OVER-EXPRESSION OF HUMAN Y-BOX BINDING PROTEIN-1: A COMPARISON BETWEEN MAMMALIAN AND BACTERIAL EXPRESSION SYSTEM

by

WONG JIAN YUAN

This report is submitted as a partial fulfillment of the requirement for the degree of Bachelor of Health Sciences (Biomedicine)

University Science Malaysia

2008

CERTIFICATE

This is to certify that the dissertation entitled

"OVER-EXPRESSION OF HUMAN Y-BOX PROTEIN-1: A COMPARISON BETWEEN MAMMALIAN AND BACTERIAL EXPRESSION SYSTEM"

is the bonafide record of research work done by

Mr Wong Jian Yuan

during the period from July 2008 to October 2008 under my supervision.

Supervisor,

.....

Dr. Shaharum Shamsuddin
Lecturer
School of Health Sciences
Universiti Sains Malaysia
Health Campus
16150 Kubang Kerian
Kelantan, Malaysia

Date: 24/12/08

DR SHAHARUM SHAMSUDDIN
Penyarah Kanan
Pusat Pengajian Sains Kesihatan
Universiti Sains Malaysia
Kampus Kesihatan
16150 Kubang Kerian
Kelantan

Co-supervisor,

.....

Dr. Nik Norliza Nik Hassan
Lecturer
School of Health Sciences
Universiti sains Malaysia
Health Campus
16150 Kubang Kerian
Kelantan, Malaysia

Date: 21 DEC 2008

DR. NIK NORLIZA BT. NIK HASSAN
Senior Lecturer
School of Health Sciences
Health Campus, Universiti Sains Malaysia
16150 Kubang Kerian, Kelantan

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my main supervisor, Dr. Shaharum Shamsuddin, for his guidance, encouragement, motivation and excellent technical assistance towards the completion of this thesis. I would like to thank my co-supervisor, Dr. Nik Norliza Nik Hassan for her assistance, guidance and supervision throughout this research. They greatly supervised me when facing technical problems. Throughout their suggestion, creative idea, technical supports and cooperation, they had highly improved my knowledge of molecular biology in terms of principles, techniques and ideas.

Special thanks to Mr. Venugopal and Mr. Tan Ghim Cheong, postgraduate study students who demonstrate the molecular techniques, provide their technical support and helping in troubleshooting the problems. Their guidance has leaded me to solve a lot of technical problems during the critical moment.

I am thankful to Professor Asma Ismail, the director of Institute for Research in Molecular Medicine (INFORMM) and Dr. See Too Wei Cun for allowing me to use the cell culture facilities at INFORMM and Cell Culture Laboratory School of Health Sciences.

This thesis is dedicate to my beloved parents and family members for their love, encouragement and support in completing this thesis. Their supports allow me to fully concentrate to complete this research.

