Identification and Molecular Characterization of the cDNA Encoding *Cucumis melo* Allergen, Cuc m 3, a Plant Pathogenesis-Related Protein

Mojtaba Sankian¹, Jafar Hajavi¹, Malihe Moghadam¹, Abdol-Reza Varasteh*²

**Abstract**

**Background:** Melon (*Cucumis melo*) allergy is one of the most common food allergies, characterized by oral allergy syndrome. To date, two allergen molecules, Cuc m 1 and Cuc m 2, have been fully characterized in melon pulp, but there are few reports about the molecular characteristics of Cuc m 3.

**Methods:** The Cuc m 3 cDNA has been characterized by rapid amplification of cDNA ends (RACE), which revealed a 456 base-pair (bp) fragment encoding a 151-amino acid polypeptide with a predicted molecular mass of 16.97 kDa, and identified 79 and 178 bp untranslated sequences at the 5’ and 3’ ends, respectively.

**Results:** In silico analysis showed strong similarities between Cuc m 3 and other plant pathogen-related protein 1s from cucumber, grape, bell pepper, and tomato.

**Conclusion:** Here we report the identification and characterization of the Cuc m 3 cDNA, which will be utilized for further analyses of structural and allergenic features of this allergen.

**Keywords:** Allergen, Cuc m 3, Melon, Plant pathogenesis-related protein 1

**Introduction**

Melon (*Cucumis melo*) belongs to the gourd family, Cucurbitaceae, which also includes cucumber, pumpkin, squash, and watermelon. Honeydew, cantaloupe, and muskmelon represent some of the most common hybrids and cultivars of this family (1, 2). Recent studies have revealed that up to 7% of young children and about 4% of adults suffer from some type of food allergy (3). Melon (*C. melo*) allergy is one of the most common food allergies, characterized by oral itching, lip swelling, and labial edema (4). The initial report of the melon sensitivity was closely linked to the early description of the oral allergy syndrome (OAS). In 1970, Anderson et al. reported a case series of patients with ragweed allergy who experienced oral symptoms after eating various melons (eg, watermelon, cantaloupe, and honeydew) and bananas (5). Ortolani et al. described an association between allergy to grass pollen and some vegetable hypersensitivity, such as tomato, melon, and watermelon (6). In a study of patients with pollen allergies, about one fifth of the patients showed IgE sensitivities to melon and pollen (7). Specific IgE assays suggest that some common antigenic epitopes exist between melon and grass pollen allergens (8, 9). Additionally, there are anecdotal reports of anaphylactic reactions to melon and ethanol-induced anaphylaxis after the ingestion of overripe melon (10, 11). More recently, Brehler et al. detected specific IgE antibodies to a wide variety of fruits, including melon, in 69% of serum samples from 136 patients with latex allergies (12).
To date, three melon allergens have been identified in melon pulp including Cuc m 1, a subtilisin-like protease (13), Cuc m 2, a profilin (4, 14), and Cuc m 3 (15), a plant pathogen-related protein 1 (PR-1). Cuc m 1 and Cuc m 2, which were recognized by more than 50% of the patients’ sera, have been characterized by several research groups (15). Cuc m 3 belongs to the pathogen-related family and exhibits some amino acid similarities with the other members of this family in grape and cucumber. This allergen, as a PR-1 protein, is highly stable and insensitive to proteases that are found in the melon fruit juices (15). PR-1 is a dominant group of PRs induced by pathogens. Since their discovery in 1970, numerous researchers have attempted to assess the function of PR-1 proteins in plants, but with little success (15, 16). The aims of this study were to clone and characterize the cDNA encoding Cuc m 3, and use it to analyze structural and allergenic features of this allergen.

Materials and Methods

Isolation of RNA, first-strand cDNA synthesis, and amplification of melon Cuc m 3 cDNA

Total RNA was extracted from 1 g of fine powder ground from the inner layer of melon in liquid nitrogen by means of the guanidine thiocyanate extraction method [19]. First-strand cDNA was synthesized from 2 µg of total RNA using a first-strand cDNA synthesis Kit (Fermentas, Lithuania) with an oligo (dT) 18-containing primer (5'-CGCTACGTAACGGCATGACAGTGTTTTTTT TTTTTTTTTTTT-3'). Amplification of the unknown sequence of the 3' end from melon Cuc m 3 cDNA was carried out by 3' rapid amplification of cDNA ends (RACE) primer (GeneRacerTM kit, Invitrogen, San Diego, CA, USA) and degenerate primers (Table 1) that were designed based on highly-conserved sequences in the coding regions of several members of PR-1 family found in Genebank. The amplification was performed with the following conditions: 3 min at 95 °C, then the first 5 cycles of denaturation at 95 °C for 30 s, annealing at 53 °C for 120 s and elongation at 72 °C for 60 s; then 25 continuous cycles of denaturation at 95 °C for 30 s, annealing at 58 °C for 60 s and elongation at 72 °C for 45 s; and finally elongation for at 72 °C for 3 min on a Corbett Research thermocycler (Sydney, Australia). Ten microliters of the polymerase chain reaction (PCR) product were electrophoresed on a 2% ethidium-stained agarose gel and documented on a G-Box gel documentation system (Syngene, Cambridge, UK). After amplification, PCR products were cleaned using Gene Clean II kit (Q-Biogen, Illkirch, France), sub-cloned into the vector pTZ57R/T with the TA cloning kit (Fermentas, Lithuania), and sequenced (MWG Biotech AG, Ebersberg, Germany).

