

Sweet Potato Leaf Curl Virus: Coat Protein Gene Expression in *Escherichia coli* and Product Identification by Mass Spectrometry

Dina Lida Gutierrez Reynoso, Rodrigo A. Valverde, Norimoto Murai

Department of Plant Pathology & Crop Physiology, College of Agriculture and LSU Agricultural Center, Louisiana State University, Baton Rouge, USA

Email: nmurai@lsu.edu

Received 17 September 2015; accepted 4 December 2015; published 7 December 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Sweet potato is one of the first natural GMOs, genetically modified 8000 years ago by *Agrobacterium rhizogenes* as reported recently by Kyndt *et al.* A section of 10 kbp long DNA (Transferred-DNA or T-DNA) of the Ri (Root-inducing) plasmid was transferred to the plant genome by *A. rhizogenes* and has been maintained in all 291 hexaploid sweet potato cultivars of the world. The maintenance in the sweet potato genome and expression of two T-DNA genes for tryptophan-2-monooxygenase (*iaaM*) and for indole-3-acetamide hydrolase (*iaaH*) are likely to be physiologically significant since these enzymes convert tryptophan to indole-3-acetic acid, a major plant growth hormone auxin. Sweet potato (*Ipomoea batatas* (L.) Lam) is ranked the third most important root crop after potato and cassava, and the seventh in global food crop production with more than 126 million metric tons. Although sweet potato originated in Central or South America, China currently produces over 86% of world production with 109 million metric tons. In the United States, North Carolina is the leading producer with 38.5% of the 2007 sweet potato production, followed by California, Mississippi, and Louisiana with 23%, 19%, and 15.9%, respectively. Leaf curl virus diseases have been reported in sweet potato throughout the world. One of the causal agents is *Sweet potato leaf curl virus* (SPLCV) belonging to the genus *Begomovirus* (family *Geminiviridae*). Although SPLCV does not cause symptoms on Beauregard, one of the most predominant sweet potato cultivars in the US, it can reduce the yield up to 26%. Serological detection of SPLCV is not currently available due to the difficulties in obtaining purified virions that can be used as antigen for anti-serum production. In attempts to obtain the coat protein (CP) of SPLCV for antibody production, primers were designed to amplify the CP gene. This gene was cloned into the expression vector pMAL-c2E as a fusion protein with maltose-binding protein, and transformed into *Escherichia coli* strain XL1-Blue. After gene induction, a fusion protein of 72 kDa was purified by amylose affinity chromatography. The yield of the purified fusion protein was approximately 200 µg/liter of bac-

terial culture. Digestion with enterokinase cleaved the fusion protein into a 42.5 kDa maltose-binding protein and a 29.4 kDa protein. The latter protein was identified by mass spectrometry analysis as the coat protein of SPLCV based on the fact that the mass spectrometry elucidated the sequences corresponding to 37% of amino acid positions of the SPLCV coat protein.

Keywords

Affinity Chromatography Purification, Coat Protein, *Escherichia coli*, Mass Spectrometry, Maltose Binding Protein, Sweet Potato Leaf Curl Virus

1. Introduction

Sweet potato is one of the first natural Genetically Modified Organisms (GMOs), genetically modified 8000 years ago by *Agrobacterium rhizogenes* [1]. A section of 10 kbp long DNA (Transferred-DNA or T-DNA) of the Ri (Root-inducing) plasmid was transferred to the plant genome by *A. rhizogenes* and has been maintained in all 291 hexaploid sweet potato cultivars of the world. The maintenance in the sweet potato genome and expression of two T-DNA genes for tryptophan-2-monooxygenase (*iaaM*) and for indole-3-acetamide hydrolase (*iaaH*) are likely to be physiologically significant since these enzymes convert tryptophan to indole-3-acetic acid, a major plant growth hormone auxin.

With more than 126 million metric tons (FAO, 2007) in annual production, sweet potato (*Ipomoea batatas* (L.) Lam) is the third most important root crop after potato and cassava and is ranked seventh in global food crop production [2]. Due to its nutritional qualities (rich in carbohydrates, dietary fiber, β -carotene, vitamin C, and vitamin B6), sweet potato is considered as a crop with great nutritional value not only for human consumption but also for animal feeding and industrial use [3]. Although sweet potato originated in Central or South America, the current world production is centered in Asia, with China as the major producer with 109 million metric tons that counts for over 86% of the sweet potato world production (FAO, 2007). The United States is one of the few developed countries that produce sweet potatoes (836,970 metric tons) with less than 1% of the 2007 world production (FAO, 2007). North Carolina is the leading producer with 38.5% of the 2007 US sweet potato production, followed by California, Mississippi, and Louisiana with 23%, 19%, and 15.9%, respectively (USDA, 2008).

Sweet potato belongs to the family *Convolvulaceae* (the morning glory family) and grows widely in tropical, subtropical, and warm temperate climates. Sweet potato is especially valued because it is highly adaptable, tolerates high temperatures and low fertility soils, is easy to propagate and maintain, and yields well in adverse conditions [4]. Many diseases caused by fungi, bacteria, nematodes, viruses, and mycoplasma have been described to affect sweet potato production [5]. Because of its vegetative propagation, sweet potato is prone to accumulate systemic pathogens in propagating materials [6].

Sweet potato leaf curl virus (SPLCV), a whitefly-transmitted geminivirus, was originally found in an ornamental sweet potato and some breeding lines in the United States [7]-[12]. Today, this virus is now found infecting sweet potato throughout the world [13]-[15]. Leaf curling is a common symptom on infected plants (*I. batatas* W-285, *I. nil*, *I. setosa*) (Figure 1(a)), but yellow vein symptoms are observed on some hosts (*I. aquatica*, *I. cordatotriloba*) (Figure 1(b)). Although SPLCV does not cause symptoms on Beauregard, one of the most predominant sweet potato cultivar in the US, it can reduce the yield up to 26% [6]. Since SPLCV could become an important constraint for sweet potato production, diagnosis, identification, and characterization are essential to develop an appropriate management strategy. Methods available to detect SPLCV are based on symptomatology of indicator hosts (grafting assays) and on the detection of viral DNA (polymerase chain reaction (PCR), and molecular hybridization) [16] [17].