TABLE OF CONTENTS

ACKNOWLEDGEMENT iii

CONTENT iv

LIST OF TABLES viii

LIST OF FIGURES ix

LIST OF SYMBOLS x

LIST OF ABBREVIATIONS xi

ABSTRAK xii

ABSTRACT xiii

1.0 LITERATURE REVIEW 1

1.1 Calcium Phosphate Transfection 1

1.2 Transformation 4

1.3 HeLa cancer cells 5

1.4 Glutathione-S-Transferase (GST) gene fusion system 5

2.0 INTRODUCTION 7

2.1 Y-box binding protein 1, YB-1 7

2.1.1 dbpA : organization, structure and function 10

2.1.2 dbpB (YB-1) : organization, structure and function 12

2.1.3 Relationship of YB-1 in cancer 14

2.2 Objectives 15

3.0 MATERIAL AND METHOD 18

3.1 General equipments 18

3.2 General reagents, solutions, buffers and media 19

3.3 DNA Material and vectors 20

3.4 Recombinant DNA Methods 21

3.4.1 Extraction of the plasmid DNA with the QIAprep Spin Miniprep Kit Protocol (QIAGEN, Germany) 21

3.4.2 Restriction digests of plasmid DNA 22

3.4.3 Agarose gel electrophoresis of plasmid DNA 22

3.5 Microbiology methods 23

3.5.1 Media and solution 23

3.5.2 Bacteria strains 24

3.5.3 Freezing and storage of *Escherichia coli* (*E. coli*) cells 24

3.5.4 Preparation of competent *Escherichia coli* (*E. coli*) cells and transformation 25

3.5.5 Transformation of bacterial cell using XL1-Blue *E. coli* (Stratagene, Germany) and DH5α *E. coli* 26

3.6 Bacterial protein expression and analysis 27

3.6.1 Expression of the full length YB-1 protein using GST systems (Amersham- Pharmacia, Sweden) 27

3.6.2 Double digestion of pGEX-2T~YB1 28

3.6.3 Identification of GST fused YB-1 expressed in *E. coli* 28

3.7 Cellular Biology Techniques 29

3.7.1 Cell lines 29

3.7.2 Preparation of medium 29

3.7.3 Thawing cells 29

3.7.4 Sub-culturing of cells 30

3.7.5 Cryopreservation of cells 30

3.7.6 Preparation of total cell lysate 31

3.7.7 RIPA lysis method 31

3.7.8 Transfection of adherent cells using Calcium Phosphate method 32

3.8 Protein analysis 33

3.8.1 Sodium Dodecyl Sulphate–Polyacrylamide Gel Electrophoresis (SDS-PAGE) 33

3.8.2 Coomassie Brilliant Blue Staining 36

3.8.3 Western blot analysis 36

4.0 RESULT 38

4.1 Plasmid confirmation of pGEX~YB-137 38

4.2 Plasmid confirmation of pREP9~YB-1) 39

4.3 Expression of pGEX-2T~YB1 in *E. coli* (BL–21) 40

4.4 Identification of 42 kDa nuclear YB-1 protein in HeLa cells 41

4.5 Over-expression assay with pREP9~YB-1 42

5.0 DISCUSSION 44

5.1 Plasmid confirmation of pREP9~YB-1 44

5.2 Expression of pGEX-2T~YB1 in *E. coli* (BL–21) 45

5.3 Over-expression assay with pREP9~YB-1 46

5.4 Different expression size of YB-1 protein in bacterial system and mammalian system 47

5.5 Calcium phosphate transfection in over-expression assay 48

5.5.1 Factors affecting efficiency of transfection 48

5.5.2 Troubleshooting on poor expression in HeLa cells 50

6.0 CONCLUSION 51

7.0 REFERENCES 52

LIST OF TABLES

Table 2.1: Proposed functions for Y-Box proteins in different organisms..... 8

Table 3.1: General reagents, solutions, buffers and media 19

Table 3.2: DNA and vectors used in this research 20

Table 3.3: Components of restriction enzyme digestion 22

Table 3.4: Standard media and solution used in research 23

Table 3.5: List of bacterial strains and their features 24

Table 3.6: Stock solution for SDS-PAGE preparation 34

Table 3.7: Formulation for SDS-PAGE Separating Gels 35

Table 3.8: Stacking gel recipe (for 2 gels preparation) 35

Table 3.9: List of antibodies used in this research 37

LIST OF FIGURES

Figure 1.1: Schematic diagram of parental vector pGEX series used in this research 6

Figure 2.1: Domain structure of human dbpA 11

Figure 2.2: Domain structure of human dbpB 13

Figure 2.3: Schematic diagram of pGEX-2T~YB1 16

Figure 2.4: Schematic diagram of vector pREP9 used in this research 17

Figure 4.1: Verification of the pGST~YB-1 expression vector, by double digestion of pGEX2TK~YB-1 with *Eco R1* and *Bam H1* 38

Figure 4.2: Verification of the pREP9~YB-1 expression vector, by double digestion of pREP9~YB-1 with λ *Hind III* and *Xho I* 39

Figure 4.3: Expression of the GST~YB-1 protein in *E. coli* BL-21 40

Figure 4.4: Detection of YB-1 in HeLa cell lysate 41

Figure 4.5: Western analysis result of over-expression assay with pREP9~YB-1 in HeLa cells 42

Figure 4.6: Coomassie brilliant blue staining result of over-expression of the pREP9~YB-1 protein in HeLa cells 43

