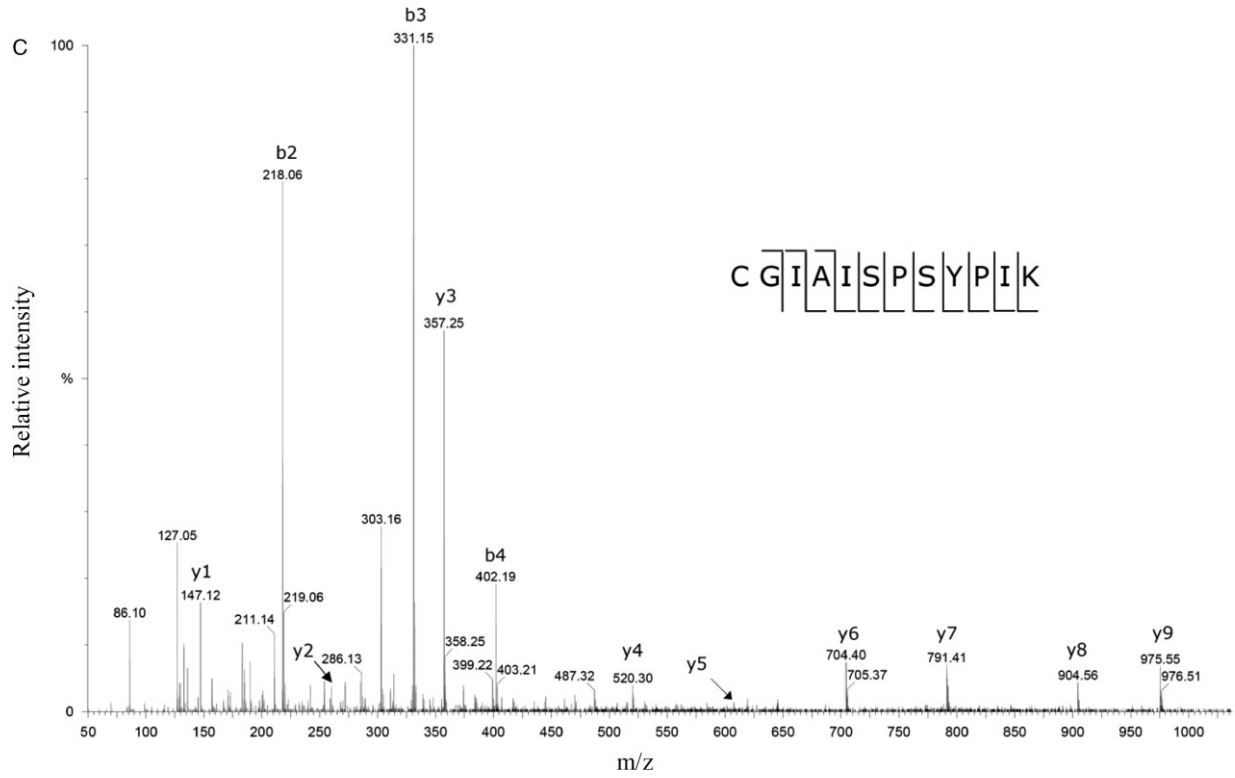

Figure 4. Continued.

Table 2. Protein identification of 1D gel bands from *Curcuma aromatica* by MS-BLAST

Band[a]	Accession number	Protein	Organism	Mass (kDa)	MW (expt.)	Peptide sequence	HSP score[b]	% Pos.[c]
S1/a	Q40417	Dimeric mannose-specific lectin	*Narcissus* hybrid cultivar	18.5	1606.644	NEALWSSDTDGLGSR N A+W S+TDG NKAIWASNTDG	57	81
					1555.644	MQTDGNFLVFNDR MQTDGN V+ MQTDGNLVVY	54	80
					1421.704	TVTNNNLLYAGDK +N+LY GD DNILYSGD	44	87
					1606.644	DQALWSSDTDGLGSR D +W +T GL S DKPIWATNTGGLSS	42	57
S1/b	Q7NTB6	Probable mannose-binding lectin	*Chromobacterium violaceum*	8.84	1454.584	MQTDGNFLLFN MQTDGNF L++ MQTDGNFVLYD	66	90
					1454.664	MQTDGNFLLFGGR MQ DGN ++ MQDDGNLVIY	40	70
S1/c	Q41624	Mannose-binding lectin	*Tulipa* hybrid cultivar	19	1403.715	ETEGNFVLVLQR EGN+VLVLQ EGNYVLVLQ	56	100
					1620.638	QDALWSSDTDQLSR LW +TDQ S LWATNTDQFS	46	70
					1898.8390	VVLYGNPLFTLPNAEPR VV+YG VVIYG	35	100
S1/d	P92932	Lectin-related protein	*Allium sativum*	32.8	1512.724	MQTDGNFLLFDGR MQ DGNF ++	54	80

