CLONING AND PROMOTER ANALYSIS OF HUMAN CHOLINE KINASE ISOFORMS

YEE YOKE HIANG

UNIVERSITI SAINS MALAYSIA

2012

CLONING AND PROMOTER ANALYSIS OF HUMAN CHOLINE KINASE ISOFORMS

BY

YEE YOKE HIANG

Thesis submitted in fulfillment of the requirements

for the degree of

Master of Science

September 2012

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisors, Dr. See Too Wei Cun and Dr. Few Ling Ling for their trust in giving me a chance to complete my study in the Master of Science. They gave me scientific guidance and advice, assisted me in solving problems and encouraged my research progress. I am also very appreciative of their help throughout the preparation of this thesis.

I would also like to extend my sincere appreciation to Dr. Yvonne Tee Get Bee for her willingness to be my co-supervisor and for her invaluable help in this research study especially in the area of data analysis. In addition, I would also like to express my gratitude to Dr. Khoo Boon Yin for kindly providing me with the human breast adenocarcinoma, MCF-7 cells and human hepatoma, HepG2 cells.

I would like to record special thanks to the officers and technologists from the School of Health Sciences. They gave me assistance in the form of technical advice and guidance to my whole study.

I am very thankful to Universiti Sains Malaysia for its generous financial support in the form of the USM Fellowship and Research University Grant Scheme (1001/PPSK/813034) and in all aspects of my Master of Science research.

Finally, I would like extend my gratitude to all my laboratory members for their friendship that enlightened my laboratory life making it so special and memorable.

With their co-operation, moral support and the sharing of ideas, my MSc. study was completed smoothly.

Last but not least, I would like to appreciate my family and my husband for their continuous support and love.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ..... **ii**

TABLE OF CONTENTS ..... **iv**

LIST OF TABLES ..... **viii**

LIST OF FIGURES ..... **ix**

LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMNS ..... **xi**

ABSTRAK ..... **xv**

ABSTRACT ..... **xvii**

CHAPTER 1.0 INTRODUCTION **1**

1.1 Structures and functions of phospholipids 1

1.1.1 Phosphoglycerides 1

1.1.1(a) Phosphatidic acid 3

1.1.1(b) Phosphatidylcholine 5

1.1.1(c) Phosphatidylethanolamine 5

1.1.1(d) Phosphatidylserine 6

1.1.1(e) Phosphatidylglycerol 6

1.1.1(f) Diphosphatidylglycerol 7

1.1.1(g) Phosphatidylinositol 7

1.1.2 Sphingolipids 8

1.2 Biosynthesis of major phospholipids in eukaryotic membranes 10

1.2.1 Synthesis of phosphatidic acid 12

1.2.2 Synthesis of phosphatidylcholine 12

1.2.2(a) CDP-choline pathway 13

1.2.2(b) Phosphatidylethanolamine methylation pathway 14

1.2.3 Synthesis of phosphatidylethnolamine 14

1.2.4 Synthesis of phosphatidylserine 15

1.3 Choline kinase 16

1.3.1 Biochemical characteristic of choline kinase isoforms 17

1.4 Choline kinase in cell transformation and carcinogenesis 19

1.5 Promoter 22

1.5.1 Transcription factors 22

1.5.2 Transcriptional regulation of choline kinase 25

1.6 Statement of problem and rationale of study 31

1.7 Aims of the study 32

CHAPTER 2.0 MATERIALS AND METHODS 35
2.1 Materials 35
2.1.1 Chemicals and reagents 35
2.1.2 Enzymes 35
2.1.3 Kits 35
2.1.4 Consumables 35
2.1.5 General instruments 35
2.2 Methods 43
2.2.1 Sterilization of buffers and consumables 43
2.2.2 Maintenance of *Escherichia coli* cell 43
2.2.2(a) Preparation of bacterial culture medium 43
2.2.2(ai) LB (Luria-Bertani) medium 43
2.2.2(aii) LB agar with ampicilin plates 44
2.2.2(aiii) Medium A 44
2.2.2(aiv) Medium B 44
2.2.2(b) Preparation of *E. coli* glycerol stock 44
2.2.2(c) Preparation of XL1-Blue competent cells 45
2.2.3 Culture of human cancer cell lines 45
2.2.3(a) Preparation of medium and buffers 46
2.2.3(ai) Cell culture medium 46
2.2.3(aii) Heat inactivation of fetal bovine serum 46
2.2.3(aiii) Preparation of Phosphate buffered saline 46
2.2.3(aiv) Trypsin-EDTA 47
2.2.3(av) Cryopreservation medium 47
2.2.3(b) Cell passage of human cancer cell lines 47
2.2.3(c) Cryopreservation of human cancer cell lines 47
2.2.3(d) Revival of cryopreserved human cancer cell lines 48
2.2.3(e) Cell counting 48
2.2.4 Software for promoter analysis and other bioinformatics analysis 51
2.2.5 Polymerase chain reaction (PCR) 51
2.2.5(a) PCR primers 51
2.2.5(b) Oligonucleotides preparation 55
2.2.5(c) Human genomic DNA 55

2.2.5(d) DNA sequencing 55
2.2.6 Plasmid for cloning of *ckα* and *ckβ* promoters 55
2.2.7 Plasmid extraction 56
2.2.8 Agarose gel electrophoresis 60
2.2.8(a) Preparation of 1% agarose gel 61
2.2.8(b) Tris-acetate-EDTA buffer (10X) 61
2.2.8(c) 6X DNA loading dye 61
2.2.8(d) 2X RNA loading dye 62
2.2.8(e) Sample preparation for the electrophoresis 62
2.2.8(f) Ethidium bromide (EtBr) DNA staining solution 62
2.2.9 DNA gel extraction 63
2.2.10 Determination of DNA and RNA concentration 63
2.2.11 DNA digestion with restriction endonucleases 64
2.2.12 Purification of restriction enzyme digested DNA 64
2.2.13 DNA ligation 65
2.2.14 Heat-shock transformation of *E. coli* competent cells 65
2.2.15 Determination of transcriptional start site 66
2.2.15(a) Total RNA extraction 66
2.2.15(b) 5' RLM-RACE 67
2.2.16 Identification of transcription factor binding element 71
2.2.17 Transient transfection of MCF-7 cells 71
2.2.18 Dual-Glo luciferase assay 72
2.2.19 Site directed mutagenesis by PCR 75
2.2.20 Treatment of MCF-7 cells with Phorbol-12-Myristate-13-Acetate (PMA) 75
2.2.21 Statistical analysis 76

CHAPTER 3.0 RESULTS 77
3.1 Cloning of *ckα* and *ckβ* promoters 77
3.2 Determination of transcriptional start site for *ckα* and *ckβ* promoters 82
3.3 Identification of transcription factor binding elements of *ckα* and *ckβ* promoters 87
3.4 Construction of *ckα* promoter 91
3.5 Construction of *ckβ* promoter 96
3.6 Optimization of transient transfection 99

3.7 Identification of highly functional regulatory regions of *ckα* promoter 101
3.8 Identification of highly functional regulatory regions of *ckβ* promoter 106
3.9 Effect of ETS and GATA sites mutations on *ck*β promoter activity 109
3.10 Effect of PMA treatment on *ckβ* promoter 111

CHAPTER 4.0 DISCUSSION 113
4.1 Cloning of *ckα* and *ckβ* promoters 113
4.2 Transcriptional start site determination 114
4.3 Determination of transcription factor binding sites 117
4.4 Gene reporter assays 118
4.5 Transcription factors that are involved in the regulation of *ckα* and *ckβ* gene expression 122
4.6 ETS transcription factor negatively regulates the *ckα* gene expression..... 123
4.7 ETS and GATA transcription factors negatively regulates the *ckβ* gene expression 126