LIST OF SYMBOLS

α	Alpha
β	Beta
bp	Base pair
λ	Lambda
°C	Degree celsius
Da	Dalton
kDa	Kilo dalton
kb	Kilo base pair
g	Gram
L	Liter
mL	Mili liter
mM	Mili molar
mA	Mili ampere
nm	Nano meter
rpm	Revolutions per minute
µg	Micro gram
µL	Micro liter
(v/v)	Volume per volume
V	Volt
(w/v)	Weight per volume
%	Percentage

LIST OF ABBREVIATIONS

AIS	Adenocarcinoma *in situ*
APS	Ammonium persulfate
cDNA	Complementary deoxyribonucleic acid
CIN	Cervical intraepithelial neoplasia
CSD	Cold Shock Domain
CTCF	CCCTC-binding factor (zinc finger protein)
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
E. Coli	*Escherichia coli*
EDTA	Ethylene diamine tetra protein
FBS	Fetal bovine serum
HeLa	Henrietta Lack. (an immortal cell line)
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria-Bertani
OD	Optical density
PBS	Phosphate buffered saline
PVDF	Polyvinylidene fluoride
RNA	Ribonucleic acid
SIL	Squamous intraepithelial lesion
SDS	Sodim dodecyl sulfate
SDS-PAGE	Sodim dodecyl sulfate-polyacrylamide gel electrophoresis
TEMED	N,N,N',N',- tetramethylethylenediamine
YB-1	Y-box binding protein 1

"OVER-EXPRESSION" PROTEIN-1 Y-BOX IKATAN MANUSIA - PERBANDINGAN SISTEM PENGEKSPRESAN ANTARA MAMALIA DAN BAKTERIA

ABSTRAK

Protein "Y-Box 1" manusia (YB-1) merupakan sejenis protein yang terlibat dalam pengendalian transkripsi dan translasi dalam rangkaian gen yang besar. YB-1 melibatkan regulasi fungsi sel yang berbagai-bagai. Kajian dahulu menunjukkan pengekspresan YB-1 ditingkatkan dalam banyak kes kanser dan menunjukan hubungan yang rapat dengan prognosis kanser serta memberi rintangan terhadap ubat kanser. Bagi lebih memahami dan mendalami peranan YB-1 dalam regulasi transkripsi and translasi perkembangan barah dan kanser, satu ujian "over-expression" telah dilakukan. Dua plasmid gabungan yang berbeza, pGEX-2TK~YB-1 and pREP~YB-1 telah digunakan bagi melakukan ujian "over-expression" dalam bakteria (*Escherichia coli,* BL-21) dan sistem mamalia (sel-sel HeLa). Perbandingan telah dilakukan dengan "Blot Western" menggunakan antibody poliklonal α YB-1. Keputusan "over-expression" menunjukan perbezaan pergerakan protein YB-1 dalam sistem pengekspresan bakteria dan mamalia jika dibandingkan dengan saiz nilai teorinya. Sementara itu, ujian "over-expression" melibatkan transfeksi pREP~YB-1 ke dalam sel-sel HeLa dengan menggunakan kaedah transfeksi kalsium fosfat. Disebabkan kekurangan hasil, dan kurangnya kajian kuantitatif terhadap mekanisma dalam transfeksi kalsium fosfat, pengenalpastian masalah telah dilakukan dan langkah optima telah dicadangkan dalam perbincangan kajian ini.

OVER-EXPRESSION OF HUMAN Y-BOX BINDING PROTEIN-1: A COMPARISON BETWEEN MAMMALIAN AND BACTERIAL EXPRESSION SYSTEM

ABSTRACT

The human Y-box protein 1 (YB-1) is a class of proteins involved in transcriptional and translational regulation of a wide range of genes. YB-1 is implicated in regulation of multiple cellular functions. Previous studies found that the YB-1 expression is upregulated in many tumours and is highly correlated with adverse cancer prognosis and resistance to cancer drugs. In order to further explore and understand the role of YB-1 in terms of transcriptional and translational regulation in development of malignancy tumours, an over-expression assay has been carried out. Two different recombinants YB-1 plasmid, pGEX-2TK~YB-1 and pREP~YB-1 were used to carry out over-expression assay in bacterial (*Escherichia coli,* BL-21) and mammalian system (HeLa cancer cells) respectively. The comparison was analyzed by western blot analysis using α YB-1 polyclonal antibody. The over-expression results show a different migration of YB-1 protein in bacterial and mammalian expression system compared to their actual theoretical size. Meanwhile, over-expression assay involved transfection of pREP~YB-1 into HeLa cells using calcium phosphate transfection method. Due to poor reproducibility and rare quantitative studies on the mechanisms of the calcium phosphate transfection, trouble shooting was done and the optimization steps were suggested in the discussion of this research.

