FUNCTIONAL CHARACTERIZATION OF THE PROMOTER REGION OF THE ZEBRAFISH ELOVL FAMILY MEMBER 5

by

CHEW YEN SHAN

Thesis submitted in fulfillment of the requirements for the degree of Master of Science

October 2015

ACKNOWLEDGEMENT

First and foremost, I wish to thank my supervisor, Professor Dr. Alexander Chong Shu Chien for his constant guidance, valuable advices, encouragement and support in solving problems arose when my master project is in progress. His wisdom, vast knowledge and commitment to the highest standards inspired and motivated me.

Besides, I would also like to thank my senior Meng Kiat and Shu Shen for their constant guidance and giving me useful ideas and suggestions when I faced difficulties in my project. Also, I would like to send my sincere gratitude to all the seniors from Lab 218 especially Ann, Boon Kai, Wan Yin, Michelle who had lent their hand whenever I need help. I am also grateful to all the staffs in School of Biological Sciences for their help and supports. In addition, I would like to express my gratitude to the Ministry of Higher Education for the financial assistance throughout the research.

Heartfelt acknowledgements are expressed to my friends and family. Without their moral support and encouragement, this thesis would not have been possible. When I felt down, their love always give me strength to face all the problems happened.

Thank you!

Chew Yen Shan

October 2015

TABLE OF CONTENTS

		Page
Acknowledgement		ii
Table of contents		iii
List of tables		viii
List of figures		ix
List of abbreviation		x
Abstrak		xiii
Abstract		xv
CHAPTER 1 - INTRODUCTION		
1.1	Introduction	1
1.2	Problem statement	3
1.3	Research objectives	3
CHAPTER 2 - LITERATURE REVIEW		
2.1	Polyunsaturated fatty acid (PUFA)	4
2.2	PUFA biosynthesis pathway	6
2.3	Elongases	8
2.4	Elongation of very long-chain fatty acid family member 5 (*Elovl5*)	9
2.5	Zebrafish (*Danio rerio*)	11
2.6	Transcription	15
2.7	Transcriptional regulation mechanism	19
2.8	The gene promoter	19
2.9	Transcription factors	20

CHAPTER 3 - MATERIALS AND METHOD

3.1 General Method 22

3.1.1 Sterilization 22

3.1.2 Optical density 22

3.1.3 Centrifugation 22

3.1.4 Storage of materials 22

3.2 Materials 23

3.3 Culture media and stock solutions

3.3.1 Media 25

3.3.2 Stock solutions 26

3.3.3 Antibiotic 26

3.3.4 0.1 M IPTG solution 26

3.3.5 Agarose gel preparation 26

3.3.6 Ethanol preparation 27

3.4 Host strain and vectors 27

3.5 Isolation of total RNA 27

3.5.1 DNase treatment of total RNA 28

3.6 Rapid amplification of cDNA ends (RACE) 29

3.6.1 Primer design 29

3.6.2 Generating RACE-ready cDNA 29

3.6.3 Amplification of cDNA ends (RACE) 30

3.6.4 Identification of potential transcription factors binding site 30

3.7 Cloning of zebrafish *elovl5* upstream 2.0 kb promoter 31

3.7.1 CTAB DNA Extraction 31

3.7.2 Agarose gel electrophoresis of DNA 31

3.7.3 Primer design 32

3.7.4 Amplification of promoter region using PCR 33

3.7.5 Gel purification of PCR products 34

3.7.6 Ligation of PCR products to pGEM®-T Easy Vector 34

3.7.7 *E.coli* DH5 α competent cell preparation 34

3.7.8 *E.coli* DH5 α transformation 35

3.7.9 Screening of recombinant colonies 36

3.7.10 Extraction of recombinant plasmid 36

3.7.11 Restriction endonuclease digestion of recombinant plasmid 37

3.7.12 Ligation of purified DNA fragments to pGL3-Basic Vector 37

3.7.13 Sequencing of recombinant plasmids 38

3.8 Cell line 38

3.8.1 Maintenance of cells in culture 38

3.8.2 Subculturing of cells 38

3.8.3 Cell storage and recovery 39

3.9 Transient transfection 39

3.9.1 Lipofectamine® 2000 DNA transfection 39

3.9.2 Preparation of cell lysates 40

3.9.3 Dual-glo luciferase reporter assay 40

3.10 Site-directed mutagenesis 41

3.11 Nuclear extraction from cells 42

3.11.1 Cell culture preparation 42

3.12 Biotin 3' end DNA labeling 43

3.13 Annealing complementary pairs of oligonucleotides 44

3.14 Electrophoretic Mobility Shift Assay (EMSA) 44

3.14.1 Binding reactions 44

3.14.2 Electrophoresis of DNA-protein complexes 45

3.14.3 Electrophoretic transfer of binding reactions to nylon membrane 45

3.14.4 Crosslink transferred DNA to membrane 45

3.14.5 Detection of biotin-labeled DNA by chemiluminescence 45

3.15 Competitive EMSA 46

3.16 Zebrafish regular care and maintenance 46

3.17 Zebrafish breeding 47

3.18 Microinjection of GFP plasmids into zebrafish embryos 47

CHAPTER FOUR – RESULTS AND DISCUSSION

4.1 Identification of transcription start sites of zebrafish *elovl5* promoter 49

4.2 Identification of putative *cis*-acting elements in zebrafish *elovl5* promoter 49

4.3 Zebrafish *elovl5* promoter deletion analysis in ZFL cell line 53

4.4 Mutagenesis of SREs in *elovl5* promoter 56

4.5 Binding of SREBP-1 SRE sites *in vitro* 59

4.6 Assessment of *elovl5* promoter-GFP expression in zebrafish embryos 61

CHAPTER FIVE - CONCLUSION AND FUTURE STUDIES

6.1 Conclusion 65

6.2 Future studies 66

REFERENCES 67

APPENDICES

Appendix 1 80

Appendix 2 81

Appendix 3 84

LIST OF TABLES

		Page
3.1	Materials used and their suppliers	23
3.2	Compositions of LB medium (per liter)	25
3.3	Compositions of SOC medium (per liter)	25
3.4	Compositions of LB agar (per liter)	25
3.5	Compositions of CTAB buffer (per liter)	26
3.6	Compositions 10X TBE buffer (per liter)	26
3.7	Genotype of *E.coli* strain used in this study	27
3.8	The nucleotide sequences used in 5' RACE	29
3.9	Forward and reverse primers used in promoter-reporter constructs and the expected size of the amplicon	33
3.10	PCR parameter of promoter fragments amplification	33
3.11	The sequence of potential SREBP-1 binding site used in both mutagenesis and DNA- Protein binding reaction	42
3.12	Forward and reverse primers used in promoter-GFP reporter constructs	48
4.1	Percentage of GFP expression of 72 hpf embryos injected with 884 bp of *elovl5* promoter fragment	63

