

Cloning, expression, and purification of recombinant major mango allergen Man i 1 in *Escherichia coli*

Wen-Che Tsai^a, Tzee-Chung Wu^b, Bor-Luen Chiang^{c,d,e}, Hsiao-Wei Wen^{a,*}

^a Department of Food Science and Biotechnology, National Chung Hsing University, Taichung, 402, Taiwan

^b Department of Pediatrics, Taipei Veterans General Hospital, and School of Medicine, National Yang Ming University, Taipei, 112, Taiwan

^c Department of Pediatrics, National Taiwan University Hospital, National Taiwan University, Taipei, 100, Taiwan

^d Graduate Institute of Clinical Medicine, College of Medicine, National Taiwan University, Taipei, 100, Taiwan

^e Department of Medical Research, National Taiwan University Hospital, Taipei, 100, Taiwan

ARTICLE INFO

Article history:

Received 25 March 2016

Received in revised form

22 June 2016

Accepted 23 June 2016

Available online 24 June 2016

Keywords:

Mango

Allergen

Man i 1

Escherichia coli

ABSTRACT

In recent years, the number of people around the world who suffer from fruit allergies has increased. Mango can induce anaphylaxis, and two major mango allergens have been identified – Man i 1 and Man i 2. Apart from their molecular weights and pI values, no other information about them is known. This work identifies the DNA and amino acid sequences of Man i 1 and constructs an expression system for recombinant Man i 1 (rMan i 1). Firstly, 3' and 5' RACE assays were used to identify the cDNA fragment of Man i 1. Subsequently, the full length of Man i 1 cDNA was inserted into a pET-21a(+) vector, and the inserted plasmid was transformed to *Escherichia coli* BL21 (DE3) to express rMan i 1. The conditions for the expression of rMan i 1, including IPTG concentration, induction temperature, and induction time, were optimized. The highest amount of soluble rMan i 1 was obtained after induction with 0.1 mM IPTG at 16 °C for 20 h. The His-tagged rMan i 1 was purified using Ni-NTA agarose and its identity was verified using an anti-histidine antibody and the serum of a mango-allergic person. Additionally, rMan i 1 was identified as glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and shared 86.2% identity in amino acid sequence of GAPDH from wheat. Finally, an *E. coli* expression system of rMan i 1 was established, with the potential to be used in immunotherapy against mango allergy or the development of assays for detecting the residues of mango allergens.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Mango (*Mangifera indica* L.) is an economically favorable fruit that is planted in the tropical and subtropical areas of Asia and Africa. Mango is often processed into a variety of products, such as juice, shakes, ice cream, jams, pickles, dried goods, cookies or cakes [1]. The demand for mango has recently increased possibly owing to its bright golden color, sweet taste, favorable flavor, and high nutritional value. For instance, mango contains high levels of polyphenols, carotenoids, and ascorbic acid. In particular, β-carotene accounts for 48–84% of mango flesh by dry weight [2]. Although mango is a popular fruit globally, it is also an important allergic source and can induce a life-threatening anaphylactic reaction [3,4]. Some people are allergic to not only mango flesh, but

also its pollen and seeds [5,6]. Regarding prevalence, mango-allergic populations constitute 0.3%, 6%, and 16% of all food-allergic patients in Switzerland, France, and Thailand, respectively [7–9]. This variation in the prevalence of mango allergy may be associated with differences of cultivars, eating habits, and heredity.

The symptoms of mango allergy usually present in the gastrointestinal, skin and respiratory systems, and include diarrhea, itchy mouth or throat, swollen lips, erythema, urticaria, rhinoconjunctivitis, cough and dyspnea [10,11]. Mango-allergic reactions can be generally identified as manifesting immediate hypersensitivity and delayed hypersensitivity. The former occur with a few minutes to 30 min of ingestion, while the latter may occur after 8–12 h of exposure, with the allergic reactions reaching a peak at 24–72 h, before slowly decaying [12]. Mango proteins of 14, 16, 18, 20, 25, 27, 30, 37, 40, 43, 50, 60, and 67 kDa have been recognized using sera from mango allergic patients, and two have been identified as major allergens, named Man i 1 (40 kDa) and Man i 2 (30 kDa) [11,13–15]. The amino acid sequences of Man i 1 or Man i 2 have not been determined.

* Corresponding author. Department of Food Science and Biotechnology, National Chung Hsing University, 145, Xingda Rd., Taichung 402, Taiwan.

E-mail address: hwwen@dragon.nchu.edu.tw (H.-W. Wen).

Escherichia coli, an enteric bacterium, is extensively used as an expression host for the production of recombinant protein, because its genetic and physiological characteristics are well known. Additionally, the expression system of *E. coli* has several favorable characteristics for producing large-scale recombinant proteins, including inexpensive carbon sources for the growth, rapid bacterial growth and easy handling [16]. However, the *E. coli* expression system is associated with such issues as a lack of post-translational modification and the degradation of native protein [17]. Moreover, the expressed proteins are sometimes aggregated as inclusion bodies, which must undergo complicated processes, such as unfolding, dialysis, and refolding, to yield soluble and active proteins. Therefore, optimizing entire isolation process from inclusion bodies takes time and the process commonly does not provide high yields [18,19]. Thus, an efficient way is to optimize the expression conditions to obtain directly a soluble and active recombinant protein, instead of a denatured recombinant protein form the inclusion bodies.

This study determines the amino acid sequence of Man i 1 and optimizes the expression and purification system for obtaining a recombinant Man i 1 (rMan i 1). Firstly, the full-length cDNA sequence of Man i 1 was determined using the 3' and 5' RACE assays, cloned into the pET-21a(+) vector; and transformed to *E. coli* BL21 (DE3) for expression. To obtain a soluble rMan i 1, the expression conditions of temperature, induction time, and IPTG concentrations were optimized. Finally, the soluble rMan i 1 was purified using Ni-NTA resins and identified using immunoassays with an anti-histidine antibody and a serum from mango-allergic patient.

2. Materials and methods

2.1. Materials

Mangos of different varieties (Irwin, Yuwen, Jinhuang, Keitt, and Tu) were bought from local markets in Taiwan. Bovine serum albumin (BSA), β -Nicotinamide adenine dinucleotide (NAD^+), DL-Glyceraldehyde 3-phosphate, phenyl methyl sulfonyl fluoride (PMSF), isopropyl β -D-1-thiogalactopyranoside (IPTG), Triton X-100, Tween 20, chloroform, β -mercaptoethanol, diethylpyrocarbonate (DEPC), and Coomassie Brilliant Blue R250 (CBB) were purchased from the Sigma-Aldrich Corporation (USA). Ethylenediaminetetraacetic acid (EDTA), triethanolamine, polyvinyl pyrrolidone (PVPP), pET-21a(+) vector, and *E. coli* BL21 (DE3) were purchased from the Merck Millipore Corporation (Germany). TRIzol® reagent, pCR™II-TOPO® Vector, 5' and 3' rapid-amplification of cDNA ends (RACE) kit, a cDNA synthesis kit, and Taq DNA polymerase were purchased from Invitrogen (USA). The restriction enzymes and T4 DNA ligase were bought from New England Biolabs (USA). A mango allergic patient was recruited by National Taiwan University Hospital (NTUH) and the other three patients were recruited by Taipei Veterans General Hospital (TVGH) and the sera used in this study were obtained therefrom. Luria-Bertain (LB) broth and agar were purchased from Acumedia Manufacturers, Inc. (USA).

2.2. Mango protein extract

The method of Singh et al. for extracting mango protein was slightly modified for use in this study [20]. Briefly, mango fruits were cut into small pieces, frozen with liquid nitrogen, and immediately ground using a grinding apparatus (IKA, Germany). The fine mango powders were mixed using the extraction buffer (20 mM sodium phosphate buffer, 20 mM EDTA, 20 mM cysteine, 0.5% Triton X-100, 0.1% β -mercaptoethanol, 1 mM PMSF, and 66 mg

PVPP, pH 7.4) at 4 °C for 20 min with gentle stirring. After being centrifuged twice at 8,000 rpm and 4 °C for 20 min, the supernatant of the extraction solution was dialyzed overnight against 10 mM sodium phosphate buffer (pH 7.4). The protein concentration of the extraction sample was measured using the Bio-Rad protein assay with BSA as the standard.