1.0 LITERATURE REVIEW

1.1 Calcium Phosphate Transfection

The introduction of desirable genetic sequences into mammalian cells is an essential tool for analysis of gene structure. Transfection, the introduction of functional foreign DNA into the cell nucleus with the aim of repairing missing cell function and to provide means to enhance or silence gene expression is currently used extensively in the laboratory and is fast becoming a therapeutic reality. Viral infection is the oldest method for gene transfer, which was first demonstrated on *Salmonella* in 1952 (Zinder and Lederberg, 1952). The ability to introduce nucleic acids into cells has enabled the advancement of our knowledge of genetic regulation and protein function within eukaryotic cells, tissues and organisms. For gene transfer into cells, different viral vectors based on retroviruses, adenoviruses and other viruses were used. It is the most efficient method for transfering of DNA into cells, but it has serious drawbacks such as the risk of recombination, strong immunogenicity and carcinogenity (Burand *et al.*, 1980; Crystal, 1995; Tripathy *et al.*, 1996).

Progress in transfection technology was relatively slow until the advent of molecular biology techniques for cloning plasmid DNA. These techniques provided the means to prepare and manipulate DNA sequences and the ability to prepare virtually unlimited amounts of relatively pure DNA for transfection experiments. Cloned sequences could also be used to generate RNA *in vitro* with phage RNA polymerase using DNA templates with the corresponding polymerase promoter (Melton *et al.*, 1984). As the ability to prepare DNA and RNA for transfection became easier, additional methods, such as

electroporation (Wong and Neumann, 1982), liposome-mediated transfer (Fraley *at el.*, 1980), polymeric [for example dendrimers (Haensler and Szoka, 1993) and poly ethylenimine (Godbey *et al.*, 1999)] and inorganic vectors (Chowdhury and Akaike, 2005) were developed to enable more efficient transfer of the nucleic acids to a broad range of cultured mammalian cells.

Out of the latter group, calcium phosphate is by far the most widely used and studied biomaterial because it can transfect a wide variety of mammalian cells *in vitro* (Jordan *et al.*, 1996) and avoids issues such as immune responses (Crystal, 1995) and high toxicity associated with other techniques (Kim *et al.*, 2005). Besides that, an *in situ* precipitation of the inorganic salt in the presence of DNA gives nano-particles which cells can immediately take up (Graham and van der Eb 1973a). The method is very easy and inexpensive, but the transfection efficiency is inferior to commercially available transfection agents, which are based on liposomes (Ishii *et al.*, 1997) and polymers (Tang *et al.*, 1996).

Calcium phosphate transfection utilizes the basic procedure of Graham and van der Eb (1973b). The basic feature of this procedure include mixture of DNA with calcium chloride and sodium phosphate in buffered saline, formation of calcium phosphate-DNA complex, and incubation with cultured cells. After removal of the DNA containing medium, a brief incubation with dimethyl sulfoxide or glycerol is usually employed to enhance DNA uptake (Graham and van der Eb, 1973a).

For most research involving transfection, the transfected gene is only transiently expressed. The DNA introduced in the transfection process is usually not inserted into the nuclear genome, the foreign DNA is lost at the later stage when the cells undergo mitosis. Meanwhile for stable transfection, the transfected gene actually remains in the genome of the cell and its daughter cells. To accomplish this, another gene is co-transfected, which gives the cell some selection advantage, such as resistance towards a certain toxin. Some of the transfected cells will, by chance, have inserted the foreign genetic material into their genome. If the toxin, towards which the co-transfected gene offers resistance, is then added to the cell culture, only those few cells with the foreign genes inserted into their genome will be able to proliferate, while other cells will die. After applying this selection pressure for some time, only the cells with a stable transfection remain and can be cultivated further.