Despite the availability of sensitive nucleic acid-based assays for detecting plant viruses, serological assays are still the methods of choice for screening large numbers of plant materials. Furthermore, these methods are particularly valuable in developing countries since they require few resources. Serological detection of SPLCV is not currently available due to the difficulties in obtaining purified virions that can be used as antigen for anti-serum production. It is possible that either SPLCV virus particles occur at low concentration in plant tissues or

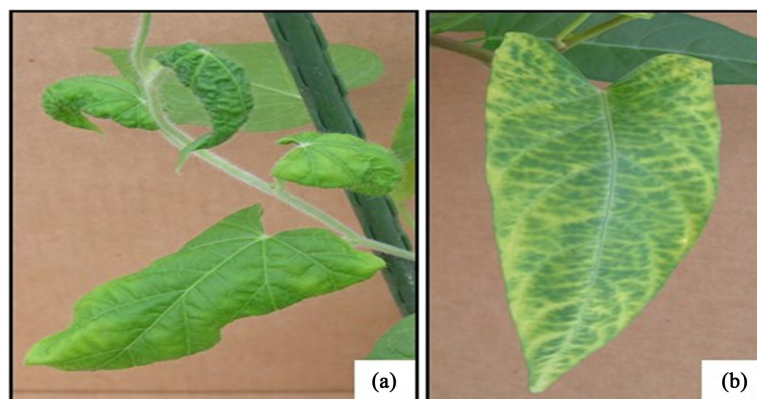


Figure 1. Leaf curl and yellow vein symptoms caused by *Sweet potato leaf curl virus* on *Ipomoea nil* (a) and *I. aquatica* (b), respectively.

viral particles are not stable following standard purification procedures. This problem is not unique to SPLCV. Purification of begomovirus virions can be very difficult; the success of purification methods depends greatly on the virus and host plant [18]. Using partially purified virions, serological relationships of SPLCV to *Bean golden mosaic virus* (BGMV) and to *Mungbean yellow mosaic virus* (MYMV) by Western blot was reported by Onuki *et al.* [19]. Serological relationship between SPLCV and *Tomato yellow leaf curl virus* (TYLCV) was not observed using antisera to TYLCV and partially purified preparations of SPLCV in ELISA on nitrocellulose membranes and Western blot [20].

Cloning and expressing viral coat protein genes in bacteria can overcome the difficulties in obtaining purified plant virus preparations for antiserum production [21]–[23]. Polyclonal antisera prepared to four begomovirus coat proteins (BGMV, TYLCV, *Cabbage leaf curl virus*, and *Tomato mottle virus*) expressed in *Escherichia coli* have been prepared and used for begomovirus detection in different immunological assays (ELISA, Western blots, leaf imprint blots, tissue blots, and immunogold labeling in electron microscopy) [24].

The availability of the complete genome sequences of most begomoviruses, including SPLCV [8] should allow the expression of a recombinant coat protein in bacteria. Thus, we attempt to express and purify the coat protein of SPLCV in *E. coli* in order to produce the SPLCV antisera for serological assays.

2. Materials and Methods

The overall flow of experiments from gene cloning to protein expression is illustrated in **Figure 2**.

2.1. Cloning of the SPLCV Coat Protein Gene

Total DNA was extracted from *I. setosa* leaf tissue infected with the US isolate of SPLCV using the DNeasy plant mini kit (QIAGEN, Valencia, CA). Based on the known sequence of SPLCV, specific primers SPB-1 (5'-CAGAGTCGGTACCTATGACAGGGCGAA-3') and SPB-2 (5'-TACTCTGCAGTTAATTGTTGTGCGAATC-3') were used to amplify the full-length coat protein (CP) gene (a *Kpn* I restriction site on SPB-1 and a *Pst* I restriction site on SPB-2 are underlined). SPB-2 primer sequence was kindly provided by Dr. P. Lotrakul (Chulalongkorn University, Bangkok, Thailand), while SPB-1 primer sequence was modified from a primer sequence (5'-GGTCAAGCTTTATGACAGGGCGAATTCC-3') also provided by Dr. P. Lotrakul. PCR reaction mixtures were prepared as described by Lotrakul *et al.* [7].

Amplification was performed in a Genius Thermocycler (Techne, Cambridge, UK) with 38 cycles of 94°C for 1 min, 58°C for 1 min, and 72°C for 3 min, followed by a final extension cycle of 72°C for 10 min. PCR products were separated by electrophoresis on 1.2% (w/v) agarose gel and stained with ethidium bromide. Bands corresponding to the expected PCR products were excised from agarose gels and further purified using MiniElute (Qiagen) DNA purification kit and ligated into the pGEM-T Easy vector (Promega, Madison, WI). Recombinant plasmids were transformed into *E. coli* JM 109 competent cells following manufacturer's instructions (Promega). Plasmids were isolated from transformed colonies using the FlexiPrep kit Miniprep procedure (Amersham Pharmacia Biotech Inc., Piscataway, NJ). The successful insertion of PCR products were confirmed

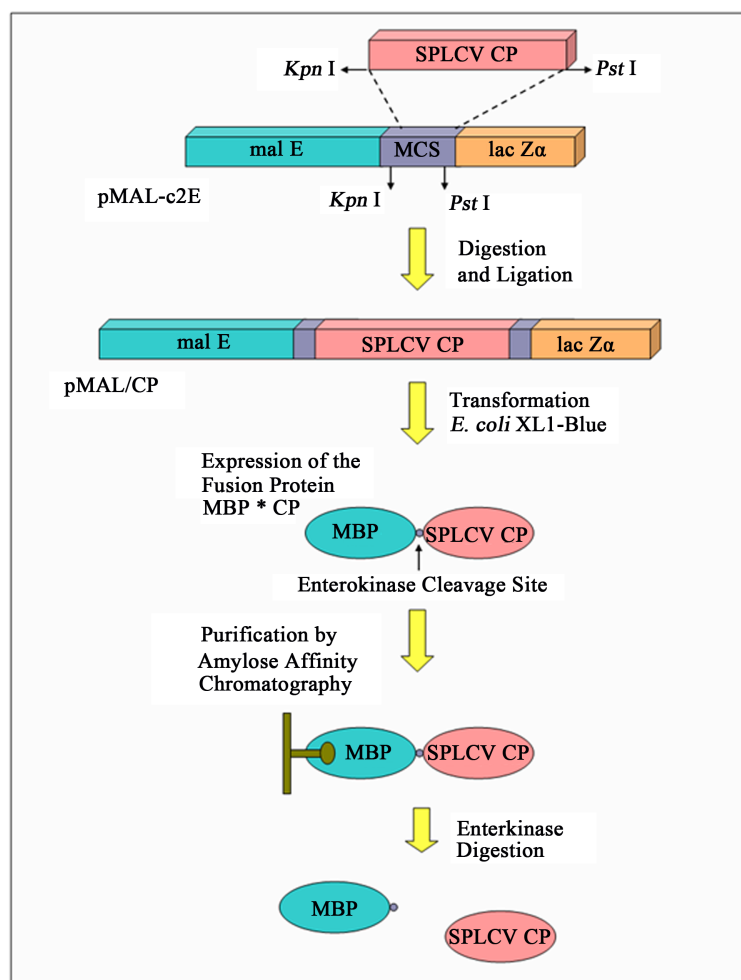


Figure 2. Diagram showing the expression in *E. coli* and purification of the fusion protein MBP * CP [Maltose-binding protein (MBP) and coat protein (CP) of *Sweet potato leaf curl virus* (SPLCV)]. MCS: Multiple cloning site in pMAL-c2E expression vector. mal E: gene encoding the MBP. lacZ α : gene encoding the β -galactosidase α fragment.

by enzyme digestions (*Eco* RI, *Kpn* I, and *Pst* I). In order to verify the identity of PCR products as the corresponding SPLCV coat protein gene, recombinant plasmids were sequenced at the Research Technology Support Facility of Michigan State University, East Lansing using a Perkin Elmer/Applied Biosystems 3100 capillary electrophoresis sequencer (Perkin Elmer, Foster City, CA).