LIST OF FIGURES

		Page
2.1	The chemical structures of Arachidonic acid (ARA), Eicosapentaenoic acid (EPA) and Docosahexaenoic acid (DHA)	5
2.2	The potential PUFA biosynthesis pathway in fish using linoleic and α-linolenic acid as precursors through the n-3 and n-6 pathway	7
2.3	Zebrafish (*Danio rerio*)	14
2.4	An overview of the transcription and translation in eukaryotes	17
2.5	Transcription initiation	18
4.1A	Schematic representation of zebrafish *elovl5* 5' flanking region and the location of the two putative transcription start site	52
4.1B	Comparison of *elovl5* promoter sequence between GenBank database and 2.8 kb cloned sequence	52
4.2	Structure and deletion analysis of zebrafish *elovl5* gene promoter	55
4.3	Effect of mutation on zebrafish *elovl5* gene promoter	60
4.4	Electrophorectic mobility shift assay (EMSA) of potential *elovl5* SRE-1 site using ZFL cell line as nuclear extract	58
4.5	The 884 bp zebrafish *elovl5* promoter directs tissue-specific expression of GFP in zebrafish embryos	64

LIST OF ABBREVIATIONS

ARA	arachidonic acid
ATCC	American Type Culture Collection
bHLH-ZIP	basic helix-loop-helix-leucine zipper
bp	base pair
$CaCl_2$	calcium chloride
cDNA	complementary DNA
CoA	coactivators
CTAB	cetyltrimetyl ammonium bromide
dH_2O	distilled water
DHA	docosahexaenoic acid
DMEM	Dulbecco's minimum essential medium
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside triphosphate
EDTA	ethylene diaminetetraacetic acid
Elovl	elongation of very long-chain
EtBr	ethidium bromide
EPA	eicosapentaeoic acid
Fads	fatty acid desaturase
FBS	fetal bovine serum
et al.	and others
g	gram
GTF	General Transcription Factors
HCl	Hydrochloric acid
hpf	hour post-fertilization

IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase pair
L	liter
LC-PUFA	Long-chain polyunsaturated fatty acid
LB	Luria-Bertani
LXR	liver X receptor
mA	milliamps
ml	milliliter
mM	milimolar
$MgCl_2$	magnesium chloride
mRNA	messenger RNA
NaCl	Sodium chloride
NF-Y	Nuclear factor Y
OD	optical density
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PLB	passive lysis buffer
pmol	picomole
PTU	phenylthiourea
PUFA	polyunsaturated fatty acid
RE	restriction site
RNA	ribonucleic acid
rpm	rotation per minute
S/MUFA	saturated/monounsaturated fatty acid
SREBP	Sterol regulartory element binding protein

Sp1	Specificity Protein 1
TBE	Tris/Borate/EDTA
TSS	transcription start site
U	unit
UTR	untranslated region
UV	ultraviolet
V	volt
v/v	volume/volume percentage
w/v	weight/volume percentage
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
µg	microgram
µl	microliter
°C	degree Celsius

PENCIRIAN FUNGSI

KAWASAN PROMOTER ELOVL AHLI FAMILI 5 IKAN ZEBRA

ABSTRAK

Pemanjangan rantaian asid lemak yang sangat panjang protein 5 (*elovl5*) adalah enzim yang penting dalam memangkinkan tindak balas kondensasi semasa langkah pemanjangan dalam biosintesis PUFA. Bersama-sama dengan asid lemak desaturase, ia menghasilkan asid lemak yang penting seperti asid eicosapentaenoic (EPA, 20:5n-3), asid docosahexaenoic (DHA, 22:6n-3) dan asid arakidonik (ARA, 20:4n-6) yang banyak terdapat dalam minyak ikan. Sebelum ini, fungsi *elovl5* manusia dan tikus telah dikaji dan ini mencadangkan ia terlibat dalam proses metabolisme yang sangat penting. Kawalan utama pengekspresan gen *elovl5* berlaku terutamanya pada peringkat transkripsi, oleh itu, pencirian promoter *elovl5* adalah penting untuk penghasilan PUFA. Dalam kajian ini, promoter *elovl5* dengan anggaran saiz 2.8 kb telah berjaya diklonkan. Dengan menggunakan pendekatan 5'-RACE. dua TSS telah dikenalpasti dalam gen *elovl5*. Analisis bioinformatik dalam 2.8 kb rantau promoter mendedahkan kehadiran tapak pengikatan untuk faktor transkripsi seperti SREBP, Sp1 dan NF-Y yang penting dalam pengaturan gen yang berkaitan dengan asid lemak. Analisis pemansuhan promoter hujung 5' telah menunjukkan bahawa 884 pasangan bes serpihan promoter mengandungi tapak pengikatan SRE berkebolehan untuk menjalankan aktiviti transkripsi yang asas manakala mutagenesis pada tapak pengikatan SRE menunjukkan bahawa tapak pengikatan SRE amat penting untuk pengaktifan transkripsi dalam gen *elovl5*. Pengasaian perubahan anjakan elektroporetik menunjukkan bahawa SRE proksimal

yang dikenalpasti pada rantau promoter adalah tapak utama bagi pengikatan SREBP-1. Akhirnya, kajian transgenesis secara sementara telah dijalankan dan GFP telah diekspreskan dalam lapisan sinsitium kuning telur embrio ikan zebra, membuktikan bahawa promoter yang diklonkan adalah promoter yang berfungsi. Kesimpulannya, 884 pasangan bes serpihan promoter *elovl5* ikan zebra berkebolehan untuk menerajui ungkapan asas dan tapak pengikatan proksimal SRE adalah penting dalam mengawal aktiviti transkripsi dalam *elovl5* ikan zebra.