CHAPTER 5.0 CONCLUSION 131
REFERENCES 135
APPENDICES 161
LIST OF PUBLICATIONS 174

LIST OF TABLES

Table 1.1 Major types of phosphoglycerides ... 4

Table 1.2 Some major transcription factor families and functions ... 24

Table 2.1 List of chemicals and reagents ... 36

Table 2.2 List of enzymes ... 38

Table 2.3 List of kits ... 39

Table 2.4 List of consumables ... 40

Table 2.5 List of general instruments ... 41

Table 2.6 Software for promoter analysis and other bioinformatics analysis ... 52

Table 2.7 PCR primers. ... 53

Table 2.8 Plasmid constructs of *ckα* and *ckβ* promoters produced in this study. ... 59

LIST OF FIGURES

Figure 1.1	The structure of plasma membrane..	2
Figure 1.2	Core structure of sphingolipids, long chain bases form the core structure of sphingolipids	9
Figure 1.3	Pathways for the biosynthesis of major eukaryotic membrane phospholipids	11
Figure 1.4	The possible regulatory elements for mouse *ckα* and *ckβ* genes.	28
Figure 1.5	Schematic diagram of the putative HRE binding sites in human *ckα* promoter.	30
Figure 1.6	Flow chart of study	34
Figure 2.1	Schematic diagram of Haemocytometer and cells counting system	50
Figure 2.2	Map for the pGL4.10[luc2] vector..	57
Figure 2.3	Map for the pGL4.73[hRLuc/SV40] vector..	58
Figure 2.4	Schematic diagram of 5’ RLM-RACE procedure.	70
Figure 2.5	Schematic diagram of Dual-Glo reporter assay system	74
Figure 3.1	Promoter sequences of *ckα*..	78
Figure 3.2	Promoter sequences of *ckβ*..	79
Figure 3.3	PCR amplified *ckα* and *ckβ* promoters.	80
Figure 3.4	Restriction digestion analysis of *ckα* and *ckβ* promoters..	81
Figure 3.5	Total RNA extracted from the human liver hepatocellular carcinoma..	83
Figure 3.6	PCR products of TSS *ckα* and *ckβ* promoters..	85
Figure 3.7	The DNA sequence alignment result of TSS for *ckα* and *ckβ* promoters..	86
Figure 3.8	The potential transcription factor binding sites of *ckα* promoter..	89
Figure 3.9	The potential transcription factor binding sites of *ckβ* promoter..	90
Figure 3.10	Double restriction enzymes analysis of *ckα* deletion constructs.	93
Figure 3.11	PCR products of second and third series of *ckα* deletion constructs.	94
Figure 3.12	Double digestion analysis of second and third series of *ckα* deletion constructs.	95
Figure 3.13	PCR products of *ckβ* deletion constructs.	97
Figure 3.14	Double digestion analysis of *ckβ* deletion constructs.	98
Figure 3.15	Effect of varying concentrations of internal control (*Renilla* luciferase) vector on test reporter (firefly luciferase) activity..	100
Figure 3.16	Deletion constructs of *ckα* promoter.	102
Figure 3.17	Second series deletion constructs of *ckα* promoter.	103

Figure 3.18 Third series deletion constructs of *ckα* promoter. 105
Figure 3.19 Deletion constructs of *ckβ* promoter.. 107
Figure 3.20 Second series deletion constructs of *ckβ* promoter........................... 108
Figure 3.21 Luciferase activity of full length and mutant *ckβ* promoter constructs. 110
Figure 3.22 The effect of PMA treatment on full length and mutant *ckβ* promoter constructs. 112
Figure 4.1 The suggested regulation pathway of *ckβ* promoter by ETS and GATA transcription factors. 130

LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMNS

~	Approximately
%	Percentage
^{o}C	Degree Celcius
µg	Microgram
µl	Microliter
µM	Micromolar
AhR/Ar	Aryl hydrocarbon receptor/androgen receptor
AP-1	Activator protein 1
ATP	Adenosine-5'-triphosphate
bp	Base pair
BSA	Bovine serum albumin
CCl_4	Carbon tetrachloride
cDNA	Complementary DNA
CL	Cardiolipin
CREB	cAMP responsive element binding
cm	Centimeter
DBTSS	Database of transcriptional start sites
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DNTP	Deoxyribonucleotide triphosphate
E2F	Electro-acoustic 2 factor
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EK	Ethanolamine kinase
Elk	E-twenty six-like
ER	Endoplasmic reticulum
EtBr	Ethidium bromide

ETS	E-twenty six
FBS	Fetal bovine serum
g/l	Gram per liter
g	Gram
H_2O	Water
ckα	Choline kinase alpha
ckβ	Choline kinase beta
HepG2	Human liver hepatocellular carcinoma
HRG	Heregulin
HSF	Heat shock factor
Ik	Ikaros
IPTG	Isopropyl-β-D-thiogalactopyranoside
kb	Kilobase
kDa	Kilo Dalton
L	Liter
LB	Luria-Bertani
Lyf-1	Lymphoid transcription factor 1
luc2	Luciferase
MCF-7	Human breast adenocarcinoma
mg	Milligram
ml	Milliliter
mm	Millimeter
mM	Millimolar
mRNA	Messenger RNA
MZF1	Myeloid zinc finger 1
NCBI	National Center for Biotechnology Information
NF-kappa	Nuclear factor kappa
ng	Nanogram
nm	Nanometer
p300	Protein 300
PA	Phosphatidic acid
PAH	Polycyclic aromatic hydrocarbon

PBS	Phosphate buffered saline
PC	Phosphatidylcholine
PCR	Polymerase chain reaction
PDGF	Platelet-derived growth factor
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PI	Phosphatidylinositol
PKC	Protein kinase C
PMA	Phorbol-12-Myristate-13-Acetate
PS	Phosphatidylserine
psi	Pound-force per square inch
RLM-RACE	RNA Ligase Mediated Rapid Amplification of cDNA Ends
Rluc	*Renilla* luciferase
RNA	Ribonucleic acid
rpm	Revolutions per minute
rRNA	Ribosomal RNA
SD	Standard deviation
SM	Sphingomyelin
Sp1	Specificity protein 1
SRY	Sex-determining region Y
SV40	Simian virus 40
TAE	Tris-acetate-EDTA
TE	Tris-EDTA
TFSEARCH	Searching transcription factor binding sites
T_m	Melting temperature
Tris	Tris(hydroxymethyl)-aminomethane
tRNA	Transfer RNA
TSS	Transcriptional start site
U	Unit
USF	Upstream stimulatory factor
UV	Ultraviolet
v/v	Volume to volume

