CHARACTERIZATION OF THE NOVEL TRANSCRIPTIONAL REGULATOR HUMAN MESODERM INDUCTION EARLY RESPONSE GENE 1 (hMI-ER1): IT PROMOTERS, INTERACTING PROTEINS AND TRANSCRIPTIONAL REGULATORY FUNCTIONS

ZHIHU DING

Your file Votre référence
ISBN: 0-612-99031-1
Our file Notre référence
ISBN: 0-612-99031-1

CHARACTERIZATION OF THE NOVEL TRANSCRIPTIONAL REGULATOR HUMAN MESODERM INDUCTION EARLY RESPONSE GENE 1 (hMI-ER1): ITS PROMOTERS, INTERACTING PROTEINS AND TRANSCRIPTIONAL REGULATORY FUNCTIONS

By

© Zhihu Ding

A thesis submitted to the

School of Graduate Studies

in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

Division of Basic Medical Sciences

Faculty of Medicine

Memorial University of Newfoundland

May, 2004

St. John's **Newfoundland**

ABSTRACT

Mesoderm induction early response 1 (mi-er1), previously called *er1*, was first isolated as a novel fibroblast growth factor-regulated early-response gene from *Xenopus* embryonic cells induced to differentiate into mesoderm. The human orthologue of *xmi-er1, hmi-er1,* was found to have ubiquitous, but low level, expression in normal human tissues (Paterno *et al.*, 1998; Paterno *et al.*, 2002). Breast carcinoma cell lines and tumours, on the other hand, showed elevated levels (Paterno *et al.*, 1998), suggesting that *hmi-er1* expression is associated with the neoplastic state in human breast carcinoma. Structurally, hMI-ER1 has conserved domains found in a number of transcriptional regulators, including an acid activation domain (Paterno *et al.*, 1997), an ELM2 domain (Solari *et al.*, 1999) and a signature SANT domain (Aasland *et al.*, 1996).

My hypothesis is that hMI-ER1 may be a potent transcriptional regulator, and deregulation of its expression and/or functions may contribute to tumorigenesis. The purpose of this project was to: (1) isolate and characterize the *hmi-er1* promoters; (2) investigate the role of hMI-ER1 in transcriptional regulation; and (3) identify and characterize hMI-ER1-interacting proteins.

Cloning and sequence analysis of one of the *hmi-er1* promoters, P2, showed the absence of a TATA box, but presence of a CpG island, multiple Sp1 binding sites, and other factor binding sites which may be important in the regulation of gene expression. Functional characterization of the promoter revealed that Sp1 binds to the promoter and plays a positive role in the regulation of the P2 promoter activity of *hmi-er1*. hMI-ER1

was found to function as transcriptional repressor on the G5TKCAT promoter by recruitment of HDAC1 *via* the ELM2 domain. hMI-ER1α and β also function as transcriptional repressors on their own P2 promoter, through interaction with Sp1 and interference with the Sp1-DNA binding *via* the SANT domain. In order to better understand the function of hMI-ER1, an attempt was made to identify more hMI-ER1 interacting proteins by yeast two-hybrid cDNA library screening. HSP40 and TRABID were identified as proteins that interact specifically with hMI-ER1α and β. Furthermore, hMI-ER1α and β were found to interact with the tumour suppressor RB and may participate in cell growth regulation. These results suggest that hMI-ER1 may function as a transcriptional regulator, through distinct mechanisms, to regulate cellular functions.

ACKNOWLEDGEMENT

I am indebted to my supervisor, Dr. Gary Paterno, for his supervision, support, help and encouragement throughout the entire course of this study. I am also very grateful to Dr. Laura Gillespie for her constant support and help. Without her, my thesis would not be the same. I would also like to extend my thanks to my committee members, Dr. Ken Kao and Dr. Helene Paradis for their expertise and invaluable comments, which helped put this work into perspective. My thanks also are due to Dr. Alan Pater, who was supervisor of my M.Sc. Degree, for his help. I am grateful to Paula, Corinne, and Yuan for their excellent technical assistance. GST pull-down assays with the RB deletion mutants in Chapter 5 was carried out by Corrine. The colony formation efficiency assays of HeLa Tet-On cells were performed by Yuan. I wish to extend thanks to Artee, Janine, Yoella, Marianne, Ivy, Tina and Gord for their help and making work fun.