The SPLCV coat protein gene was excised from the pGEM-T Easy vector by restriction enzymes (*Kpn* I and *Pst* I) and subcloned into the expression vector pMAL-c2E (New England Biolabs, Beverly, MA). The ligation mixture in 10 μ l volume consisted of 1.7 μ l containing 264 ng of pMAL-c2E vector (*Kpn* I and *Pst* I digested), 0.5 μ l containing 90 ng of the coat protein gene (*Kpn* I and *Pst* I digested), 5 μ l of 2 \times Rapid ligation buffer (Promega), 1 μ l of 10 mM ATP, 1 μ l of T4 DNA ligase (Promega) and 0.8 μ l of nuclease-free water. The ligation solutions were mixed and incubated for 1 hr at room temperature (24°C to 25°C), and then overnight at 4°C. Recombinant plasmids were transformed to *E. coli* XL1-Blue competent cells according to the procedures described by Dyer [25].

Plasmids were recovered from transformed colonies using the Wizard *Plus* SV Minipreps DNA Purification System (Promega). Recombinant plasmids pMAL-CP were identified by PCR using primers: malE and M13/pUC (New England Biolabs, #S1237S and #S1224S, respectively). PCR reactions were conducted as described previously using the following parameters: 1 cycle of 94°C for 1.5 min, 35 cycles of 94°C for 1 min, 65°C for 1 min, and 72°C for 1.5 min, and 1 cycle of 72°C for 5 min. Recombinant plasmids were characterized

by restriction enzyme digestions (*Mfe* I, *Nco* I, and *Nhe* I). The coding sequence of the SPLCV coat protein gene was ligated in-frame to that of the maltose-binding protein gene in the pMAL-c2E vector as confirmed by DNA sequencing of pMAL-CP as described previously.

2.2. Overexpression Experiments

Bacterial colonies carrying the recombinant plasmid pMAL-CP were grown at 37°C in LB broth containing 0.2% (w/v) glucose and 100 µg/ml ampicillin. When cultures reached an OD600 value of 0.5, cells were induced by the addition of 0.3 mM IPTG (Isopropyl-β-D-1-thiogalactopyranoside) and followed by incubation at either 26 or 37°C for the duration of 15 or 30 min, 1 hr, 2 hr, 4 hr, or 6 hr. For analysis of the total cell protein extracts, 1 ml of both induced and non-induced cells were harvested and re-suspended in 25 µl of ddH₂O, then 25 µl of 2× loading buffer (0.125 M TrisHCl, pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, 10% (v/v) 2-mercaptoethanol, 0.2% (w/v) bromophenol blue) was added, mixed and kept at -20°C. For samples induced more than 2 hr, 1 ml of bacterial cultures were harvested and re-suspended in 50 µl of ddH₂O, then 50 µl of 2 × loading buffer was added, mixed and kept at -20°C. When all samples were collected, proteins were denatured by keeping the samples in boiling water for 5 min. After microfuging for one min, 20 µl of supernatant of each sample was loaded on SDS-PAGEs (stacking gel 4% (w/v) and resolving gel either 8 or 10% (w/v)) as described by Laemmli [26]. Proteins were visualized by staining in a Coomassie Brilliant Blue solution (45% (v/v) methanol, 0.1% (w/v) Coomassie Brilliant Blue R-250, 10% (v/v) acetic acid). Selected bacterial colonies harboring the recombinant plasmid pMAL-CP were chosen for a time course expression experiment. Bacterial colonies carrying the pMAL-c2E plasmid (expression vector without insert) were used as controls in all expression experiments.

2.3. Expression and Purification of the Fusion Protein

Expression and purification procedures were performed according to Dyer [25] with some modifications. One liter of LB broth containing 0.2% (w/v) glucose and 100 µg/ml ampicillin was inoculated with 10 ml of an overnight culture of XL1-Blue cells containing the recombinant plasmid pMAL-CP, and incubated at 37°C. After the OD600 reached a value of 0.5, IPTG was added to a final concentration of 0.3 mM and incubation was continued at 26°C for an additional one hour in a gyrotory water bath shaker at 200 rpm. Cells were harvested by centrifugation (4800 rpm, 20 min, 4°C, in a Beckman J2-21M Centrifuge, JA 10 rotor). The pellet was kept overnight at -20°C and then re-suspended in 50 ml of lysis buffer [10 mM phosphate, 30 mM NaCl, 0.25% (v/v) Tween 20, 10 mM EDTA, 10 mM EGTA (ethylene glycol tetraacetic acid), pH 7.0]. Freshly prepared lysozyme (1 mg/ml) was added, and the suspension was kept on ice for 20 min. Cells were disrupted by sonication (6 times, 15 s, short pulses, 50% output), and NaCl was added to a final concentration of 0.5 M. After centrifugation (9000 rpm, 30 min, 4°C, in a Beckman Coulter Avanti J-25 Centrifuge, JA 25.5 rotor), the supernatant was diluted 1:5 with column buffer (10 mM phosphate, 0.5 M NaCl, 1 mM sodium azide, 1 mM EGTA, pH 7.0) plus 0.25% (v/v) Tween 20 and kept at 4°C.

The fusion protein MBP * CP comprising the SPLCV CP and the Maltose-binding protein (MBP) was purified using an amylose-resin affinity column chromatography. The column was prepared by adding amylose resin (New England Biolabs) to a 30 ml-syringe with a final bed volume of 4 ml. The column was washed with one volume of column buffer and three volumes of column buffer plus 0.25% (v/v) Tween 20. The diluted extract was added to the column at a flow rate of 1 ml/min. Then, the column was washed with three volumes of column buffer plus 0.25% (v/v) Tween 20 and five volumes of column buffer. The fusion protein MBP * CP was eluted with 10 mM maltose in the column buffer, and one ml each of column eluate was collected sequentially in 14 successive fractions. Column fractions were monitored for the presence of protein by measuring UV absorption at 280 nm. Six fractions containing the fusion protein were pooled and concentrated by filtration using Amicon® Ultra-4 30,000 NMWL and Microcon YM-30 filter units (Millipore, Billerica, MA), and reconstituted to about 1 mg/ml in 1× PBS (per liter 8 g NaCl, 1.15 g Na₂HPO₄, 0.2 g KCl, 0.2 g KH₂PO₄, pH 7.4).