w/v	Weight to volume
X	Times
x *g*	fold gravity

PENGKLONAN DAN ANALISIS PROMOTER ISOFORM KOLINA KINASE MANUSIA

ABSTRAK

Kolina kinase (CK) merupakan enzim pertama di dalam laluan CDP-kolina untuk biosintesis fosfatidilkolina dalam semua tisu haiwan. CK memfosforilasikan kolina kepada fosfokolina dengan kehadiran ATP dan Mg^{2+}. Dalam manusia, enzim ini dikodkan oleh dua gen yang berbeza iaitu *ckα* dan *ckβ* yang menghasilkan sekurang-kurangnya tiga isoform protein yang dikenali sebagai CKα1, CKα2 dan CKβ. CK memainkan peranan yang penting dalam sintesis fosfolipid, karsinogenesis dan distrofi otot serta kecacatan tulang. Kawasan promoter CK adalah penting dalam pengawalaturan ekspresi gen *ck*. Kajian ini melaporkan pengklonan promoter *ckα* dan *ckβ* manusia dan penggunaan satu system pelapor untuk menilai aktiviti promoter tersebut. Jujukan promoter *ckα* (2009 bp) and *ck*β (2000 bp) manusia pada bahagian hulu gen masing-masing, telah diklonkan ke dalam vektor pelapor lusiferase tanpa promoter, pGL4.10[*luc2*]. Plasmid rekombinan ini ditransfeksikan bersama dengan vektor kawalan dalaman *Renilla* lusiferase, pGL4.73[*hRluc*/SV40] ke dalam sel adenokarsinoma payudara manusia, MCF-7. Aktiviti promoter diukur dengan menggunakan asai lusiferase. Pelbagai mutan penghapusan terminal-5' juga telah dihasilkan melalui teknik PCR dan diklonkan ke dalam vektor pGL4.10[*luc2*] untuk mengenalpasti kawasan promoter *ckα* dan *ckβ* yang penting untuk transkripsi gen. Keputusan kajian menunjukkan bahawa faktor transkripsi ETS merupakan pengawalatur negatif yang penting terhadap promoter *ckα* manakala factor-faktor transkripsi ETS and GATA merupakan pengawalatur negatif yang penting terhadap aktiviti promoter *ckβ*. Untuk mengesahkan kepentingan ETS dan GATA terhadap

pengawalaturan transkripsi gen *ckβ*, beberapa mutasi telah diperkenalkan pada tapak pelekatan ETS dan GATA di dalam promoter *ckβ*. Aktiviti promoter bagi binaan mutan didapati telah meningkat secara mendadak. Seterusnya, sel-sel MCF-7 yang ditransfeksikan dengan vektor pelapor promoter *ckβ* telah dirawat dengan forbol 12-myristat 13-asetat (PMA) untuk meninjau peranan PMA dalam mengawaltur ekspresi gen *ckβ* melalui factor-faktor transkripsi ETS dan GATA. PMA merupakan pengaktif PKC. Pengaktifan PKC telah meningkatkan pengawalaturan negatif, ETS dan GATA dan akhirnya mengurangkan aktiviti transkripsi promoter *ckβ*. Keputusan menunjukkan bahawa aktiviti promoter *ckβ* telah dikurangkan dengan ketara selepas dirawat dengan PMA. Oleh itu, kajian ini telah mengenal pasti bahawa faktor transkripsi ETS dan GATA merupakan represor yang penting dalam pengawalaturan transkripsi gen *ckβ*.

CLONING AND PROMOTER ANALYSIS OF HUMAN CHOLINE KINASE ISOFORMS

ABSTRACT

Choline kinase (CK) is the first enzyme in the CDP-choline pathway for the synthesis of phosphatidylcholine in all animal tissues. CK phosphorylates choline into phosphocholine in the presence of ATP and Mg^{2+}. In humans, this enzyme is encoded by two separate genes, *ckα* and *ckβ* which produce at least three protein isoforms known as CKα1, CKα2 and CKβ. CK plays an important role in phospholipid synthesis, carcinogenesis and muscular dystrophy as well as bone deformities. The CK promoter region plays a significant role in the regulation of the *ck* gene expression. This study reports the cloning of *ckα* and *ckβ* promoters and the use of a reporter system for evaluating the promoter activity. Promoter sequences of human *ckα* (2009 bp) and *ckβ* (2000 bp), located upstream of their respective genes, were cloned into a promoterless luciferase reporter vector, pGL4.10[*luc2*]. The recombinant plasmid was co-transfected with *Renilla* luciferase internal control vector, pGL4.73[*hRluc*/SV40] into the human breast adenocarcinoma, MCF-7, cell line. Its promoter activity was measured using the luciferase assay. Various 5'-terminal deletion mutants were constructed by PCR technique and cloned into pGL4.10[*luc2*] vector in order to identify the region of the *ckα* and *ckβ* promoters that are important for gene transcription. The results showed that the ETS transcription factor is a crucial negative regulator for the *ckα* promoter while ETS and GATA transcription factors are important negative regulators for the *ckβ* promoter activity. To confirm the importance of ETS and GATA on the regulation of *ckβ* gene transcription, several mutations were introduced to the ETS and GATA

binding sites in *ckβ* promoter. The promoter activities of the mutant constructs were dramatically increased. Subsequently, MCF-7 cells transfected with *ckβ* promoter reporter vector were treated with phorbol 12-myristate 13-acetate (PMA) to explore the role of PMA in *ckβ* gene regulation via ETS and GATA transcription factors. PMA is the activator of PKC. The activation of PKC increases the binding of negative regulators, ETS and GATA and hence decreases the transcriptional activity of *ckβ* promoter. Results showed that the activity of the *ckβ* promoter was significantly reduced after treatment with PMA. Thus, this study has identified ETS and GATA transcription factors as the important repressors in the regulation of *ckβ* gene transcription.

CHAPTER 1.0 INTRODUCTION

1.1 Structures and functions of phospholipids

Phospholipids are the major structural components of membranes (Figure 1.1) (Campbell and Reece, 2002). In almost all cells, the structure is a phospholipid bilayer that surrounds and contains the cytoplasm (Silhavy *et al.*, 2010). Phospholipids have only two fatty acid tails and the third hydroxyl group of glycerol is joined to a phosphate group and a simple organic molecule such as choline (Campbell and Reece, 2002). All phospholipids have hydrophobic and hydrophilic domains. The hydrophobic domain is composed largely of the hydrocarbon chains of fatty acids; the hydrophilic domain, called a polar head group, contains phosphate and other charged or polar group (McKee and McKee, 2003). Furthermore, several phospholipids are emulsifying agents and surface active agents (Kuroki and Voelker, 1994) as they are amphipathic molecules (Metzler and Metzler, 2001). Phospholipids consist of two major types: phosphoglycerides (also known as glycerophospholipids) and sphingomyelins (also known as sphingolipids) (Newsholme and Leech, 2010).

1.1.1 Phosphoglycerides

Phosphoglycerides, which are derived from glycerol, are the most abundant types of phospholipids found in cell membrane. All phosphoglycerides are derived from sn-glycerol-3-phosphate (L-Glycerol 3-phosphate). Two fatty acid chains are esteried to the first and second carbons of the glycerol molecule, denoted as positions C-1 and C-2, respectively. Saturated fatty acids normally occur at C-1 of glycerol while the fatty acid substituent at C-2 is normally unsaturated. The C-3 hydroxyl group of the glycerol backbone is esterified to form different phosphoglycerides.

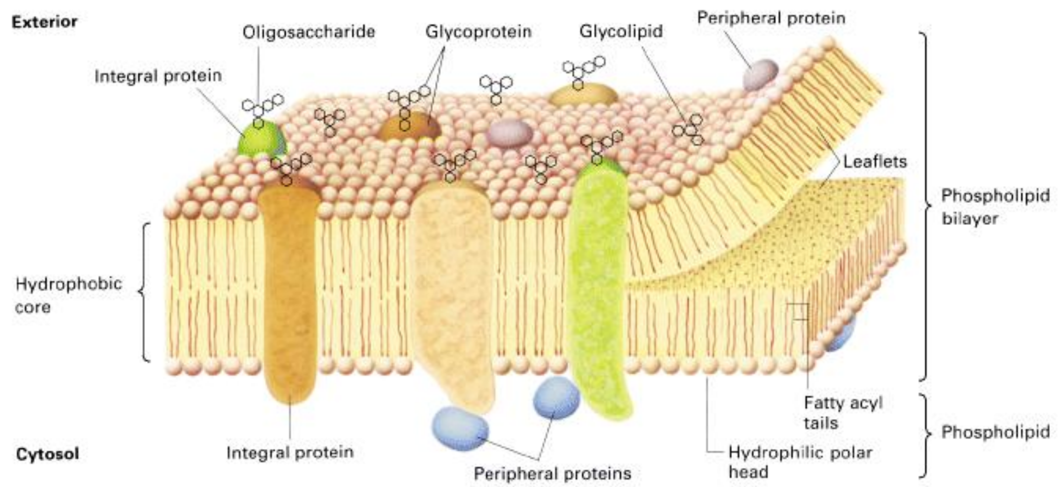


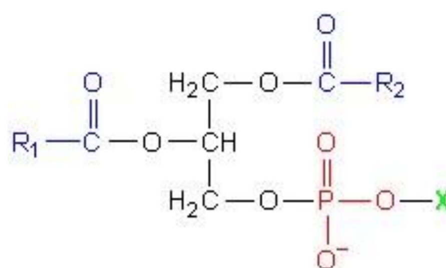
Figure 1.1 The structure of plasma membrane. It is made up of phospholipid bilayer and proteins with different functions are embedded inside it (Henderson et al., 1990).