FUNCTIONAL CHARACTERIZATION OF THE PROMOTER REGION OF THE ZEBRAFISH ELOVL FAMILY MEMBER 5

ABSTRACT

Elongation of very long-chain fatty acids protein 5 (*elovl5*) is an enzyme that is crucial in performing the condensation reaction in elongation step in the PUFA biosynthesis pathway. Together with fatty acid desaturases, it produces physiologically important fatty acids such as eicosapentaenoic acid (EPA, 20:5n-3), docosahexaenoic acid (DHA, 22:6n-3) and arachidonic acid (ARA, 20:4n-6) that are found abundant in fish oil. The function of *elovl5* has been studied previously in human and mouse, suggesting that it is involved in distinct metabolic processes. The control of *elovl5* gene expression mainly occurs at transcription level, therefore, characterization of *elovl5* promoter is vital for the PUFA production. In this study, *elovl5* promoter with a size of 2.8 kb was successfully cloned. By using 5'-RACE approach, two putative transcription start sites were identified in *elovl5* gene. Bioinformatic analysis of the 2.8 kb promoter region revealed the presence of SREBP, Sp1 and NF-Y cis-elements which are important in regulating fatty acids related genes. The 5'end promoter deletion analysis demonstrated that 884 bp promoter fragment contained the putative SRE binding sites were able to drive basal transcriptional activity while mutagenesis on the binding sites further showed that the putative SRE binding sites were important in modulating high level of basal transcriptional activity of *elovl5* promoter. Electrophoretic mobility shift assay (EMSA) indicates that proximal SRE identified in the promoter region is the SREBP-1 primary binding site. Lastly, transient transgenesis was carried out and

green fluorescent protein (GFP) was expressed in yolk syncytial layer of the microinjected zebrafish embryos, showing that the cloned promoter is a functional promoter. In conclusion, 884 bp of zebrafish *elovl5* promoter fragment is able to drive basal expression and proximal SRE binding site is crucial in regulating zebrafish *elovl5* transcriptional activity.

CHAPTER ONE

INTRODUCTION

1.1 Introduction

The long-chain polyunsaturated fatty acids (LC-PUFA), particularly eicosapentaenoic acid (EPA, 20:5n-3), docosahexaenoic acid (DHA, 22:6n-3) and arachidonic acid (ARA, 20:4n-6), which are found abundant in fish oil (Sargent *et al*., 2002) are essential in maintaining normal cellular function and affecting membrane fluidity and activities of membrane proteins (Wallis *et al*., 2002; Stillwell and Wassall, 2003). Besides, PUFA are also known to serve as ligands for nuclear receptors and transcription factors in the nucleus and thus these activated receptors will regulate the expression of the genes (Kliewer *et al*., 1997)

Desaturases and elongases are enzymes which are essential in catalyzing a series of desaturation and elongation reactions in the PUFA biosynthesis pathway respectively. Fatty acid desaturases (*Fads*) catalyze the introduction of double bonds whereby the position of the double bonds is dependent on the type of the desaturases. On the other hand, members of the elongation of very long-chain fatty acids (*Elovl*) gene family encode the fatty acid elongases that are crucial in catalyzing the addition of two carbon atoms on the fatty acyl-CoA intermediates in a four-step reaction-cycle of the fatty acid chain elongation system. Identification and characterization of the mammal's fatty acids elongases, ELOVL2 and ELOVL5, showed that they are important enzymes in PUFA biosynthesis pathway. Recent studies also suggested that *Elovl*5 mainly elongates carbon 18 and carbon 20 (C_{18} and C_{20}) PUFA (Inagaki *et al*., 2002; Moon *et al*., 2001) and it is highly expressed in testis, adrenal glands and livers in mammals (Leonard *et al*., 2000). Conversely, zebrafish *elovl5* has

broader substrate specification for C_{18}, C_{20}, and C_{22} PUFA as well as monosaturated fatty acids (Agaba *et al*., 2004)

The zebrafish, *Danio rerio*, has become an excellent model used to study gene expression. It is a tropical freshwater fish, which is native to the streams of the southeastern Himalaya region (Talwar and Jhingram, 1991) and usually inhabits streams, canals and ponds. Zebrafish is able to produce a large number of offspring in each clutch of eggs in every 4 to 5 days. Besides, the embryos are large, robust, transparent and have rapid development during embryonic stage. Furthermore, external fertilization allows scientist to manipulate and observe embryos under a dissecting microscope easily (Kimmel *et al*., 1995). Last but not least, its genome has been fully sequenced and available since 2009. Due to these attractive characteristics, it is gradually being recognized as a model organism for studies of vertebrate development and gene function (Mayden *et al*., 2007).

Transcription is a process where DNA is used as a template to create a complementary RNA, with the activity of RNA polymerase, which will be subsequently translated into protein. The expression of a particular gene is regulated by a number of factors such as the transcription factors, nuclear receptors, enhancers, activators and so on which bind to the *cis*-acting elements which are located in the promoter region. These factors will therefore activate or repress the transcription of the particular gene.

1.2 Problem statement

Elongases play a crucial role in PUFA biosynthesis pathway which may be optimized by understanding of the molecular basis of PUFA biosynthesis pathway. However, to date, very little is know about the regulation of *elovl5* at transcriptional level as no study has been performed in zebrafish *elovl5* promoter region since year 2004 when the zebrafish *elovl5* gene has been first isolated (Agaba *et al*., 2004). Without the understanding of its transcriptional regulation, it will remains as a huge challenge to modulate the *elovl5* expression in order to improve LC-PUFA pathway. Characterization of the promoter region will thus reveal the crucial regulatory elements that responsible in activating the expression of zebrafish *elovl5* gene. Zebrafish, as a popular genetic studies tool is able to recapitulate gene expression programs and realtime expression can be easily visualized and tracked. Hence, study and manipulation on zebrafish *elovl5* gene will provide better understanding of the promoter regulation of *elovl5* gene. Therefore, the aims of this study are to characterize zebrafish *elovl5* promoter region and determine its functional activity *in vivo*.

1.3 Research objectives

1. To clone and sequence the promoter region of the zebrafish *elovl5* gene.
2. To determine the functional activity of the promoter fragments *in vitro* and *in vivo*.
3. To identify the essential *cis*-acting elements which are involved in the regulation of zebrafish *elovl5* gene.