A reporter gene or plasmid is needed in order to observe the successful of transfection. In 1982, Gorman *et al.* (1982) initiated the reporter gene concept with the bacterial chloramphenicol acetyltransferase (CAT) gene and associated CAT assay system (Gorman *et al.*, 1982). By using a reporter gene that is not endogenous to the cell coupled with a sensitive assay system for that gene product, it allows researcher to clone regulatory sequences of interest upstream of the reporter gene to study expression of the reporter gene under various conditions. Nowadays a common reporter gene is GFP (green fluorescent protein), which is a powerful "tool" to express the protein in small sets of specific cells. This allows researchers to optically detect specific types of cells *in vitro* (in a dish), or even *in vivo* (in the living organism).

1.2 Transformation

Transformation involved transfer of genetic information between bacteria by means of "naked" intracellular DNA fragments derived from bacterial donor cells and incorporated into a competent recipient cell. Transformation was first demonstrated in 1928 by Frederick Griffith, an English bacteriologist searching for a vaccine against bacterial pneumonia. Griffith discovered that a non-virulent strain of *Streptococcus pneumoniae* could be transformed into a virulent one by exposure to strains of virulent *S. pneumoniae* that had been killed with heat. Meanwhile, transformation of *E. coli* was first described by Mandel and Higa (1970). Subsequent modifications to improve transformation efficiencies have included prolonged exposure of cells to calcium chloride (Dagert and Ehrlich, 1979), substitution of calcium with other cations such as Mn^{2+}, and K^{+}, and addition of other compounds such as dimethyl sulfoxide, dithiothreitol, and cobalt hexammine chloride (Hanahan, 1983).

Transformation is always involved with the uptake of DNA by a competent recipient cell. Competence refers to the state of being able to take up exogenous DNA from the environment in transformation. Two different forms of competence should be distinguished: natural and artificial. Natural competence involve bacteria which naturally capable of taking up DNA under laboratory conditions; many more may be able to take it up in their natural environments. Such species carry sets of genes specifying machinery for bringing DNA across the cell membrane or membranes (Chen and Dubnau, 2004). Meanwhile, artificial competence is not encoded in the cell genes. Instead it is induced by laboratory procedures in which cells are passively made permeable to DNA, using conditions that do not normally occur in nature (Chen and Dubnau, 2004).

1.3 HeLa cancer cells

A HeLa cell (also Hela or hela cell) is an immortal cell line used in medical research. The cell line was derived from cervical cancer cells taken from Henrietta Lacks, who died from her cancer on October 4, 1951. Henrietta's cells, named HeLa after the first letters in Henrietta and Lack. It became the first human cells to live indefinitely outside the body. HeLa cells helped researchers understand the differences between cancerous and normal cells, and quickly became a standard laboratory tool for studying the effects of radiation, growing viruses and testing medications. HeLa is still one of the most widely used cell lines (Skloot, 2001).

1.4 Glutathione-S-Transferase (GST) gene fusion system

GST expression system was originally established to overcome problems with coupling, purification and detection of proteins. This system employs a pGEX vector (Figure 1.1), a vector based on the original pSj5 plasmid, which directs the synthesis of foreign proteins in *E. coli* as a fusion with the C domain of *Shcistosoma japonicum* 26 (Sj 26) protein, a 26 kDa Glutathione S transferase (GST) (Smith and Johnson, 1988). Expression of all pGEX vector series is under the control of IPTG inducible *trp-lac* (*tac*) promoter (Smith and Johnson, 1988).

The first pGEX vector ever developed was pGEX-1, which contains: *tac* promoter (Amann *et al.*, 1983; De Boer *et al.*, 1983) followed by the complete coding sequence of Sj26 (Smith *et al.*, 1986), in which the normal termination codon is replaced by a polylinker containing a unique recognition site for *Bam* HI, *Sma* I and *Eco* RI, and followed by TGA translation termination codons in all three reading frames. There are 10 different pGEX

pGEX vectors available at present (Amersham-Pharmacia Biotech Inc, USA) and all of them contain the reading frame at different location in relation to multiple cloning site to allow introduction of specific sequences such as sites for protein cleavage with Factor Xa (pGEX-3X, pGEX-5X-1, pGEX-5X2, pGEX-5X3) or thrombin (pGEX-2T, pGEX-2TK).

