P-ISSN: 2305-6622; E-ISSN: 2306-3599
International Journal of Agriculture and Biosciences
www.ijagbio.com; editor@ijagbio.com

Research Article

Cloning and Bioinformatics Analysis of *accD* Gene from Bell Pepper (*Capsicum annuum*)

Samiee M, Baghban Kohnehrouz B* and Norouzi M

Department of Plant Breeding and Biotechnology, Faculty of Agriculture, University of Tabriz, Tabriz-51666, Iran
***Corresponding author:** bahramrouz@yahoo.com

Article History: Received: December 02, 2015 Revised: March 13, 2016 Accepted: March 22, 2016

ABSTRACT

Acetyl-CoA carboxylase (EC 6.4.1.2) plays an important role in biosynthesis of malonyl-CoA in plastids. Plastidic ACCase is devided into four subunits including β-carboxyl transferase encoded by *accD* gene in plastidial genome. As the importance of this gene, we have isolated and cloned the *accD* gene from chloroplast genome of *Capsicum annuum*. According to the multiple alignment result, the cloned fragment ~2770bp-long comprises of *accD* promoter, mRNA coding region with 1655bp-long, which is encoded a deduced 551aa polypeptide. Furthermore, the sequence data of pepper's *accD* gene region shows a maximum homology of 93% with Tobacco, Tomato and Potato, followed by an average homology with Phoenix and Glycine with efficient performance of 81%, and 49%, respectively. The -35-like (TTGACA), -10-like (TATCAA) regions of *accD* promoter for assembling plastid-encoded RNA polymerase (PEP) were identified using bioinformatics analysis. Sequence characterization of this promoter revealed that this region contained several important motifs such as AT-rich, CAAT box, G-box, and ATTAAT.

Key words: ACCase, *accD* gene, Bell pepper, Bioinformatics analysis, Cloning, Chloroplast

INTRODUCTION

Chloroplasts are one of the most important cytoplasm organelles found in free-living photosynthetic microorganisms and higher plants that conduct photosynthesis to produce carbohydrates and free Oxygen molecules. According to resources, cyanobacteria known as primary plastids are considered as the ancestors of chloroplasts, which were adapted by a single endosymbiotic event in eukaryotic cell during its early evolution (Verma and Daniell, 2007). Contrary to cyanobacteria, the plastid genome is reduced in size (120 to 220 kb), but the existing genomic sequences still show clear similarities (Martin *et al*., 2002). Chloroplasts possess their own genomic chloroplast DNA (cpDNA) combined into a single, large and circular DNA molecule, typically 120,000 –170,000 bp in length. They can have a contour length of around 30–60 micrometers, and have a mass of about 80–130 million Daltons. Plastid DNA presents a highly conserved structure of two identical copies of an inverted repeat region (IRA and IRB) separating a large single copy (LSC) region and a small single copy (SSC) region. The chloroplast genome in land plants typically contains 110 to 120 unique genes, whereas cyanobacteria contain more than 1500 genes. Therefore, it shows that many of the missing genes were transferred into the nuclear genome of the host (Martin *et al*., 2002).

Plastids perform some important biochemical activities including evolution of oxygen, production of starch, synthesis of amino acids, and pigments, and key aspects of sulfur and nitrogen metabolism, as well as de novo fatty acid synthesis (Verma and Daniell, 2007). Plastidic acetyl-CoA carboxylase (ACCase; EC 6.4.1.2) catalyzes the formation of malonyl-CoA, the first committed, rate-limiting step of fatty acid biosynthesis (Madoka *et al*., 2002). In plants, this gene represents two forms of ACCase including a homomeric ACCase in cytosol and a heteromeric ACCase in plastids, which are so similar to animal-type ACCase and *E.coli*-type ACCase, respectively (Konishi and Sasaki, 1994; Sasaki *et al*., 1993). Since, cell membranes are impermeable to acyl-CoAs especially malonyl-CoA (Banhegyi *et al*., 1996; Jacobson and Stumpf, 1972), the plastidic ACCase makes an important contribution for producing malonyl-CoA in plastids. The plastidic ACCase are devided into four subunits comprising of biotin carboxylase (BC), biotin carboxyl carrier (BCCP), and α and β subunits of