2.3. SDS-PAGE analysis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis was performed as described by Laemmli [21]. Each protein sample was mixed with 5X sample buffer (10% SDS, 50% glycerol, 0.25 M Tris-HCl pH 6.8, β -mercaptoethanol, and 0.05% bromophenol blue) prior to the denaturation in boiling water for 5 min. The protein samples were analyzed by SDS-PAGE with a 5% (w/v) stacking gel and a 12% (w/v) resolving gel at a constant voltage of 100 V. Proteins in the gel were detected by staining with CBB, and then destained using a destaining solution that contained 10% acetic acid and 25% methanol.

2.4. Dot blot and Western blot

Dot blot and Western blot were conducted as described elsewhere with some modifications [22,23]. One microgram of mango extract was spotted on each PVDF membrane and dried for 10 min at 60 °C for dot blotting. In the Western blot, proteins were separated by SDS-PAGE analysis and then transferred to the PVDF membrane. For both assays, the PVDF membrane was blocked with 5% skimmed milk powder in phosphate buffer saline with 0.1% Tween 20 (PBST, pH 7.4) for one hour at 37 °C. To investigate the specific interaction between sample proteins and the serum of a mango-allergic patient, the membrane was firstly incubated overnight with a ten-fold dilution of the patient serum (1:10 in 0.3% BSA, PBST) at 4 °C. After it had been washed with PBST three times, the membrane was incubated with goat anti-human IgE antibody (1:2500 in PBST; KPL, USA) at 37 °C for one hour. After it had been washed with PBST three times, biotin-labeled anti-goat antibody (1:5000 in PBST; Abcam, USA) was added and incubated at 37 °C for one hour. The membrane was washed three times with PBST, and then incubated with peroxidase-labeled streptavidin (1:10000 in PBST, KPL, USA) at 37 °C for 30 min. The results were obtained by staining with ECL reagents and analyzing using a Luminescence Image System (Hansor, Taiwan). To study the specific interaction between the protein sample and the anti-histidine antibody, the blocked membrane was incubated overnight with anti-histidine antibody (1:1000 in PBST; Miniport, Germany) at 4 °C, and then reacted for one hour at RT with anti-mouse antibody that was conjugated with horseradish peroxidase (1:5000 in PBST; GeneTex, USA). After the membrane had been washed, the target protein was detected with ECL reagents, as described above.

2.5. LC/MS/MS and MALDI-TOF analysis

For the peptide mapping of the target, the protein sample was firstly separated out using SDS-PAGE analysis and stained with CBB. The band that contained Man i 1 (40 kDa) was cut out and digested by trypsin in the gel of SDS-PAGE. The peptide mapping of the trypsinized sample was carried out by LC/MS/MS (LC: 1200 Series Binary Pump, Agilent; Mass Spectrometry: LTQ-XL, Thermo Scientific) analysis or a MALDI-TOF assay (Voyager DE PRO, Applied Biosystems).

2.6. RNA extraction and cDNA synthesis

Mango fruit was frozen in liquid nitrogen and immediately

homogenized to a fine powder. One hundred micrograms of ground mango powder was reacted with 1 mL TRIzol[®] reagent for 20 min at RT. Chloroform (200 μ L) was added, and the mixture was shaken vigorously for 15 s, before being incubated for three minutes. After the mixture had been centrifuged at 12,000 g for 15 min at 4 °C, the supernatant (500 μ L) was collected and mixed with 500 μ L isopropanol, before being incubated at –20 °C for one hour. The mixture was centrifuged again at 12,000 g for 10 min at 4 °C, and the obtained pellet (total RNA precipitate) was washed twice with 75% ethanol. Finally, the pellet was re-suspended in 30 μ L DEPC-treated water, and the concentration of total RNA in the sample was measured using a NanoDrop ND2000 (Thermo Scientific, USA). A cDNA synthesis kit was used to obtain cDNA from this RNA sample, following the manufacturer's instructions.

2.7. Cloning

To obtain the full-length cDNA sequence of Man i 1, 5' RACE and 3' RACE kits were utilized, according to the manufacturer's instructions. The primers for the 3' RACE reaction were GSP1 (5'-CTGGAGTGTCTCACTGACAAGGAG-3') and GSP2 (5'-TGCTCACTTGAAGGGT GGTGC-3'), and those for the 5' RACE reaction were GSP3 (5'-CAACCTTCTTGGCACCACC-3') and GSP4 (5'-CCTTCAAGTGAGCAGCAGC-3'). The PCR products of 5' RACE and 3' RACE were firstly separated using an agarose gel, and the target fragments were then extracted by a Wizard[®] SV Gel and a PCR Clean-Up System (Promega, USA). The extracted target DNA fragment was cloned into a TOPO vector for DNA sequencing to identify the full cDNA of Man i 1. A primer set for Man i 1 was designed according to the full length of Man i 1 cDNA. The PCR amplification of Man i 1 was conducted using a forward primer with an *Nde*I site (5'-GCCA-TATGCGCCGAGGCAAGATTAAG-3') and a reverse primer with an *Xho*I site (5'-CTCGAGAGCAGTTGATGCCATGTG-3'), where underlined sequences in parentheses are the restriction enzyme sites. The PCR conditions were 94 °C for five minutes, 35 cycles of 94 °C for 45 s, 55 °C for 45 s, 68 °C for one minute and a final extension step at 72 °C for seven minutes. The PCR product of Man i 1 was verified using 1% agarose gel and cloned into a TOPO vector to form TOPO/Man i 1 plasmid, which was then transformed to *E. coli* DH5 α . After the TOPO/Man i 1 plasmid was confirmed by DNA sequence analysis, the TOPO/Man i 1 plasmid and pET-21a(+) plasmid were double-digested by *Nde*I and *Xho*I. The digested Man i 1 and pET-21a(+) vector were extracted from DNA agarose gel and ligated together with T4 DNA ligase, to form an expression plasmid, pET-21a(+)/Man i 1. The recombinant pET-21a(+)/Man i 1 expression plasmid was transformed to *E. coli* BL21 (DE3), which were cultured in LB broth that contained ampicillin. The colonies that contained the insert fragment of Man i 1 were identified by a PCR reaction.

2.8. Expression and purification of recombinant Man i 1 (rMan i 1)

To induce the expression of rMan i 1, *E. coli* BL21 (DE3) that contained pET-21a(+)/Man i 1 plasmid was cultured in LB medium that contained 50 μ g/mL ampicillin at 37 °C for 12 h with shaking (200 rpm). Subsequently, the culture was added to another fresh LB medium at 37 °C with shaking (200 rpm). When the OD₆₀₀ of the culture reached 0.5 AU, IPTG was added to the LB medium to initiate expression for the protein. To optimize the conditions for the expression of rMan i 1, various induction conditions were optimized, including temperature (16, 20, 25 and 30 °C), time (0, 4, 8, 12, 16, 20 and 24 h), and IPTG concentration (0.1, 0.25, 0.5, 0.75 and 1.0 mM).

Having been cultured under optimal expression conditions, *E. coli* BL21 (DE3) was collected with centrifugation at 4 °C for ten minutes at 8,000 rpm. The cell pellets were resuspended with lysis buffer (phosphate buffer, pH 8.0), stored in ice, and then lysed with

sonication. The supernatant of the cell lysate was obtained by centrifugation at 4 °C for 20 min at 16,000 g and stored at 4 °C. rMan i 1 was purified by an affinity column that had been packed with Ni-NTA agarose (Qiagen, Germany) and equilibrated with five column volumes (CVs) of lysis buffer. The supernatant was loaded into the column, and then washed with nine CVs of wash buffer (PBS, containing 20 mM imidazole, pH 8.0) to remove unwanted cellular constituents or contaminants. The bound rMan i 1 protein was eluted with six CVs with gradient of 100–300 mM imidazole in PBS (pH 8.0). The protein concentration of rMan i 1 was measured using the Bio-Rad protein assay with BSA as the standard. The purity of the rMan i 1 sample was estimated using SDS-PAGE analysis, and rMan i 1 was identified by Western blot with the anti-histidine antibody and the serum of a mango -allergic patient. Finally, the amino acid sequence of rMan i 1 was determined by LC/MS/MS and MALDI-TOF analysis.

2.9. GAPDH activity of rMan i 1

To investigate the GAPDH activity of purified rMan i 1, the method from Egea's study was utilized with some modification [24]. GAPDH activity was performed in a reaction containing rMan i 1, glyceraldehyde-3-phosphate (50 mg/mL), NAD⁺ (10 mM) and assay buffer (40 mM triethanolamine, 50 mM Na₂HPO₄, and 5 mM EDTA, pH 8.6) at 25 °C for 20 min. The absorbance at 340 nm was measured due to the formation of NADH.