To determine the 5' end of Cuc m 3 cDNA, a GeneRacerTM RNA oligo tail was ligated to the 5' end of the purified cDNA. These reactions created first-strand cDNA with known priming sites at the 5' end. Then, the 5' untranslated region (UTR) of the cDNA was amplified with a GeneRacerTM 5' primer and a gene-specific primer (5'-GGCAAGATGTTGCGCCATTACAC-3'), which was designed based on the Cuc m 3 3' cDNA sequence. The cDNA 3' end was also amplified by PCR using a gene-specific primer (3'-CACAGAATTTCTCCACACCACATTGGAG-3') and a GeneRacerTM 3' primer. All RACE-PCRs were carried out under the following program: 3 min at 94 °C, then 38 cycles of denaturation at 95 °C for 60 s, annealing at 68 °C for 60 s, and elongation at 72 °C for 60 s; then a final elongation at 72 °C for 5 min on a Corbett Research thermocycler (Sydney, Australia). After amplification of the 5' and 3' cDNA ends, the PCR products were sub-cloned into the vector pTZ57R/T with the TA cloning kit (Fermentas, Lithuania). Clones carrying inserts were characterized by restriction analyses and sequencing. All of the amplifications for T/A cloning were performed with Pfu DNA polymerase, and poly-A tails were added using Taq DNA polymerase.

Table 1. The degenerate primers designed based on highly conserved sequence in the coding regions of several plant PR-1 sequences

<table>
<thead>
<tr>
<th>Name</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>5' GCNGTGAVAVTTGCTGGGT 3'</td>
</tr>
<tr>
<td>F2</td>
<td>5' CGKGCMAAGTYGGVYGG 3'</td>
</tr>
<tr>
<td>F3</td>
<td>5' GGTGTTTYYWCHTGCAACTA 3'</td>
</tr>
<tr>
<td>F4</td>
<td>5' CAAAGTGAGGTGCCACAAAATACG 3'</td>
</tr>
</tbody>
</table>

Analysis of predicted protein sequence

The predicted molecular mass and isoelectric point were determined by the Gene Runner program v 3.05 (Hustington software). The deduced protein sequence of Cuc m 3 was next subjected to a protein-protein BLAST similarity search. Multiple sequence
alignment was performed by BioEdit software. The deduced amino acid sequences of PR-1s were obtained from the NCBI Protein Database with the following accession numbers: Q8S3W2 (Cucumis sativus), XP002273416 (Vitis vinifera), Q7XAJ6 (Vitis vinifera), ADB54823 (Vitis pseudoreticulata), EEE87889 (Populus trichocarpa), BAB78476 (Solanum torvum), AAK30143 (Capsicum annuum), ACB88202 (Solanum lycopersicum), and AAU15051.1 (Cynodon dactylon).

**Results**

**cDNA cloning and sequencing of the Cuc m 3 gene**

PCR of the melon cDNA containing the defined sequence at the 3′ end by the degenerate 5′ forward primer (F2) and the 3′ primer (5′-CGCTACGTAAACGGCATGACAGTG-3′) resulted in a 680-bp product (Fig. 1) that was cloned into pTZ57R/T. Four of these clones were sequenced (Table 1). After sequencing, two primers were designed based on the known sequence to obtain the full sequence of Cuc m 3, including the 5′ and 3′ ends. The amplified 3′-end sequence confirmed the 3′-end sequence generated by degenerate forward and defined reverse primers. By the 5′-RACE technique, a 480-bp fragment was amplified with an overlapping sequence with the obtained 3-end sequence (Fig. 1).

Finally, a 456 bp nucleotide sequence representing the complete cDNA sequence of Cuc m 3, including 79 and 178 bps of the 5′ and 3′ UTRs, respectively, was obtained by the cluster analysis of all the above fragments (Fig. 2).

The nucleotide sequences for the coding region and 5′ UTR of Cuc m 3 were deposited in the GenBank database under the accession numbers of EU556704.1 and EU679402.1, respectively. The amino acid sequence of this protein can be accessed through the NCBI Protein Database, Accession Number ACB45874.1.

**Nucleotide and amino acid sequence analysis of Cuc m 3 cDNA**

Sequencing analysis revealed a 453-bp open-reading frame (EU556704), which encodes Cuc m 3, a 151-amino-acid polypeptide (ACB45874.1) with a predicted molecular mass of 16.97 kDa and a theoretical isoelectric point value of 9.47 (Fig. 2).
potential N-glycosylation site was identified at amino acids 33-36 (NKTV) using the Gene Runner program.

The deduced protein sequence was subjected to a protein-protein BLAST similarity.