Cite This Article as: Samiee M, Baghban Kohnehrouz B and Norouzi M, 2016. Cloning and bioinformatics analysis of *accD* gene from bell pepper (*Capsicum annuum*). Inter J Agri Biosci, 5(2): 67-72. www.ijagbio.com

carboxyl transferase (α and β-CT), which are encoded by *accC*, *accB*, *accA*, and *accD*, respectively (Ohlrogge and Browse, 1995; Sasaki *et al*., 1995). In higher plants, the *accD* gene is located in LSC region of plastid genome.

So far, a study associated with *accD* gene retrieved from common sweet pepper has not been reported. Therefore, the main aim of this work was to isolate and *in silico* analysis of *accD* gene and its promoter sequence from chloroplast genome of Bell Pepper (*Capsicum annuum*). Moreover, an effort has been made to identify important regulatory elements in *accD* promoter using computational approach.

MATERIALS AND METHODS

Plant, plasmid, enzymes and chemicals

Bell papper (*Capsicum annuum* L. cv California Wonder) plants were grown in a pod in a greenhouse condition. The pTG19-T cloning vector (Vivantis, USA) derived from pTZ19-R vector (Accession no. Y14835.1) was used as a plasmid backbone for all cloning purposes. *Escherichia coli* strain DH5α (Invitrogen, USA) was used as a host strain for molecular cloning. Restriction enzymes, *Taq* DNA polymerase enzyme, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and Isopropyl β-D-1-thiogalactopyranoside (IPTG) were obtained from Thermo Fisher Scientific Inc. (USA). All other chemicals were molecular biology grade (Merck, Germany).

Genomic DNA extraction

The presence of polysaccharides hinders enzymatic activity of *Taq* DNA polymerase. Hence, followed by 24 hours darkness treatment for full breaking down the cellular starch content, total genomic DNA was isolated from 2-3 leaves growth stage using Cetyl Trimethyl Ammonium Bromide (CTAB) reagent method (Saghaei *et al*,. 1984). The concentration of genomic DNA was measured spectrophotometrically and its quality was checked by 0.8% agarose gel electrophoresis.

Primer design and PCR amplification

The specific primer sets for the *accD* gen region in plastid genome was synthesized from accession number of NC_018552. All of primers were designed using Primer3 software at NCBI. The *Eco*RI and *Pst*I restriction sites (underlined letters) were added at the 5'-ends of forward and reverse primer set listed as follows:

accD_F: 5'- AAGAGCTCTGCAATTAAACTCGGCCCAA-3'
accD_R: 5'-AACTGCAGTCGTTCGAATCCGACCCAAC-3'

PCR amplification conducted in 20µl volume, consisting of 50ng total genomic DNA, 400nM of each of two primers. The reactions were run in standard thermal cycler (Astec, Japan) with following common program: an initial denaturation step of 94ºC for 5min, 35 cycles of 94ºC for 1min, 57.2ºC for 30 sec and 72ºC for 60 sec; then a final extension step of 4 min at 72ºC. The PCR-amplified fragments were checked out in 0.8% agarose gel electrophoresis and subjected to electrophoresis at 85 volts.

Molecular cloning

First PCR-amplified fragment were purified by gel purification kit (Bioneer, South Korea; cat. no. k-3035) according to the manufacturer's instructions and used in ligation reaction with pTG19-T vector in 48 ng/µl concentration of insert DNA at 4°C for 24 hrs. Then it was TA-cloned into linear pTG19-T cloning vector by T4 DNA ligase (200u/µl, Vivantis, USA). *E.coli* competent bacterial cells were prepared using TSS protocol and transformation was done by 5µl of ligation reaction using heat shock procedure.