Table 2. Continued

Band[a]	Accession number	Protein	Organism	Mass (kDa)	MW (expt.)	Peptide sequence	HSP score[b]	% Pos.
						MQVDGNFVIY		
					1606.704	DQALWSSDTDGATLK +A+WSS TD HAIWSSHTD	48	88
					1580.076	DGHVVLFPEDV DG VV++ DV DGNVVIYGPDV	39	72
S2/a	Q39906	Lectin	*Galanthus nivalis*	17.3	1898.984	VVLYGPARFTLPGQEPR VV+YGPAR VVIYGPAR	57	100
					1454.584	MQTDGNFLLFNR MQTDGN +N MQTDGNLVVYN	55	72
S2/b	Q40232	Mannose-binding protein	*Listera ovata*	17.7	1260.584	AVWASNTDGKGR AVWAS T+G G AVWASGTNGRG	56	81
					1260.584	GLWASNTDGKGR +WASNT+ IWASNTN	47	100
					1562.624	QEALWSSGTDGLGSR A+W GT+ AIWATGTN	39	75
					1085.924	GHYTLVLQR G Y L+LQR GNYYLILQR	38	77
					1555.704	MQTDGNFLVFNDR MQ DGN ++ MQRDGNLVIY	36	70
S2/c	Q8U2Y6	Prismane protein	*Pyrococcus furiosus*	51.3	1901.084	YLGAKPPLFTAPNLFPR YLG KPP F PN F YLGPKPPEFLTPNVF	63	66
S2/d	P82473	Cysteine proteinase GP-I	*Zingiber officinale*	24.6	1404.584	ENDHVVSLDSYK EN HVVS+DSY ENAHVVSIDSY	71	90
					1432.724	GCLALSPSYPLKK +A+SPSYP+K IAISPSYPIK	66	100
					1099.384	FWGDSGYLR WG+SGY+R WGESGYIR	56	100
S2/e	Q7XYU7	Senescence-associated cysteine protease	*Anthurium andraeanum*	50.5	1304.664	NCLALSPSYPLK +A+ PSYP+K IAIEPSYPIK	56	90
					1276.567	DDLPDXXDWR DDLP DWR DDLPQKVDWR	56	70
					1099.504	FWGDSGYRL WG+ GY WGEAGY	37	83
					1511.885	GGAAAPVSVLGGCNVL GA AP+ GGC GAVAPIKDQGGC	29	66

[a] Band name according to position on 1D gel (Fig. 3).
[b] MS BLAST match was defined as statistically significant if the score of the HSP was higher than the threshold value that scoring scheme are described in (23).
[c] % positive.

Table 3. Comparison of the amino acid sequences of query peptides from *Curcuma aromatica* and sequences from database