3. Results and discussion

3.1. Analysis of protein extracts of various mango varieties

The allergen contents of five common mango fruits – Irwin, Yuwen, Jinhuang, Keitt, and Tu – were analyzed. Tu is a variety indigenous to Taiwan. Fig. 1A reveals that these five varieties had similar protein patterns, which were also similar to those of four mango species (Eden, Ngowe, Osteen and Tommy Atkins) that were examined by Kinder [13]. All five mango varieties yielded a detectable band at 40 kDa and the corresponding protein has been identified as a major mango allergen, Man i 1 [11]. To investigate the allergic activity of the five mango fruits, their protein extracts were analyzed using a dot blot (Fig. 1B) and western blot (Fig. 1C) with the sera from mango-allergic patients. In the dot blot assay, all spots (numbers 1–5) reacted with the serum with similar intensities, indicating that the patient was sensitive to all five mango species with similar immunoactivities. In the Western blot assay, all mango varieties include a band at 40 kDa. Patient 1 had strong reaction with Man i 1, followed by Patient 3 and Patient 2. The results of Kinder's study [13] and the present study demonstrate that Man i 1 is a major allergen in numerous mango species, including Eden, Ngowe, Osteen, Tommy Atkins, Irwin, Yuwen, Jinhuang, Keitt, and Tu.

3.2. The full cDNA length of Man i 1

The amino acid sequence of Man i 1 has not yet been identified. Hence, in this study, peptide mapping is utilized to obtain the part of the amino acid sequence of Man i 1. The protein extract from Irwin was separated in SDS-PAGE, and the 40 kDa protein band was excised from the gel. After it had been treated with trypsin, the sample was analyzed by LC/MS/MS to identify the amino acid sequence of its fragments. As shown in (Supplementary Fig. 1), about 50% of the amino acid sequence of Man i 1 was identical to that of glyceraldehyde 3-phosphate dehydrogenase (GAPDH) in *Medicago truncatula*, which is a small annual legume that originated from the Mediterranean region. GAPDH is one of the

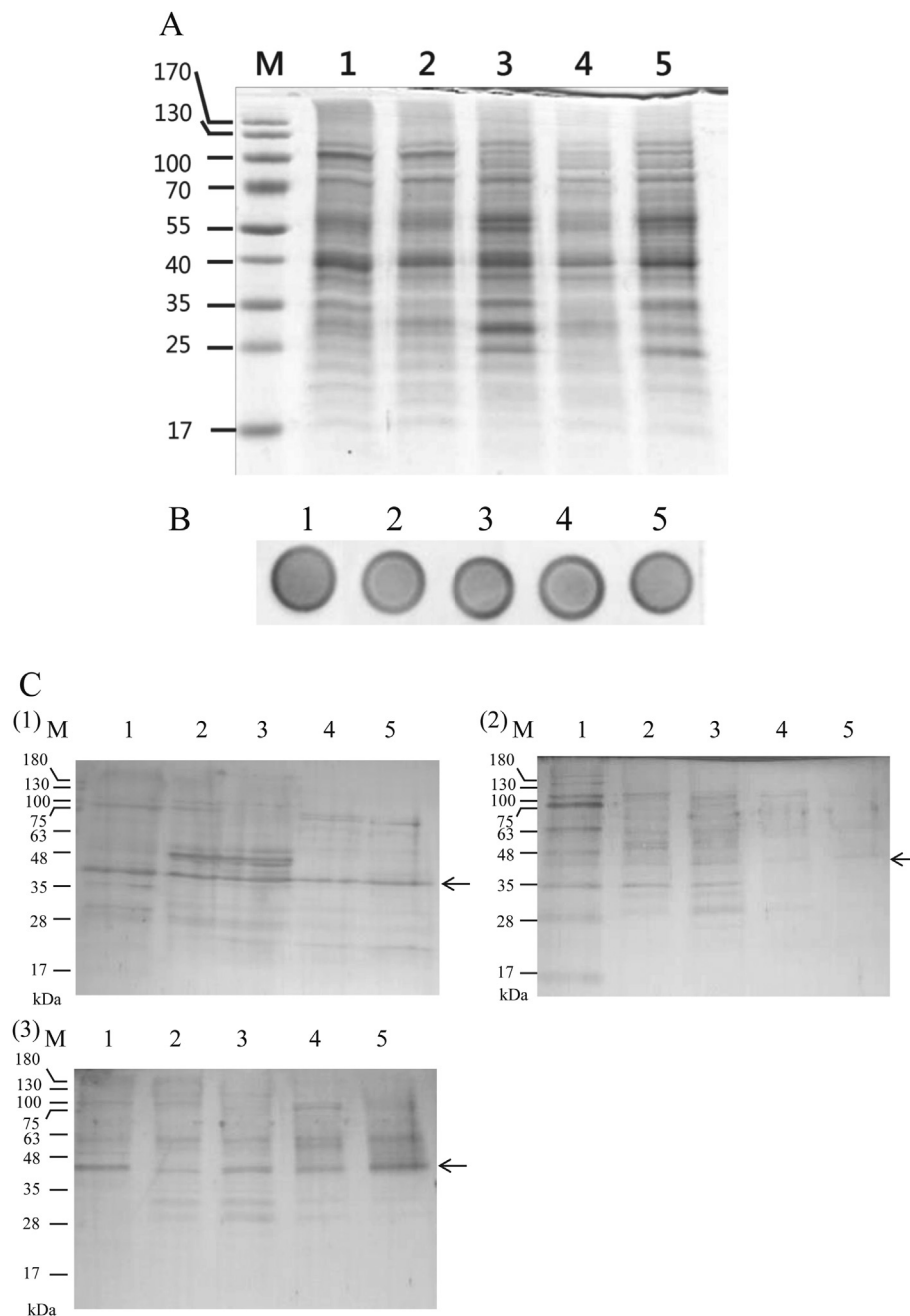


Fig. 1. (A) SDS-PAGE of protein extracts from various mango varieties. M: marker, Lane 1: Irwin, Lane 2: Yuwen, Lane 3: Jinhuang, Lane 4: Keitt, Lane 5: Tu. Dot blot (B) and Western blot (C) of different mango varieties with the sera of mango-allergic patients. C-1: Patient 1; C-2: Patient 2, and C-3: Patient 3. 1: Irwin, 2: Yuwen, 3: Jinhuang, 4: Keitt, 5: Tu.

enzymes that is involved in glycolysis and has been proved to be an allergen in kiwifruit, wheat and rambutan [25–27]. Based on the results of peptide mapping, two degenerate primers were constructed; the forward primer (5'-GAYACNGTNCAYGGNCARTGG-3') was derived from DTVHGQW while the reverse primer (5'-NCCCCAYTCRTTTRTCRTACCA-3') was derived from WYDNEWG. A PCR reaction yielded an 816 bp amplicon (Supplementary Fig. 2), which was verified using a nest PCR with a 189 bp amplicon (Supplementary Fig. 3). Finally, the DNA sequences of 816 bp and 189 bp amplicons were utilized in designing the primers for the 3' RACE and 5' RACE assays to obtain the full sequence of Man i 1. As displayed in Fig. 2, an open reading frame of the full cDNA length of Man i 1 has 1,023 base pairs that can be translated into a polypeptide chain with 341 amino acids with a predicted N-

glycosylation site at the 154th amino acid. However, the result of signal peptide sequence prediction showed that there was no signal peptide within Man i 1. This obtained sequence has been submitted to GenBank with the accession number KU975060.1.