CHAPTER TWO

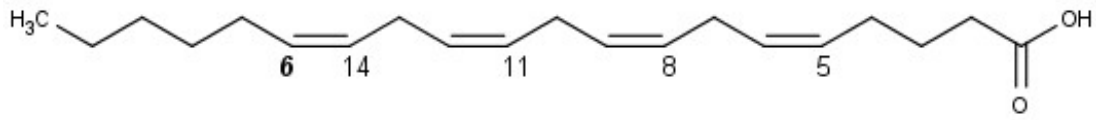
LITERATURE REVIEW

2.1 Polyunsaturated fatty acids (PUFA)

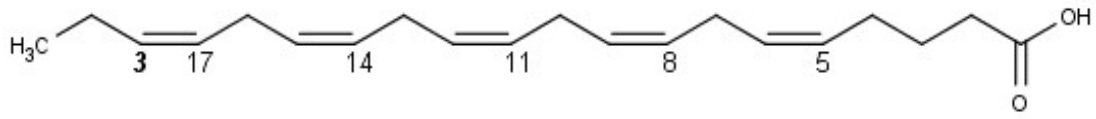
Fatty acids are carboxylic acids that have a carbonyl group at the end of the aliphatic chain and a methyl group on the opposite end. Fatty acids can be designated as X: Yn-Z, where X represents the number of carbon atoms, Y represents the number of double bonds and Z represents the position of the first double bond counting from the methyl end. Fatty acids are either saturated (without double bond in the chain) or unsaturated (with double bonds) (Vance and Vance, 1985). Polyunsaturated fatty acids (PUFA) are fatty acids with aliphatic tails longer than 12 carbons that have more than one double bonds between carbon atoms (See Figure 2.1)

Eicosapentaenoic acid (EPA; 20:5n-3), arachidonic acid (ARA; 20:4n-6) and docosahexaenoic acid (DHA; 22:6n-3) are LC-PUFA that are important constituents of membrane phospholipids, which determine the membrane fluidity, activities of membrane proteins, various enzymes reactions and signal transduction (McMurchie, 1988). This is especially critical in neurotransmission and photoreception reactions as DHA-rich phospholipids provide a unique degree of fluidity and compressibility of cell membranes which allow rapid conformational changes of the neurons (Salem *et al*., 2001).

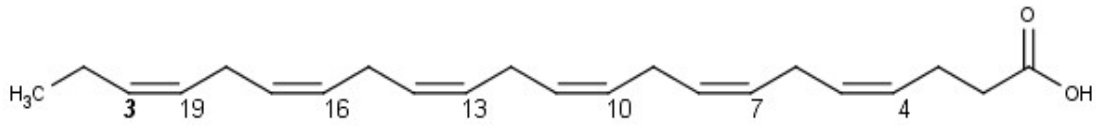
Studies also showed that there are several beneficial effects on inflammatory and pathological conditions, such as cardiovascular and neurological diseases by increasing daily consumption of EPA and DHA (Brouwer *et al*., 2006; Eilander *et al*., 2007; Ruxton *et al*., 2007).



Arachidonic acid (20:4n-6)



Eicosapentaenoic acid (20:5n-3)



Docosahexaenoic acid (22:6n-3)

Figure 2.1: The chemical structures of Arachidonic acid (ARA), Eicosapentaenoic acid (EPA) and Docosahexaenoic acid (DHA).

2.2 PUFA biosynthesis pathway

In general, PUFA biosynthesis pathway requires sequential elongation and desaturation activities catalyzed by two groups of enzymes namely the elongation of very long-chain fatty acids (*Elovl*) and fatty acyl desaturases (*Fads*) (See Figure 2.2). They are responsible in converting the essential fatty acids 18:2n-6 (linoleic acid, LA) and 18:3n-3 (α-linolenic acid, ALA) to longer chain and higher degree of unsaturated fatty acids such as EPA, DHA and ARA (Sprecher *et al*., 2000; Nakamura *et al*., 2001).

Elongases are enzymes that catalyze the condensation of fatty acids with malonyl-coA in the elongation cycle. Previous studies have reported that there are few elongase family members involved in mammal's PUFA biosynthesis pathway which are generally specific in certain substrate (Jakobsson *et al*., 2006).

On the other hand, *Fads* are responsible in introducing a double bond in the fatty acyl chain and the position of the double bonds is dependent on the type of the desaturases. To date, a few Δ6 *fads* have been characterized in several fish species (Zheng *et al*., 2004, 2005a, 2009; Tocher *et al*., 2006) except for zebrafish and rabbitfish which possess bifunctional Δ6/Δ5 *fads* (Hasting *et al*., 2001; Li *et al*., 2010).

Animals generally consume ALA and LA synthesized by plants in their diet. The EPA and ARA are synthesized from ALA and LA respectively which first involves an insertion of a double bond at the Δ6 position by Δ6 desaturase, followed by a 2-carbon units elongation and further desaturation at the Δ5 position by Δ5 desaturase (Cook, 1996) whereas synthesizing DHA required further elongation and desaturation of EPA (Sprecher *et al*., 1995).