For screening, after recovery of bacteria on antibiotic-free Luria-Bertani (LB)-liquid culture, the cells were plated on LB-agar medium (peptone 1% (w/v), yeast extract 0.5% (w/v), NaCl 1% (w/v), agar 1.2% (w/v)) containing ampicillin (100µg/ml), X-gal (100µg/ml) and IPTG (1mM). The resultant white colonies on media containing x-gal and IPTG were further confirmed by direct colony PCR technique before inoculation of liquid bacterial culture. The recombinant plasmid DNA was extracted from positive colonies incubated in the liquid LB medium containing ampicillin (100µg/ml) for 16hrs using plasmid extraction kit (Bioneer, South Korea; cat. no. k-3034) according to the manufacturer's instructions. The fragment was validated by *Eco*RI restriction digestion analysis and verified by sequencing. T7 promoter primer (TAATACGACTCACATTAGGG) was used for sequencing.

Bioinformatics analysis

The nucleotide Sequence of *accD* gene insert was confirmed by BLAST analysis at national center of biotechnology information (NCBI) (http://www.ncbi.nlm.nih.gov/BLAST/) and aligned using CLUSTAL Omega program. PLANTCARE software (http://bioinformatics.psb.ugent.be/webtools/plantcare/html) was used to identify regulatory elements in the sequence data of *accD* promoter region.

RESULTS

Cloning of *accD* gene

The *accD* gene fragment was obtained by PCR amplification using accD_F and accD_R primer set. The size of PCR amplified fragment with ~2770 bp in length (Lane 1) was confirmed by gene ruler 1kb DNA ladder (Thermo Scientific Co., USA) (Lane M) (Fig. 1).

The fragment was cloned into pTG19-T cloning vector as illustrated in Figure 2. In the cloning of the fragment, the correct colonies were screened out by standard blue/white screening system. The DNA plasmid with ~2770bp fragment was confirmed by *Eco*RI restriction digestion and confirmed by gene ruler 1kb DNA ladder (Lane M) (Fig. 3). The restriction analysis result showed that the desired insert was cloned into pTG19-T cloning vector in opposite orientation, which is led to excise the fragment with ~2770bp-long from the cloning vector with ~2900bp in length (Lane 2) as shown in Figure 3.

Sequencing analysis of *accD* gene

The size of the pepper's plastid genome is 156,781 bps, which is the largest among known Solanaceous plastomes. The quadripartite structure includes in 87,366

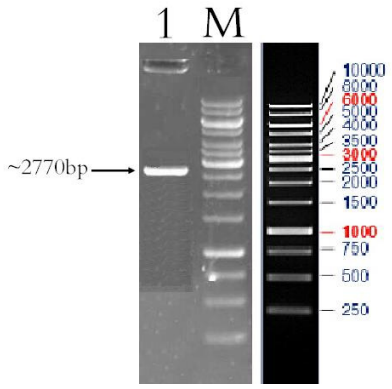

Fig 1: Illustration of the PCR amplified fragment of *accD* gene region retrieved from plastome of bell pepper. DNA concentrations and size of the PCR product (lane 1) was confirmed by gene ruler 1kb DNA ladder (Thermo Scientific Co., USA) (Lane M) loaded on 0.8% agarose gel.

bps of LSC and 25,783 bps of SSC that are separated by a pair of 17,849 bp of IR copies. According to chloroplast gene mapping, *accD* gene sequence is located at the large single copy (LSC) region (Jo *et al*., 2011). The *accD* coding sequence of pepper was compared with some of plastomes from other plants Using BLAST software at NCBI. As the nucleotide alignment results, the sequence data of pepper's *accD* gene region shows a maximum homology of 93% with tobacco, tomato and potato, followed by an average homology with Phoenix and Glycine with efficient performance of 81%, and 49%, respectively (for more detail see Table 1).

