

# A cDNA Sequence Coding for a Glutamic Acid-Rich Protein Is Differentially Expressed in Cassava Storage Roots

Cláudia Regina Batista de Souza<sup>1,2</sup>, Luiz Joaquim Castelo Branco Carvalho<sup>1</sup>,  
Elionor Rita Pereira de Almeida<sup>1</sup> and Eugen Silvano Gander<sup>1,\*</sup>

<sup>1</sup>EMBRAPA – Recursos Genéticos e Biotecnologia, Caixa Postal 02372, Brasília-DF, 70770-900, Brazil; <sup>2</sup>Current Address: Universidade Federal do Pará (UFPA), Rua Augusto Correa n. 01, Departamento de Genética, Guamá, Belém-PA, 66075-110, Brazil

**Abstract:** We report the isolation and characterization of a cDNA sequence (Mec1) coding for a glutamic acid-rich protein (Pt2L4) from cassava storage roots. Comparative sequence analysis showed a high identity of Pt2L4 with cassava protein C54, which is expressed in vascular tissues of storage roots. Northern blot analysis showed that the Mec1 transcription expression pattern might be related to the maturation of the storage parenchyma cells.

**Keywords:** Cassava, gene expression, Pt2L4 protein, secondary growth, storage root formation.

## INTRODUCTION

Cassava (*Manihot esculenta* Crantz) belongs to the Euphorbiaceae and ranks fourth, after rice, sugar cane and maize [1], as a source of calories and is the major staple food for more than 600 million people in Africa, Latin America and Asia [2].

The cassava storage root (CSR) contains about 85% carbohydrates, 1% proteins (dry weight) and a few other nutrients [2]. The CSR results from the swelling of adventitious roots through secondary growth. Anatomical studies of the CSR differentiate three major tissue systems [3, 4] and a model for tissue organization, suitable for gene expression analyses, has been proposed [5, 6]. According to this model; tissue system I is composed of phellogen and phellogen, tissue system II of phloem and vascular cambium, and tissue system III of secondary xylem with its highly specialized parenchyma cells packed with starch granules [5, 6]. The few approaches to study and/or alter protein content and polymorphism in CSR have been undertaken either by conventional breeding [7, 8] or by biochemical and biotechnological methods [9, 10, 11]. So far, only few investigations bear on the isolation of proteins and their respective genes in cassava storage roots [6, 12-15] and in our laboratory two major, alcohol soluble proteins, designated Pt2L4 and Pt3L4, which are differentially expressed in CSR compartments, were isolated and partially characterized [6, 13].

In the present communication, we report the isolation and characterization of a cDNA sequence (Mec1), coding for the Pt2L4 protein. The deduced amino acid sequence shows that Pt2L4 is a glutamic acid rich protein with a molecular weight of 18.8 kDa and a predicted isoelectric point of 3.75. The high identity of Pt2L4 with cassava protein C54 (14), and the

Mec1 expression pattern that is related to the maturation of the storage parenchyma cells, indicate a function related to storage root formation.

## MATERIALS AND METHODS

### Plant Material

Plants of the commercial cultivar IAC 12-829 were propagated by stem cuttings and planted in field plots of EMBRAPA – Genetic Resources and Biotechnology and grown for 10 months. Storage roots were harvested and roots of uniform 5 cm diameter were sampled to dissect tissue layers according to our compartmentalization model (Fig. 1). Dissected tissue layers were processed on ice and immediately placed in liquid nitrogen. Young stem, stem peel, fruits with seeds, leaves, and cotyledons were treated equally and were used for total RNA extraction. If not immediately processed samples for total RNA and protein extraction were stored at -80°C until use. Young leaves were harvested in the field, immediately immersed in liquid nitrogen and used for genomic DNA extraction.

### Isolation of cDNA and Nucleotide Sequence Analysis

Total RNA from manually dissected storage root layers L1 to L5, and from cotyledons, young stems, petioles and leaves of cassava plants were extracted as described previously [13].

Based on the partial sequence information of the Pt2L4 described before [13], specific primers were designed and used in a RACE PCR amplification system (GIBCO BRL). The primers were: Pr1: 5'acagtagagaaaactgaagaa3' (forward), Pr2: 5'cacaacaacggctctctgctgctgg3' (reverse) Pr3: 5'atggctactgctgagtagta3' (forward) and Pr4: 5'ctcagtctctcagcttcaac3' (reverse). PCR products were cloned into the pGEM-T Easy Vector System (Promega) and sequenced in an ABI 377 automated sequencer (Perkin-Elmer, Foster City, CA). The cDNA sequence was aligned to the Gene Bank nucleotide sequence databank using the BLAST algorithm [16].