PUFA Biosynthesis Pathway

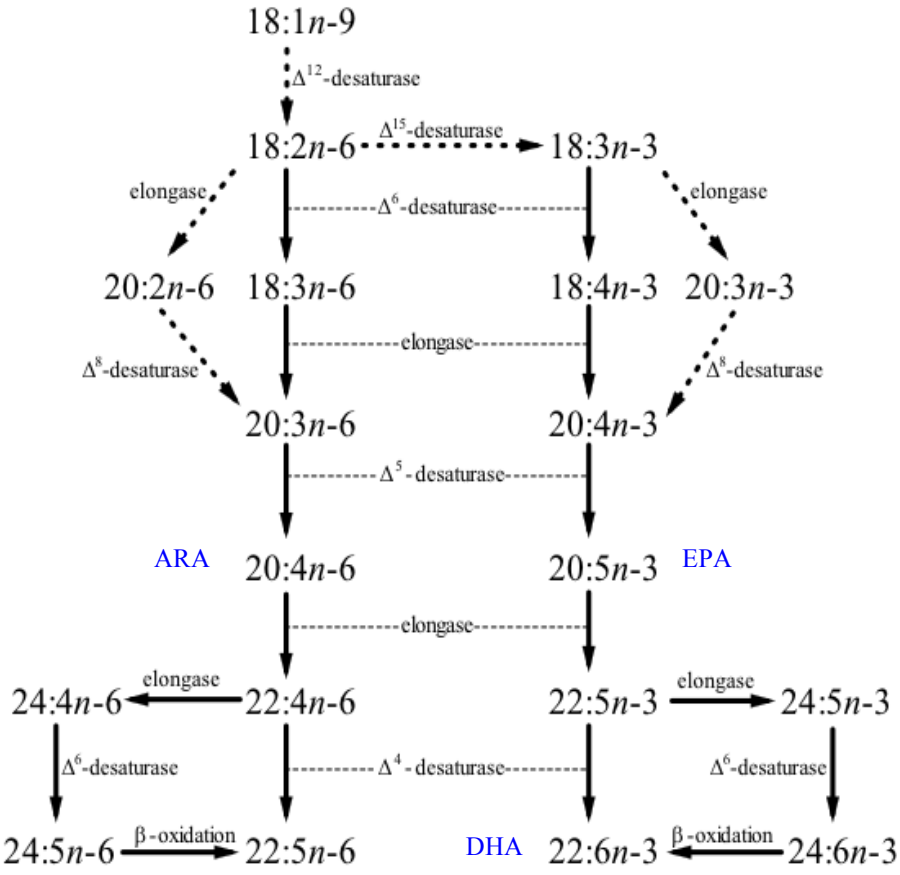


Figure 2.2: The potential PUFA biosynthesis pathway in fish using linoleic and α-linolenic acid as precursors through the n-3 and n-6 pathway. The PUFA are synthesized through series of desaturation and elongation process catalyzed by desaturases and elongases (Figure adapted from Monroig *et al*., 2011).

2.3 Elongases

The elongases catalyze the first, rate-limiting step of PUFA elongation pathway. In general, elongases can be divided into 2 subfamilies depending on their substrate preferences. Saturated or monounsaturated fatty acid (S/MUFA) elongases are responsible for elongating saturated and monounsaturated fatty acids while PUFA elongases are accounts for elongating PUFA (Meyer *et al*., 2004). To date, seven mammalian elongases have been characterized termed ELOVL 1-7 which comprises both enzymes that are ubiquitously expressed and tissue-specific enzymes. They are distinguished based on the similarity of motifs in their respective protein sequences (Jakobsson *et al*., 2006; Naganuma *et al*., 2011).

Elovl1, *Elovl3*, *Elovl6* and *Elovl7* are enzymes that responsible in elongating saturated and monounsaturated fatty acids. Previous studies showed that mutation in ELOVL3 and ELOVL6 cause abnormalities in sebaceous lipid composition, metabolic irregularities in brown adipose tissue and skin defects (Westerberg *et al*., 2004, 2006) while ELOVL7 is found to play a role in prostate cancer growth (Tamura *et al*., 2009). Additionally, *Elovl1* is found to be highly expressed in highly myelinated parts of the central nervous system in mouse and *Elovl1*-derived fatty acid products served to maintain the integrity of the membrane (Tvrdik *et al*., 2000). Nevertheless, to date, these enzymes have not been functionally characterized in fish species.

Elovl2, *Elovl4* and *Elovl5* are subfamilies of PUFA elongases that show preference towards C_{18} - C_{20} PUFA (Leonard *et al*., 2000; Tvrdik *et al*., 2000; Wang *et al*., 2005; Cameron *et al*., 2007). Studies on mammalian *Elovl2* revealed that it has highest mRNA expression in testis and liver compared to brain, kidney and white adipose tissue (Tvrdik *et al*., 2000). Meanwhile, in zebrafish, *elovl2* is highly

detected in liver, followed by intestine and brain (Monroig *et al*., 2009). Previous studies also showed that human ELOVL2 is only active towards C_{20} and C_{22} PUFA while *elovl2* of zebrafish and Atlantic Salmon are able to elongate C_{18} and C_{20} PUFA, although it showed preference towards C_{22} PUFA. It is thus suggested that *elovl2* is the key component in synthesizing DHA (Monroig *et al*., 2009; Morais *et al*., 2009). But neither of them showed distinct activity towards saturated fatty acid and monounsaturated fatty acid (Leonard *et al*., 2002).

In humans, recent studies also showed ELOVL4 play an important role in synthesizing C_{28} and C_{30} saturated very long chain fatty acid (VLCFA) in skin and C_{28} to C_{38} polyunsaturated VLCFA in retina, a tissue with high content of DHA, and to a lesser extent in the brain and testis (Agbaga *et al*., 2008; Mandal *et al*., 2004). Similarly, recent studies on fish *elovl4* such as zebrafish and Atlantic salmon showed that *elovl4* is expressed in eye, brain and testis and functional characterization on *elovl4* suggests that it effectively convert EPA and ARA to polyenoic products up to C_{36} (Monroig *et al*., 2010; Carmona-Antonanzas *et al*., 2011). Besides, studies also showed that deletion in *Elovl4* gene in mouse result in highly impaired skin barrier function (Vasireddy *et al*., 2007).

2.4 Elongation of very long-chain fatty acids family member 5 (*Elovl5*)

The ability of *Elovl5* in elongating PUFA is important in maintaining DHA level. ELOVL5 has been found to be highly expressed in human especially in the testis and adrenal gland as these tissues was found to have high level of DHA. Besides, research has shown that *Elovl5* is important in mouse's liver development during postnatal stage (Wang *et al*., 2005).

To investigate the *elovl5* activities in biosynthesizing long chain PUFA (LC-PUFA) in fish species, zebrafish elongase was first identified and isolated and named as zfELO in 2004 (Agaba *et al*., 2004). However, due to its function and sequence similarity with human ELOVL 5 (Leonard *et al*., 2000) and rat rELO1 (Inagaki *et al*., 2002), it is later designated as *elovl5*. The work of Agaba and collegues indicated that *elovl5* of zebrafish has the ability to elongate C_{18}, C_{20}, and C_{22} PUFA, with high activity towards C_{18} PUFA and a lower extend towards C_{22} PUFA. Other than PUFA, zebrafish elongase was also capable to lengthen monounsaturated and saturated fatty acids (Agaba *et al*., 2004). This variations occurs can be exemplified by the zebrafish desaturase which also posses both Δ6 and Δ5 function which coded from a single gene (Hastings *et al*., 2001). However, in contrast to zebrafish *elovl5*, mammalian *Elovl5* was found to elongate C_{18} and C_{20} PUFA but was unable to elongate PUFA substrates beyond C_{22} (Inagaki *et al*., 2002; Moon *et al*., 2001; Parker-Barnes *et al*., 2000; Wang *et al*., 2005).