The promoters of chloroplast genes are typically composed of two hexamer sequences, ctpl and ctp2, separated on average by 16-18 nucleotides and resembling -35 and -10 prokaryotic core promoter, respectively (Handley-Bowdoin and Chua, 1987). The sequence data of *accD* promoter region were analyzed using PLANTCARE bioinformatics software and regulatory elements as well as conserved motifs in promoter region were identified (Table 2). In figure 4, the -35 and -10 (TATA box) regions of *accD* promoter respectively with the conserved sequences of TTGACA and TATCAA (bold made and underlined letters) was identified using bioinformatics analysis. Moreover, the conserved ribosome-binding region of *accD* promoter with sequence data of AGGAGAGGA (bold made and underlined letters) was illustrated in figure 4.

With regard to amino acid composition of *accD* gene, the nucleotide sequence of *accD* gene contains 1655nt mRNA coding region, which is encoded into a 551aa polypeptide. The nucleotide and amino acid composition of *accD* gene in ORF2 origin were presented in Figures 4 and 5, respectively. As the results, the conserved sequences of prokaryotic transcriptional termination signals led to form Hairpin-loop structure at the end of transcription process were identified at the untranslated-3´region of mRNA of *accD* gene with the sequences of TTTGTAGCAAA and TGATTACGAATCA (bold made and underlined letters) (Fig. 4). Furthermore, the binding sites corresponding to acetyl-CoA, carboxy biotin and carboxyl transferase were identified in amino acid sequence of *accD* gene in the range of 368-382, 386-404 and 360-370 (Fig. 5).

Table 1: The nucleotide alignment results among the sequence data of *accD* gene region from Bell papper (*Capsicum annuum*) compared to some other species.

	CaccD	*PaccD*	*TaccD*	*NaccD*	*GaccD*	*PhaccD*	*EaccD*
CaccD	-	93	93	93	49	81	51
PaccD			100	100	53	86	86
TaccD				100	54	86	86
NaccD					53	95	86
GaccD						66	-
PhaccD							54

CaccD (*Capsicum annuum*), NC_018552; *PaccD* (*Solanum tuberosum*) NC_008096.2; *TaccD* (*Solanum licopersicum*) NC_007898.3; *NaccD* (*Nicotiana tabacum*) NC_ 001879.2; *GaccD* (*Glycine max*) NC_007942.1; *PhaccD* (*Phoenix dactylifera*) NC_013991.2; *EaccD* (*Epifagus virginiana*) M81884.1.

Table 2: Regulatory elements in *accD* promoter sequence from bell pepper

Site	Position	Strand	Sequence	Function
AT-rich element	409	-	ATAGAAATCAA	Binding site for DNA binding protein (ATBP-1)
	642	-		
Box 4	560	+	ATTAAT	Part of a conserved DNA module involved in light responsiveness
Box I	107	-	TTTCAAA	Light responsive element
CAAT-box	179	+	CAAT	Common cis-acting element in promoter and enhancer regions
	462	-	CAAAT	
G-box	468	+	CACGAC	Cis-acting regulatory element involved in light responsiveness
GATA-motif	662	+	GATAGGA	Part of a light responsive element
GCC box	482	+	AGCCGCC	-
GCN4_motif	248	+		Cis-regulatory element involved in endosperm expression
HSE	399	-	AAAAAATTTC	Cis-acting element involved in heat stress responsiveness
	593	-		
	400	-		
Sp1	484		GGGCGG	Light responsive element
TATA-box	67	+	TATA	Core promoter element around -30 of transcription start
TC-rich repeats	395	-	ATTTTCTTCA	Cis-acting element involved in defense and stress responsiveness
Unnamed__4	423	-	CTCC	
	477	-		-
	431	-		