\*Address correspondence to this author at the EMBRAPA – Recursos Genéticos e Biotecnologia, Caixa Postal 02372, Brasília-DF, 70770-900, Brazil; E-mail: egander@cenargen.embrapa.br

## RNA Gel-Blot Analysis

Total RNA samples from different tissue of roots and from cotyledons, young stems, petioles and leaves of cassava plants were isolated as described above. Northern analysis with the *Mec1* cDNA as probe was performed as described before [13].

## Genomic DNA Gel-Blot Analysis

Genomic DNA from young cassava leaves was extracted as described by Deblaere *et al.* [17]. Samples containing 10 µg of DNA were restriction digested with *Bgl II*, *Eco RI*, *Hind III* and *Nsi I* and separated by electrophoresis on 0.7% agarose gel. The DNA fragments were blotted on a Hybond N<sup>+</sup> membrane and probed with [<sup>32</sup>P] labeled *Mec1* cDNA.

## Microscopic Analysis

Hand sections of fresh storage root were treated with O-toluidine blue to identify tissue systems and layers as previously described [5]. 200 µm cross sections from storage roots with advanced secondary growth were observed under the light microscope (magnification: x100) and photographed on an automatic record system on 100ASA color films. Pictures were converted to digital images and processed for area quantification using the dt-Scan software.

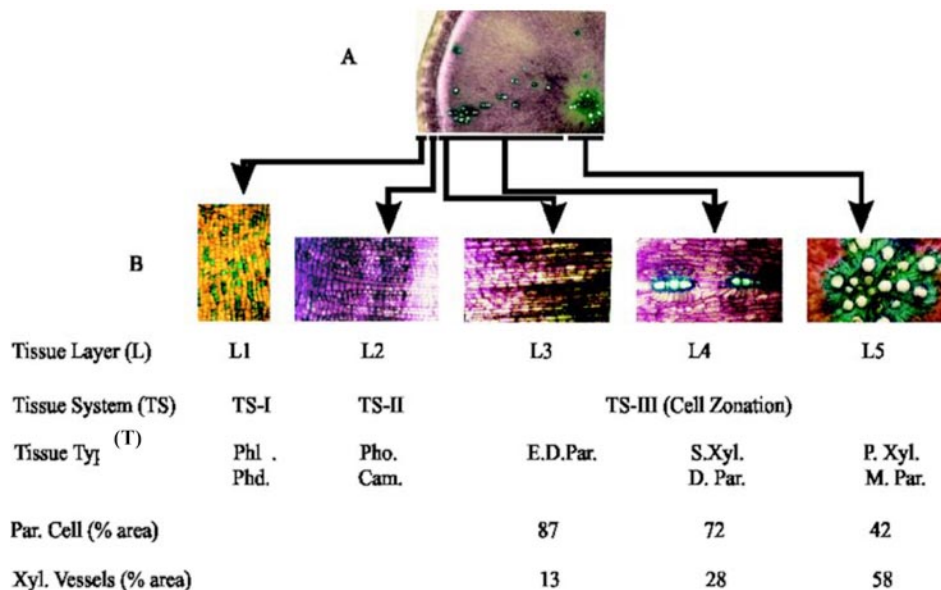
## RESULTS

### The Tissue Organization Model of Cassava Storage Roots

Fig. 1 shows an O-Toluidine Blue stained [5] cross section through a 5 cm diameter cassava storage root (Fig. 1A).

This type of staining permits the histological distinction of the five tissue system layers L1 to L5 (Fig. 1B), organized into the three tissue system TS-I, TS-II and TSIII. Tissue System I comprises the white outer layer composed of phellogen and phelloderm, next follows the dark-red colored Tissue System II composed of secondary phloem and vascular cambium and finally, the second largest white and dark-red central cylinder of Tissue System III of L3 to L5, composed of the secondary xylem with its array of vessels and parenchyma cells (Fig. 1B).