Later, *elovl5* cDNAs of several marine and freshwater species have been isolated, including atlantic salmon, cobia cod, gilthead sea bream, turbot, Nile tilapia, striped snakehead and African catfish where the amino acid sequences were highly conserved among fish species (Agaba *et al*, 2005; Morais *et al*., 2009; Zheng *et al*., 2009, Kuah *et al*., 2015). Studied indicated that the encoded elongases posses C_{18} to C_{22} PUFA elongation activities but no activity towards saturated fatty acids.

In zebrafish, *elovl5* was found highly expressed in liver and intestine which are organs that involved in biosynthesis of LC-PUFA. Previous studies on follicle maturation in zebrafish showed high expression of *elovl5* in pre-vitellogenic follicles, suggesting that *elovl5* plays a role in oocyte LC-PUFA biosynthesis (Ishak *et al*., 2008). Additionally, Monroig and colleague reported that *elovl5* was expressed

in zebrafish embryos yolk syncytial layer, which suggest it is important in remodeling of yolk fatty acids during zebrafish early embryogenesis (Monroig *et al*., 2009). Furthermore, recent studies of *elovl5* during embryo development in zebrafish found that *elovl5* showed correlation in the embryonic pronephros, with blood filtration and osmoregulation function in zebrafish (Tan *et al*., 2010).

2.5 Zebrafish (*Danio rerio*)

The zebrafish, *Danio rerio*, is a tropical freshwater fish belonging to the minnow family. Scientific classification of zebrafish is as below:

Kingdom : Animalia

Phylum : Chordata

Class : Actinopterygii

Order : Cypriniformes

Family : Cyprinidae

Genus : *Danio*

Species : *D. rerio*

The fish is given the name zebrafish as it has five to seven uniform, pigmented, dark blue longitudinal stripes on the side of the body, which extended to the end of the caudal fin (Barman, 1991). Zebrafish is a relatively small fish, which is about 3cm in length and life span is around 2-3 years. Zebrafish is native to the streams of the southeastern Himalaya region (Talwar and Jhingram, 1991) and usually inhabits canals, streams, ponds and slow-moving to stagnant water bodies (Spence *et al*., 2008).

Males are characterized as torpedo-shaped and have gold stripes between the blue stripes. Females usually are larger in size, rounded body shape with whitish belly and have a small genital papilla in front of the anal fin origin (See Figure 2.3) (Laale, 1977; Wixon, 2000). Zebrafish are omnivorous, whereby they consume primarily on zooplankton and insects. The generation time for zebrafish takes about 3-4 months, where its growth rate is the most rapid for first three months and decreasing after that (Spence *et al*., 2007b). In the presence of male, females are able to spawn at intervals of 2-3 days, laying hundreds of eggs in each clutch. Fertilized eggs are transparent in nature, which makes zebrafish a useful model in biological research (Spence *et al*., 2008).

In the scientific world, zebrafish is gradually recognized as a model organism for studies of vertebrate development and gene function (Mayden *et al*., 2007) as it possesses a number of attributes which make it tractable to experimental manipulation. For instance, its ability to produce a large number of offspring in each clutch holds on advantage over other models as large-scale genetic approaches can be used in identifying novel genes and defining gene function (Pelegri, 2002). Besides, zebrafish are relatively easy to maintain compared to other fish stock (Brand *et al*., 2002). The zebrafish embryos are large, robust and transparent, which allows scientists to see the expression of the gene by simple staining techniques. Additionally, the rapid development of the transparent embryos, where the major organs are developed within 36 hours, can be monitored under a dissecting microscope (Kimmel *et al*., 1995).

In term of PUFA biosynthesis, it is demonstrated that there are elongation and desaturation activities in zebrafish oocyte thus allow scientists to study the role of PUFA in maturation and ovulation process. Recently, zebrafish has been used as a

model organism for nutritional genomics studies to be applied on aquaculture fishes (Ulloa *et al*., 2011). Since many of the gene functions of this model similar to that of human, zebrafish is thus extensively used as model for inflammation, blood clotting, heart and kidney diseases in cardiovascular research (Berghmans *et al*., 2005; Guyon *et al*., 2006).

Figure 2.3: Zebrafish (*Danio rerio*). The well-understood physiology and developmental behavior makes it intensively been used as a model organism (Figure adapted from http://www.socmucimm.org/introduction-zebrafish-danio-rerio/).

2.6 Transcription

The central dogma of molecular biology stated that the coded genetic information is transcribed into the messenger RNA (mRNA) during transcription, each transcript contains the information of a particular protein to be synthesized and eventually, the mature mRNA will be translated in the ribosome. As such, the information cannot flow back from protein to protein or nucleic acid (Crick, 1970). The DNA sequence that encodes for a protein not only contains the coding sequence but also the regulatory sequences which can regulate the synthesis of the particular protein. The regulatory sequence that is upstream of the coding sequence is known as 5' untranslated region (5' UTR) while the regulatory sequence that is downstream of the coding region, is known as 3' untranslated region (3' UTR) (See Figure 2.4)

During transcription, the gene sequence is read by RNA polymerase and produces the complementary, antiparallel RNA strand. The DNA is read from 3' to 5' during transcription while the complementary RNA is created from 5' to 3'. A core promoter sequence in the DNA is required to initiate the transcription process, where RNA polymerase is able to bind to it in the presence of various specific transcription factors. During the elongation process, RNA polymerase traverses the template strands to create an RNA copy using complementary base pairs with the DNA template. The transcription of mRNA is terminated through the polyadenylation process, a template-independent addition of as at the new 3' end (Lykke-Andersen and Jensen, 2007).

The initiation of transcription is an important control point in eukaryotic gene expression (Latchman, 1998). It is initiated in the core promoter (see Figure 2.5), the TATA box (consensus TATAA/TAA/T) which is found approximately 30 bp upstream of the transcription start site in most of the genes. The binding of RNA

polymerase II alone to the TATA box is incapable of initiating transcription, therefore, the assembly of the basal transcription factors such as TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH, and TFIIJ are needed to initiate the transcription.