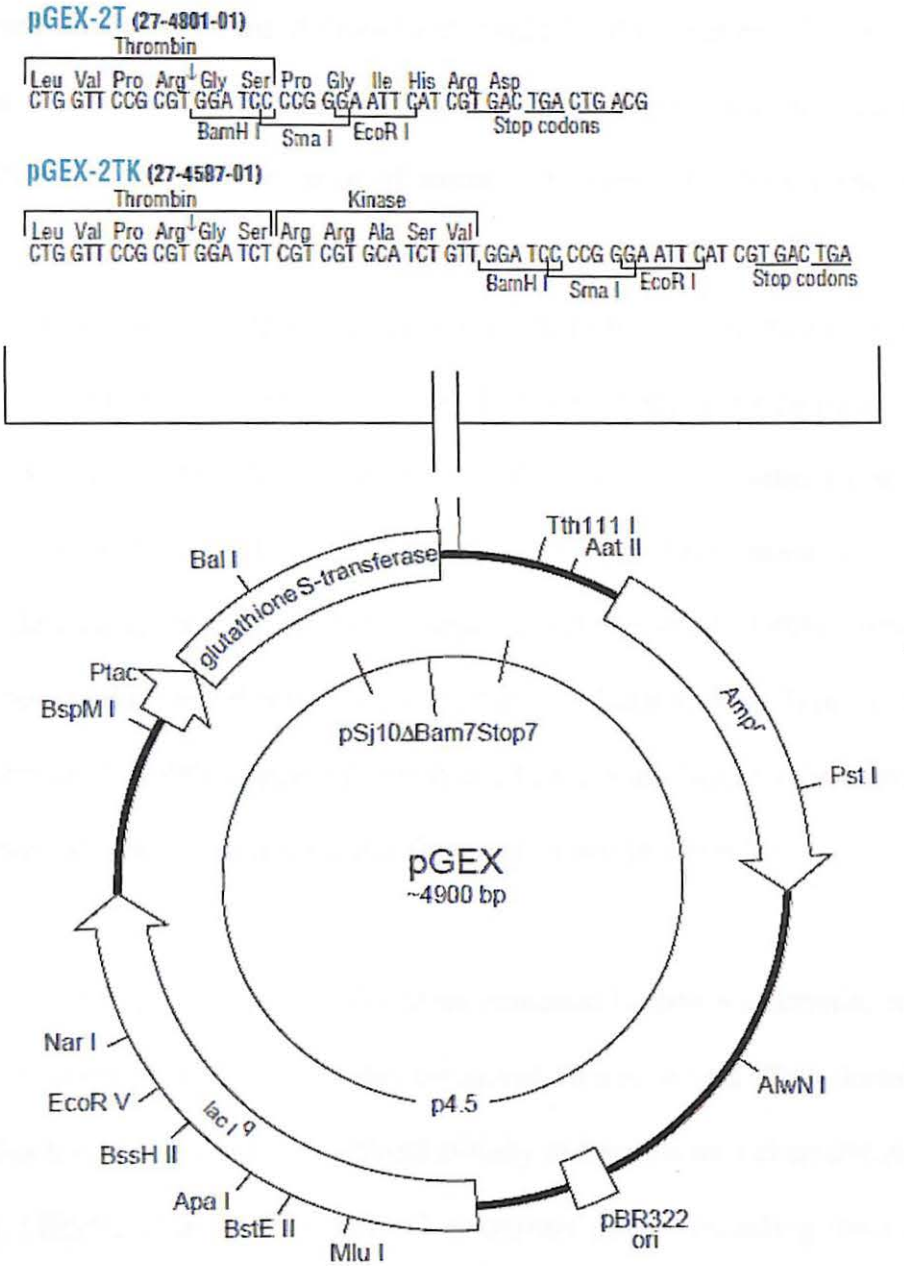

Figure 1.1: Schematic diagram of parental vector pGEX series used in this research