The phosphoglycerides are named and classified according to the hydroxyl is esterified to the phosphate group of the phosphatidate. Phosphatidate have four common subtituents, serine, ethanolamine, choline and inositol (Lodish *et al.*, 2000; Berg *et al.*, 2002; McKee and McKee, 2003; Plopper, 2011). The major classes of phosphoglycerides are shown in Table 1.1.

1.1.1(a) Phosphatidic acid

Phosphatidic acid (PA) is the simplest phosphoglyceride, it is also known as 1,2-diacyl-sn-glycerol-3-phosphate. PA is progressively being recognized as a crucial lipid second messenger in animal systems as well as in plants (Munnik, 2001; Testerink *et al.*, 2004). Hydrolysis of PA by the enzyme phosphatidate phosphatase yields diacylglycerols, which are the precursors for the biosynthesis of triglycerides, phosphatidylcholine and phosphatidylethanolamine via the Kennedy pathway (Athenstaedt and Daum, 1999; Kooijman *et al.*, 2005). PA also has significant roles in cell signaling and metabolic regulation in all organisms. Surprisingly, recent findings have demonstrated that PA also has an unprecedented role as a pH biosensor by coupling changes in pH to intracellular signaling pathways (Young *et al.*, 2010; Shin and Loewen, 2011). Several proteins, including protein kinase C and small G-proteins are activated by this diacylglycerol (Yang and Kazanietz, 2007).

Table 1.1 Major types of phosphoglycerides (McKee and McKee, 2003)



X Substituent		Name of Phospholipid
Name of X-OH	**Formula of X**	
Water	$-H$	Phosphatidic acid
Choline	$-CH_2CH_2\overset{+}{N}(CH_3)_3$	Phosphatidylcholine (lecithin)
Ethanolamine	$-CH_2CH_2\overset{+}{N}H_3$	Phosphatidylethanolamine (cephalin)
Serine	$-CH_2-CH(\overset{+}{N}H_3)(COO^-)$	Phosphatidylserine
Glycerol	$-CH_2CH(OH)CH_2OH$	Phosphatidylglycerol
Phosphatidylglycerol	$-CH_2CH(OH)-CH_2-O-P(=O)(O^-)-O-CH_2-CH(OCOR)-CH_2OCOR$	Diphosphatidylglycerol (cardiolipin)
Inositol	OH OH OH OH OH	Phosphatidylinositol

1.1.1(b) Phosphatidylcholine

Phosphatidylcholine (PC) (also known as lecithin) has choline bonded with phosphate group of the phosphatidate (Scholfield, 1981). The PC generally is the most abundant phospholipid class in a membrane, in both animal and plant. It amounts to almost 50% of the total constituents and is the key building block of membrane bilayers. Many prokaryotes lack PC, but it is estimated that more than 10% of bacteria possess PC (Sohlenkamp *et al.*, 2003). In animal tissues, PC tends to exist mainly in the diacyl form. In mammals, it is well known that PC is found in lipoproteins, biliary lipid aggregates and lung surfactant (Kanno *et al.*, 2007). PC is synthesized through two distinct enzymatic pathways, namely, the *de novo* CDP-choline pathway and the methylation pathway (Walkey *et al.*, 1998). The CDP-choline pathway involves the conversion of choline to CDP-choline through a phosphorylcholine intermediate for condensation with diacylglycerol to produce PC. This enzymatic pathway is present in all cells. In the methylation pathway, PC is produced via sequential methylation of phophatidylethanolamine by phosphatidylethanolamie N-methyltransferase, which is expressed only in the liver cells (Fernandez-Murray and McMaster, 2005; Witola and Ben Mamoun, 2007).

1.1.1(c) Phosphatidylethanolamine

Phosphatidylethanolamine (PE) (also referred as cephalin), is the second most abundant phospholipid in animal and plant, and constitutes for more than 50% of the total phospholipid species in eukaryotic membranes (Gibellini and Smith, 2010). 20% of phospholipids in the liver are PE and as much as 45% in the brain. Higher proportions of PE are found in mitochondria than in other organelles (Vance, 2008). Besides being the major constituent of membrane, the PE is a zwitterionic

phospholipid that involves in a number of biological events, such as in the blood coagulation, blood pressure regulation, and in the visual process (Song *et al.*, 2012). The PE is synthesized via the *de novo* CDP-ethanolamine pathway in which the final reaction occurs on the endoplasmic reticulum and nuclear envelope (Vance and Vance, 2009).

1.1.1(d) Phosphatidylserine

Phosphatidylserine (PS) (1,2-diacyl-*sn*-glycero-3-phospho-L-serine) (Gaus *et al.*, 2005) is a quantitatively minor membrane phospholipid that makes up to 10% of the total phospholipids of mammalian cell membranes (Yamaji-Hasegawa and Tsujimoto, 2006). The cytoplasmic leaflet of the plasma membrane is generally rich in PS (Kay and Grinstein, 2011). Any disruption on the asymmetric distribution of phospholipids will lead to the appearance of PS on the cell surface. The PS is known to play a central role in both apoptosis and blood coagulation (Chaurio *et al.*, 2009). Furthermore, PS is a required cofactor for protein kinase C (Vance and Steenbergen, 2005).

1.1.1(e) Phosphatidylglycerol

Phosphatidylglycerol (PG) is a ubiquitous lipid in biological membranes (Sakurai *et al.*, 2007). The PG is found in bacteria, such as *E.coli*, which made up to 20% in bacterial membranes. It is also mainly presents in thylakoid membranes of chloroplasts, which perform specific roles in the photosynthetic process in cyanobacteria and higher plants (Henin *et al.*, 2009; Wada and Murata, 2009). The PG is synthesized only in the mitochondria of non-photosynthetic eukaryotes such as *Saccharomyces cerevisiae* and it is used as a precursor for the biosynthesis of

cardiolipin (Hagio *et al.*, 2002). The PG is the major end product of the Kennedy pathway. PG forms the bilayer of the inner membrane and the inner leaflet of the outer membrane (Lu *et al.*, 2011). The PG is crucial for viability of wild-type cells not only because it is a major component of the inner membrane but it also plays important roles in the initiation of DNA replication of *oriC* and *Sec*A-dependent protein translocation (Dowhan, 2009).

1.1.1(f) Diphosphatidylglycerol

Diphosphatidylglycerol (also known as cardiolipin, CL) is located and synthesized in the mitochondrial inner membrane (Schlame *et al.*, 2000). It is exclusively present in membranes of bacteria and of mitochondria of eukaryotes, and hence is found in virtually all organisms of the three domains of life: eubacteria, archaebacteria and eukaryote (Schlame, 2008). Furthermore, CL is most abundantly found in human tissues especially those that are rich in mitochondria, such as heart, skeletal muscle and liver tissue (Lu *et al.*, 2006). The CL interacts with a large number of mitochondrial proteins and effects functional activation of certain enzymes, especially those involve in oxidative phosphorylation and photo-phosphorylation that result in ATP production (Chen *et al.*, 2006).

