Cloning of Isochorismate Synthase cDNA into Expression Vector

Abdul Attermidzi Bin Matali

This project is submitted in partial fulfillment of the requirements for the degree of

Bachelor of Science with Honours

(Resource Biotechnology)

Department of Molecular Biology
Faculty of Resource Science and Technology
UNIVERSITI MALAYSIA SARAWAK
2010

Acknowledgement

First of all, I am heartily thankful and express my gratitude to my supervisor, Dr Hairul Azman Roslan, for the valuable guidance, advice and support from the initial to the final level of this project.

Secondly, my deepest gratitude to all Postgraduate students of Genetic Engineering Laboratory, whose motivate, help and guidance me to develop an understanding of the study. Thank you so much brothers and sisters.

Special thanks also to all my laboratory friends especially Aiesah Udin, Muhammad Mustaqim, Siti Hadizah Mumin and Judith Liew for sharing literature and offered invaluable assistance. Not forgetting to Alina Sali and Rafiuddin Rasman for sharing their bioinformatics guidance.

I also would like to convey to the Faculty of Resource Science and Technology, Universiti Malaysia Sarawak for providing me excellent laboratory facilities. An honorable mention goes to my family and blessing to all of those who supported me in any respect during the completion of the project.

Thank you everyone.

Abdul Attermidzi bin Matali

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AGE	Agarose Gel Electrophoresis
CaCl2	Calcium Chloride
cDNA	Complementary DNA
DHBA	2,3-dihydroxylbenzoic acid
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
Etbr	Ethidium bromide
ICS	Isochorismate synthase
IPTG	Isopropyl-β-D-thiogalactopyranoside
LB	Luria broth
MgCl	Magnesium chloride
NaCl	Sodium chloride
OD	Optical density
PCR	Polymerase Chain Reaction
PEG	Polyethylene glycol
RNA	Ribonucleic acid
Rpm	Rotor per minute
TAE	Tris-Base, Acetic acid, EDTA

List of Tables

List of Figures

Cloning of Isochorismate Synthase cDNA into Expression Vector

Abdul Attermidzi bin Matali

Resource Biotechnology Programme
Department of Molecular Biology
Faculty of Resource Science and Technology
Universiti Malaysia Sarawak

ABSTRACT

Isochorismate Synthase (ICS) enzyme has its own function to catalyze chorismate to isochorismate and finally produces anthraquinones as a secondary metabolite via the Shikimate Pathway. *ICS* cDNA has been extracted from *Morinda citrifolia* species and had been cloned into pGEM-T plasmid. In this study, *ICS* cDNA gene was re-extracted from pGEM-T plasmid. Several verifications were done to identify the isolated plasmid and gene. Then, the gene was cloned into two types of digested expression vector, which are pET-41a(+) and pUC19. Cloning analysis was done for pET-41a(+) by direct colony Polymerase Chain Reaction (PCR). While cloning analysis for pUC19 was done by blue-white colony screening method. The analyses showed that an *ICS* cDNA gene was successfully clone to pET-41a(+). Meanwhile, blue white colonies of inserted pUC19 to E.coli XL1 blue competent cells were formed in plates of agar containing LAIX (LB Ampicillin IPTG X-gal).

Keywords: ICS cDNA, Morinda citrifolia, PCR, pET-41a(+), pGEM-T, pUC19

ABSTRAK

Enzim Isokorimat Sintase (ICS) mempunyai fungsi tersendiri iaitu mengkatalisiskan korismat ke isokorismat dan akhirnya menghasilkan antrakinon sebagai metabolit kedua dalam laluan Shikimate. Gen ICS cDNA telah diekstrakkan daripada spesis Morinda citrifolia dan telah diklonkan ke dalam plasmid pGEM-T. Dalam kajian ini, ICS cDNA akan diekstrakkan semula daripada plasmid pGEM-T. Beberapa ujian pengesahan telah dilakukan untuk mengesahkan dan mengenalpasti plasmid yang telah diekstrakkan dan juga gen nya. Kemudian, gen yang telah diekstrakkan akan diklonkan ke dalam 2 jenis vektor ekspresi, iaitu pET-41a(+) dan pUC19. Analisis pengkloman telah dijalankan ke atas pET-41a(+) melalui PCR kelangsungan koloni. Manakala analisis pengklonan untuk pUC19 telah dilakukan melalui kaedah penapisan koloni biru dan putih. Analisis menunjukkan bahawa gen ICS cDNA berjaya diklonkan ke dalam pET-41a(+). Sementara itu, koloni biru dan putih pUC19 dan gen ICS cDNA di dalamnya dihasilkan melalui plat agar yang mengandungi LAIX (LB Ampicillin IPTG X-gal).