3.3. Cloning of Man i 1 gene

Based on the Man i 1 cDNA sequence, a primer set was constructed with *Nde*I and *Xho*I sites. After a PCR reaction, the amplicon of Man i 1 was observed as a band around the 1000 bp (Fig. 3A). This amplicon was ligated with a TOPO vector to obtain the TOPO/Man i 1 plasmid for DNA sequencing and the preservation of the Man i 1 DNA fragment. For expression, the Man i 1 DNA fragment was cut from the TOPO/Man i 1 plasmid with *Nde*I and *Xho*I. As shown in

```

1  atg gcc gga ggc aag aag att aag atc gga atc aat ggt ttt gga
   M  A  G  G  K  K  I  K  I  G  I  N  G  F  G
46  agg atc ggt cgc ttg gtt gct aga gtt gct ctg cag agg aac gat
   R  I  G  R  L  V  A  R  V  A  L  Q  R  N  D
91  gtt gag ctt gtg gct gtt aac gat ccc ttc atc aca act gat tac
   V  E  L  V  A  V  N  D  P  F  I  T  T  D  Y
136 atg aca tac atg ttt aag tac gat act gtt cac ggt cag tgg aaa
   M  T  Y  M  F  K  Y  D  T  V  H  G  Q  W  K
181 cac cat gag gtg aag att aag gac caa aag acc ctt ctc ttc gat
   H  H  E  V  K  I  K  D  Q  K  T  L  L  F  D
226 gag aag cca gtc act gtt ttc ggt atc aga aac cct gag gag att
   E  K  P  V  T  V  F  G  I  R  N  P  E  E  I
271 cca tgg gct gag acc ggt gcc gag ttt gtt gtt gag tca act gga
   P  W  A  E  T  G  A  E  F  V  V  E  S  T  G
316 gtg ttc act gac aag gag aag gct gct gct cac ttg aag ggt ggt
   V  F  T  D  K  E  K  A  A  A  H  L  K  G  G
361 gcc gag aag gtt gtt att tcg gcc cca agc aaa gat gct ccc atg
   A  E  K  V  V  I  S  A  P  S  K  D  A  P  M
406 ttt gtt gtg ggt gta aat gag aaa gaa tac aag cca gaa ctt aac
   F  V  V  G  V  N  E  K  E  Y  K  P  E  L  N
451 att gtt tca aat gca agc tgcc aca aca aac tgc ctt gct cct ctc
   I  V  S  N  A  S  C  T  T  N  C  L  A  P  L
496 gca aag gtt atc aat gac aga ttt gga att gtt gag ggt ctt atg
   A  K  V  I  N  D  R  F  G  I  V  E  G  L  M
541 acc act gtc cac tcc atc act gcc acc cag aag act gtt gat gga
   T  T  V  H  S  I  T  A  T  Q  K  T  V  D  G
586 cca tca atg aag gac tgg aga ggt ggt aga gcc gca tcc ttc aac
   P  S  M  K  D  W  R  G  G  R  A  A  S  F  N
631 atc atc ccc agc agc act gga gct gcc aag gct gtc gga aaa gtt
   I  I  P  S  S  T  G  A  A  K  A  V  G  K  V
676 ctg cct tcc ctc aat ggc aaa ttg act ggc atg tct ttc cgt gtt
   L  P  S  L  N  G  K  L  T  G  M  S  F  R  V
721 cct aca gtt gat gtt tca gtt gtg gac ctt act gtc aga ctt gag
   P  T  V  D  V  S  V  V  D  L  T  V  R  L  E
766 aag gag gca acc tat gag gaa atc aaa gct gcc atc aaa gaa gag
   K  E  A  T  Y  E  E  I  K  A  A  I  K  E  E
811 tct gag ggc aag ctt aag ggc att ttg gga tac act gaa gaa gat
   S  E  G  K  L  K  G  I  L  G  Y  T  E  E  D
856 gtt gta tcc act gat ttc gtg ggt gac agc agg tca agc att ttc
   V  V  S  T  D  F  V  G  D  S  R  S  S  I  F
901 gat gcc aag gct ggt att gct ttg aac aag aaa ttt gtg aaa ctt
   D  A  K  A  G  I  A  L  N  K  K  F  V  K  L
946 gtc tct tgg tat gac aat gaa tgg gga tac agt tca cgc gtg att
   V  S  W  Y  D  N  E  W  G  Y  S  S  R  V  I
991 gac ttg att gtc cac atg gca tca act gct taa
   D  L  I  V  H  M  A  S  T  A  *

```

Fig. 2. Full cDNA sequence of Man i 1. Start codon is underlined. Symbol “*” indicates stop codon. Gray region is the predicted glycosylation site.

Fig. 3B, digestion yielded two fragments. The large one was the TOPO vector (~4000 bp) and the small one was the Man i 1 DNA fragment, which was subsequently ligated with pET-21a(+) vector. To confirm a successful insertion of the Man i 1 DNA fragment, this newly constructed plasmid was digested with *NdeI* and *XhoI*, resulted in a pET-21a(+) fragment (~5500 bp) and a Man i 1 DNA fragment (Fig. 3C), revealing that the pET-21a(+)/Man i 1 plasmid has been successfully constructed with the structure that is displayed in Supplementary Fig. 4.

3.4. Optimization of conditions for expression of rMan i 1

The conditions for the expression of rMan i 1 were optimized by

varying three major parameters, which were induction temperature, IPTG concentration, and induction time. Four commonly used induction temperatures (16, 20, 25, and 30 °C) were used to obtain the soluble form of rMan i 1. Effort was made to avoid the formation of an inclusion body, because such an inclusion body would require a complicated purification process with the potential of the formation of a misfolded structure after a refolding step. Among many methods for obtaining a soluble expressed protein, reducing the induction temperature is easy and frequently used. At low temperatures, bacteria typically exhibit a low rate of protein synthesis, and a correspondingly low probability of misfolding and aggregation of proteins, favoring a relatively high yield of the expressed protein [28]. In Fig. 4A, large amounts of the inclusion body of rMan

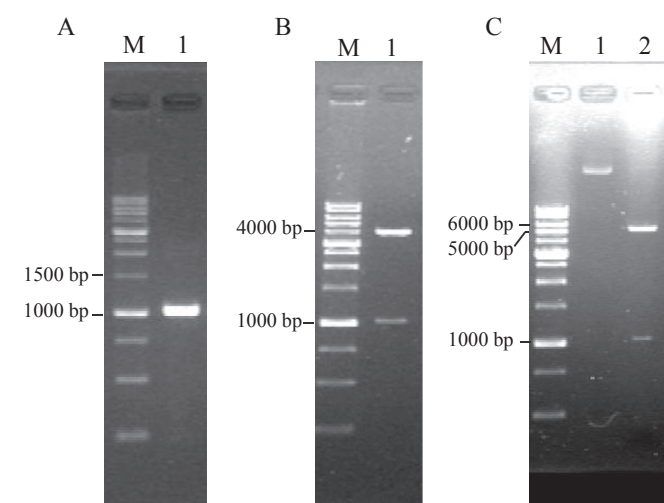


Fig. 3. (A) PCR product of Man i 1 with *NdeI* and *XhoI* sites was analyzed using agarose gel. (B) TOPO/Man i 1 plasmid was digested using *NdeI* and *XhoI* enzymes. (C) pET-21a(+)/Man i 1 plasmid was digested using *NdeI* and *XhoI* enzymes. Lane 1: undigested pET-21a(+)/Man i 1 plasmid. Lane 2: digested pET-21a(+)/Man i 1 plasmid.

i 1 are present in the pellet at high induction temperatures (20, 25, and 30 °C), but this amount is significantly lower at 16 °C. Therefore, 16 °C was selected as the optimal temperature for expressing the soluble rMan i 1.

The IPTG concentration in culture broth was optimized herein. IPTG is generally an effective inducer of the expression of a high level of recombinant proteins in the T7 expression system. However, IPTG is costly and toxic to *E. coli*, so its concentration must be optimized on a case by case basis [29]. As displayed in Fig. 4B, among the various IPTG concentrations used herein, 0.1 mM IPTG yielded the highest expression level of rMan i 1. The expression level of rMan i 1 gradually decreased as the IPTG concentration increased, indicating the toxicity of IPTG to *E. coli*. The situation was similar when *E. coli* was used to express acetolactate synthase. The amount of expressed acetolactate synthase declined significantly noticed as the IPTG concentration increased over 0.6 mM [30]. To reduce the cost and obtain a large amount of rMan i 1, 0.1 mM IPTG was chosen for the subsequent optimization of induction time. In Fig. 4C, the protein expression level increases with induction time from 0 to 20 h, reaching a plateau between 20 and 24 h. The same trend was evident when the T7 expression system was used to express ascorbate peroxidase and interleukin-37b [31,32]. Therefore, the optimizing the expression time is an effective way to increase the amount of recombinant protein obtained. Based on the results herein, the highest expression level of soluble rMan i 1 was obtained using with 0.1 mM IPTG at 16 °C for 20 h.