The increase in diameter of the storage root is the result of the secondary growth pattern, occurring mainly in the outermost part of the root cylinder. Consequently the inner cell layers in the central cylinder are older than those in the outer part and results in distinct cell zones with respect to parenchyma maturation and specialization, i.e., to permanent starch storage cells in the central cylinder. This zoning pattern is shown in Fig. 1B and it can be seen that the proportion of early developing parenchyma cells corresponds to 87% of total cells in tissue layer 3, decreases to 72% in layer 4 and reaches a minimum of 42% in tissue layers 5. This pattern of aging of parenchyma cells has been related to the accumulation of amylose, a starch granule constituent that accumulates in later stages of starch granule formation in cassava storage roots [18] and a similar pattern of distinct tissue composition and parenchyma cell differentiation has also been reported by others [3, 4, 5, 19] and has been used to study sugar and starch accumulation patterns [18, 20]. The quantitative analysis of tissue system III also clearly demonstrates an increase of the area occupied by vessels from 13% in L3 to 58% in L5 (Fig. 1B).



**Figure 1.** Identification of tissue systems and layers in a cassava storage root.

(A) Cross section with O-Toluidine Blue staining.

(B). (L): Tissue layers (TS): Tissue systems and (T): Tissue Type). Note that layers L3,L4 and L5 belong to the same tissue system.

Phl: phellogen; Phd: phelloderm; Pho: phloem; Cam: cambium; Y. Par: young developing parenchyma; D. Par: intermediately developed parenchyma; M. Par: mature parenchyma; P. Xyl: primary xylem; S. Xyl: secondary xylem; Par. Cell (% area): % parenchyma cells/ tissue layer; Xyl. Vessels (% area): % xylem vessels / tissue layer.

### Isolation and Characterization of Pt2L4 Coding Sequences

Two PCR amplification products of 649 and 426 bp obtained with 3' and 5' RACE PCR reactions were cloned and sequenced. A fragment of 649 bp, obtained with the primer pair Pr1/primer-T in the 3' direction, and a fragment of 426 bp obtained with the primer pair Pr2/poly C in the 5' direction, matched the sequence of the 220 bp Pt2L4 fragment previously identified [13]. With a 105 bp overlap in the two clones it was possible to isolate a full-length cDNA (Fig. 2) coined *Mec1*. This cDNA sequence of about 970 bp's consists of a 531 bp open reading frame (ORF), a 66 bp 5' untranslated region and a 368 bp 3'-noncoding region containing a polyadenylation signal (Fig. 2). A cDNA clone corresponding to a 531 bp ORF was obtained by PCR amplification using primers Pr3 and Pr4. The deduced amino acid sequence consists of 177 residues, accounting for a molecular weight of 18.8 kDa and a predicted isoelectric point of 3.75. A BLAST search of Pt2L4 showed the highest similarity (62%) with cassava protein C54 (AAP57707) as described previously by Zhang *et al.* [14] (Fig. 3). The deduced amino acid composition of the Pt2L4 protein shows that the

most abundant amino acids are glutamic acid (31.6%), alanine (16.94%), valine (13.55%) and proline (11.29%). Protein P2L4 contains four repeats of an ETPKEE motif (Fig. 3). In one of these repeats the lysine residue is substituted by alanine, and in another the threonine is substituted by valine. Only one of these repeats (italic in Fig. 3) is conserved in the C54 protein and so far no function has been related to these elements.

### Gene Expression Analysis and Gene Copy Number Determination

Total RNA blots from different root layers, leaves, stem peels, cotyledons and petioles were hybridized with a 531 bp *Mec1* probe. In leaves and stem peels no signals were detected. In the other RNA samples, however, a transcript of about 1.0 kb was observed with varying intensities. In cotyledons the signal was weak, while in young stems and petioles the signal was strong and a pattern, suggesting developmentally related expression, was observed in the different storage root layers (Fig. 4).

To determine the *Mec1* copy number in the cassava genome, genomic DNA was digested with restriction enzymes