Kata Kunci: ICS cDNA, Morinda citrifolia, PCR, pET-41a(+), pGEM-T, pUC19

1.0 INTRODUCTION

Morinda citrifolia L., commonly known as noni or 'mengkudu' is a plant that is grows widely in tropics regions. It is belong to Rubiaceae family and consists of three varieties such as citrifolia, bracteata and potteri (Ahmed *et al.*, 2008). The genus *Morinda* in Peninsular Malaysia comprises nine species (Abdullah *et al.*, 1998) which consists of 3 species of trees; *M. citrifolia*, *M. elliptica*, and *M. corneri*, -while others consist of 6 species of climbers. The fruit contains several bioactive compound that have been identified including insecticidal fatty acid, flavonoids, terpenoids, sitosterol and anthraquinones (Siddiqui *et al.*, 2007; Ahmed *et al.*, 2008).

Morinda citrifolia L. has its medicinal properties for over 2000 years in Polynesia region (Ahmed *et al.*, 2008). There are many compounds in noni leaves such as β-sitosterol, campesta-5,7,22-trien-3β-ol, stigmasterol, *E*-phytol and citrifolinoside (West and Bing-nan, 2008). These leaves can be used to treat headaches (West and Bing-nan, 2008), typical inflammation, poisonous fish and insect stings (West *et al.*, 2009). Previous study reported that leaves extract from Morinda have antioxidant and anti-inflammatory component which can be used to reduce UV-induced erythema in human skin (West *et al.*, 2009). The fruit juice itself has good effective in protecting the liver from toxin that is induced by carbon tetrachloride (CCl_4) which inhibits inflammatory response and suppressing liver enzyme activities that can prevent cell membrane damage (Mian-Ying *et al.*, 2008).

Nowadays, recent research is mainly focused on secondary metabolite products which have important role such as pharmaceutical drugs, dyes, flavors, antioxidants, insecticides and pheromones (Abdullah *et al.*, 1998; Borroto *et al.*, 2008; Mustafa and Verpoorte, 2005).

An important secondary product such as Anthraquinones (AQ) has been used in many applications nowadays and this product is derived from chorismate by the conversion of isochorismate synthase (ICS) enzyme. Modern Biotechnology such as cloning technique is a good tool to analyze this *ICS* gene by using an expression vector. Complementary DNA (cDNA) or DNA copies of *ICS* can be made actually from mRNA molecules that are commonly isolated from noni's cell and these cDNA can be cloned later (Russell, 2006).

The objectives of this study are:

- To clone a near complete *isochorismate synthase* cDNA that have been isolated from 'mengkudu' into an expression vector.
- To analyze the cloned product via Polymerase Chain Reaction (PCR) technique.

2.0 LITERATURE REVIEW

2.1 Isochorismate Synthase (ICS)

2.1.1 Background of ICS

Isochorismate synthase (ICS) enzyme plays an important role in catalyzing chorismate to form isochorismate which is then continued in the synthesis of secondary metabolites such as phylloquinone and anthraquinone. Historically, *ICS* was first extracted from *Aerobacter aerogenes* and this gene was cloned from bacteria such as *E. coli, Pseudomonas aeruginosa, Aeromonas hydrophila, Flavobacterium, Hemophilus influenza* and *Bacillus subtilis* (Tegelen *et al.*, 1999).

2.1.2 Role of *ICS* in plants and bacteria

In plants, isochorismate is generally a precursor for biosynthesis of anthraquinones, naphtoquinones, catalpalactone and alkaloid in orchids while in bacteria there are two *ICS* genes found in *E. coli* and *B. subtilis* which is involved in production of menaquinone and siderophore biosynthesis (Tegelen *et al.*, 1999).