1.1.1(g) Phosphatidylinositol

Phosphatidylinositol (PI) can be phosphorylated to form phosphoinositide. PI participates in essential metabolic processes in animals through a number of metabolites (Kim *et al.*, 2010). Phosphorylation of PI by a number of different kinases produces phosphatidylinositol 4-phosphate (PIP) and phosphatidylinositol 4,5-biphosphate (PIP_2). The PIP_2 is a substrate for phospholipase C which lead to the

production of inositol 1,4,5-triphosphate and diacylglycerol that serve as signaling molecules in animals. Inositol 1,4,5-triphosphate binds to specific receptors and induces the release of calcium from intracellular stores, while diacylglycerol activates protein kinase C, which in turn controls many key cellular functions, including differentiation, proliferation, metabolism and apoptosis (Desrivieres *et al.*, 1998). The PI is synthesized from the phosphatidate and the reaction involves CTP to produce CDP-diacylglycerol which then combines with myo-inositol to form PI and CMP (Berridge and Fain, 1979).

1.1.2 Sphingolipids

Sphingolipid (also known as sphingomyelin) molecules consist of a long-chain amino alcohol (Figure 1.2). In animals, this alcohol is known as sphingosine (Chaurio *et al.*, 2009). The sphingolipid is different from phosphoglycerides because it does not contain glycerol backbone although they are similar in shape (Lodish *et al.*, 2000). The core of each type of sphingolipid is ceramide, contain a fatty acid linked by an amide bond derivative of sphingosine (McKee and McKee, 2003). A sphingomyelin is formed by the esterification of a phosphorylcholine or a phosphoethanolamine to the 1-hydroxyl group of a ceramide (McKee and McKee, 2003; Gibellini and Smith, 2010). Sphingosine-1-phosphate is a bioactive lipid and served as the primary precursor of ceramides which are involved as mediators of cellular events, for example cytoskeleton rearrangements, growth, motility and survival (Spiegel and Milstien, 2003). The ceramides are also precursors for the glycolipids (also referred to as glycosphingolipids) with an immense range of function in tissues (McKee and McKee, 2003). Furthermore, sphingomyelin and other sphingolipids together with cholesterol are located in an intimate association in

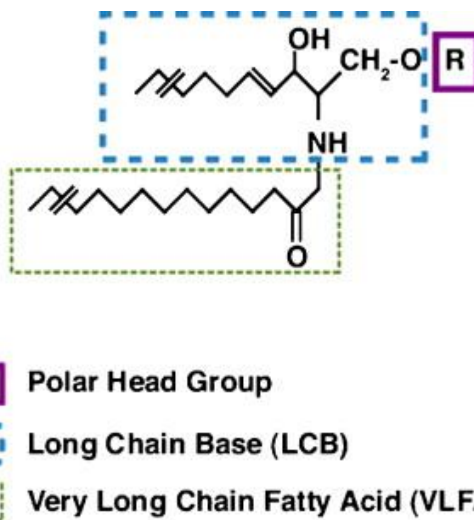


Figure 1.2 Core structure of sphingolipids, long chain bases form the core structure of sphingolipids (Rao and Acharya, 2008).

specific sub-domains to form lipid rafts, which appear to be associated with cellular signaling and the signal transduction process (Dowhan, 2009).

1.2 Biosynthesis of major phospholipids in eukaryotic membranes

The molecular basis for phospholipid function on eukaryotic systems is more complex and difficult than in prokaryotes. It is essential to study the roles of phospholipids in cell physiology and specific cellular processes (Dowhan, 1997). The studies for pathway of the major eukaryotic phospholipid biosynthesis were established mainly by Eugene Kennedy (Kennedy, 1957). Besides the study on the components of the cell membrane, the implication of phospholipid metabolites and derivatives in different mitogenic signaling pathways has also been identified (Gallego-Ortega *et al.*, 2009). Figure 1.3 shows the pathways for the biosynthesis of major eukaryotic membrane phospholipids.

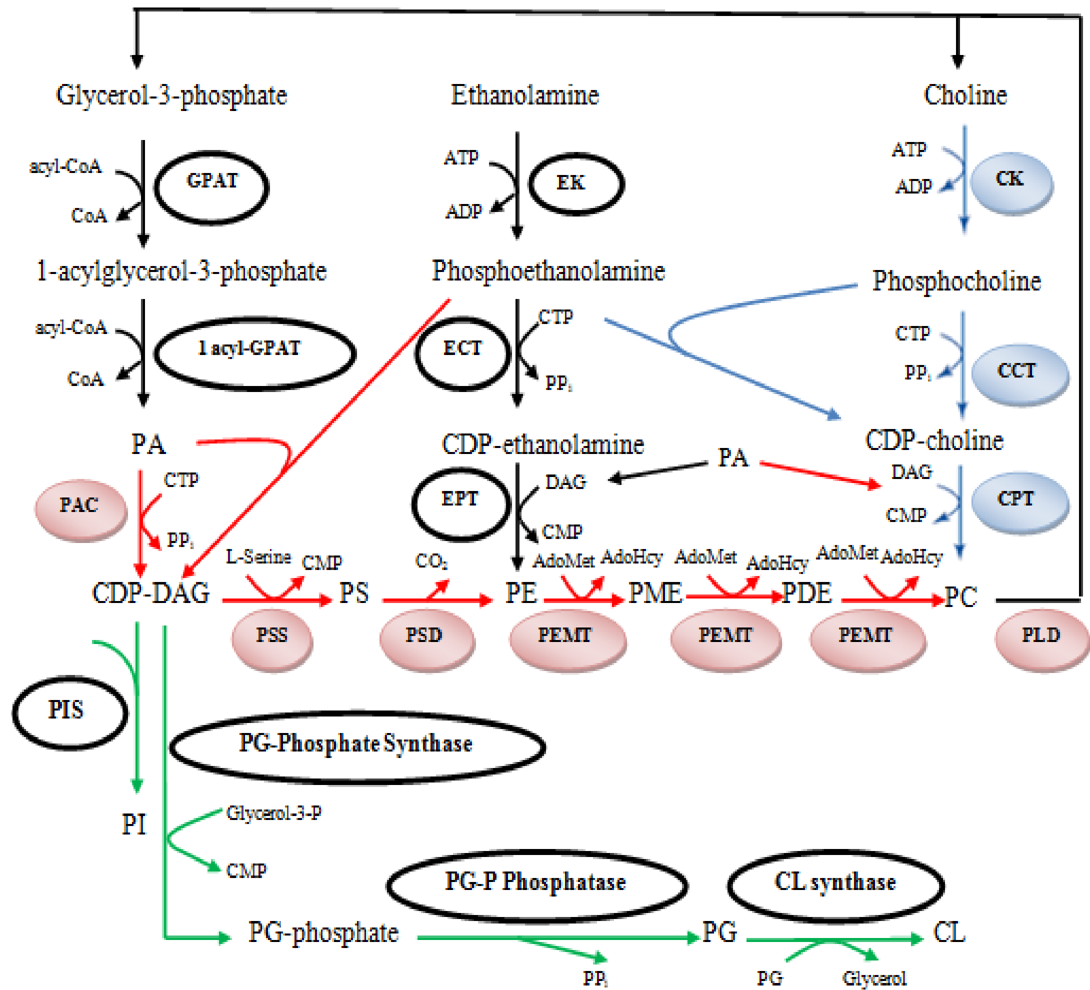


Figure 1.3 Pathways for the biosynthesis of major eukaryotic membrane phospholipids (Kent and Carman, 1999). Enzymes involved in individual steps are circled. CDP-choline pathway is shown in blue. PEMT pathway is shown in red. PLD and EK reactions are shown in black. PI and CL reactions are shown in green.