2.0 INTRODUCTION

2.1 Y-box binding protein 1, YB-1

The human Y-box protein 1 (YB-1) is a member of the Y-box protein family, a class of proteins involved in transcriptional and translational regulation of a wide range of genes. Proteins of this family are active as transcriptional and translational regulators for messenger RNA and for a wide range of genes that contain the Y-box sequence (50-CTGATTGGCCAA-30) in their promoters (Ladomery, 1997). The name Y-box protein comes from the ability of the cold shock domain (CSD) to bind to the Y-box sequence [5'-CTGATTGG – 3'] of DNA, which is an inverted CCAAT box, in the promoter region of many genes (Wolffe, 1994). The human YB-1 cDNA was first isolated by screening a human lymphoblastoid cell cDNA library expressed in *Escherichia coli* with an oligodeoxynucleotide containing the Y-box sequence (Didier *et al.*, 1988). Subsequently, additional Y-box binding proteins have been identified in human, frog (Tafuri and Wolffe, 1990), rat (Ozer *et al.*, 1990), mouse (Tafuri *et al.*, 1993), and chicken (Grant and Deeley, 1993). Examples of other proteins from this family are given in Table 2.1.

The Y-box protein family contains three domains: N-terminal domain, cold shock domain and C-terminal domain. The highly conserved 70 amino acid DNA domain, the so called 'cold shock domain' CSD was defined initially in bacteria as a characteristic feature of this family (Wolffe *et al.*, 1992). Several eukaryotic genes, including those encoding major histocompatibility complex class II antigens, thymidine kinase, proliferating cell nuclear antigen, DNA polymerase a, epidermal growth factor receptor, DNA topoisomerase

IIa, and multidrug resistance 1 protein (MDR1) contain a Y-box in their regulatory regions (Ladomery and Sommerville, 1995; Wolffe, 1994; Wolffe *et al.*, 1992).

Table 2.1: Proposed functions for Y-Box proteins in different organisms

Organism	Protein	Proposed function	Reference
Man	YB-1	Transcriptional repressor DNA repair	(Didier *et al.*, 1988; Hasegawa *et al.*, 1991)
Rat	EF1	Transcription activator	(Ozer *et al.*, 1990)
Mouse	NSEP1 MUSYB1 MSY1	Chromatin structural protein/ss DNA recognition Transcriptional activator Translational masking protein/general mRNA packaging	(Gai *et al.*, 1992; Kolluri *et al.*, 1992; Tafuri *et al.*, 1993)
Xenopus	FRGY2 MRNAP4 YB3 FRGY1	Transcriptional activator Translational masking protein/general mRNA packaging Translational repressor Transcriptional activator Transcriptional activator	(Tafuri and Wolffe, 1990; Tafuri and Wolffe, 1992) (Cohen and Reynolds, 1991; Murray *et al.*, 1992; Tafuri and Wolffe, 1990)
E. coli	CS 7.4	Transcriptional activator (cold shock stress)	(La Teana *et al.*, 1991)
B. subtilis	CspB	Cold shock response	(Willimsky *et al.*, 1992)

The Y-box proteins were first identified and isolated by binding to a DNA probe containing the Y-box sequence (Dorn *et al.*, 1987). It has been demonstrated that binding of the Y-box proteins to the intact Y-box sequence involves specific interactions (Didier *et al.*, 1988). Subsequent binding studies have shown that the Y-box proteins can bind to a whole range of nucleic acids; namely, single-stranded DNA (ssDNA), RNA, damaged DNA, RNA and even triple-stranded DNA (Horwitz *et al.*, 1994; Koike *et al.*, 1997). Although no consistent picture emerged from the different binding studies, it appears that Y-box proteins have a preference for binding ssDNA, especially when it contains the Y-box sequence, evidence exists that Y-box proteins have a preference for binding pyrimidine rich DNA sequences (Wolffe, 1994).