DHBA (2,3-dihydroxybenzoic acid) is synthesized mainly from isochorismate in bacteria same as in plant such as *Catharanthus roseus* which also produces DHBA and Trp-derived indole alkaloids (Moreno *et al.*, 1994; Tegelen *et al.*, 1999; Mustafa and Verpoorte, 2005). In plant *C. roseus*, DHBA (2, 3-dihydroxylbenzoic acid) play a major role in the defense mechanism of which only detected in elicited cells with induction of ICS enzyme activity (Moreno, 1994) and ICS protein that was obtained has 57% homology with the *ICS1* of *A. thaliana* and 20% homology with bacterial ICS (Mustafa and Verpoorte, 2005).

In previous work, the ICS enzyme has been partially purified from *Galium mollugo* cell cultures and biochemical characteristics such as substrate affinity, pH optimum and cofactor requirement were found to be similar with *C. roseus* and *Rubia Tinctorum* (Tegelen *et al.*, 1999). In *Rubia peregrine*, an expression of bacterial *isochorismate synthase* cDNA in root culture are twice rather than *isochorismate synthase* activity in control roots (Lodhi *et al.*, 1996).

Other research also found that isochorismic acid is the first intermediate biosynthesis of *ent c* gene in *E. coli* that coding for *ICS* and the PCR product is then cloned into a binary vector to express chimeric gene into root cultures of *R. peregrine* (Lodhi *et al.*, 1996).

2.2 Secondary Metabolite of Anthraquinones

2.2.1 Synthesis of Anthraquinones from Shikimic Pathway

Anthraquinones are most abundant quinones in plant kingdom and can be commercialized in food processing industry based on their properties such as light and heat resistant (Lodhi *et al.*, 1996). It is synthesized mainly either from acetate-malonate pathway or Shikimic acid pathway (Lodhi *et al.*, 1996) and this pathway provides chorismate that involves synthesis of aromatic component such as Phe, Tyr and Trp and then involves in protein synthesis of aromatic compounds (Tegelen *et al.*, 1999). In *Morinda citrifolia* cells, approximately 90% of anthraquinones are glycosylated as O-glucosylxylosyl and then stored in the vacuole (Hagendoorn *et al.*, 1994).

Secondary metabolite of anthraquinones is also occurred in bacteria, fungi and lichens (Borroto *et al.*, 2008). It is dependent to the primary metabolism because most precursors and

energy necessary are derived from it such as the production of various enzymes (Hagendoorn *et al.*, 1994).

2.2.2 Applications of Anthraquinones

In Cuba, Anthraquinones product from *Morinda royoc* is important as a diet supplement. It is also propagated via *in vitro* culture because of the sporadic seed germination and slow growth development (Borroto *et al.*, 2008). Therefore, in order to produce secondary metabolites in high scale, the cultures were be maintained in large-scale bioreactors designed with controlled and physical factor such as aeration, temperature, gaseous composition, pH and light (Ahmed *et al.*, 2008).

2.3 Expression Vector

2.3.1. Background of Expression Vector

According to Russell (2006), expression vector is defined as a cloning vector containing the regulatory sequences that is necessary that allow the transcription and translation of a cloned gene or genes. It is used to produce the targeted protein that is encoded by a cloned gene in the transformed host such as *E. coli*.

Proteins such as antibiotics generally have been used for several decades in order to identify, localize and analyze the protein. Hence, high production of protein needed for further application use and analysis. However, there is a drawback where sometimes the protein yield production is low and impurity that is often occurs due to several reasons. It depends on protein

stability, solubility, size, post-translational and modifications (Glover and Hames, 1995). Therefore, since the production of protein is limited, it is possible for development of wide variety of vectors in both prokaryotic and eukaryotic host system for the efficient transcription of cloned DNA and the expression of targeted protein (Glover and Hames, 1995).

For immunological purposes, it is available for designing expression vector due to immunogenicity and antigenicity of recombinant polypeptide that is less dependent in the host and expression system (Glover and Hames, 1995).