2.4. Western Blot

Proteins electrophoresed on SDS-PAGE were transferred to nitrocellulose membranes. Polyacrylamide gels were soaked in blotting buffer (0.025 M TrisHCl, 0.192 M glycine, 20% (v/v) methanol) before assembling the transfer cassette. Electroblothing was conducted at 4°C, 35 V overnight in a TE 22 Transphor electrophoresis

unit (Hofer Scientific Instruments, San Francisco, CA). Hereafter, all incubations and washings were performed at room temperature in a shaker with gentle agitation. Blotted membranes were blocked for 1.5 hr with TBS (0.02 M TrisHCl, 0.5 M NaCl, pH 7.5) containing 3% (w/v) nonfat dry milk and 1% (v/v) Triton X-100. Blocking solution was discarded, and membranes were rinsed with TBS. Membranes were probed using either Anti MBP polyclonal antibody (1:10,000 dilution) (New England Biolabs) overnight, or with TYLCV polyclonal antiserum #1214 (1:1000 dilution) (provided by Dr. J. Polston, University of Florida, Gainesville) for 5 hr. The TYLCV antiserum #1214 was produced against the TYLCV coat protein expressed in *E. coli* [24]. Membranes were washed four times, 5 min each with TBS containing 0.05% (v/v) Tween 20, and then incubated for 1 hr with alkaline phosphatase-conjugated goat anti-rabbit IgG (1:10,000 dilution) (Bio-Rad, Hercules, CA). Antisera were diluted in TBS containing 2% (w/v) nonfat dry milk. Protein bands were visualized by the addition of color development solution (BCIP/NBT, 5-bromo-4-chloro-3-indolyl phosphate/nitroblue tetrazolium) (Bio-Rad).

2.5. Enterokinase Digestion

The fusion protein MBP * CP was digested with 0.013% (w/v) enterokinase (New England Biolabs) in a 20.5 μ l volume as follows: 10 μ l containing 7.5 μ g of fusion protein (in 1x PBS), 10 μ l of 2 \times digestion buffer (40 mM Tris-HCl pH 8.0, 100 mM NaCl, 4 mM CaCl₂), and 0.5 μ l containing 1 ng of enterokinase. The solution was incubated at room temperature for 3 hr. Digested proteins were separated by 10% (w/v) SDS-PAGE.

2.6. Mass Spectrometry

A 30 kDa polypeptide was obtained after enterokinase digestion and separation by 10% (w/v) polyacrylamide gel after SDS-PAGE. To facilitate identification and confirmation of 30 kDa peptide as the SPLCV coat protein, trypsin digestion and mass spectrometry analyses were performed at the Proteomics Core Facility at Pennington Biomedical Research Center, Louisiana State University, Baton Rouge. According to the procedures of DeLany *et al.* [27] and Zvonic *et al.* [28], excised gel plug was deposited into a 96-well plate and transferred to the MassPrep (Waters/Micromass) station. Proteins within the gel plug were automatically destained, reduced, alkylated, dehydrated, rehydrated, and digested with trypsin. The resulting peptides were extracted and separated by capillary liquid chromatography coupled to an Electrospray Ionization Mass Spectrometer (ESI MS/MS) Micromass Quadrupole-Time of Flight (Q-TOF) Mass Spectrometer (Waters). MassLynx 4.0 software package (Waters) was used to identify individual mass spectrograms. Database searches for protein identification were done in NCBI database using the ProteinLynx Globalserver 1.1 software (Waters). The number of peptides analyzed and the percentage of coverage of the total amino acid sequence were determined for the identification of the SPLCV coat protein.

3. Results

A DNA fragment of 789 base pairs (bp) containing the SPLCV coat protein gene was amplified with forward primer SPB-1 and reverse primer SPB-2, cloned into pGEM-T Easy vector and ligated into pMAL-c2E expression vector (Figure 2). After transformation into *E. coli* XL1-Blue, the recombinant plasmid pMAL-CP (SPLCV coat protein gene cloned into the pMAL-c2E) was identified by PCR. A 937 bp fragment corresponding to the SPLCV coat protein gene including flanking regions of the recombinant plasmid was amplified with primers malE and M13/pUC, while a shorter 196 bp fragment was amplified from the pMAL-c2E (plasmid without insert) using the same primers. Recombinant plasmids (pMAL-CP) were also characterized by restriction enzyme digestions and size determination of digestion products. *Mfe* I, *Nco* I, and *Nhe* I restriction sites are present in the SPLCV coat protein gene but not in the pMAL-c2E. DNA sequencing confirmed that the coding sequence of SPLCV coat protein gene was in frame with that of Maltose-Binding Protein gene in the recombinant plasmid pMAL-CP.

Preliminary expression tests to optimize the expression conditions for the fusion protein MBP * CP were conducted using two different IPTG concentrations (0.3 and 0.6 mM), two different temperatures (26°C and 37°C) and two bacterial host strains (XL1-Blue and TB1). When total protein extracts from IPTG-induced and non-induced cells were compared after SDS-PAGE analysis, expected protein bands of 72 kDa (42.5 kDa for MBP plus 29.4 kDa for the SPLCV CP) were not detectable in induced samples (Figure 3). A time course of protein expression was studied after the addition of 0.3 mM IPTG and subsequent incubation at 26°C, for 15 or 30 min, 1 hr, 2 hr, 4 hr or 6 hr (Figure 3 lanes 2 to 7) resulting in the similar negative results. The predicted