My thanks are due to Dr. Richard D. Neuman, Dr. Penny Moody-Corbett, the Faculty of Medicine, and the School of Graduate Studies for their cordial administrative help and support throughout this program. Acknowledgement is also due to the F.A. Aldrich Fellowship Foundation, the Albert George Hatcher Memorial Scholarship Foundation, and the Colman Graduate Student Award Foundation for the support.

My love and appreciation go to my wife, Nan Li, who did most of the housework when I was working in the evenings and on weekends. My love also goes to Lee J. Ding, my son for being his own little self. His lovely smile always made me forget the frustration of some unexpected experimental results and the weariness of thesis writing.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

µg	micrograms
µl	microlitres
3-AT	3-amino-1,2,4-triazole
aa	amino acid
AD	transcriptional activation domain
ADA	the adaptor coactivator
ALT	alternative telomere
AP1	activator protein-1
APAF-1	apoptotic proteinase activating factor-1
ARC	activator-recruited cofactor
β-gal	β-galatosidase
BHC	BRAF-histone deacetylase complex
BMP	bone morphogenetic protein
bp	base pair
BSA	bovine serum albumin
C	cytosine
CaM	calmodulin
CAT	chloramphenicol transferase
CBP	CREB-binding protein
CDK	cyclin dependent kinases
CDP	CAAT-displacement protein
cDNA	complementary deoxyribonucleic acid
ChIP	chromatin immunoprecipitations
CpG	5'-carbon of a cytosine preceding a guanine
CRE	cAMP response element
CREB	cAMP response element binding protein
CRSP	cofactor required for Sp1 activation
CPSF	cleavage and polyadenylation specificity factor
CTD	carboxyl-terminal domain of the largest subunit of RNA polymerase II
CstF	cleavage stimulation factor F
DAG	diacylglcerol
DBD	DNA-binding domain
DMEM	Dulbecco s modified Eagle s medium
DMSO	dimethl sulfoxidae
DNA	deoxyribonucleic acid
DNMT	DNA cytosine methyltransferases
dNTPs	deoxyribonucleotide triphosphates
DPE	downstream promoter element
DREAM	downstream regulatory element antagonist modulator

DRIP	vitamin D receptor-interacting protein
DSE	downstream sequence element
DTT	dithiothreitol
ECs	endothelial cells
ECL	enhanced chemiluminescence
ECM	extracellular matrix
EDTA	ethyldiamine tetraacetic acid
EGF	epidermal growth factor
ELM2	EGL27 and MTA1 homology 2
EMSA	electophoretic mobility shift assay
ER	estrogen receptor
ERα	estrogen receptor α
ER1	early-response gene 1
ERE	estrogen response element
ERK	extracellular signal-regulated kinase; ERK1 = p44; ERK2 = p42
EST	expressed sequence tag
FACS	fluorescence activated cell sorting
FADD	Fas-associating protein with death domain
FCS	fetal calf serum
FGF	fibroblast growth factor
FGFR	fibroblast growth factor receptor
G	guanine
g	gram
G1	first gap cell cycle phase
GAL4-α	pM plasmid containing full-length *hmi-er1α* in frame with GAL4 DNA binding domain
GAL4-β	pM plasmid containing full-length *hmi-er1β* in frame with GAL4 DNA binding domain
GRB2	growth factor receptor bound protein 2
GST	glutathione S-transferase
GST-α	GST plasmid (pGEX-4T-1) containing full-length *hmi-er1α*
GST-β	GST plasmid (pGEX-4T-1) containing full-length *hmi-er1β*
H	histone
h	hour
HAT	histone acetyltransferase
HDAC	histone deacetylase
HP1	heterochromatin protein 1
hmi-er1	human homologue of *xmi-er1*
HRP	horseradish peroxidase
iMI-ER1	inhibitory proteins of hMI-ER1
Inr	initiator
IP	immunoprecipitation
IP_3	inositol 1,4,5-trisphosphate
ISWI	imitation switch