2.3.2. Prokaryotic Expression Vector

Expression vectors are designed to transcribe the cloned gene and translate it into protein. They have some features such as multiple cloning sites and the selectable marker. Certain expression vectors have a lower copy number within cell and rarely have a screenable marker (Carson and Robertson, 2006). Besides that, other modifications of expression vector include the specific promoter to allow transcription of the cloned gene, ribosome binding site and ATG start codon (Carson and Robertson, 2006; Russell, 2006). Expression vector such as pET, pGEX and pMAC vectors have their own design advantages and experiences of these expression strategies (Glover and Hames, 1995).

Most plasmid expression vector utilizes either the promoter of the *lac* or of the *trp* operons from *E.coli* or the β-lactamase promoter from pBR332 (Glover, 1984; Goeddel, 1990).

2.3.3. Previous Study of Using Expression Vector

Expression vector has been used widely in many previous researches. For example, pBI121 expression vector is chosen in cloning of *Fragaria ananassa* Duch (strawberry) *FaEtr2* gene and this expression vector was transformed into *Agrobacterium tumefeciens* LBA4404 for further analysis (Chunli *et al.*, 2009). Besides that, the pGTPIL expression vector is used for expression of *IL2* cDNA from the Tibet pig (Li *et al.*, 2005).

Another previous study reported that a novel expression plasmid, pTO-N has been constructed for the production of large quantities of foreign protein of bovine pancreatic phospholipase A_2 (Tiliang *et al.*, 1990). It consists of strong and tightly regulated T7 gene 10 promoter and the *ompA* Shine-Dalgarno (SD) sequence, *ompA* sequence and a cloning linear region (Tiliang *et al.*, 1990).

Escherichia coli are widely used in recombinant DNA-based technologies as vectors for overproduction of proteins with theoretical and practical values. Furthermore, the advanced knowledge concerning the genetics and physiology of *E.coli* has accounted for the preferential use of *E.coli* as a host for gene expression (Goeddel, 1990).

The advantage of *E.coli* is based expression systems include rapid generation of biomass due to high rates of cell growth and availability of low cost culture condition. Controlled gene expression may be difficult in high copy-number of plasmid and if the protein of interest is deleterious to the cell, undesired instability or cell death may result in poor yield (Goeddel, 1990).

2.3.4. pET Expression Vector

The pET expression vector is an expression plasmid via the use of T7 RNA polymerase system is said to be the prokaryotic system of choice for expressing the foreign genes (Glover and Hames, 1995). It is designed especially for cloning and high level expression of peptide sequence in protein. It utilizes the strong bacteriophage T7 RNA polymerase promoter for high level expression of cloned gene in *E. coli* (Glover and Hames, 1995).

The promoter for the *E.coli* tryptophan operon (*trp*) is used for the expression of heterolegous genes and at least 100 proteins have been successfully synthesized with this promoter (Goeddel, 1990).

The cloning cassettes of expression vector are different between prototype series. For instance in pET-3 (pET-3a, pET-3b, pET-3c); the pET-3a series contains a short leader peptide, unique *Nde*I, *Nhe*I and *BamH*I restriction enzyme site that can be utilized to insert the cloned sequence and finally will be transcribed and translated under T7 polymerase control.

Another generation of pET expression vectors is constructed with extensive cloning cassetts such as in pET-17b and pET-16b (Glover and Hames, 1995). While pET-23a-c vectors are available which contain C-terminal histidine repeats.

Another important factor is the T7 RNA polymerase protein does not exists in laboratory stocks of *E.coli* and therefore all DNA manipulation was performed in any RecA, K12 *E.coli* strand for cloning such as HB101 and JM109 (Glover and Hames, 1995).

2.3.5 pUC Expression Vector

The pUC plasmid is also expression vector. This is because the lac operon is active when isopropyl-β -D-thiogalactopyranoside, IPTG is supplied in the growth media (Dominic, 2006).

The cloning site of pUC19 are clustered in *lacZ* region, thus the expressed product is a fusion protein which carrying a short segment of β-galactosidase enzyme (Williams *et al.*, 1993; Dominic, 2006).

3.0 MATERIALS AND METHODS

3.1 Preparation of Media

3.1.1 Preparation of Luria Broth Media

The Luria Broth media was prepared where the materials using 10 g of Tryptone, 5 g of yeast extract, 5 g of NaCl and made up to 1 litre with distilled water. The media was autoclaved at 121°C and stored at 4°C until used.