CK, choline kinase; EK, ethanolamine kinase; CCT, choline phosphate cytidyltransferase; ECT, ethanolamine phosphate cytidyltransferase; CPT, choline phosphotransferase; EPT, ethanolamine phosphotransferase; GPAT, glycerol-3-phosphate acyltransferase; 1 acyl-GPAT, 1 acyltransferase-GPAT; PAC, phosphatidate cytidyltransferase; PSS, phosphatidylserine synthase; PSD, phosphatidylserine decarboxylase; PEMT, phosphatidylethanolamine methyltransferase; PLD, phospholipase D; PIS, phosphatidylinositol synthase; PG-P synthase, phosphatidylglycerol phosphate synthase; PG-P phosphatase, phosphatidylglycerol phosphate phosphatase; CL synthase, cardiolipin or diphosphatidylglycerol synthase; PA, phosphatidic acid; PS, phosphatidylserine; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PME, phosphatidylmonoethanolamine; PDE, phosphatidyldiethanolamine; PI, phosphatidylinositol; PG, phosphatidylglycerol; CL, cardiolipin (diphosphatidylglycerol); CTP, Phosphocholine cytidyltransferase; DAG, diacylglycerol; AdoMet, S-adenosylmethionine; AdoHcy, adenosylhomocystein.

1.2.1 Synthesis of phosphatidic acid

The synthesis of PA started with the acylation of glycerol-3-phosphate. The PA biosynthesis plays an important step in the activation of fatty acids by acyl-CoA synthetase to yield acyl-CoA. Five different forms of acyl-CoA synthetase have been identified. The synthetases function differently in the biosynthesis of phospholipid and triacylglycerol (TG) (Vance and Vance, 2002). The PA can be further hydrolyzed to the central liponucleotide intermediate, cytidine diphosphate-diacylglycerol (CDP-DAG) via phosphatidate cytidyltransferase (PAC) (Figure 1.3), which is also known as CDP-DAG synthase (CDS) (Nigou and Besra, 2002). The CDP-DAG is a key intermediate in phospholipid synthesis in all organisms because of its location at the branchpoint of the biosynthetic pathway (Icho *et al.*, 1985). The CDP-DAG plays a role as the donor of phosphatidyl moieties to the primary hydroxyl groups of serine or glycerol-3-phosphate in the generation of PI (Zuker, 1996) or CL (Sandoval-Calderon *et al.*, 2009), respectively (Figure 1.3).

1.2.2 Synthesis of phosphatidylcholine

The PC biosynthetis consists of two pathways, the CDP-choline pathway (Kennedy pathway) and the phosphatidylethanolamine methylation (PEMT) pathway (Aoyama *et al.*, 2004). The PEMT pathway predominates in yeast and fungi. For mammals, PEMT is significant only in mammalian liver cells for ~ 30% of hepatic PC biosynthesis while the CDP-choline pathway is the major route in other mammalian cells (Lopez-Lara and Geiger, 2001; Li and Vance, 2008).

PC can be found in significant amounts in membranes of rather diverse bacteria especially in symbionts, pathogens and photosynthetic bacteria (Sohlenkamp *et al.*,

2003; Comerci *et al.*, 2006). The PC is constantly turned over to yield free choline and PA. This process is catalyzed by PLD, and the latter can be further hydrolyzed to DAG by the action of PA phosphatase (shown in black in Figure 1.3) (Kent and Carman, 1999; Zhou *et al.*, 2000).

1.2.2(a) CDP-choline pathway

The CDP-choline pathway (shown in blue in Figure 1.3) is also known as Kennedy pathway, named after Eugene Kennedy who elucidated it in 1956 (Gibellini and Smith, 2010). The choline is mainly used for the synthesis of PC through the CDP-choline pathway, which accounts for up to 95% of the total choline pool in most tissues. The remaining 5% include free choline, phosphocholine, glycerophosphocholine, CDP-choline and acetylcholine (Li and Vance, 2008). There are three reactions involved in the CDP-choline pathway. Free choline is first catalyzed by an enzyme choline kinase (CK), and is the first reaction of choline phosphorylation by ATP (Ishidate, 1997). The second reaction is the phosphocholine cytidylyltransferase (CCT) catalyzes the formation of CDP-choline from phosphocholine and CTP (Kent, 1997). The third reaction is that cholinephosphotransferase (CPT) catalyzes the final condensation reaction of CDP-choline with diacylglycerol (McMaster and Bell, 1997; Vance and Vance, 2002).

For the CDP-choline pathway in mammals, CK activity is the first step in biosynthesis, but the CCT is considered as the regulatory as well as the rate limiting step (Lykidis and Jackowski, 2001). However, several reports have indicated that the step of CK is also slow and can be a regulatory step for the net biosynthesis of PC in certain situations (Aoyama *et al.*, 2004). The PC biosynthesis is also required for

normal very low-density lipoprotein secretion from hepatocytes. When choline supply is deficient, recent studies indicate that choline is recycled into the liver and redistributed from the kidneys, lungs (where saturated PC is an essential lipid component of the pulmonary surfactant), and intestines to the liver and the brain (Li and Vance, 2008). However, choline itself is not synthesized *de novo* in animal cells and must be obtained from dietary sources or by degradation of existing choline-containing lipids (Ishidate, 1997).

1.2.2(b) Phosphatidylethanolamine methylation pathway

Phosphatidylethanolamine methylation (PEMT) pathway is another pathway from which PC can be synthesized *de novo* in mammalian cells (shown in red in Figure 1.3). In mammalian cells, phospholipid N-methyltransferase catalyzes the three successive AdoMet-dependent methylations of phosphatidylethanolamine (Figure 1.3) (Watkins *et al.*, 2003; Verkade *et al.*, 2007). The PE used in the PEMT pathway is derived from PS (Kent and Carman, 1999). The PEMT is only expressed in significant levels in the liver and testis (Jackowski and Fagone, 2005).

1.2.3 Synthesis of phosphatidylethnolamine

The major pathway for the biosynthesis of PE in mammals is known to be a CDP-ethanolamine pathway (Lykidis and Jackowski, 2001) which was first described by Kennedy and Weiss in 1956 (Vance and Vance, 2002). The CDP-ethanolamine pathway consists of three enzymatic steps (Figure 1.3). Firstly, ethanolamine kinase (EK) catalyzes the ATP-dependent phosphorylation of ethanolamine, forming phosphoethanolamine and the byproduct ADP. In the second step which is considered to be the rate-limiting step of the Kennedy pathway, it is the CTP:

phosphoethanolamine cytidylyltransferase (ECT) uses phosphoethanolamine and CTP to form the high-energy donor CDP-ethanolamine with the release of pyrophosphate (Leonardi *et al.*, 2009). The final step is catalyzed by CDP-ethanolamine: 1, 2-diacylglycerol ethanolaminephosphotransferase (EPT) using CDP-ethanolamine and a lipid anchor such as diacylglycerol (DAG) or alkyl-acylglycerol (AAG) to form PE and CMP as byproducts (Bleijerveld *et al.*, 2004).