Firstly, TFIID also known as TATA-binding protein is bound to the TATA box facilitated by TFIIA. TFIIB will then bind to the carboxyl terminal of TFIID and bind to the promoter downstream of TATA box. TFIIB will also bind to RNA polymerase II to recruit it to the complex in association with TFIIF. Subsequently, another factor which is TFIIH, is responsible to unwind double-stranded DNA. It will then phosphorylate the C-terminal domain of RNA Polymerase II (Weaver and Hendrick, 1997). The phosphorylated RNA polymerase II are now capable to elongate the mRNA. During the elongation process, TFIIF remain associated with DNA polymerase while TFIIA and TFIIB will remain bound to TATA box, allowing other polymerases to bind and initiate transcriptions.

The central roles of transcription in the process of gene expression provide a control point for regulating the expression of genes in only in certain cell types or response to particular signal (Darnell, 1982; Latchman, 2002). For instance, the gene encoding antibody molecules are transcribed at high level only in the antibody-producing B cells.

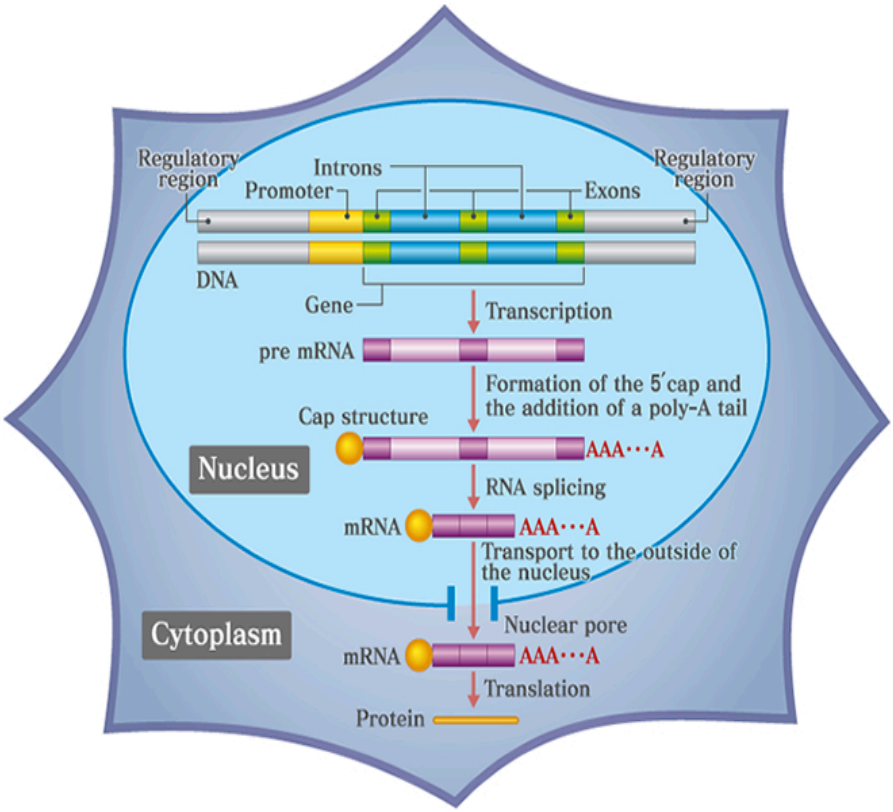


Figure 2.4: An overview of the transcription and translation in eukaryotes. There is a regulation mechanism at each level. Figure adapted from (http://csls-text.c.u-tokyo.ac.jp/)

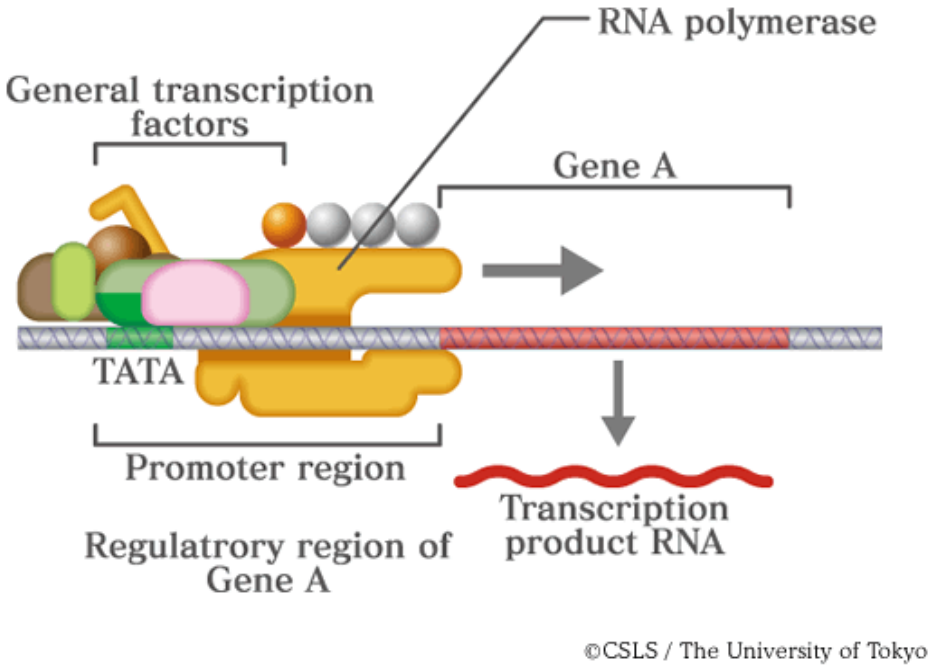


Figure 2.5: Transcription initiation. RNA polymerase is binding to a core promoter sequence with the aid of transcription factors. Figure adapted from (http://csls-text.c.u-tokyo.ac.jp/)

2.7 Transcriptional regulation mechanism

Cis-acting elements are particular DNA sequences that are involve in transcriptional regulation while *trans*-acting factor are proteins that bind to *cis*-elements to control gene expression. In some genes, there are transcriptional regulatory sequences known as enhancers and silencers which will cause changes in the gene expression when there is intra- or extracellular signaling (Ishiura, 2008).

The binding of the specific trans-acting factors to the enhancer sequence promotes the binding of RNA polymerase to the promoter region, hence encouraging the expression of the gene. On the other hand, binding of the specific proteins on silencers suppresses gene expression. Different from a promoter, enhancers and silencers may be located either upstream or downstream of the gene. However, promoters are only present in upstream of the gene, as they represent points where RNA polymerase bind to and allow the initiation of transcription. In addition to this, enhancers and silencers are still able to function even their sequences are reversed. Conversely, a promoter with reversed sequence cause the RNA polymerase proceeds in the opposite direction thus preventing the promoter from functioning (Ishiura, 2008).