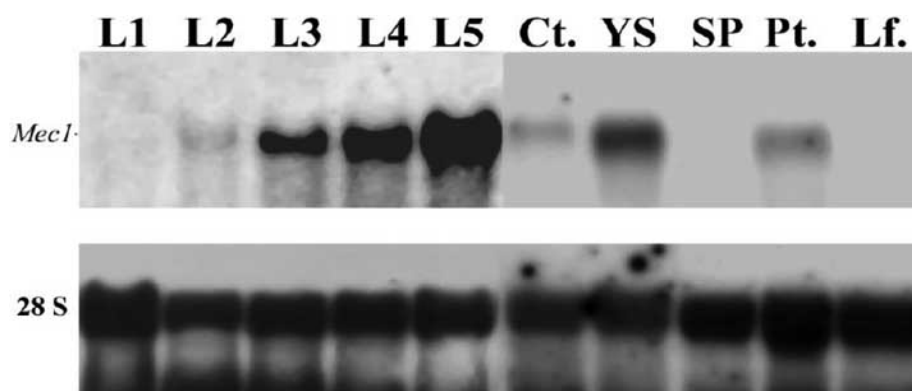
ctttctgtctctccttctggttctcagctttatTTTTtag	45
tttctatttcttggctgactatggctactgctgaggtagtaaca	90
M A T A E V V T	
gcacagactgcacttctgaggaaaaaccagctgaagaagtgaag	135
A Q T A L P E E K P A E E V K	
gtttcagagattgtaacagaagaggcagcaccagcagtagagcca	180
V S E I V T E E A A P A V E P	
gttgctgaagagcccaaggaagcagagccagttgagtagtctgaa	225
V A E E P K E A E P V A V S E	
gagccaaaggaggctgatgatgctccggctgaagtagcgggtgaa	270
E P K E A D D A P A E V A V E	
actaaagaggtttagaagtgaagaggccaagactgtgacagaa	315
T K E V V E V E E A K T V T E	
gagccaacagtagagaaaaactgaagaagaagaagagactcctaag	360
E P T V E K T E E E E E T P K	
gaagaaacaccagagcctgtggttgtaaggagactcctaaagag	405
E E T P E P V V V K E T P K E	
gaaccagcagcagagaccgttgttggaggctcccaagagaca	450
E P A A E T V V V E A P K E T	
accgaggctgcaactgaggcagaagcaccggcaccggaatccgca	495
T E A A T E A E A P A P E S A	
ccagcatcagcatcgaaaactccagctgaagaagaagttccaaag	540
P A S A S E T P A E E E V P K	
gaggaagaaggtgatgagaagaatctgaagcagaagttgaagct	585
E E E G D E K K S E A E V E A	
gagaagactgag <b>ta</b> atgagatagctcaggggttaattgggtttt	630
E K T E *	
gcatgccgtgctgtaattttcgtattgttaggttggtgcttaata	675
agagttttattactttctanggttctgggtctgggttaacggggaa	720
tgtcccacctgaattgaaccaatgggaactcaccacatggcagac	765
attgcatttgcaacagatggcgatgttttgggtactcatttata	810
ctacgtggcaactagcattgtgttcctattgggtccaagcagtt	855
ccatgaattttatatacatgtgcttgggtgatgct <b>ggtaataa</b>	900
<b>tt</b> acttgtagcgttctgtagtacttgggttaataataattcctag	945
cactaagttattttaaaaaaaaa	965

**Figure 2.** *Mec1* nucleotide sequence and deduced amino acid sequence for Pt2L4 protein. Start codon, stop codon and potential poly (A<sup>+</sup>) site are bold printed.

\*The nucleotide sequence data reported here appear in the NCBI GenBank under the accession number AY101376 (c1).

Pt2L4	1	MATAEVVTAQTALPEEKPAEEVKVSEIVTEEAAP-AVEPVAAEPEKAEVP-AVSEEPKEA	58
C54	1	MATVEVVSAQTALPEEK-AEEVKVPE-----AAPVADEKPEEPEKAEAVDQVSEEPVAP	54
		*** ** *	
Pt2L4	59	D-DAPAEVAVETKEVVEVEEAKTVTEEPTVEKTEEEEEETPKKEETPEPVVVKETPKKEEPAA	117
C54	55	ESDAPAEVAVETKEVVEVEDPKDVKEEPEVEKTAEEEE-TPEEETSQ-----EEPAP	104
		***** *	
Pt2L4	118	ETVVVEAPKETTEAAATEAEAPAPESAPASAS <b>ETPAEEVVPKEE</b> EGDEKKSEAEVEA--EK	175
C54	105	ETVVVEAPKETTEPATDAEAPAPE-----AP <b>VEAPKKEE</b> EGGKPKPEAEAAVEK	154
		***** *	
Pt2L4	176	TE	177
C54	155	PE	156
		*	

**Figure 3.** Alignment of deduced amino acid sequence of Pt2L4 with the C54 protein. Identical amino acids are indicated by asterisks under the alignment. ETPKKE repeats are in bold. Dashes represent gaps for alignment optimization. Sequence alignment was performed with BLAST program.