S1/a	*NH* 1 MAKTSFLILATIFLGVITLPSCLSDNILYSGDTLSTGQFLSYGSYVFIMQ Query NNLLYAGD *NH* 51 EDCNLVLYDVDKPIWATNTGGLSSDCHLSMQTDGNLVVYSPQNKAIWASN Query DQALWSSDTDGLGS MQTDGNFLVF NEALWSSD *NH* 101 TDGENGHFVCILQKDRNVVIYGTDRWATGTYTGAVGIPESPASEKYPTSG Query TDG
S1/b	*CV* 1 MLYRMANHHPLWASNTNGKDAMRAIMQTDGNFVLYDFHGKPLWASGTNGK Query MQTDGNFLLFN *CV* 51 PGCFVTMQDDGNLVIYEPKIPVWASNTAQ Query MQTDGNFLLF
S1/c	*TH* 101 KDGNLVIYSKSGNSVWASQTHQAEGNYVLVLQKDRNVVIYGPSLWATNTD Query EGNFVLVLQ VVLYG LWSSDTD *TH* 151 QFSLTSNSTTESGSGMANEGKIAMVTK Query QLS
S1/d	*AS* 201 VGAHAGCRAAMQVDGNFVIYFNLHAIWSSHTDRENGNYVLVLQQDGNVVI Query MQTDGNFLLF QALWSSDTD DGHVVL *AS* 251 YGPDVWSTGTHVKSGGGRAVVTAMNGTVGGRGSVNQKHVTAIRKVGTSAL Query FPEDV
S2/a	*GN* 51 LVLYDVDKPIWATNTGGLSRSCYLNMQTDGNLVVYNPSNKPIWASNTGGQ Query MQTDGNFLLFN *GN* 101 NGNYVCILQKDRNVVIYGPARWATGTNIHGAGIVGVPGSAPQNSTAEMIK Query VVLYGPAR
S2/b	*LO* 51 IIQGDCNLVLYDNNRAVWASGTNGRGSNCILSMQRDGNLVIYSSGRAIWA Query AVWASNTDGKG MQTDGNFLVF LWA *LO* 101 SNTNRQNGNYYLILQRDRNVVIYDNSNNAIWATGTNVGNAAIAVIPHNNG Query SNTD GHYTLVLQR ALWSSGTD
S2/c	*PF* 401 KGIYLGPKPPEFLTPNVFEALRKQFDLRLISDPERDLRDMLSKGISVEES Query YLGAKPPLFTAPNLF
S2/d	*ZO* 101 KENAHVVSIDSYRNVPSNDEKSLQKAVANQPVSVTMDAAGRDFQLYRNGI Query ENDHVVSLDSY *ZO* 151 FTGSCNISANHYRTVGGRETENDKDYWTVKNSWGKNWGESGYIRVERNIA Query WGDSGYLR *ZO* 201 ESSGKCGIAISPSYPIKE Query LALSPSYPLK
S2/e	*AA* 101 LGVKPGQVRPRRANRAPGRGRDLSANGDDLPQKVDWREKGAVAPIKDQGG Query DDLPQKVDWR GAAAPVSVLGG *AA* 151 CGSCWAFSTVAAVEGINQIVTGDLIVLSEQELVDCDTAYNEGCNGGLMDY Query C *AA* 301 SGKDYWIVRNSWGKSWGEAGYIRMERNLPSSSSGKCGIAIEPSYPIKKGQ Query WGDSGY LALSPSYPLK

Note: Identities are highlighted in dark grey and similarities in light grey.

was hit significantly with Mascot ion score for both of them. Peptides from the rest of the protein bands failed to give a significant match with Mascot. To overcome this problem, peptides were further sequenced *de novo* and matched for homology-based identification with the modified MS-driven BLAST searching method. This search uses an alternative scoring scheme, based on threshold scores that are set conditionally on the number of retrieved high-scoring segment pairs and the total number of fragmented precursors. Table 2

J Sci Food Agric **88**:1025–1034 (2008)
DOI: 10.1002/jsfa

lists the positive hits of identified *Curcuma aromatica* proteins by MS-BLAST searching. They were significant with respect to the threshold values. The score of the high-scoring segment pair (HSP) of the protein hit was compared with the threshold score.