Another crucial route to synthesize PE is the decarboxylation of PS, which is described by Borkenhagen and colleagues (Bleijerveld *et al.*, 2007). In this pathway, PS synthesized from PC or PE by phosphatidylserine-1 and -2, respectively, is decarboxylated by phosphatidylserine decarboxylase (PSD) to generate PE (Vance and Vance, 2004). To date, only one mammalian PSD has been cloned, and it was shown to be active in mitochondria (Vermeulen *et al.*, 1997; Bleijerveld *et al.*, 2007).

1.2.4 Synthesis of phosphatidylserine

PS is made in prokaryotes, some plants and in yeasts through the CDP-diacylglycerol pathway. However this pathway does not exist in mammalian cells. Instead, PS is synthesized by a base-exchange reaction, in which the head group of a PC or PE is exchanged for L-serine (Vance and Vance, 2002; Moldave, 2003). This pathway was first described by Hubscher in 1959 (Vance and Vance, 2002). The PS is synthesized by two distinct PS synthases, PS synthase-1 (PSS1) and PS synthase-2 (PSS2) (Arikketh *et al.*, 2008). The PSS1 exchanges serine for the choline head group of PC, while PSS2 uses PE as a substrate (Vance and Steenbergen, 2005).

1.3 Choline kinase

Dietary choline is an essential nutrient which is absorbed by the intestine in the form of lysophosphatidylcholine and choline, and, the uptake by the cell is carried out either by passive diffusion or diffusion facilitated by specific choline transporters (Michel *et al.*, 2006; Li and Vance, 2008). Upon entry into the cell, the choline is immediately phosphorylated to phosphocholine or oxidized to betaine. These processes take place mainly in hepatocytes (Katz-Brull *et al.*, 2002). The phosphorylation of choline is catalyzed by choline kinase (CK) which constitutes the first step of Kennedy pathway (Ishidate, 1997).

The enzymatic reaction of CK in Brewer's yeast was first described by Wittenberg and Kornberg in 1953 (Wittenberg and Kornberg, 1953). In human, CK (ATP: choline phosphotransferase, EC 2.7.1.32) (Aoyama *et al.*, 2000) is encoded by two separate genes, *CKA* and *CKB*, located on chromosomes 11q13.2 and 23q13.33, respectively (Gallego-Ortega *et al.*, 2011). Alternative splicing of the *CKA* transcript results in two functional isoforms, CKα1 and CKα2; *CKB* encodes CKβ (Wu and Vance, 2010). CKα1 has 457 amino acids while CKα2 has 439 amino acids. They are only different in an additional stretch of 18 amino acids that begin at the position 155. CKβ has 395 amino acids and its sequence is approximately 40% different from CKα1 and CKα2 (Gallego-Ortega *et al.*, 2011). Active CK has been proposed to consist of either homo- or hetero-dimeric (oligomeric) forms, and neither isoform is active in its monomeric form (Liao *et al.*, 2006).

Nowadays, different cDNAs of CK from yeasts, mammals and plants have been cloned and characterized (Strausberg *et al.*, 2002; Kent, 2005). The CK also carries

the Ethanolamine Kinase activity in several organisms and has been reported to be involved in the synthesis of PE using ethanolamine as substrate to produce phosphoethanolamine (Aoyama *et al.*, 2000; Gallego-Ortega *et al.*, 2011). Furthermore, CK is a vital protein that involves in cancer, and the specific interference with the CK activity constitutes a new efficient strategy for cancer treatment (Ramirez de Molina *et al.*, 2004b).

In mammals, the CK cDNA was first obtained from a rat liver λgt11 library by screening using the antibody against rat liver CK. The open reading frame encoded 435 amino acids with a molecular mass of 49743 Da (Uchida and Yamashita, 1992). Later, the liver enzyme was found to exist in two isozymes, termed CKRI and CKRII (Porter and Kent, 1990). The 5' region of the *CKR* gene exhibited the features of not only a housekeeping gene but also of a gene regulated via a variety of putative cis-acting elements (Uchida, 1994).

1.3.1 Biochemical characteristic of choline kinase isoforms

Recently, the enzymatic properties of CKα1 and CKβ have been characterized *in vitro* and *in vivo* (Gallego-Ortega *et al.*, 2011). *In vitro*, the Michaelis constant (K_m) of CKβ for choline was higher than CKα1 while the K_m of CKβ for ethanolamine was lower than CKα1. *In vivo*, CKβ showed only an EK role while CKα1 displayed both CK/EK roles, suggesting that each CK isoform plays distinct role in biochemical pathways under *in vivo* conditions and behave differently *in vitro* (Gallego-Ortega *et al.*, 2009; Gallego-Ortega *et al.*, 2011).

Homozygosity for the mutant CKα mice are lethal in embryonic stage (Wu *et al.*, 2008) whereas mice with a knocked-out CKβ are viable to the adult stage but causes rostrocaudal muscular dystrophy (*rmd*) and bone deformities (Sher *et al.*, 2006; Wu *et al.*, 2009; Wu *et al.*, 2010). The *rmd* mice show normal PC levels in all tissues except in skeletal muscle of the forelimbs with an unusual rostral-to-caudal gradient (Sher *et al.*, 2006). The heterozygous mutant of CKα mice appeared normal in embryonic development and gross anatomy and was fertile (Wu *et al.*, 2008). It suggests that CKα plays a crucial role under physiological conditions but CKβ is insufficient to compensate the PC biosynthesis or even to provide the minimal requirements of phosphocholine (Gallego-Ortega *et al.*, 2011). In addition, heterozygous CKα knockout mice showed decreased PE levels (Wu *et al.*, 2008). This suggests that there is a dual role for the CKα isoform in the biosynthesis of PC and PE (Gallego-Ortega *et al.*, 2011).

In humans, disruption of PC synthesis by loss of function due to mutation of CKβ activity in muscle causes congenital muscular dystrophy (CMD) (Mitsuhashi *et al.*, 2011a; Mitsuhashi *et al.*, 2011b). The feature of the muscle pathology in humans and mice is characterized by distinct mitochondrial morphological abnormalities where mitochondria were absent in the center of muscle fibers and significantly enlarged at the periphery (Sher *et al.*, 2006; Wu *et al.*, 2010; Mitsuhashi *et al.*, 2011a). PC levels are significantly decreased and CK activity is absent in the skeletal muscle from CMD human and *rmd* mice. However, in other tissues, where there was only a mild decrease in CK activity, PC levels are not altered and there are no obvious pathological changes. This is supported by the fact that CKα was not detected in muscle (Mitsuhashi *et al.*, 2011b). Furthermore, loss of function through mutation of

CKβ causes mitochondria to be abnormally large and exhibit decreased inner membrane potential in hindlimb muscle of mice whereas presence of mega mitochondria was in insignificant amounts in forelimb muscle and functioned normally. Hence, the distinction in forelimb versus hindlimb muscle of CKβ deficient mice is predominantly due to CKα, but less contribution of CKβ, to total CK activity in forelimb muscle (Wu *et al.*, 2010).

1.4 Choline kinase in cell transformation and carcinogenesis

Overexpression of CK, which leads to higher level of phosphocholine can potentially lead to carcinogenesis. CK is suggested to be involved in the *ras* signaling pathway mediated by the *ras* oncogene (Ramirez de Molina *et al.*, 2004b). There are two well known *ras* oncogene effectors, PI3K and Ral-GDS (Ramirez de Molina *et al.*, 2002a). Several *ras* oncogenes such as *src*, *raf* and *mos* could also increase the endogenous CK activity and phosphocholine level when expressed in the murine fibroblasts (Ratnam and Kent, 1995; Hernandez-Alcoceba *et al.*, 1997; Ramirez de Molina *et al.*, 2001).

Production of phosphocholine also plays an essential role in cell growth that are induced by growth factors both in murine fibroblasts (Kiss and Chung, 1996) and in different human cells. Treatment of human cells with growth factors such as EGF, PDGF or HRG results in blockage of DNA synthesis. The increase in growth factors induced the production of phosphorylcholine (Ramirez de Molina *et al.*, 2004a).