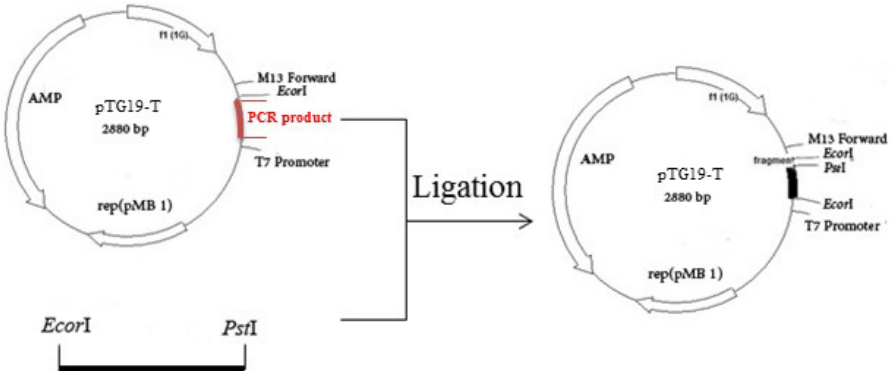

Fig. 2: Diagram of the cloning procedure into pTG19-T cloning vector

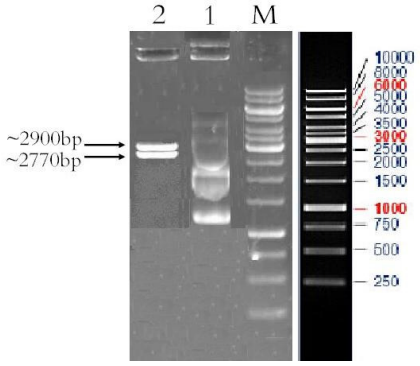

Fig. 3: Screening for recombinant plasmid DNA harboring PCR amplified fragment of *accD* gene region. The PCR product was excised from the recombinant DNA plasmid (Lane 1) by *Eco*RI restriction digestion (Lane 2). The exact size of the plasmid DNAs with the insert was validated by gene ruler 1kb DNA ladder (Thermo Scientific, cat. no. sm0313) (Lane M).

DISCUSSION

Although, AT-rich DNA sequences have been reported to stimulate transcription in yeast, but still their function in plant genes is less understood (Chen *et al*., 1987; Russell *et al*., 1983; Struhl, 1985). According to several studies, AT-rich elements in the plant promoter possess transcriptional activity such as the AT-rich element of β-phaseolin promoter from French bean fused on the cauliflower mosaic virus 35s (CaMV) promoter, which is led to increase in the expression of β-glucuronidase (GUS) reporter gene in transgenic tobacco (Bustos *et al*., 1989).

The CAAT-box motif, a common cis-acting element in promoter and enhancer regions, is often located at position of -80. This motif plays an important role in determining the efficiency of promoter and is found in *Brassica rapa, A.thaliana*, *Glycine max*, *Petunia hybrid* and *H. vulgare* (Shirsat *et al*., 1989). G-Box in *Pisum sativum* and *A. thaliana* is involved in light responsibility. The GATA motif is a binding transcription factors bind to the DNA consensus sequence GATA (Molkentin, 2000; Patient and McGhee, 2002).

In different plants species, the GCC-Box is a conserved element presents in the promoters of several ethylene-induced pathogenesis-related genes, such as PRB-1b, chitinase β-1,3-glucanase (Eyal *et al.,* 1993; Hart *et al*., 1993; Ohme-Takagi and Shinshi, 1990). In the bean chitinase gene, this box has a function as a part of the minimal promoter responsive to ethylene induction (Brogfie *et al*., 1989; Roby *et al*., 1991). The GCN4 motif in rice is necessary for gene expression, especially in aleorune cells of the endosperm (Wu *et al*., 1998).

According to the reports associated with in vitro interactions of a tomato (*Lycopersicom esculentum*) HSF with the *apx1* promoter and its mutational analysis, the HSE is responsible for the heat-shock induction of the gene and partially contributes to the induction by oxidative stress (Storozhenko *et al*., 1998). A TATA box sequence has been found in almost all plant genes in recent studies (Mesing *et al*., 1983). In eukaryotic promoters, between 10 and 20% of all genes (Gershenzon and Ioshikhes, 2005) contain a TATA box (sequence TATAAA), which provides a TATA binding protein and assists the formation of the RNA polymerase transcriptional complex (Smale and Kadonaga, 2003). The TATA box typically lies very close in the transcription initiation site (often within 50 bases), and tends to be surrounded by GC rich sequences. A guanidine (G) and cytidine (C) rich sequence usually present in multiple locations in the promoter region and normally surround the TATA box and CAP site.