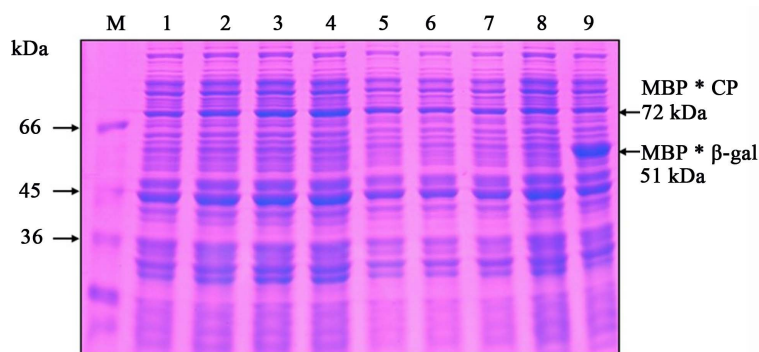


Figure 3. Time course study of the expression of the fusion protein MBP * CP [Maltose-binding protein (MBP) and coat protein (CP) of *Sweet potato leaf curl virus*] in an *E. coli* strain XL1-Blue. Cells were incubated in the presence of 0.3 mM IPTG at 26 °C. Induced cells harboring the recombinant plasmid pMAL-CP expressed the fusion protein MBP * CP (72 kDa), while induced cells harboring the pMAL-c2E vector expressed the MBP fused with the β -galactosidase α fragment protein (MBP * β -gal, 51 kDa). Total proteins were separated by 10% (w/v) SDS-PAGE and stained with Coomassie Brilliant Blue. Lane 1, non-induced cells harboring pMAL-CP. Lanes 2-7, induced cells for the duration of 15 min, 30 min, 1 hr, 2 hr, 4 hr, and 6 hr, respectively) harboring pMAL-CP. Lane 8, non-induced cells harboring pMAL-c2E. Lane 9, induced cells (2 hr) harboring pMAL-c2E. Line M, molecular weight marker.

amino acid sequence of the fusion protein MBP * CP was analyzed by ProtParam tool [29] to calculate some of the physico-chemical properties that could give us a clue to improve the expression tests. To determine if the *E. coli* strain (XL1-Blue) was the factor involved in the lack of expression of the fusion protein due to the presence of some proteases that could degrade the fusion protein, recombinant plasmids pMAL-CP were transformed into *E. coli* TB1 strain, but the expression was not detectable.

Fusion protein yields of 5 to 6 mg/liter of culture have been reported even when an expected protein band was not detectable after SDS-PAGE analysis (New England Biolabs, 2004). We decided to scale up to the one liter culture followed by purification of the fusion protein MAL * CP by the amylose-resin column affinity chromatography. After induction of *E. coli* XL1-Blue cells harboring the recombinant plasmid pMAL-CP, a 72 kDa polypeptide was purified from the crude protein extract of *E. coli*. The 72 kDa polypeptide was eluted from the amylose-resin column with 10 mM maltose (six fractions of 1 ml each) monitored by UV absorption at 280 nm (Figure 4, fractions 5 to 10; Figure 5, lanes 6, 7 and 8). This polypeptide was identified as the expected fusion protein based on its size 72 kDa on SDA-PAGE analysis, and its reaction with the Anti-MBP polyclonal antiserum by Western blot analysis. The 72 kDa polypeptide reacted with the polyclonal antiserum #1214 of tomato yellow leaf curl virus (TYLC) on Western blot analysis. Although this antiserum was not specific (it reacted with other bacterial proteins), it did not react with the MBP. Therefore, it is possible that some antibodies reacted with some epitopes of the SPLCV coat protein. Comparable expression levels of MAL fusion proteins with the β -galactosidase α fragment protein (MBP * β -gal, 51 kDa) (Figure 5, lane 3), and with the SPLCV coat protein (MBP * CP, 72 kDa) (Figure 5, lane 4) were observed after induction by 0.3 mM IPTG for one hr, suggesting that the SPLCV coat protein appears to behave as a normal passenger protein in this expression vector.

Yield of the purified MBP * CP fusion protein was approximately 200 μ g/liter of bacterial culture. After the fusion protein was digested with enterokinase, a 30 kDa protein was separated from maltose-binding protein by 10% (w/v) SDS-PAGE. The purified 30 kDa protein was digested by trypsin, and trypsin digestion products were separated by capillary liquid chromatography. The 30 kDa protein is expected to produce 40 peptide fragments when digested with trypsin, on a basis of prediction of the PeptideCutter software. Trypsin-digested fractionated polypeptides were analyzed by the Electrospray Ionization Mass Spectrometer (ESI MS/MS) Micromass Quadrupole-Time of Flight (Q-TOF) Mass Spectrometer (Waters).

The 30 kDa protein from the MBP * CP was identified as the SPLCV coat protein based on the amino acid sequence determination by mass spectrometry of trypsin digestion products. The mass spectrometry analyses elucidated in part the overlapping peptide sequences of 26 trypsin digestion products. The peptide sequences of eight polypeptides were listed in Table 1, corresponding to the 95 amino acid positions out of 256 residues of SPLCV coat protein. The longest peptide sequence identified by mass spectrometry is 21 amino acid residues

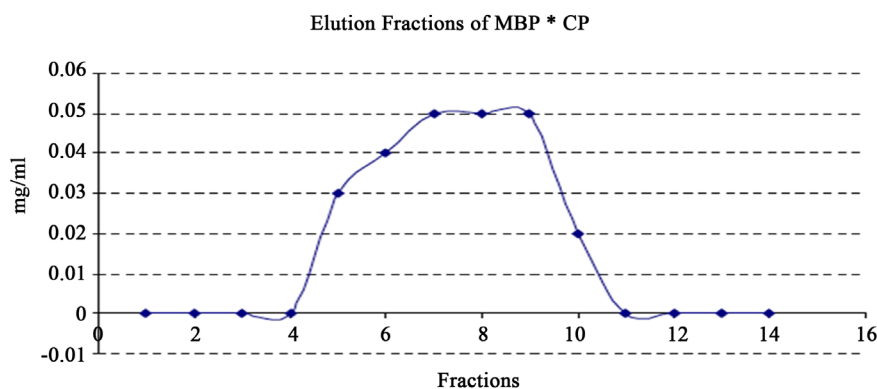


Figure 4. Elution profile of amylose-resin column affinity chromatography after elution with 10 mM maltose. Total proteins were extracted from one liter culture of *E. coli* XL1-Blue cells containing pMAL-CP, induced and grown for one hr, and applied to the amylose-resin column. The column was washed and eluted with 10 mM maltose, and one ml fraction was collected successively. Fractions were monitored by UV absorption at 280 nm for the detection of proteins.