Apart from its role in transcription, YB-1 is a multifunctional protein that affects splicing, translational control, and repair of damaged DNA by interacting with several repair proteins (Gaudreault *et al.*, 2004; Kohno *et al.*, 2003). Sakura *et al.* (1988) further isolated two homologues of this gene, dbpA and dbpB, from human placental cDNA expression libraries using the *c-erb2* gene promoter as the probe. Sequence analysis revealed that dbpB was identical to the YB-1. Kudo *et al.* (1995) isolated two clones identical to dbpA and dbpB (YB-1) cDNA after screening a λ-phage expression library with a DNA probe containing the human leukosialin CD43 promoter sequence which were found to have expression pattern are not restricted to only hematopoietic cells. The dbpA was located on the chromosome 12p13 (Kudo *et al.*, 1995) and dbpB on chromosome 1p34 (Makino *et al.*, 1996). The dbpA is expressed in skeletal and heart muscles and the dbpB (YB-1) is found in skeletal muscle, kidney, lung and liver cells (Kudo *et al.*, 1995). YB-1 protein is mainly located in the cytoplasm of the cell lines (Koike *et al.*, 1997). Interestingly

it was detected in both the cytoplasm and nucleus in various types of human cancers (Ohga *et al.*, 1996).

2.1.1 *dbpA* : organization, structure and function

All Y-Box proteins contain three domains: the N terminal domain, CSD and C terminal domain. The dbpA gene (Figure 2.1) consists of 10 exons spanning 24 kb of the genomic DNA (Kudo *et al.*, 1995). Its N terminal domain is encoded by the first exon, the CSD is encoded by the 2^{nd} to the 5^{th} exon and the C terminal domain by the 5^{th} to 8^{th} exons (Figure 2.1). However, the exon sequences of the N terminal domain of dbpA are found to have only 15% homology with the dbpB's (YB-1) N-terminal domain, whereas its C terminal has 30% homology compared to CSD, which shows 96% sequence homology with its YB-1 counterpart (Toh *et al.*, 1998).

Despite the significant differences between the N terminal domain and the C terminal domain of YB-1 and the dbpA protein, the dbpA protein binds to the consensus 5' - CTGATTGG-3' Y-Box element since the binding occurs via the highly conserved CSD domain. The dbpA in HeLa cells is localized mostly in the cytoplasm, especially in perinuclear region (Kudo *et al.*, 1995). However, despite extensive biochemical and structural characterization of dbpA, the function of this protein is still not clear (Kudo *et al.*, 1995).

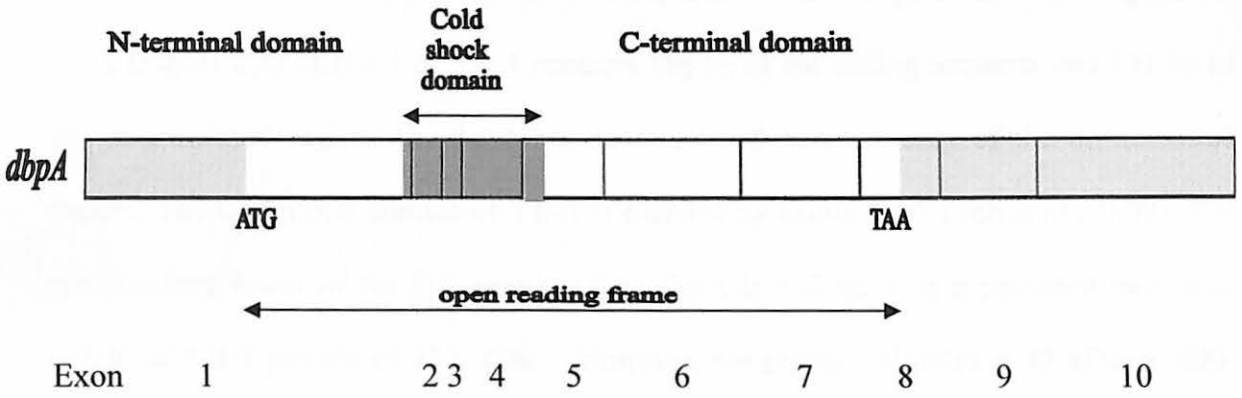

Figure 2.1: Domain structure of human *dbpA*. Light and dark shading represent non-coding sequences and the cold shock domains (CSD), respectively. Adapted from Toh *et al.,* (1998)