Conclusions

The aim of this study was to isolate, cloning and in silico characterization of *accD* gene from chloroplast genome of common sweet pepper by specific primers. Amplified fragment was cloned into pTG19-T vector. Then it was transformed into *E. coli* strain DH5α. In the cloning of the fragment, the correct colonies were screened out by standard blue/white screening system. The DNA plasmid with ~2770bp fragment was confirmed by *Eco*RI restriction digestion and the sequence data was validated by sequencing. The *accD* coding sequence of pepper was compared with some of plastomes from other plants Using BLAST software at NCBI. As the nucleotide alignment results, the sequence data of pepper's *accD* gene region shows a maximum homology of 93% with Tobacco, Tomato and Potato, followed by an average homology with Phoenix and Glycine with efficient performance of 81%, and 49%, respectively by gene ruler 1kb DNA ladder. The promoter region was analyzed and found that the motifs like CAAT-box, TATA-box, HSE and Box 4 is present in bell pepper plastome.

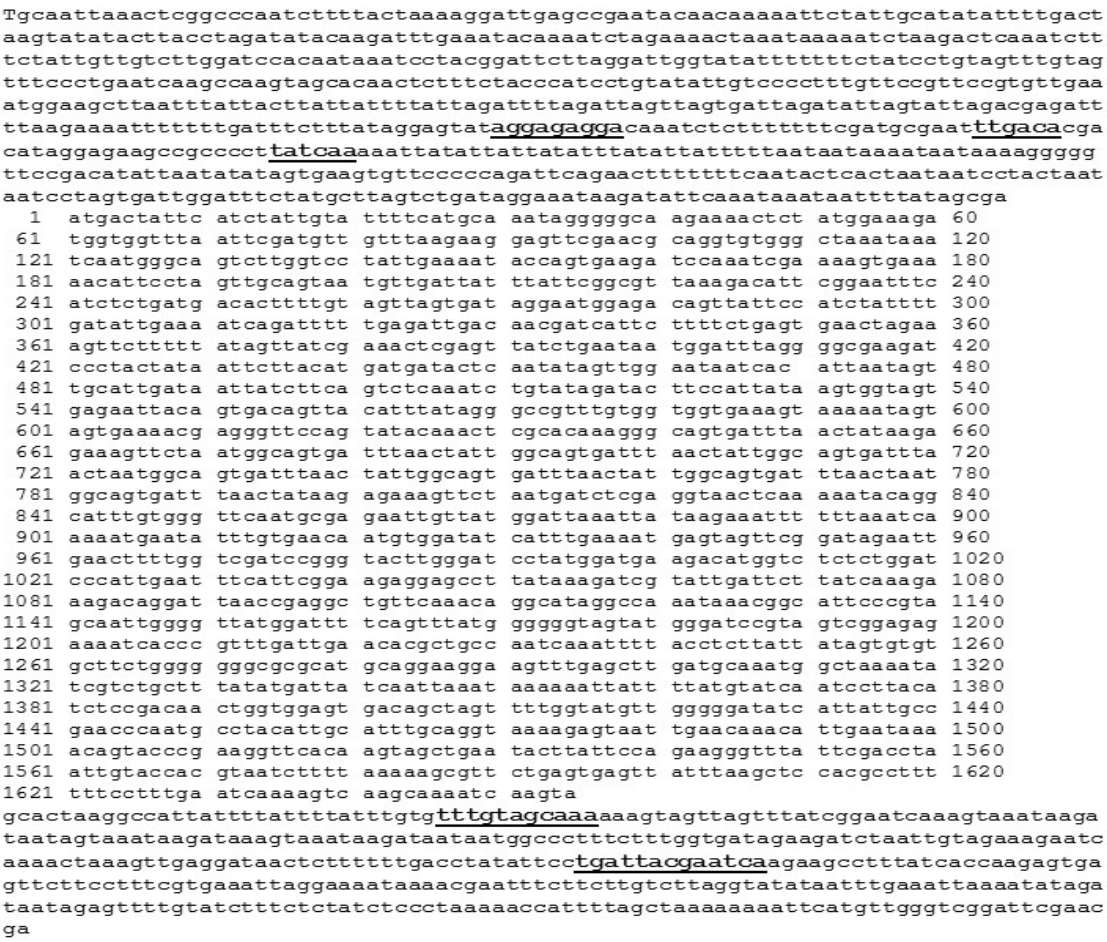

Fig. 4: Illustration of the sequence data of *accD* gene and its promoter sequence. The prokaryotic motives including -35 (TTGACA), -10 (TATCAA) and the conserved ribosome-binding (AGGAGAGGA) regions of *accD* promoter are shown in bold made and underlined letters.

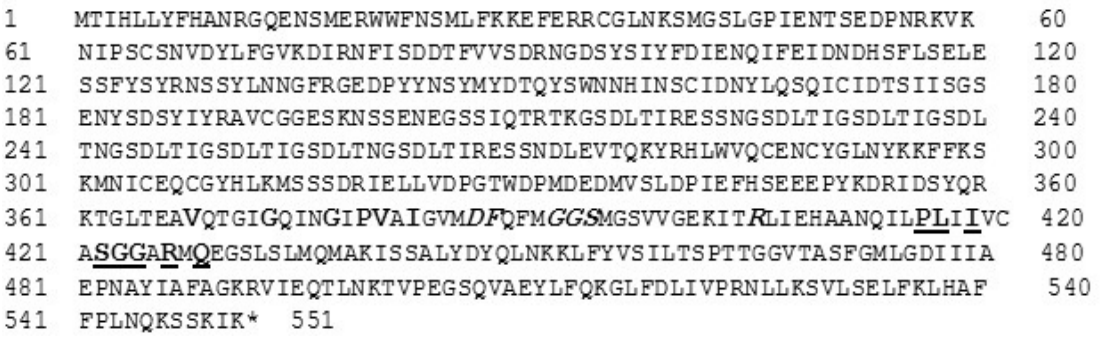