3.1.2 Preparation of Luria Broth Agar

About 200 mL of Luria Broth and 3 g of agar bacteriological were added in a 500 ml volume bottle. The media was autoclaved and stored at 4°C.

3.2 Preparation of Overnight Bacteria Culture

An overnight bacteria culture was prepared by using a single colony of *Escherichia coli* XL1 Blue.

Firstly, 5 ml of Luria Broth media was added into bijou bottle. Then, a single colony of *Escherichia coli* XL1 Blue was inoculated from the agar plate by using autoclaved tips. After that, the tip was added into bijou bottle containing LB media. The culture was incubated in a shaker incubator for overnight at 37 C.

3.3 Preparation of Competent Cell

A bacterial culture was inoculated and culture was verified by using spectrophotometer, where the absorbance; OD_{600} was obtained 0.567. The cells were cooled on ice within 20 minutes. After that, the cells were centrifuged at 3500 rpm, 4°C for 5 minutes. The supernatant was discarded. The pellet was resuspended in 100 mM $CaCl_2$ The pure glycerol (20% v/v) was added to the cell suspension and mixed them gently. Approximately 200 µl of the suspension was aliquoted into 1.5 eppendorf tube. Finally, the tubes were snap-frozen in liquid nitrogen and stored in -80°C fridge.

3.4 Transformation of Inserted pGEM-T Plasmid into Competent Cells

The pGEM-T plasmid vector containing *isochorismate synthase* cDNA was transformed into *Escherichia coli* XL1 Blue competent cells.

Firstly, about 1 µl of pGEMT-T plasmid containing *ICS* was aliquoted into pre-chilled 1.5 ml microcentrifuge tube. Next, about 50 µl of competent cells were added and mixed by flicking gently. The mixture was cooled on ice for 20 minutes. Then, the mixture was then heat shocked at 42°C in water bath for 45 seconds. The tube was immediately cooled on ice for 2 minutes. After that, about 1 ml of Luria Broth was added into the mixture and flicked gently. Finally, the mixture was incubated at 37°C incubator shaker for 90 minutes.

The next step, the transformed culture was spread on agar plate containing ampicillin. About 200 µl of ampicillin (50 ng ml) was added into 200 ml of melted luria agar. The bottle was swirled to mix the media well. Then, 20 ml of the media was poured into petri dish. After that, approximately 15 – 20 µl of the transformed culture was spread on agar plate with hockey

stick and sealed the plates with parafilm. Finally, the plates were incubated for overnight at 37°C.

3.5 Plasmid cDNA Isolation

3.5.1 Selection of Transformed Culture

The plates of transformed culture were observed to be selected. A single colony with plasmid was selected and inoculated into a bijou bottle containing 10 ml of LB broth for subculturing method. The colony was incubated with shaking 150 rpm for overnight at 37°C.

3.5.2 Isolation of Plasmid Using TENS Solution Buffer

Firstly, about 9 ml of overnight transformed cultures was transferred into a sterile falcon tube. Next, the cultures were centrifuged at 6000 rpm, 4°C for 5 minutes. Then, the supernatant was discarded and 300 µl of TENS buffer was added and vortexed gently.

After that, 150 µl of 3.0 M sodium acetate (pH 5.2) was added. The mix was vortexed and spin for 6 minutes at 13,000 rpm. Then, the supernatant was transferred into a fresh tube containing 0.9 ml of 100% EtoH precooled to -20°C. Again, the solution was centrifuge for 4 minutes at 13,000 rpm. The supernatant was discarded and the pellet was rinsed a time with 1 ml of 70% EtoH and was centrifuge again for 2 minutes. The remaining supernatant was then discarded and the pellet was air-dried for 15 minutes. Finally, the pellet was resuspended and dissolved in 50 µl of TE buffer and stored at -20 °C.