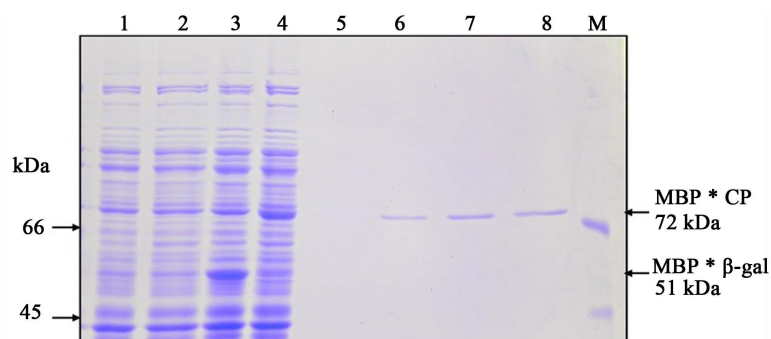


Figure 5. SDS-PAGE characterization of proteins expressed in *E. coli* and purified by amylose resin affinity column chromatography. *E. coli* cells were induced in the presence of 0.3 mM IPTG, at 26 °C, for one hr. Induced cells harboring the pMAL-c2E vector expressed the MBP fused with the β -galactosidase α fragment protein (MBP * β -gal) (lane 3), and induced cells with the recombinant plasmid pMAL-CP expressed the fusion protein MBP * CP (lane 4). Total protein products were separated by 8% (w/v) SDS-PAGE and stained with Coomassie Brilliant Blue. Lane 1, non-induced cells harboring pMAL-c2E. Lane 2, non-induced cells harboring pMAL-CP. Lane 3, induced cells harboring pMAL-c2E. Lane 4, induced cells harboring pMAL-CP. Lanes 5 - 8, MBP * CP fusion protein purified by amylose-resin affinity chromatography in fractions 6, 7, 8, and 9, respectively. Lane M, molecular weight marker.

Table 1. Peptide sequence of the coat protein of Sweet potato leaf curl virus identified by mass spectrometry. The coat protein gene was expressed as a fusion protein of maltose-binding protein in *E. coli*, the affinity-purified protein product was digested by trypsin, digestion products were fractionated by capillary liquid chromatography and peptide sequences were identified by mass spectrometry analysis.

Amino acid residue positions	Peptide Sequence	<i>m/z</i>	Peptide Molecular Weight
13 to 20	PYGGRPVR	451.223	900.43
22 to 42	RLNFETAIVPYTGNVPIAAR	758.347	2272.02
86 to 94	FVCVSDFTR	565.725	1129.44
117 to 125	VWMDDNVAK	539.226	1076.44
127 to 138	DHTNIITYWLIR	772.854	1543.69
179 to 191	FSVTVSGGPYSHK	683.304	1364.59
204 to 216	YNHVTYNHKEEAK	544.894	1631.66
246 to 254	AYFYDShNN	565.712	1129.41

corresponding to the amino acid positions 22 to 42 of the coat protein with the molecular weight of 2272.02 at m/z 758.347. The shorter peptide sequences in **Table 1** correspond to two peptide sequences of 13 amino acid residues (positions 179 to 191, and position 204 to 216), one peptide sequence of 12 amino acid residues (positions 127 to 138), three peptide sequences of 9 amino acid residues (positions 86 to 94, positions 117 to 125, positions 246 to 254), and one sequence of 8 amino acid residues (positions 13 to 20). All together the mass spectrometry analyses identified sequences corresponding to 37% of amino acid positions of the SPLCV coat protein.

4. Discussion

Many systems for expression of foreign genes in bacteria have been developed. The pMAL-c2E expression vector of the pMAL Protein Fusion and Purification System from New England Biolabs was selected to express the coat protein gene of SPLCV in *E. coli*. Previously, the system was used successfully to express the bean seed storage protein phaseolin with a final yield of purified phaseolin protein at 3 to 5 mg per liter of bacterial culture [25]. The target gene (SPLCV coat protein) was inserted downstream from the mal E gene of *E. coli*, which encoded a maltose-binding protein (MBP), resulting in a cytoplasmic expression of a fusion protein (MBP * CP). This method uses the strong “tac” promoter and the mal E translation initiation signals to express large amounts of fusion protein (New England Biolabs, 2004). The fusion protein was then isolated by affinity chromatography using an amylose-resin column, and the target passenger protein of 30 kDa was cleaved from the MBP by enterokinase, and separated from MBP by 10% (w/v) SDS-PAGE.

Subcloning SPLCV coat protein gene into the pMAL-c2E was not an easy task considering the size of the insert (766 bp) and the size of the vector (6626 bp). High quality DNA as well as an appropriate insert:vector ratio (1:3) were required for the efficient ligation. Overexpression tests with bacterial cells harboring the recombinant plasmid pMAL-CP did not produce a clear-cut result since a distinct protein band of 72 kDa was not observed after the induction of gene expression by IPTG. Attempts were made unsuccessfully to improve the yield of gene expression products for the fusion protein MBP * CP, using two different IPTG concentrations (0.3 vs 0.6 mM), two incubation temperatures (26°C vs 37°C), six incubation durations (15 min, 30 min, 1 hr, 2 hr, 4 hr and 6 hr), and two bacteria host strains (XL1-Blue vs TB1). We suspected that the low expression of fusion protein MBP * CP might have been masked by bacterial proteins of similar size (72 kDa) because of the limitation of resolution in the single-dimensional separation by SDS/PAGE.

We therefore decided to scale up the bacterial culture to one liter and to introduce the protein purification step by the addition of the amylose-resin column affinity chromatography. These new steps yielded a single band of the 72 kDa fusion protein (42.5 kDa for MBP plus 29.4 kDa for the SPLCV CP) by the SDS-PAGE even though the protein expression level was unexpectedly low. Yield of the MBP * CP fusion protein after purification by the amylose affinity chromatography was approximately 200 µg/liter of bacterial culture. According to the information provided by the manufacturer (New England Biolabs, 2004), the yield of fusion protein from the affinity purification can be up to 200 mg/liter culture, with typical yields in the range of 10 to 40 mg/liter. However, yields can vary greatly depending upon the sequences fused to malE. Nikolaeva *et al.* [21] used the pMAL system to produce polyclonal antisera to the coat protein of *Citrus tristeza virus* obtaining 100 mg/liter culture of purified fusion protein, whereas Meng *et al.* [23] using the same system to produce a polyclonal antiserum to the coat protein of *Rupestris stem pitting associated virus* obtained only 12 mg/liter culture of purified fusion protein.

The low expression level of the MBP * CP fusion protein in *E. coli* might have been predicted based on the low abundance of SPLCV particles in plants. Similar low expression level of Tomato yellow leaf curl virus (TYLCV) was also observed in tomato after the coat protein gene was expressed using the expression vector pETh-3b in *E. coli* for anti-serum production [24]. Several factors could have influenced the yield of MBP * CP fusion protein. The efficient expression of foreign genes in *E. coli* can be affected by the physico-chemical properties of the target protein [29], the stability of the mRNA, degradation of the target protein by *E. coli* proteases, the presence of rare codons [30], or the toxicity of the expressed target protein. The solubility of the fusion protein could also be involved in the low expression. It is common for recombinant proteins expressed in *E. coli* to be produced as aggregates (inclusion bodies). Growth at 37°C can cause some proteins to accumulate as inclusion bodies, while incubation at 30°C may lead to soluble proteins [31] [32]. The grand average of hydrophobicity (GRAVY) index [33] were estimated for both the MBP * β -gal (-0.483) and MBP * CP (-0.537), indicating that both proteins were soluble (positive GRAVY: hydrophobic, negative GRAVY: hydrophilic).