<table>
<thead>
<tr>
<th>Accession</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAB78467</td>
<td>Cucumis melo</td>
</tr>
<tr>
<td>Q8S3W2</td>
<td>Cuc m 3 fragments Ref. 12</td>
</tr>
<tr>
<td>XPO2273416</td>
<td>Vitis vinifera</td>
</tr>
<tr>
<td>BAF59881</td>
<td>Vitis hybrid</td>
</tr>
<tr>
<td>QT7A67</td>
<td>Vitis vinifera</td>
</tr>
<tr>
<td>ADB54823</td>
<td>Vitis pseudoreticulat</td>
</tr>
<tr>
<td>EEE87889.1</td>
<td>Populus trichocarpa</td>
</tr>
<tr>
<td>BAF78476</td>
<td>Solanum torvum</td>
</tr>
<tr>
<td>AA30143</td>
<td>Capsicum annuum</td>
</tr>
<tr>
<td>ACD88202</td>
<td>Solanum lycopersicum</td>
</tr>
<tr>
<td>AAU15051.1</td>
<td>Cydonia dactylon</td>
</tr>
<tr>
<td>P35782</td>
<td>Vespa cabro</td>
</tr>
</tbody>
</table>

Comparison of amino acid sequence deduced from Cuc m 3 cDNA (ACB45874) with other plant PR-1 from Cucumis sativus (Q8S3W2), Vitis vinifera (XPO2273416), Vitis hybrid (VB59881), Vitis vinifera (QT7A67), Vitis pseudoreticulat (ADB54823), Populus trichocarpa (EEE87889), Solanum torvum (BAF78476), Capsicum annuum (AA30143), Solanum lycopersicum (ACD88202) and Cydonia dactylon (AAU15051.1) and Vespa cabro (P35782). The amino acid sequence identity and similarity with Cuc m 3 were indicated at the end of the alignment. Black shading indicates identical amino acids.

Fig. 3. Comparison of amino acid sequence deduced from Cuc m 3 cDNA (ACB45874) with other plant PR-1 from Cucumis sativus (Q8S3W2), Vitis vinifera (XPO2273416), Vitis hybrid (VB59881), Vitis vinifera (QT7A67), Vitis pseudoreticulat (ADB54823), Populus trichocarpa (EEE87889), Solanum torvum (BAF78476), Capsicum annuum (AA30143), Solanum lycopersicum (ACD88202) and Cydonia dactylon (AAU15051.1) and Vespa cabro (P35782). The amino acid sequence identity and similarity with Cuc m 3 were indicated at the end of the alignment. Black shading indicates identical amino acids.
Molecular characterization of Cuc m 3

search, which identified a sperm-coating protein (SCP)-like extracellular protein domain in the deduced Cuc m 3 amino acid sequence. A protein homology search using BLAST software demonstrated that Cuc m 3 shared 63, 57, 55, 53, 50, 50, and 37% identical residues with pathogenesis-related protein PR-1 of cucumber (Cucumis sativus), grape (Vitis vinifera), black cottonwood (Populus trichocarpa), Turkey berry (Solanum torvum), bell pepper (Capsicum annuum), tomato (Solanum lycopersicum), and Bermuda grass (Cynodon dactylon), respectively. The highest degrees of 74% to 55% amino acid sequence similarities were with proteins in the PR-1 family (Fig. 3). All proteins with significant matches had sizes similar to the Cuc m 3 protein, ranging from 141 to 176 amino acids.

Discussion
In this work, we report the complete cDNA sequence of Cuc m 3 from Cucumis melo using 5’ and 3’RACE and degenerate primers designed based on the partial amino acid sequences of members of the PR-1 family (15). Previously reported partial amino acid sequences (40 amino acids) indicate 10% differences with the deduced amino acid derived from the complete Cuc m 3 cDNA. The complete amino acid sequence of Cuc m 3 showed a strong similarity (more than 60% sequence identity) with the other members of PR-1 family from grape and cucumber. This fact, combined with the presence of an SCP-like extracellular protein domain in the deduced Cuc m 3 amino acid sequence, implies that this allergen belongs to the plant pathogenesis-related protein 1 (PR-1) family. This PR protein accumulates after infections with pathogens, and may act as an anti-fungal agent or be involved in cell wall loosening. Many plant allergens from food and pollen have been found to be PR proteins (15-18). Because PR proteins can be induced by stress conditions, such as pathogen infection (19), they could be potential cross-reacting allergens of plant foods and their quantitative presence can vary with growing, harvesting, and storage conditions. Their weak homology to group 5 allergens from insect venoms could link food allergy and hypersensitivity to insect stings in some patients (20). In conclusion, in this report we identified and characterized the cDNA encoding Cuc m 3, which can be utilized in in vitro expression and potentially aid in more detailed analyses of structural and allergenic features of this allergen than were previously possible.

Acknowledgements
This study is conducted and financially supported by the research administration of Mashhad University of Medical Sciences. This article is derived from the master's thesis of the last author (Thesis No.A-234).

References