3.5.3 Isolation of Plasmid via GF-1 Plasmid DNA Extraction Kit Vivantis

Another method for isolation of inserted pGEM-T plasmid was by using GF-1 Plasmid DNA Extraction Kit from Vivantis. Firstly, the falcon tube containing an overnight transformed culture was centrifuged at 6000 rpm for 5 minutes and the supernatant was discarded. The pellet was resuspended completely in 250 µl of S1 solution by vortexing and pipetting. Next, 250 µl of S2 solution was added followed by addition of 400 µl of buffer NB and mixed gently. Then, the suspension was centrifuged at 14,000 rpm for 10 minutes. The supernatant was then transferred to a column and again centrifuged at 10,000 rpm for a minute and was discarded flow through. After that, about 700 µl of wash buffer was added and re-centrifuged twice at 10,000 rpm for a minute to remove flow through. Finally, a column was transferred to a new microcentrifuge and was added with 50 µl of elution buffer.

3.6 Analysis of Isolated Plasmid by Agarose Gel Electrophoresis

Analysis of isolated plasmid was done by running 1% of agarose gel electrophoresis. About 0.3 g of agarose was mixed well to 30 ml of TAE buffer. Then, the agarose was melted in an oven and let cooled for several minutes. After that, 3 µl of Ethr was added and mixed well. The gel was poured into the gel cast and put the gel comb until hardened. The gel cast was placed properly and the sample was prepared by mixing about 1 µl of plasmid with 1 µl of 6X loading dye. The plasmid band size was estimated by using GeneRuler^{TM} 1 kb ladder from Fermentas. The sample was run at 100 volt for 20 minutes. Finally, the expected isolating plasmid band was visualized on UV transluminator.

3.7 Plasmid cDNA Verification Analysis

3.7.1 Plasmid cDNA Verification by Direct Polymerase Chain Reaction (PCR)

Verification of plasmid was conducted by performing direct PCR on *ICS* gene with ICSha 1F, ICSha 2F and ICSha 3R internal fragment primer and also SP6/T7 primer. The PCR reaction was performed after reaction mixtures have been prepared. The PCR products were then verified through AGE analysis where the expected *ICS* gene sizes were identified by using 100bp and 1 kb ladder.

The PCR reaction mixtures (including Master Mix preparation) and also the PCR reaction profile are shown in tables below:

Table 1. The volume of 1X PCR mixture and 7X of PCR master mix

Reagents	PCR Reaction 1 ICSha 1F and ICSha 3R Volume (µl)	PCR Reaction 2 ICSha 2F and ICSha 3R Volume (µl)	PCR Reaction 3 SP6 and T7 Volume (µl)	7X Master Mix Volume (µl)
10 X Buffer A (PCR)	2	2	2	14
Distilled Water	11.1	11.1	11.1	77.7
2mM dNTP	2	2	2	14
25mM MgCl	2	2	2	14
Primer	ICSha 1F: 1 ICSha 3R: 1	ICSha 2F: 1 ICSha 3R: 1	SP6: 1 T7: 1	- -
Taq Polymerase	0.4	0.4	0.4	2.8
DNA Template	0.5	0.5	0.5	-
Total Volume	20	20	20	122.5

Table 2. The PCR reaction profile for all primers

Step	Temperature(°C)	Time (Sec)
Initial Denaturation	94	180
Denaturation	94	30
Annealing	59	30
Elongation	72	45
Final extension	72	60

Note that denaturation, annealing and elongation step were repeated for 35 cycles. The direct PCR reaction was performed in Mastercycler Personal PCR.

3.7.2 Plasmid cDNA Verification by Restriction Enzyme Analysis

After plasmid isolation was done, the next step was to verify the *ICS* cDNA insert via restriction enzyme analysis method. Several restriction enzymes were selected based on pGEM-T vector map derived from manufacturer. Then, further digestion was conducted to isolate the fragment.

3.7.2.1 Double Digestion of Restriction Enzyme Analysis

In this reaction, the inserted pGEM-T plasmid was restricted with *Sac*I and *Nco*I restriction enzymes, with addition of 1X Tango buffer (Fermentas) and diluted with distilled water, dH_2O. The mixture was then incubated for overnight at 37°C and thermal inactivation was done at 65°C for 2 minutes to stop enzyme activity. The double digestion reaction mixture was set up as shown in Table 3. Finally, the digested plasmid was verified by AGE analysis by using GeneRuler^{TM} 1 kb ladder to estimate the fragment size.