Moreover, the solubility of these proteins was confirmed by another sequence-based protein solubility evaluator (PROSO server, Smialowski *et al.* [34]). Consequently, the solubility of the fusion protein MBP * CP may not be related to the low expression levels of this protein in *E. coli*. It has been reported that proteins that contain highly charged domains may associate with other cellular components, for example, positively charged basic proteins may bind to negatively charged RNA or DNA. If this is the case, the protein may partition with cellular debris (Novagen, 2006). This could be a reason for the low expression of MBP * CP which is a basic protein (theoretical pI 8.88) that contains 77 negatively charged residues (Asp + Glu) and 86 positive charged residues (Arg + Lys). It is possible that the positively charged residues of the coat protein might bind to the negatively charged mRNA of nascent expressed MBP * CP, thus inhibiting the protein production. In contrast, the MBP * β -gal protein with a theoretical pI of 5.04 contains 63 negatively charged residues and 50 positive charged residues.

The SPLCV coat protein gene was expressed as a fusion protein in *E. coli*. Despite the low expression levels of MBP * CP, SPLCV coat protein can be obtained after enterokinase digestion and could be used as an antigen to produce antibodies to SPLCV. The availability of an antiserum to SPLCV will allow the development of a practical and inexpensive serological method for SPLCV detection. Moreover, an SPLCV specific antiserum may be used to study serological relationships of SPLCV with other begomoviruses.

Acknowledgements

The authors acknowledge the financial support in part from the Department of Plant Pathology and Crop Physiology, College of Agriculture, Louisiana State University, and from the Louisiana State University Agricultural Center. We also appreciate receiving a research assistantship support to DLGR from Professor Christopher A. Clark, and technical assistances on mass spectrometry from Professor Zhi-Yuan Chen in this Department.

References

- [1] Kyndt, T., Quispe, D., Zhai, H., Jarret, R., Ghislain, M., Liu, Q., Gheysen, G. and Kreuze, J.F. (2015) The Genome of Cultivated Sweet Potato Contains *Agrobacterium* T-DNA with Expressed Genes: An Example of a Naturally Transgenic Food Crop. *Proceedings of National Academy of Sciences of the USA*, **112**, 5844-5849. <http://dx.doi.org/10.1073/pnas.1419685112>
- [2] Kays, S.J. (2005) Sweetpotato Production Worldwide: Assessment, Trends and the Future. *Acta Horticulturae*, **670**, 19-25. <http://dx.doi.org/10.17660/actahortic.2005.670.2>
- [3] Bovell-Benjamin, A.C. (2007) Sweet Potato: A Review of Its Past, Present, and Future Role in Human Nutrition. *Advances in Food and Nutrition Research*, **52**, 1-59. [http://dx.doi.org/10.1016/S1043-4526\(06\)52001-7](http://dx.doi.org/10.1016/S1043-4526(06)52001-7)
- [4] Karyeija, R.F., Gibson, R.W. and Valkonen, J.P.T. (1998) The Significance of Sweet Potato Feathery Mottle Virus in Subsistence Sweet Potato Production in Africa. *Plant Disease*, **82**, 4-15. <http://dx.doi.org/10.1094/PDIS.1998.82.1.4>
- [5] Clark, C.A. and Moyer, J.W. (1988) Compendium of Sweet Potato Diseases. The American Phytopathological Society, St. Paul.
- [6] Clark, C.A. and Hoy, M.W. (2006) Effects of Common Viruses on Yield and Quality of Beauregard Sweetpotato in Louisiana. *Plant Disease*, **90**, 83-88. <http://dx.doi.org/10.1094/PD-90-0083>
- [7] Lotrakul, P., Valverde, R.A., Clark, C.A., Sim, J. and De La Torre, R. (1998) Detection of a Geminivirus Infecting Sweet Potato in the United States. *Plant Disease*, **82**, 1253-1257. <http://dx.doi.org/10.1094/PDIS.1998.82.11.1253>
- [8] Lotrakul, P. (2000) Biological and Molecular Properties of Sweet Potato Leaf Curl Virus. Ph.D. Dissertation, Louisiana State University, Baton Rouge.
- [9] Lotrakul, P., Valverde, R.A., Clark, C.A., Hurtt, S. and Hoy, M.W. (2002) Sweet Potato Leaf Curl Virus and Related Geminiviruses in Sweetpotato. *Acta Horticulturae*, **583**, 135-141.
- [10] Lotrakul, P., Valverde, R.A., Clark, C.A. and Fauquet, C.M. (2003) Properties of a Begomovirus Isolated from Sweetpotato Infected with Sweet Potato Leaf Curl Virus. *Revista Mexicana de Fitopatología*, **2**, 128-136.
- [11] Valverde, R.A., Sim, J. and Lotrakul, P. (2004) Whitefly Transmission of Sweet Potato Viruses. *Virus Research*, **100**, 123-128. <http://dx.doi.org/10.1016/j.virusres.2003.12.020>
- [12] Valverde, R.A., Clark, C.A. and Valkonen, J.P.T. (2007) Viruses and Virus Disease Complexes of Sweet Potato. *Plant Viruses*, **1**, 116-126.
- [13] Luan, Y.S., Zhang, J. and An, L.J. (2006) First Report of *Sweet potato leaf curl virus* in China. *Plant Disease*, **90**, 1111. <http://dx.doi.org/10.1094/PD-90-1111C>
- [14] Albuquerque, L.C., Inoue-Nagata, A.K., Pinheiro, B., Ribeiro, S.G., Resende, R.O., Moriones, E. and Navas-Castillo, J.