3.5. Purification of rMan i 1

Histidine (His) tagging has been extensively utilized to purify recombinant protein. Histidine residues can be tagged on either the N- or the C- terminus of the recombinant protein, which is purified using a Ni(II)-immobilized metal-affinity chromatograph owing to the high affinity between the histidine residues and Ni(II) ions [33,34]. In this study, rMan i 1 was His-tagged at its C-terminus. Following IPTG induction, soluble rMan i 1 was present in the supernatant of the cell lysate and then purified using a Ni-NTA column. Fig. 5A plots the elution profile of cell lysates through a Ni-NTA column, in which a large flow-through peak appeared initially at fraction number (FN) 5 to 25 and a signal peak appeared from FN 45 to 50 after elution. To obtain a clear protein profile

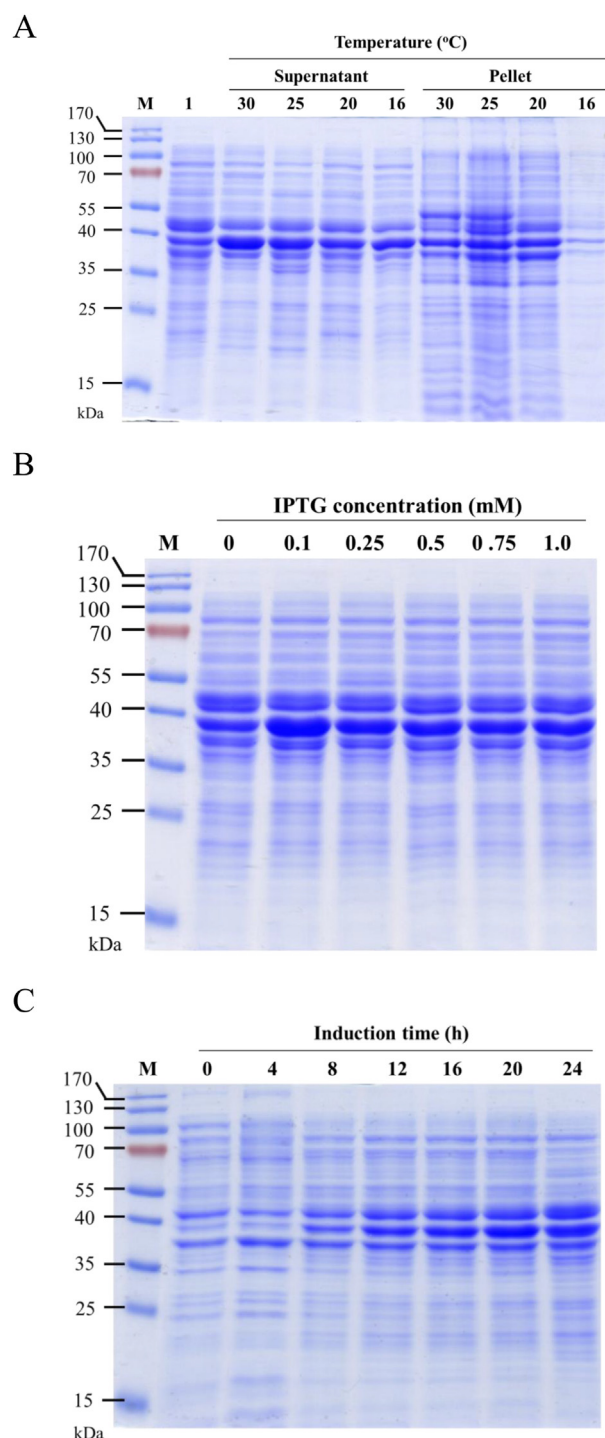


Fig. 4. SDS-PAGE analysis of optimal expression conditions for rMan i 1 protein. (A) *E. coli* BL21 (DE3) was induced with 0.1 mM IPTG at various temperatures for 16 h. M: marker. Lane 1: without IPTG. (B) *E. coli* BL21 (DE3) was induced with various concentrations of IPTG at 16 °C for 16 h. M: marker. (C) *E. coli* BL21 (DE3) was induced with 0.1 mM IPTG at 16 °C for various times. M: marker.

during elution, several fractions were selected and analyzed by SDS-PAGE. As shown in Fig. 5B, before purification, rMan i 1 was a major protein in the crude protein sample (lane 1), some of which was lost in the flow-through samples (lane 2–5). During the washing step (lane 6–9), most unbound proteins were removed. Subsequently, rMan i 1 was eluted out of the Ni-NTA column from FN 44 to 52, whose protein profiles were obtained by SDS-PAGE, as

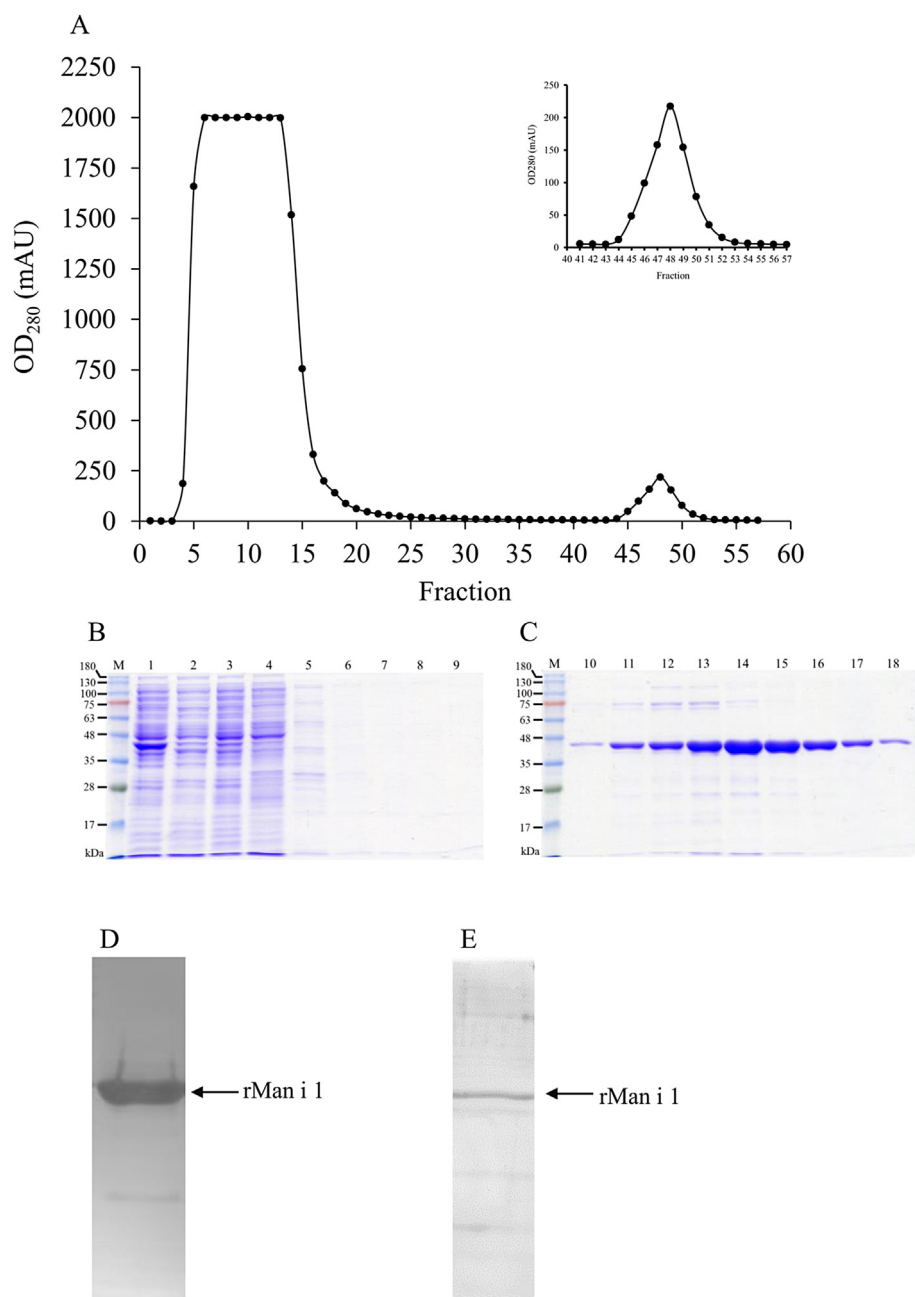


Fig. 5. Purification results of rMan i 1 using Ni-NTA column. (A) Purification profile of rMan i 1. (B) SDS-PAGE of fractions from flow-through step and wash step. M: marker. Lane 1: Before purification. Lanes 2–9: FN 5, 10, 15, 20, 25, 30, 35, and 40. (C) SDS-PAGE of fractions from elution step. M: marker. Lanes 10–18: FN 44–52. (D) Western blot of purified rMan i 1 was reacted with anti-histidine antibody. (E) Western blot of purified rMan i 1 was reacted with patient serum.

displayed in lanes 10–18 of Fig. 5C. It showed that rMan i 1 was the most abundant protein in the elution sample. To confirm the presence of His-tagged rMan i 1, the purified rMan i 1 sample was analyzed by Western blot using the anti-histidine antibody, and a strong band is observed in Fig. 5D. To evaluate the allergenicity of rMan i 1, rMan i 1 was reacted with the serum of a mango-allergic patient, and the band in Fig. 5E was thus obtained, indicating that the IgE antibodies in the patient's serum recognized rMan i 1.