2.8 The gene promoter

The gene promoter is a specific DNA sequence that initiates the transcription of a gene. In prokaryotes, these specific sequences are located at -10 and -25 positions upstream of the transcription start site. For eukaryotes, most of the gene promoters contain a TATA box (AT-rich sequence), which is usually found about 30 base pairs upstream of the transcriptional start site. This region has been defined as core promoter (Goodwin *et al*., 1990) as they play an important role in determining

the initiation point and able to produce basal levels of transcription. Those that lack of TATA box in the promoter will have lower activities. However, other elements that located upstream of the promoter are also important in regulation of gene expression. Sp1 box, a GC-rich sequence is found in a variety of genes located upstream of the TATA box. These specific sequences play a critical role in binding transcription factors that are involved in formation of transcriptional complex and were termed as upstream promoter elements (Goodwin *et al*., 1990).

2.9 Transcription factors

Transcription factors are proteins that bind to specific sequences, controlling the flow of genetic information from DNA to mRNA (Latchman, 1997). There are two different groups of transcription factors, which are activators and repressors that regulate the recruitment of RNA polymerase to the specific genes (Roeder, 1996). Transcription factors are known to have one or more DNA-binding domains (DBDs) that attach to the specific sequences of DNA adjacent to the genes they regulate (Mitchell and Tjian, 1989; Ptashne and Gann, 1997).

There are varieties of mechanisms that are used by the transcription factor to regulate the gene expression. The transcription factors will first stabilize the binding of RNA polymerase to DNA and either catalyze the acetylation process to make the DNA more accessible to transcription or deacetylation of histone proteins, hence down-regulating the transcription. Lastly, the transcription factors will recruit co-activator or corepressor to the transcription factor DNA complex. In eukaryotes, the presence of general transcription factors (GTFs) are important transcription factors which are necessary for transcription to occur.

The sterol regulatory element-binding protein (SREBP) is a member of the basic helix-loop-helix-leucine zipper (bHLH-ZIP) family transcription factors, which play a key role regulating the biosynthesis of cholesterol and fatty acids (Horton *et al*., 2002). It activates transcription by binding to sterol response elements (SREs) in the promoter region. In mammals, SREBP-1 and SREBP-2 are subfamilies that have been previously demonstrated to serve as a master regulator of *de novo* lipogenesis and cholesterol biosynthetic genes respectively (Horton *et al*., 2002). SREBP-1a and SREBP-1c are two major isoforms of SREBP-1 gene found in mammals whereby SREBP-1c is highly expressed in liver (Shimomura *et al*., 1977). However, no evidence showed that there is an alternative splicing of SREBP-1 gene in fish species (Minghetti *et al*., 2010).

CHAPTER THREE

MATERIALS AND METHODS

3.1 General method

3.1.1 Sterilization

All the apparatus and media used in this study were autoclaved under 121°C and ultra-high pressure for 15 minutes. For liquid media, the autoclaved time was extended to 20 minutes and the bottle cap loosened before autoclaving. Heat sensitive materials such as ampicillin, kanamycin and Isopropyl β-D-1-thiogalactopyranoside (IPTG) were filter-sterilized before use.

3.1.2 Optical density

The optical density of *Escherichia coli* (*E.coli*) strain DH5-α bacterial culture used for the transformation process should be between within the exponential phase of growth, indicated by optical density of 0.3g/L – 0.5g/L.

3.1.3 Centrifugation

Bacterial culture and mixtures which are not more than 1.5 ml were centrifuged by using Microcentrifuge 5415R (Eppendorf) at a speed 13,000 rpm (16,100 x g).

3.1.4 Storage of materials

All the materials and bacterial cultures were stored in the fridge at 4°C for short period of time. While for long-term storage, the materials and samples were stored in the freezer (Panasonic) at -20°C or -80°C (Thermo Scientific).

3.2 Materials

The materials used were purchased from the suppliers shown in Table 3.1.

Table 3.1: Materials used and their suppliers

Suppliers	Materials
American Type Culture Collection (ATCC)	Zebrafish liver cell line
Amresco	Agar-bacteriological
	Agarose
	Ampicillin
	EDTA
	Ethanol
	Isopropanol
	Kanamycin
	Magnesium Chloride ($MgCl_2$)
	Tryptone
	Yeast extract
BioLabs	*BamH*1 restriction enzyme
	*Kpn*I restriction enzyme
	*Xho*I restriction enzyme
Bioline	MyTaq™ Red DNA Polymerase
Clontech	SMARTer™ RACE cDNA Amplification Kit
Fisher Chemicals	Boric acid
	Chloroform
	Tris Base
Gibco	Opti-MEM® I
	Fetal Bovine Serum (FBS)
HyClone®	Trypsin EDTA
Invitrogen	Lipofectamine ® 2000
	Phosphate Buffer Saline (PBS) tablet
	SYBR ® Safe DNA Gel Stain
	DMEM
iNtRON Biotechnology	DNA-spin™ Plasmid DNA Purification Kit
	MEGAquick-spin™ Fragment DNA Purification Kit

	Muta-direct™Site Directed Mutagenesis Kit
Merck	Sodium Chloride (NaCl)
	Calcium Chloride ($CaCl_2$)
Molecular Research Center	TRI Reagent®
Promega	DNA Ladder (100 bp)
	DNA Ladder (1 kb)
	Dual Glo® Reporter Assay System
	Ethidium Bromide (EtBr)
	IPTG
	pcDNA3.1/Zeo
	pGEM®-T Easy Vector System
	pGL3-Basic Vector
	pGL3-Conrol Vector
	pRL-SV40 Vector
	pZsgreen1-1 Vector
	T4 DNA Ligase
	X-Gal
Sigma	Glycerol
	Magnesium Sulphate ($MgSO_4$)
	Phenol Red Solution
	Proteinase K
	N- Propylthiouracil (PTU)
Thermo Scientific	Chemiluminescent Nucleic Acid Detection Module
	NE-PER Nuclear and Cytoplasmic Extraction Reagents
	Biotin 3' End DNA Labeling Kit
	LightShift® Chemiluminescent EMSA Kit