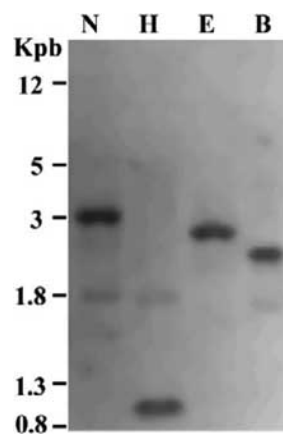
**Figure 4.** Transcript levels for *Mec1* in storage root and tissues.

Total RNA (10 µg) from: L1 – L5: the five storage root tissue layers; (Ct.): cotyledon; (YS): young stem; (SP): stem peel; (Pt.): petiole and (Lf.): leaf were separated on formaldehyde-agarose gel, transferred to a Hybond-N+ membrane and probed with *Mec1* cDNA labeled with [<sup>32</sup>P] dCTP. Hybridization with ribosomal 28S RNA is included as a standard.

that do not cut within the cDNA sequence and hybridized to the 531 bp *Mec1* probe. Our result showed that the *Mec1* probe hybridized with a single strong band, however, weak signals were also detected in all samples, suggesting that there are at least two homologous genes in cassava. (Fig. 5), as reported for C54 by Zhang *et al.* [14].

## DISCUSSION

We report the characterization of the Pt2L4 protein from cassava storage root. The Pt2L4 is a highly acidic protein with a predicted molecular weight of 18.8 kDa and a pI of 3.75. Based on SDS-PAGE analyses [13], we initially estimated its molecular weight at 32 kDa. Recent experiments, however, involving MALDI-ToF analyses, confirmed the predicted molecular weight of 18.8 kDa (data not shown). This difference in size determination using SDS-PAGE versus molecular weight prediction by nucleotide sequence is often seen with acidic proteins and has been observed by others [21, 22] and in the Hev b 5 from *Hevea brasiliensis* [23].



**Figure 5.** Southern blot of cassava genomic DNA.

Ten microgram of DNA was digested with four restriction enzymes *Nsi* I, *Hind* III, *Eco*RI and *Bam*H I, separated on a 0.7% agarose gel, transferred to a nylon membrane and hybridized with *Mec1* probe labeled with <sup>32</sup>PdCTP.

The elucidation of the function of Pt2L4 was not the goal of the present study, some hypotheses, however, based on the protein's physical-chemical characteristics as well as its expression pattern in different tissues and tissue layers, can be proposed and discussed.

As shown in Fig. 4, the *mec1* gene is differentially transcribed within tissues and storage root tissue layers. Strong transcription signals were detected in RNAs from young stems and in the different layers of the storage root transcripts increased from the outer to the inner layer (Fig. 4). This pattern parallels the increase in tissue age during secondary growth of the storage root. In leaves and stem peels the gene is silent, in stems, however, a marked activity can be seen. Taken these facts together, we propose that Pt2L4 is a protein involved in secondary growth of the underground storage organ and, in addition, also with the secondary growth of stems. The absence of transcripts in leaves and bark is consistent with this interpretation since the tissue types present in these organs are mainly vascular and palisade parenchyma in leaves, and epidermis and phloem in the case of peels. Combining the histological aspect of layers 3, 4 and 5 (Fig. 1) with the increased presence of *Mec1* transcripts (Fig. 4), one could speculate whether this protein is not also involved in the increase of the secondary xylem. These interpretations are supported by the findings of Zhang *et al.* [14], who showed for cassava protein C54, which is highly similar to Pt2L4, that the gene's promoter is active in phloem, cambium and xylem.

In addition, we have recently shown by comparative gene expression studies, that the *mec1* gene has a higher expression level in storage roots than in adventitious roots, again indicating its importance in storage root formation [15] and experiments are under way to elucidate, on the molecular level, the exact function of Pt2L4.

#### ACKNOWLEDGEMENTS

We thank Dr. Carlos Bloch Jr. for the MALDI-TOF analysis. Thanks go to the Brazilian National Research Council (CNPq) for the fellowship grant # 142760/1996-0 to CRBS, to the Inter-American Institute for Agriculture Cooperation (IICA) for grants # 244/01 and 305/00 (for CRBS) and to EMBRAPA–Genetic Resources and Biotechnology (grant # 03.1998.027 to LJCBC).