Fig 5: Illustration of *accD* amino acid composition. Bold letters: Acetyl-CoA binding site, bold letter italicized: Carboxybiotin binding site, bold letters underlines: putative catalytic site of carboxyl transferase.

Acknowledgments

This research was funded by Aras International campus and performed at genetic engineering laboratory, department of Biotechnology in Agriculture at University of Tabriz.

REFERENCES

Banhegyi G, M Csala, J Mandl, A Burchell, B Burchell, P Marcolongo, R Fulceri and A Benedetti, 1996. Fatty acyl-CoA esters and the permeability of rat liver microsomal vesicles. Biochem J, 320: 343–344.

Brogfie KE, P Biddle, R Cressman and R Broglie, 1989. Functional analysis of DNA sequences responsible for ethylene regulation of a bean chitinase gene in transgenic tobacco. Plant Cell, 1: 599-607.

Bustos MM, MJ Guiltlnan, J Jordano, D Begum, FA Kalkan and TC Hall, 1989. Regulation of fl-glucuronidase expression in transgenic tobacco plants by an ATT-rich, cis-acting sequence found upstream

of a french bean 6-phaseolin gene. Plant Cell, 1: 839-853.

Chen W, S Tabor and K Struhl, 1987. Distinguishing between mechanisms of eukaryotic transcriptional activation with bacteriophage T7 polymerase. Cell, 50: 1047-1055.

Eyal Y, Y Meller, S Lev-Yadun and R Fluhr, 1993. A basic-type PR-1 promoter directs ethylene responsiveness, vascular and abscission zone-specific expression. Plant J, 4: 225- 234.

Gershenzon NI and IP Ioshikhes, 2005. Synergy of human Pol II core promoter elements revealed by statistical sequence analysis. Bioinformatics, 21: 1295-300.