3.6. Identity of amino acid sequence of rMan i 1

The purified rMan i 1 was analyzed with MOLDI-TOF and LC/MS/MS to determine whether its amino acid sequence matched that

deduced from its cDNA sequence. Fig. 6A demonstrates that 13 peptide fragments were identified by MOLDI-TOF, and these corresponded to 42% of the deduced amino sequence; while, the sequence that was identified by LC/MS/MS covered 70% of the deduced amino acid sequence (Fig. 6B). Hence, rMan i 1 was the protein that was expressed from the clone that contained the cDNA of Man i 1.

3.7. GAPDH activity of rMan i 1

The result of LC/MS/MS in this study showed that 50% of the amino acid sequence of Man i 1 was identical to that of GAPDH in *M. truncatula*. To examine the GAPDH activity of Man i 1, the GAPDH

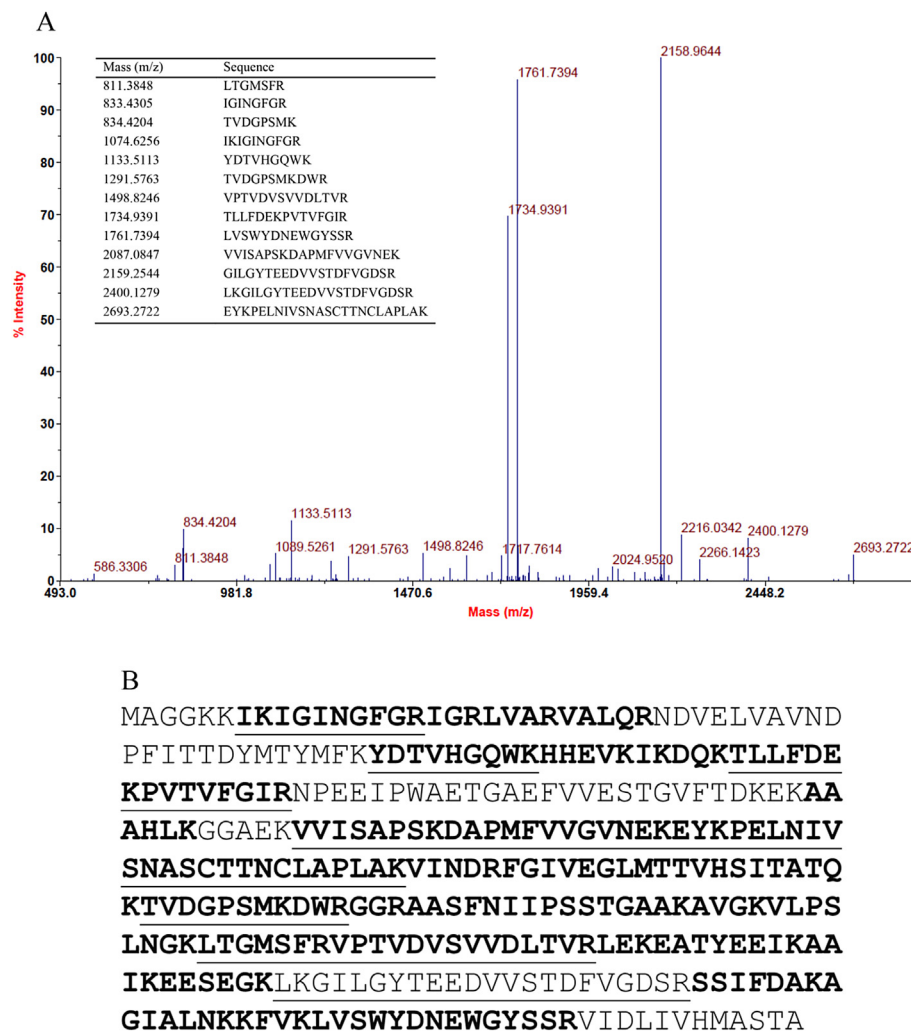


Fig. 6. (A) The analysis of rMan i 1 by MALDI-TOF. (B) Underlined sequences were the same as those in rMan i 1, based on MALDI-TOF result; and sequences in bold were the same as those in rMan i 1, based on LC/MS/MS result.

Man i 1	-----	-----	-----	-----MAGGKKIKIGINGFGR	IGRLVARVALQRND--VELV
<i>Triticum aestivum</i>	-----	-----	-----	-----MG---KIKIGINGFGR	IGRLVARVALQSPD--VELV
<i>Mangifera indica</i>	MASATFSVAKPSLQANG-KG	FTDFSGLRNSAS--LPFSRK	TSEDFLSVIAFQTSAVGNSG	-YRKGAEEAKLKVAINGFGR	IGRNFLRCWHGRKDSPLDVI
<i>Arachis hypogaea</i>	MASATLSVAKPALQNGGKG	FSEFSGLRNSSSYLPFSRK	TSDDFHSIISFQTSAVGSSG	GYRKGVTTEAKLKVAINGFGR	IGRNFLRCWHGRKDSPLDVI
Man i 1	AVNDPFITTDYMTYMFKYDT	VHGQWKHHEVKIKDQKTLLF	DEKPVTVFGIRNPEEIPWAE	TGAEFVVESTGVFTDKEKAA	AHLKGGAEKVVISAPSK-DA
<i>Triticum aestivum</i>	AVNDPFITTDYMTYMFKYDT	VHGQWKHHEVKIKDQKTLLF	GEKEVAVFGCRNPEEIPWAA	AGAEEVVESTGVFTDKEKAA	AHKGGAEKVVISAPSK-DA
<i>Mangifera indica</i>	AINDTGG-VKQASHLLKYDS	TLGIFD-ANVPVGDNAISV	DGKVIKVVTRNRPDLPLWGD	LGVDLVIEGTGVFVDRDAG	KHIQAGAKKVLITAPGKDI
<i>Arachis hypogaea</i>	AINDTGG-VKQASHLLKYDS	TLGIFD-ADVVPVEDAISV	DGKVIKVVSRNRPANLPWKE	LGVDLVIEGTGVFVDRDAG	KHIQAGAKKVLITAPGKDI
Man i 1	PMFVGVNEKEYKPELNIVS	NASCTTNCLAPLAKVINDRF	GIVEGLMTTVHSITATQKT	DGPSMKDWRGGRAASFNIIP	SSTGAAKAVGKVLPSLNGKL
<i>Triticum aestivum</i>	PMFVCGVNEKEYKSDIDIVS	NASCTTNCLAPLAKVINDRF	GIVEGLMTTVHAMTATQKT	DGPSKDWRRGGRAASFNIIP	SSTGAAKAVGKVLPELNGKL
<i>Mangifera indica</i>	PTYVGVNADAYNPDEPIIS	NASCTTNCLAPFVKVLDQKF	GIKGTMTTTHSYTDQRL	D-ASHRDLRRARAAALNIVP	TSTGAAKAVLVLPLKGL
<i>Arachis hypogaea</i>	PTYVGVNADAYSPDEPIIS	NASCTTNCLAPFVKVLDQKF	GIKGTMTTTHSYTDQRL	D-ASHRDLRRARAAALNIVP	TSTGAAKAVLVLPLKGL
Man i 1	TGMSFRVPTVDVSVVDLTVR	LEKEATYEEIKAAIKEESE	KLKGILGYTEEDVSTDFVG	DSRSSIFDAKAGIALNKKFV	KLVSWDNEWGYSSRVIDLI
<i>Triticum aestivum</i>	TGMAFRVPTVDVSVVDLTVR	LAKPATYDQIKAAIKEESE	NLKGILGYDEDLVSTDFQG	DNRSSIFDAKAGIALNDHFV	KLVSWDNEWGYSTRVVDLI
<i>Mangifera indica</i>	NGIALRVPTPNVSVVDLVVQ	VSKKTFAEVNAAFRESADK	ELKGILSVCEPLVSVDFRC	SDVSSTVDSSTLVMGDDMV	KVIAWDNEWGYSSRQVVDLA
<i>Arachis hypogaea</i>	NGIALRVPTPNVSVVDLVVQ	VSKKTFAEVNAAFRESAEK	ELNGILSVCEPLVSVDFRC	TDVSSTVDSSTLVMGDDMV	KVIAWDNEWGYSSRQVVDLA
Man i 1	VHMASTA-				
<i>Triticum aestivum</i>	RHMSTK-				
<i>Mangifera indica</i>	DIVANKWK				
<i>Arachis hypogaea</i>	DIVANSWK				

Fig. 7. A multiple alignment of amino acid sequences of Man i 1 with GAPDH from wheat (*Triticum aestivum*), mango (*Mangifera indica*), and peanut (*Arachis hypogaea*). Gray areas were the same sequences, and dashed lines were gaps.

activity of purified rMan i 1 was analyzed by measuring the formation of NADH at 340 nm. The result of Supplementary Fig. 5 shows that rMan i 1 had the GAPDH activity, and the absorbance increased by following the increase of reaction time. The GAPDH activity of rMan i 1 was estimated as 0.402 ± 0.019 U/mg. Moreover, the amino acid sequences of rMan i 1 were compared with GAPDH from other plants. As shown in Fig. 7, Man i 1 shares 86.18%, 45.59%, and 46.18% identity with GAPDH from wheat (*Triticum aestivum*), mango (*Mangifera indica*), and peanut (*Arachis hypogaea*), respectively.