The MS-BLAST results revealed that sequence coverage of S2/d can be increased by two peptide fragments. The score of the top-ranked HSP was 71 (13 peptides used; threshold score 1HSP = 66). However, S2/e was identified as having similarity to a senescence-associated cysteine protease (Q7XYU7) from *Anthurium andraeanum* with the sum of three top-scoring HSPs; 149 (four peptides used; threshold score 1HSP = 68, 2HSP = 112, 3HSP = 138) instead of cysteine proteinase GP-I (P82473). This was assumed to be due to the increase of three peptide fragments (Table 2) and their molecular mass position on 1D gel (Fig. 3). S1/a was identified as a peptide with similarity to a dimeric mannose-specific lectin from *Narcissus* hybrid cultivar. The score of the top HSP peptide sequence, 57, was lower than the threshold of statistical significance for a single matched HSP. The score of the second-ranked HSP sequence, 54, was then added and the sum (HSP = 111) compared with the threshold for two matched (HSP = 106). As the sum of the score significantly exceeded the threshold, a positive identification can be mentioned (20 peptides used; threshold score 1HSP = 67, 2HSP = 106). In the same way, S1/b was identified as a peptide having similarity to a probable mannose-binding lectin from *Chromobacterium violaceum* with the score of the top-ranked HSP; 66 (11 peptides used; threshold score 1HSP = 65). S1/c was identified as a peptide having similarity to a mannose-binding lectin from *Tulipa* hybrid cultivar[24] with the sum of two top-scoring HSPs; 102 (12 peptides used; threshold score 1HSP = 66, 2HSP = 102). S1/d was identified as a peptide having similarity to a lectin-related protein from *Allium sativum*[25] with the sum of two top-scoring HSPs; 102 (11 peptides used; threshold score 1HSP = 65, 2HSP = 101). S2/a was identified as a peptide having similarity to a lectin from *Galanthus nivalis* (common snowdrop) with the sum of three top-scoring HSPs; 112 (14 peptides used; threshold score 1HSP = 66, 2HSP = 103). S2/b was identified as a peptide having similarity to a mannose-binding protein from *Listera ovata*[26] with the sum of two top-scoring HSPs; 103 (12 peptides used; threshold score 1HSP = 66, 2HSP = 102). S2/c was identified as a peptide having similarity to a prismane protein from *Pyrococcus furiosus* with the score of the top-ranked HSP; 63 (one peptide used; threshold score 1HSP = 60). The comparison of amino acid sequences of the query from MS-BLAST and from the database is shown in Table 3.

CONCLUSIONS

This study reports the identification of proteins from *Curcuma aromatica* with a non-sequenced genome database, achieved by MS. The results showed that six of the nine gel bands analyzed were identified as lectin, which agrees with the hemagglutinating activity of fractions S1 and S2.

ACKNOWLEDGEMENTS

Financial support was provided by the Thailand Research Foundation (TRF) and Chulalongkorn University Graduate Scholarship.