Handley-Bowdoin L and NH Chua, 1987. Chloroplast promoters. Trends In BioSci, 12: 67-70.

Hart CM, F Nagy and JF Meins, 1993. A 61 bps enhancer element of the tobacco fl-1,3-glucanase B gene interacts with one or more regulated nuclear proteins. Plant Mol Biol, 21: 121-131.

Jacobson BS and PK Stumpf, 1972. Fat metabolism in higher plants: LV. Acetate uptake andaccumulation by class I and class II chloroplasts from *Spinaciaoleracea*. Arch Biochem Biophys, 153: 656–663.

Jo YD, J Park, J Kim, W Song, CG Hur, YH Lee and BC Kang, 2011. Complete sequencing and comparative analyses of the pepper (*Capsicum annuum* L.) plastom revealed high frequency of tandem repeat and larg insertion/deletions on pepper plastom. Plant Cell Reports, 30: 217-229.

Konishi T and Y Sasaki, 1994. Compartmentalization of two forms of acetyl-CoA carboxylase in plants and the origin of their tolerance towards herbicides. Proc Natl Acad Sci USA, 91: 3598–3601.

Madoka Y, K Tomizawa, J Mizoi, I Nishida, Y Nagano and Y Sasaki, 2002. Chloroplast transformation with modified accDoperon increases acetyl-coAcarboxylase and causes extension of leaf longevity and increase in seed yield in tobacco, Plant Cell Physiol, 43: 1518–1525.

Martin WRT, E Richly, A Hansen, S Cornelson, T Lins, D Leister, B Stoebe, M Hasegawa and D Penny, 2002. Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci USA, 99: 12246–12251.

Molkentin JD, 2000. The zinc finger-containing transcription factors GATA-4, -5 and -6. Ubiquitously expressed regulators of tissue-specific gene expression. J Biol Chem, 275, 38949–38952.

Ohlrogge J and J Browse, 1995. Lipid biosynthesis. Plant Cell, 7: 957–970.

Ohme-Takagi M and H Shinshi, 1990. Structure and expression of a tobacco fl-l,3-glucanase gene. Plant Mol Biol, 15: 941-946.

Patient RK and JD McGhee, 2002. The GATA family (vertebrates and invertebrates). Curr Opin Genet Dev, 12: 416–422.

Roby D, K Broglie, J Gynor and R Broglie, 1991. Regulation of a chitinase gene promoter by ethylene and elicitors in bean protoplasts. Plant Physiol, 97:433-439.

Russell DW, M Smith, D Cox, VM Williamson and ET Young, 1983. DNA sequences of two yeast promoter-up mutants. Nature, 304: 652-654.

Saghai-Maroof MA, KM Soliman, RA Jorjensen and RW Allard, 1984. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA, 81: 8014-8018.

Sasaki Y, T Konishi and Y Nagano, 1995. Thecompartmentation of acetyl- CoA carboxylase in plants. Plant Physiol, 108: 445–449.

Shirsat A, N Wilford, R Croy and D Boulter, 1989. Sequences responsible for the tissue specific promoter activity of a pea legumin gene in tobacco. Mol Gen Genet, 215: 326-331.

Smale T and T Kadonaga, 2003. The RNA polymerase II core promoter. Ann Rev Biochem, 72: 449-479.

Storozhenko S, PD Pauw, MVD MontaguInze and S Kushnir, 1998. The heat-shock element is a functional component of the Arabidopsis APX1 gene promoter. Plant Physiol, 118: 1005–1014.

Struhl K, 1985. Naturally occurring poly (dA-dT) sequences are upstream promoter elements for constitutive transcription on yeast. Proc Natl Acad Sci USA, 82: 8419-8423.

Verma D and H Daniell, 2007. Chloroplast vector systems for biotechnology applications, Plant Physiol, 145: 1129-1143.

Wu CHY, A Suzuki, H Washida and F Takaiwa, 1998. The GCN4 motif in a rice glutenins gene is essential for endosperm-specific gene expression and is activated by Opaque-2 in transgenic. Plant J, 14: 673-683.