4. Conclusion

Mango is a popular fruit but also a well-known source of allergens worldwide. This study analyzed the protein profiles of five mango species – Irwin, Yuwen, Jinhuang, Keitt, and Tu, all of which had a strong band at 40 kDa, which has been identified as a major mango allergen, Man i 1. Based on the results of this study, the full cDNA length of Man i 1 was 1,023 bp, which was consistent with a polypeptide with 341 amino acids. Additionally, this Man i 1 cDNA was inserted into a pET-21a(+) vector and transformed into *E. coli* BL21 (DE3) to produce rMan i 1. Following optimization, the highest amount of soluble rMan i 1 was obtained by inducing the transformed *E. coli* with 0.1 mM IPTG at 16 °C for 20 h. An Ni-NTA column was used to obtain a large amount of purified rMan i 1. This purified rMan i 1 was identified using the serum of a mango-allergic patient, and it was identified as GAPDH. Thus, this rMan i 1 can be used for immunotherapy against mango allergy or for producing antibodies to detect contamination by mango allergens.

Acknowledgments

The authors would like to thank the Ministry of Science and Technology, R.O.C., for financially supporting this research under Contract No. 101-2313-B-005-045-MY3. Ted Knoy is appreciated for his editorial assistance.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.pep.2016.06.009>.

References

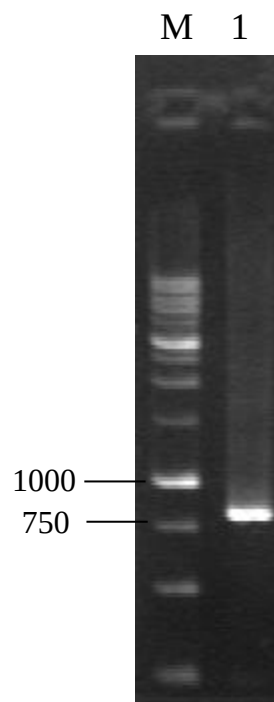
- [1] S.K. Mitra, E.A. Baldwin, Postharvest physiology storage of tropical and sub-tropical fruit, in: S.K. Mitra (Ed.), *Mango*, CAB International, Wallingford, 1997, pp. 85–122.
- [2] H.T. Codoy, D.B. Rodriguez-Amaya, Carotenoid composition of commercial mangoes from Brazil, *Leb. Wiss. Technol.* 22 (1989) 100–103.
- [3] J.M. Rubin, J. Shapiro, P. Muehlbauer, M. Grolnick, Shock reaction following ingestion of mango, *J. Am. Med. Assoc.* 193 (1965) 397–398.
- [4] J. Miell, M. Papouchado, A.J. Marshall, Anaphylactic reaction after eating a mango, *Br. Med. J.* 297 (1988) 1639–1640.
- [5] J.B. Vargas Correa, L. Sánchez Solís, J.A. Farfán Ale, H. Noguchi, M.T. Moguel Baños, M.I. Vargas de la Peña, Allergological study of pollen of mango (*Mangifera indica*) and cross reactivity with pollen of piru (*Schinus molle*), *Rev. Alerg.* 38 (1991) 134–138.
- [6] C. Fernandez, A. Fiandor, A. Martinez-Garate, J. Martinez Quesada, Allergy to pistachio: crossreactivity between pistachio nut and other Anacardiaceae, *Clin. Exp. Allergy* 25 (1995) 1254–1259.
- [7] M. Etesamifar, B. Wüthrich, IgE-mediated food allergies including oral allergy syndrome in 383 patients, *Allergologie* 21 (1998) 451–457.
- [8] F. Andre, C. Andre, L. Colin, F. Cacaraci, S. Cavagna, Role of new allergens and of allergen consumption in the increased incidence of food sensitizations in France, *Toxicology* 93 (1994) 77–83.
- [9] P. Pumhirun, P. Towiwat, P. Mahakit, Aeroallergen sensitivity of Thai patients with allergic rhinitis, *Asian Pac. J. Allergy Immunol.* 15 (1997) 183–185.
- [10] S. Duque, L. Fernández-Pellón, F. Rodríguez, Mango allergy in a latex-sensitized patient, *Allergy* 54 (1999) 1004–1005.
- [11] A. Paschke, H. Kinder, K. Zunker, M. Wigotzki, R. Weißbecher, D. Vieluf, H. Steinhart, Characterization of allergens in mango fruit and ripening dependence of the allergenic potency, *Food Agric. Immunol.* 13 (2001) 51–61.
- [12] R. Sareen, A. Shah, Hypersensitivity manifestations to the fruit mango, *Asia Pac. Allergy* 1 (2011) 43–49.
- [13] H. Kinder, B. Scharf, H. Steinhart, A. Paschke, Investigation of IgE-binding patterns and allergenicity of mango fruit allergens in different varieties, *Internet Symp. Food Allerg.* 1 (1999) 43–49.
- [14] M. Dube, K. Zunker, S. Neidhart, R. Carle, H. Steinhart, A. Paschke, Effect of technological processing on the allergenicity of mangoes (*Mangifera indica* L.), *J. Agric. Food Chem.* 52 (2004) 3938–3945.
- [15] R. Renner, C. Hippler, R. Treudler, W. Harth, A. Süß, J.C. Simon, Identification of a 27 kDa protein in patients with anaphylactic reactions to mango, *J. Invest. Allergol. Clin. Immunol.* 18 (2008) 476–481.
- [16] S.C. Makrides, Strategies for achieving high-level expression of genes in *Escherichia coli*, *Microbiol. Rev.* 60 (1996) 512–538.
- [17] F. Baneyx, M. Mujacic, Recombinant protein folding and misfolding in *Escherichia coli*, *Nat. Biotechnol.* 22 (2004) 1399–1408.
- [18] G. Georgiou, P. Valax, Expression of correctly folded proteins in *Escherichia coli*, *Curr. Opin. Biotechnol.* 7 (1996) 190–197.
- [19] H.P. Sorensen, K.K. Mortensen, Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*, *Microb. Cell Fact.* 4 (2005) 1–8.
- [20] P. Singh, U.N. Dwivedi, Purification and characterisation of multiple forms of polygalacturonase from mango (*Mangifera indica* cv. Dashehari) fruit, *Food Chem.* 111 (2008) 345–349.
- [21] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* 227 (1970) 680–685.
- [22] H. Chou, M.F. Tam, S.S. Lee, R.B. Tang, T.H. Lin, H.Y. Tai, Y.S. Chen, H.D. Shen, Asp159 is a critical core amino acid of an IgE-binding and cross-reactive epitope of a dust mite allergen Der f 7, *Mol. Immunol.* 48 (2011) 2130–2134.
- [23] J. Song, H. Zhang, Z. Liu, P. Ran, Mango profilin: cloning, expression and cross-reactivity with birch pollen profilin Bet v 2, *Mol. Biol. Rep.* 35 (2008) 231–237.
- [24] L. Egea, L. Aguilera, R. Giménez, M.A. Sorolla, J. Aguilar, J. Badía, L. Baldoma, Role of secreted glyceraldehyde-3-phosphate dehydrogenase in the infection mechanism of enterohemorrhagic and enteropathogenic *Escherichia coli*: interaction of the extracellular enzyme with human plasminogen and fibrinogen, *Int. J. Biochem. Cell Biol.* 39 (2007) 1190–1203.
- [25] A. Palacin, J. Rodriguez, C. Blanco, G. Lopez-Torrejon, R. Sánchez-Monge, J. Varela, M.A. Jiménez, J. Cumpido, T. Carrillo, J.F. Crespo, G. Salcedo, Immunoglobulin E recognition patterns to purified kiwifruit (*Actinidia deliciosa*) allergens in patients sensitized to kiwi with different clinical symptoms, *Clin. Exp. Allergy* 38 (2008) 1220–1228.
- [26] I. Sander, A. Flagge, R. Merget, T.M. Halder, H.E. Meyer, X. Baur, Identification of wheat flour allergens by means of 2-dimensional immunoblotting, *J. Allergy Clin. Immunol.* 107 (2001) 907–913.
- [27] O. Jirapongsananuruk, N. Jirattanasopa, S. Pongpruksa, P. Vichyanond, S. Piboonpocanun, Glyceraldehyde-3-phosphate dehydrogenase as a major allergen in rambutan-induced anaphylaxis, *Ann. Allergy Asthma Immunol.* 106 (2011) 545–547.
- [28] J.A. Vasina, F. Baneyx, Expression of aggregation-prone recombinant proteins at low temperatures: a comparative study of the *Escherichia coli* cspA and tac promoter systems, *Protein. Expr. Purif.* 9 (1997) 211–218.
- [29] O. Onodera, A.D. Roses, S. Tsuji, J.M. Vance, W.J. Strittmatter, J.R. Burke, Toxicity of expanded polyglutamine-domain proteins in *Escherichia coli*, *FEBS Lett.* 399 (1996) 135–139.
- [30] P. Zheng, X. Sun, L. Guo, J. Shen, Cloning, expression, and characterization of an acetolactate synthase (ALS) gene from *Anabaena azotica*, *Process Biochem.* 50 (2015) 1349–1356.
- [31] L. Lin, X. Wang, Y. Wang, cDNA clone, fusion expression and purification of the novel gene related to ascorbate peroxidase from Chinese wild *Vitis pseudoreticulata* in *E. coli*, *Mol. Biol. Rep.* 33 (2006) 197–206.
- [32] J. Gu, X. Gao, X. Pan, X. Peng, Y. Li, M. Li, High-level expression and one-step purification of a soluble recombinant human interleukin-37b in *Escherichia coli*, *Protein Expr. Purif.* 108 (2015) 18–22.
- [33] J.A. Bornhorst, J.J. Falke, Purification of proteins using polyhistidine affinity tags, *Methods Enzymol.* 326 (2000) 245–254.
- [34] K. Terpe, Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems, *Appl. Microbiol. Biotechnol.* 60 (2002) 523–533.

```

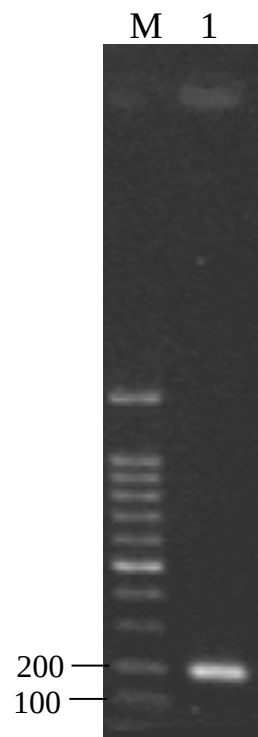
1  MASDKKIRIG  INGFGRIIGRL  VARVALQRND
31  VELVAINDPF  ITTDYMTYMF  KYDTVHGQWK
61  HFELKVKDTK  TLLFGEKEVA  VFGTRNPEEI
91  PWGEVGADYV  IESTGVFTDK  DKAAAHLKGG
121 AKKVVISAPS  KDAPMFVVG  V  NEKEYKPELN
151 IVSNASCTTN  CLAPLAKVIN  DRFGIVEGLM
181 TTVHAITATQ  KTVDGPSSKD  WRGGRAASFN
211 IIPSSTGAAK  AVGKVLPLN  GKLTGMSFRV
241 PTVDVSVVDL  TVRLEKKATY  EEIKAAIKEE
271 SEGKLGILG  YTEDDVVSTD  FIGDSRSSIF
301 DAKAGIALND  NFVKLVSWYD  NEWGYSTRVI
331 DLIVHIASVA