- (2011) A Novel Monopartite Begomovirus Infecting Sweet Potato in Brazil. *Archives of Virology*, **156**, 1291-1294. <http://dx.doi.org/10.1007/s00705-011-1016-x>
- [15] Esterhuizen, L.L., van Heerden, S.W., Rey, M.E. and van Heerden, H. (2012) Genetic Identification of Two Sweet-Potato-Infecting Begomoviruses in South Africa. *Archives of Virology*, **157**, 2241-2245. <http://dx.doi.org/10.1007/s00705-012-1398-4>
- [16] Li, R., Salih, S. and Hurtt, S. (2004) Detection of Geminiviruses in Sweetpotato by Polymerase Chain Reaction. *Plant Disease*, **88**, 1347-1351. <http://dx.doi.org/10.1094/PDIS.2004.88.12.1347>
- [17] Valverde, R.A., Kokkinos, C.D. and Clark, C.A. (2004) *Sweet potato leaf curl virus*: Detection by Molecular Hybridization. *Phytopathology*, **94**, S105.
- [18] Palmer, K.E., Schnippenkoetter, W.H. and Rybicki, E.P. (1998) Geminivirus Isolation and DNA Extraction. In: Foster, G.D. and Taylor, S.C., Eds., *Plant Virology Protocols*, Methods in Molecular Biology, Vol. 81, Humana Press Inc., Totawa, 41-52. <http://dx.doi.org/10.1385/0-89603-385-6:41>
- [19] Onuki, M., Honda, Y. and Hanada, K. (2000) Geminate Particle Morphology of Sweet Potato Leaf Curl Virus in Partially Purified Preparation and Its Serological Relationship to Two Begomoviruses by Western Blotting. *Journal of General Plant Pathology*, **66**, 182-184. <http://dx.doi.org/10.1007/PL00012942>
- [20] Gutierrez, D.L. (2008) Molecular Diversity and Coat Protein Expression of Sweet Potato Leaf Curl Virus. Ph.D. Dissertation, Louisiana State University, Baton Rouge.
- [21] Nikolaeva, O.V., Karasev, A.V., Gumpf, D.J., Lee, R.F. and Garnsey, S.M. (1995) Production of Polyclonal Antisera to the Coat Protein of Citrus Tristeza Virus Expressed in *Escherichia coli*: Application for Immunodiagnosis. *Phytopathology*, **85**, 691-694. <http://dx.doi.org/10.1094/Phyto-85-691>
- [22] Hoyer, U., Maiss, E., Jelkman, W., Lesemann, D.E. and Vetten, H.J. (1996) Identification of the Coat Protein Gene of a Sweet Potato Sunken Vein Closterovirus Isolate from Kenya and Evidence for a Serological Relationship among Geographically Diverse Closterovirus Isolates from Sweet Potato. *Phytopathology*, **86**, 744-750. <http://dx.doi.org/10.1094/Phyto-86-744>
- [23] Meng, B., Credi, R., Petrovic, N., Tomazic, I. and Gonsalves, D. (2003) Antiserum to Recombinant Virus Coat Protein Detects *Rupestris stem pitting associated virus* in Grapevines. *Plant Disease*, **87**, 515-522. <http://dx.doi.org/10.1094/PDIS.2003.87.5.515>
- [24] Abouzid, A.M., Freitas-Astua, J., Purcifull, D.E., Polston, J.E., Beckham, K.A., Crawford, W.E., Peter-senen, M.A., Peyser, B., Patte, C. and Hiebert, E. (2002) Serological Studies Using Polyclonal Antisera Prepared against the Viral Coat Protein of Four Begomoviruses Expressed in *Escherichia coli*. *Plant Disease*, **86**, 1109-1114. <http://dx.doi.org/10.1094/PDIS.2002.86.10.1109>
- [25] Dyer, J.M. (1993) Structural Analysis and Methionine Enhancement of the Bean Seed Storage Protein Phaseolin. PhD Dissertation, Louisiana State University, Baton Rouge.
- [26] Laemmli, U.K. (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, **227**, 680-685. <http://dx.doi.org/10.1038/227680a0>
- [27] DeLany, J.P., Floyd, Z.E., Zvonic, S., Smith, A., Gravois, A., Reiners, E., Wu, X., Kilroy, G., Lefevre, M. and Gimble, J.M. (2005) Proteomic Analysis of Primary Cultures of Human Adipose-Derived Stem Cells—Modulation by Adipogenesis. *Molecular and Cellular Proteomics*, **4**, 731-740. <http://dx.doi.org/10.1074/mcp.m400198-mcp200>
- [28] Zvonic, S., Lefevre, M., Kilroy, G., Floyd, Z.E., DeLany, J.P., Khetterpal, I., Gravois, A., Dow, R., White, A., Wu, X. and Gimble, J.M. (2007) Secretome of Primary Cultures of Human Adipose-Derived Stem Cells: Modulation of Serpins by Adipogenesis. *Molecular and Cellular Proteomics*, **6**, 18-28. <http://dx.doi.org/10.1074/mcp.M600217-MCP200>
- [29] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.A., Appel, R.D. and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. In: Walker, J.M., Ed., *The Proteomics Protocols Handbook*, Humana Press, Totowa, 571-607. <http://dx.doi.org/10.1385/1-59259-890-0:571>
- [30] Chen, D. and Texada, D.E. (2006) Low-Usage Codons and Rare Codons of *Escherichia coli*. *Gene Therapy and Molecular Biology*, **10**, 1-12.
- [31] Schein, C.H. (1989) Production of Soluble Recombinant Proteins in Bacteria. *Bio/Technology*, **7**, 1141-1149. <http://dx.doi.org/10.1038/nbt1189-1141>
- [32] Schein, C.H. and Noteborn, M.H.M. (1988) Formation of Soluble Recombinant Proteins in *Escherichia coli* Is Favored by Lower Growth Temperature. *Bio/Technology*, **6**, 291-294. <http://dx.doi.org/10.1038/nbt0388-291>
- [33] Kyte, J. and Doolittle, R.F. (1982) A Simple Method for Displaying the Hydropathic Character of a Protein. *Journal of Molecular Biology*, **157**, 105-132. [http://dx.doi.org/10.1016/0022-2836\(82\)90515-0](http://dx.doi.org/10.1016/0022-2836(82)90515-0)
- [34] Smialowski, P., Martin-Galiano, A.J., Mikolajka, A., Girschick, T., Holak, T.A. and Frishman, D. (2007) Protein Solubility: Sequence Based Prediction and Experimental Verification. *Bioinformatics*, **23**, 2536-2542. <http://dx.doi.org/10.1093/bioinformatics/btl623>

Abbreviations

BCIP/NBT, 5-bromo-4-chloro-3-indolyl phosphate/nitroblue tetrazolium;

bp, base pairs;

CP, coat protein;

EGTA, ethylene glycol tetraacetic acid;

ELISA, enzyme-linked immunosorbent assay;

ESI/MS, electrospray ionization/mass spectrometry;

GMO, genetically modified organism;

IPTG, isopropyl- β -D-1-thiogalactopyranoside;

PCR, polymerase chain reaction;

MBP, maltose-binding protein;

Ri-plasmid, root-inducing plasmid;

SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis;

SPLCV, sweet potato leaf curl virus;

T-DNA, transferred DNA;

TYLCV, tomato yellow leaf curl virus.