REFERENCES

1 Cao H, Sasaki Y, Fushimi H and Komatsu K, Molecular analysis of medicinally-used Chinese and Japanese *Curcuma* based on 18S rRNA gene and trnK gene sequences. *Biol Pharm Bull* **24**:1389–1394 (2001).
2 Wu WY, Xu Q, Shi LC and Zhang WB, Inhibitory effects of *Curcuma aromatica* oil on proliferation of hepatoma in mice. *World J Gastroenterol* **6**:216–219 (2000).
3 Brinkhaus B, Hentschel C, Von Keudell C, Schindler G, Lindner M, Stutzer H, *et al*, Herbal medicine with curcuma and fumitory in the treatment of irritable bowel syndrome: a randomized, placebo-controlled, double-blind clinical trial. *Scand J Gastroenterol* **40**:936–943 (2005).
4 Chearwae W, Wu CP, Chu HY, Lee TR, Ambudkar SV and Limtrakul P, Curcuminoids purified from turmeric powder modulate the function of human multidrug resistance protein 1 (ABCC1). *Cancer Chemother Pharmacol* **57**:376–388 (2006).
5 Kim JK, Jo C, Hwang HJ, Park HJ, Kim YJ and Byun MW, Color improvement by irradiation of *Curcuma aromatica* extract for industrial application. *Radiat Phys Chem* **75**:449–452 (2006).
6 Kojima H, Yanai T and Toyota A, Essential oil constituents from Japanese and Indian *Curcuma aromatica* rhizomes. *Planta Med* **64**:380–381 (1998).
7 Tohda C, Nakayama N, Hatanaka F and Komatsu K, Comparison of anti-inflammatory activities of six *Curcuma* rhizomes: a possible curcuminoid-independent pathway mediated by *Curcuma phaeocaulis* extract. *Evid Based Complement Alternat Med* **3**:255–260 (2006).
8 Pitasawat B, Choochote W, Tuetun B, Tippawangkosol P, Kanjanapothi D, Jitpakdi A, *et al*, Repellency of aromatic turmeric *Curcuma aromatica* under laboratory and field conditions. *J Vector Ecol* **28**:234–240 (2003).
9 Pant N, Jain DC, Bhakuni RS, Prajapati V, Tripathi AK and Kumar S, Zederone: a sesquiterpenic keto-dioxide from *Curcuma aromatica*. *Indian J Chem B* **40**:87–88 (2001).
10 Polkit S, Sophon K, Chantragan S, Ponpimol T, Amorn P, Apaporn B, *et al*, Hemagglutinating activity of *Curcuma* plants. *Fitoterapia* **78**:29–31 (2007).
11 Van Damme EJM, Plant lectins: a special class of plant proteins, in *Handbook of Plant Lectins: Properties and Biomedical Applications*. Wiley, Chichester, pp. 3–30 (1998).
12 Van Damme EJM, Peumans WJ, Barre A and Rouge P, Plant lectins: a composite of several distinct families of structurally and evolutionary related proteins with diverse biological roles. *Crit Rev Plant Sci* **17**:575–692 (1998).
13 Vasconcelos IM and Oliveira JTA, Antinutritional properties of plant lectins. *Toxicon* **44**:385–403 (2004).
14 Christiane B, Claus ML and John FW, Lectin-mediated drug targeting: history and applications. *Adv Drug Deliv Rev* **56**:425–435 (2004).
15 Yates JR, Database searching using mass spectrometry data. *Electrophoresis* **19**:893–900 (1998).
16 Wilm M, Shevchenko A, Houthaeve T, Breit S, Schweigerer L, Fotsis T, *et al*, Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* **379**:466–469 (1996).

17 Choi KH and Laursen RA, Amino-acid sequence and glycan structures of cysteine proteases with proline specificity from ginger rhizome *Zingiber officinale*. *Eur J Biochem* **267**:1516–1526 (2000).
18 Bradford MM, Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. *Anal Biochem* **72**:248–254 (1976).
19 Wang HX, Ng TB, Liu WK, Ooi VEC and Chang ST, Isolation and characterization of 2 distinct lectins with antiproliferative activity from the cultured mycelium of the edible mushroom *Tricholoma mongolicum*. *Int J Peptide Protein Res* **46**:508–513 (1995).
20 Laemmli UK, Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680–685 (1970).
21 Lee TD and Shively JE, Enzymatic and chemical digestion of proteins for mass-spectrometry. *Meth Enzymol* **193**:361–374 (1990).
22 David NP, Darryl JCP, David MC and John SC, Probability-based protein identification by searching sequence database using mass spectrometry data. *Electrophoresis* **20**:3551–3567 (1999).
23 Shevchenko A, Sunyaev S, Loboda A, Shevchenko A, Bork P, Ens W, *et al*, Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time of flight mass spectrometry and BLAST homology searching. *Anal Chem* **73**:1917–1926 (2001).
24 Van Damme EJM, Brike F, Winter HC, Van Leuven F, Goldstein IJ and Peumans WJ, Molecular cloning of two different mannose-binding lectins from tulip bulbs. *Eur J Biochem* **236**:419–427 (1996).
25 Smeets K, Van Damme EJM, Verhaert P, Barre A, Rouge P, Van Leuven F, *et al*, Isolation, characterization and molecular cloning of the mannose-binding lectins from leaves and roots of garlic (*Allium sativum* L). *Plant Mol Biol* **33**:223–234 (1997).
26 Van Damme EJM, Smeets K, Torrekens S, Van Leuven F and Peumans WJ, Characterization and molecular-cloning of mannose-binding lectins from the Orchidaceae species *Listera ovata, Epipactis helleborine* and *Cymbidium hybrid*. *Eur J Biochem* **221**:769–777 (1994).

J Sci Food Agric **88**:1025–1034 (2008)
DOI: 10.1002/jsfa