```

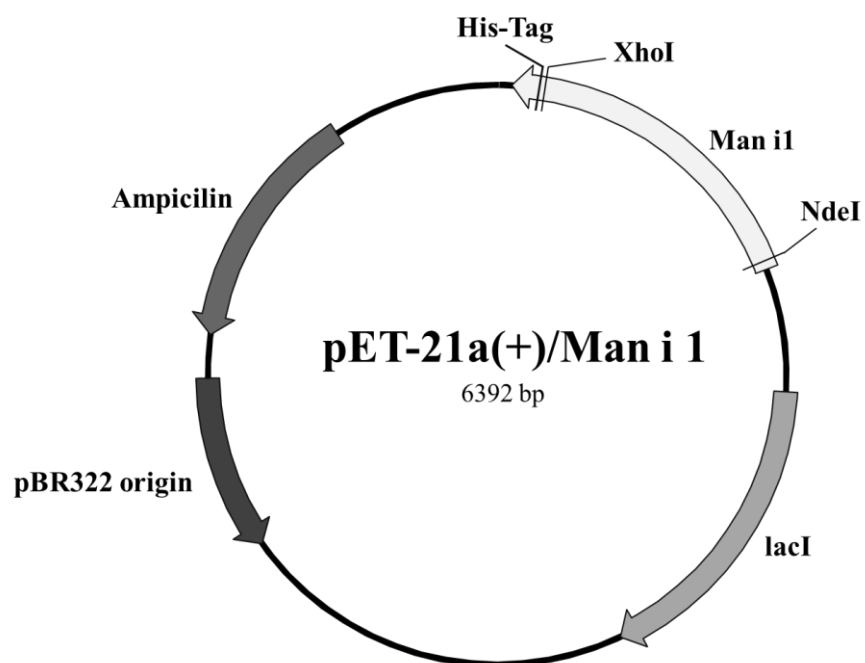
Supplementary Fig. 1. Peptide mapping analysis of Man i 1. The amino acid sequences that are same as those in glyceraldehyde 3-phosphate dehydrogenase from *Medicago truncatula*, are shaded.



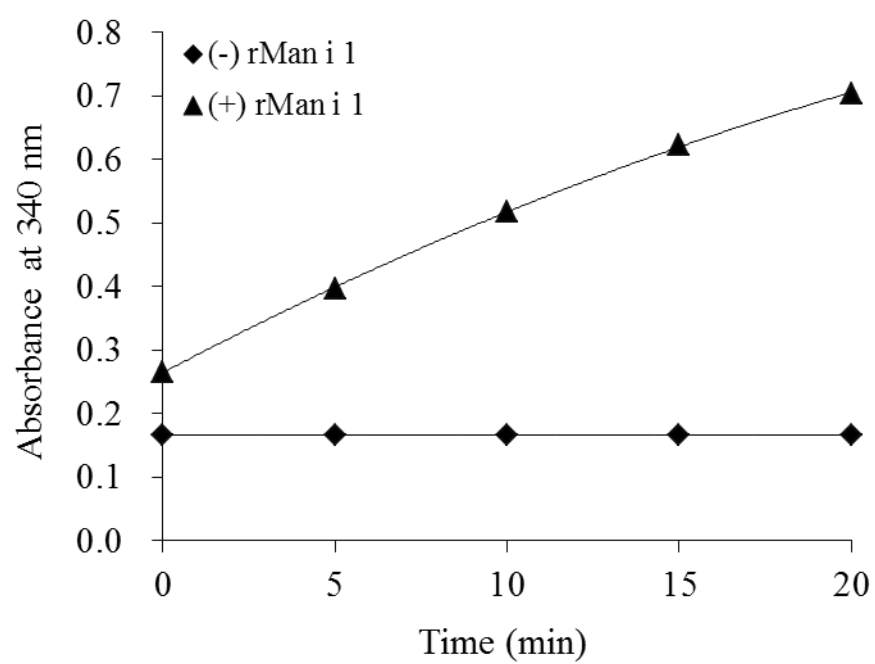
Supplementary Fig. 2. PCR product amplified using degenerate primers.



Supplementary Fig. 3. Nested PCR product amplified using degenerate primers.



Supplementary Fig. 4. Construction of a expression plasmid by inserting Man i 1 into pET-21a(+).



Supplementary Fig. 5. The glyceraldehyde 3-phosphate dehydrogenase (GAPDH) activity of rMan i 1.