#### REFERENCES

- [1] Puonti-Kaerlas, J. (1998) *Biotech. Genet. Engin. Rev.*, 15, 329–364.
- [2] Cock, J.H. (1985) In: Cassava research, production and utilization (Cock J.H. and Reyes J.A.) pp. 716-728. Centro Internacional de Agricultura Tropical, Cali, Colombia.
- [3] Rateaver, B. (1951) In: Anatomy and regeneration in the stem and root of *Manihot utilissima* Pohl. Ph.D. Thesis, University of Michigan, Ann Arbor, Michigan, USA.
- [4] Castilloa, J.J., Castilloa, A. and Pino, L.T. (1997) In La yuca frente al hambre del mundo tropical” pp. 77-100. Universidad Central de Venezuela, Caracas.
- [5] Cabral, G.B., Carvalho, L.J.C.B. and Schaal, B.A. (2000) In: Cassava Biotechnology IVth International Scientific Meeting–CBN” (Carvalho L.J.C.B., Thro A.M. and Vilarinhos A.D.) pp. 345-356. Embrapa-Recursos Genéticos e Biotecnologia, Brasília-DF, Brazil.
- [6] de Souza, C.R.B. (2001) In: Isolamento e caracterização molecular de genes e proteínas de raiz de reserva de mandioca (*Manihot esculenta* Crantz). PhD thesis, Universidade de Brasília (DF), Brazil.
- [7] Abraham, A. (1957) *Indian J. Gen. Plant Breeding*, 17, 212-217.
- [8] Bolhuis, G.G. (1953) *Euphytica*, 2, 107-112.
- [9] Clowes, A.E.E., Tatham, A.S., Beeching, J.R., Shewry, P.R. (1994) In: Proceedings of the Second International Scientific Meeting of the Cassava Biotechnology Network (Roca W.M. and Thro A.M.) pp. 716-728. Centro Internacional de Agricultura Tropical, Cali, Colombia.
- [10] Maciel, Z.G.Q., Sousa, A.M., Sousa, J.E.A., Campos, F.A.P. (1994) In: Proceedings of the Second International Scientific Meeting of the Cassava Biotechnology Network (Roca W.M. and Thro A.M.) pp. 716-728. Centro Internacional de Agricultura Tropical, Cali, Colombia.
- [11] Gomez, G. and Noma, A.T. (1986) *Nutr. Rep. Int.*, 33, 595-599.
- [12] Cabral, G.B. and Carvalho, L.J.C.B. (2001) *Braz. J. Plant Physiol.*, 13, 41-48.
- [13] de Souza, C.R.B., Carvalho, L.J.C.B., de Almeida, E.R.P. and Gander, E.S. (2002) *Plant Foods Hum. Nutr.*, 57, 353-363.
- [14] Zhang, P., Bohl-Zenger, S., Puonti-Kaerlas, J., Potrykus, I., Gruijssem, W. (2003) *Planta*, 218, 192-203.
- [15] de Souza, C.R.B., Carvalho, L.J.C.B., Cascardo J.C.M. (2004) *Protein Pept. Lett.*, 11 (6), 577-582.
- [16] Altschul, S.F., Madden, T.L., Shaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J. (1997) *Nucleic Acids Res.*, 25, 3389-3402.
- [17] Deblaere, R., Reynaerts, A., Hofte, H., Hernalsteen, J.P., Leemans, J., Van Montagu, M. (1987) *Methods Enzymol.*, 153, 277-292.
- [18] Daiuto, E.R. (2000) In: Biosíntese de amido durante o crescimento secundário da raiz de reserva da mandioca. MSc Dissertation, Universidade Estadual de São Paulo, Botucatu-SP, Brazil.
- [19] Indira, P. and Kurian, T. (1977) *J. Root Crops*, 3, 29-32.
- [20] O'Hair, S.K. (1989) *HortScience*, 24, 505-506.
- [21] Rasmussen, H.B. (1993) *APMIS Suppl.*, 39, 1-28.
- [22] Hames, B.D. (1981) in Gel Electrophoresis of Proteins (Hames B.D. and Rickwood D.) pp. 1-91. Oxford University Press, New York.
- [23] Slater, J.E., Vedvick, T., Arthur-Smith, A., Trybul, D.E., Kekwick, R.G.O. (1996) *J. Biol. Chem.*, 271, 25394-25399.