Interestingly, the elevated phosphocholine levels appear to be a common characteristic feature in cell lines derived from human tumors (Bhakoo *et al.*, 1996).

The elevation of phosphocholine level which is associated to cell malignancy is detected by nuclear magnetic resonance (NMR). It has been demonstrated that different tumors in murine and human, showed increased levels of phosphocholine compared to their corresponding normal tissues (Podo *et al.*, 2007). Phosphocholine as a malignancy marker, such as a biomarker for breast cancer, is useful for the diagnostic and monitoring of the cancer's progression (Eliyahu *et al.*, 2007).

CK generates phosphocholine in the carcinogenic process and it has been found that CK activity was increased in tissues from human tumors (Glunde *et al.*, 2005; Eliyahu *et al.*, 2007; Iorio *et al.*, 2010). Human CK, expressed in murine fibroblast, increases DNA synthesis when stimulated by insulin, insulin-like growth factor I (IGF-I) and other growth factors (Chung *et al.*, 2000). The overexpression of CK in primary human mammary epithelial cells was sufficient for growth factor-induced proliferation because its overexpression causes an increase in DNA synthesis (Ramirez de Molina *et al.*, 2004a). Furthermore, high levels of CK mRNA have been detected in colon cancer induced by 1,2-dimethylhydrazine (DMH) in rats (Nakagami *et al.*, 1999a) as well as in tumors induced by polycylic aromatic hydrocarbons (PAHs) (Mordukhovich *et al.*, 2010). Besides, carbon tetrachloride (CCl_4) was also induced CK activity 2-4 fold (Aoyama *et al.*, 2002).

Most of the studies have reported that CKα is overexpressed in a variety of cell lines derived from human tumors such as in breast, prostate, colon, brain, lung, bladder and cervix carcinoma tumor tissues (Banez-Coronel *et al.*, 2008; Nimmagadda *et al.*, 2009; Shah *et al.*, 2010) but there is no evidence that reports overexpression of CKβ in the carcinogenesis process (Gallego-Ortega *et al.*, 2009). The induction of CKα

activity in these types of cancer ranges from 40% to 60% (Gallego-Ortega *et al.*, 2006). Furthermore, a study using 167 tumor samples showed that non-small cell lung cancer (NSCLC) patients with an overexpression of CKα had a worse survival rate compared to those who did not, and this could entail important clinical consequences (Ramirez de Molina *et al.*, 2007). Recently, overexpression of CKα1 has been shown to increase invasiveness in human breast cancer as well as resistance to treatment with 5-fluorouracil (5-FU) (Shah *et al.*, 2010).

Recenly, the implication of CKβ in cell transformation and human carcinogenesis both *in vitro* and *in vivo* has been discovered. However, the effect of CKβ overexpression on cell transformation is still not very clear. Overexpression of CKβ had no significant effect on the number or the size of the cancer cell colonies. In addition, CKβ was unable to induce tumor growth under conditions where CKα1 does (Ramirez de Molina *et al.*, 2005; Gallego-Ortega *et al.*, 2009).

In gene expression profiling, mRNA of CKα1 shows significant overexpression in a panel of breast, lung and ovarian cancer cell lines when compared with their normal counterparts whereas no changes were found for CKβ mRNA levels (Eliyahu *et al.*, 2007; Gallego-Ortega *et al.*, 2009; Iorio *et al.*, 2010). These findings are important as the overexpression of CKα1 activity is able to transform cells from different origins (Ramirez de Molina *et al.*, 2005; Gallego-Ortega *et al.*, 2009; Gallego-Ortega *et al.*, 2011). Most of the findings suggest that CKα is involved in the cell growth regulation and is altered during the carcinogenic process, hence development of new families of anticancer drugs to inhibit CKα activity may provide a novel cancer therapeutic strategy (Aoyama *et al.*, 2004).

1.5 Promoter

A promoter is a DNA sequence which is directly upstream to the coding sequence that regulates transcription of a gene. Gene transcription is regulated by the binding of transcription factors specifically to the regions of the promoter (Xuan *et al.*, 2005). Promoters are the best-characterized transcriptional regulatory sequences in complex genomes because of their predictable location upstream of transcriptional start site (TSS). Promoters consist of two separate segments, the core (proximal) and the extended (distal) regions.

The core promoter is normally located within 50 bp of the TSS, where the preinitiation complex forms and the general transcription machinery assembles (Cooper *et al.*, 2006). The core promoter elements appear to interact directly with the RNA polymerase II, which is essential to begin the transcription. Most of the studies of the basal machinery have been performed with promoters that consist of the TATA box as an essential element (Smale and Kadonaga, 2003). However, the specific regulatory sequences that control spatial and temporal expression of the downstream gene are present in the extended (distal) promoter. Since promoters consists of many classes of eukaryotic transcriptional regulatory elements, hence identification and characterization of these elements are important to understand the complex network of human gene regulation (Cooper *et al.*, 2006).

1.5.1 Transcription factors

A transcription factor is a protein which acts by binding to a specific short DNA sequence that in turn regulates the gene transcription either positively or negatively (Latchman, 1997). Transcription factors are important for the regulation of gene

expression. Transcription factors are found in all living organisms and the number of transcription factors present within an organism increases with the genome size (van Nimwegen, 2003). In the human genome, there are approximately 2600 proteins that contain DNA binding domains and most of these are presumed to function as transcription factors (Babu *et al.*, 2004).

The transcription of a gene is regulated by trans-acting protein factors that interact with specific cis-acting elements (Schwemmle *et al.*, 1997). Transcription of an eukaryotic protein-coding gene is preceded by several events, such as locus decondensation, remodeling of the nucleosome, histone modifications, the binding of transcriptional activators and coactivators to enhancers, and recruitment of the basal transcription machinery to the core promoter (Smale and Kadonaga, 2003). However, only a few nucleotides are significant in its function within a promoter (Baliga, 2001).

Eukaryotic transcription factors can be categorized into various families on the basis of conserved sequences among their DNA-binding domains (Table 1.2). Several distinct trans-acting factors normally interact with a common binding site due to such structural conservation. Furthermore, some findings indicate that the interaction of distinct transcription factors with a common target site does not essentially result in similar transcriptional responses. Hence, some factors may activate the transcription,

Table 1.2 Some major transcription factor families and functions

Family	Representative transcription factors	Some functions
basic leucine zipper (bZIP)	AP-1, CREB (Cha-Molstad *et al.*, 2004; Cai *et al.*, 2008)	Liver differentiation, fat cell specification (Gilbert, 2000)
basic helix-loop-helix (bHLH)	SRY, AhR/Ar, USF (Luo and Sawadogo, 1996; Bhandari *et al.*, 2011; Pocar *et al.*, 2005)	Sex-determination, carcinogenesis (Ledent and Vervoort, 2001; Gilbert, 2000)
zinc finger	GATA, MZF1, Sp1, Ik	Carcinogenesis, expressed in hematopoietic progenitors including erythroid cells (Gaboli *et al.*, 2001; Viger *et al.*, 2008; Deniaud *et al.*, 2006; Dijon *et al.*, 2008)
Rel homology	NF-kappa	Human diseases (Ahn and Aggarwal, 2005)
winged helix-turn-helix (wHTH)	E2F, ETS, HSF, Elk-1	Carcinogenesis, response to various stresses, (Donaldson *et al.*, 1996; Hashikawa *et al.*, 2006)
grainyhead-related protein	CP2 (To *et al.*, 2010)	Erythroid gene expression (Kokoszynska *et al.*, 2008)