kb	kilobase
kDa	kilodalton
kp	kilobase pair
M	molar
mAb	monoclonal antibody
MAPK	mitogen activated protein kinase
MAPKK	MAPK kinase
MAPKKK	MAPK kinase kinase
MBD	methyl-CpG-binding domain
MBPs	methyl-CpG binding proteins
MBT	mid-blastula transition
MeCP	methyl-CpG-binding protein
MEK	MAPK/ERK kinase
mi-er1	mesoderm induction early-response 1
min	minute
ml	millitres
mm	millimeter
mM	millimolar
mRNA	messenger ribonucleic acid
MTA	metastasis-associated protein
mta1	metastasis-associated gene 1
NAD^+	nicotinamide adenine dinucleotide
N-COR	nuclear receptor corepressor
NES	nuclear export signal
ng	nanogram
NLS	nuclear localization signal
NPCs	nuclear pore complexes
NRS	neural restrictive silencer
NURD	nucleosome remodelling and histone deacetylating
nt	nucleotide
^oC	degrees Celsius
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBAF	Polybromo and BRG1-associated factor
PBS	phosphate buffered saline
PCAF	p300-CBP-associated factor
PCR	polymerase chain reaction
PI3K	phospho inositide 3 kinase
PKA	protein kinase A
PKB	protein kinase B
PKC	protein kinase C
PLCγ	phospholipase C-γ
PMSF	phenylmethylsulfonyl fluoride
pMyc	Myc-tagged plasmids (CS3+MT)

pMyc-α	Myc-tagged plasmids (CS3+MT) containing full-length *hmi-er1α*
pMyc-β	Myc-tagged plasmids (CS3+MT) containing full-length *hmi-er1β*
PTP	permeability transition pore
pTRE2-α	pTRE2 containing full-length *hmi-er1α*
pTRE2-β	pTRE2 containing full-length *hmi-er1β*
RACE	rapid amplification of cDNA ends
RB	retinoblastoma protein
RBAP	retinoblastoma associated protein
RE1	repressor element 1
RNA	ribonucleic acid
ROS	reactive oxygen species
rpm	revolutions per minute
RT	room temperature
rtTA	reverse tetracycline transaction
RT-PCR	reverse transcription PCR
S	synthesis cell cycle phase
SAGA	SPT-ADA-GCN5-acetyltransferase
SAGE	serial analysis of gene expression
SANT	domain was first found in SWI3, ADA2, N-COR and TFIIIB
SAP	Sin associated protein
SD	synthetic dropout yeast culture medium
SD/-Trp/-Leu/-His	synthetic dropout yeast culture medium lacking tryptophan, leucine, and histidine
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
sec	seconds
SET	enhancer-of-zeste
SH2	src homology domain 2
SH3	src homology domain 3
SHC	SH2 domain-containing 2 collagen related
SID	Sin3A interaction domain
SMRT	silencing mediator of retinoic acid and thyroid hormone receptor
SOS	son of sevenless
TAF	TBP-associated factor
TBP	TATA binding protein
Tet-Off	tetracycline uninduced gene expression
Tet-On	tetracycline induced gene expression
*tet*O	tetracycline repressor operator sequence
TetR	tetracycline repressor
TF	transcription factor
TK	thymidine kinase
TNF	tumour necrosis factor
TNFR1	TNF receptor-1
TR	thyroid hormone

TRABID	TRAF-binding domain
TRADD	TNFR1-associated death domain protein
TRAFs	TNF receptor associated factors
TRAP	thyroid hormone receptor-associated protein
TSA	trichostatin A
UTR	untranslated region
VDR	vitamin D_3 receptors
VEGF	vascular endothelial growth factor
Xbra	*Xenopus* homologue of the *Brachyury* gene
X-gal	5-bromo-4-chloro-3-indoyl-b-D-galactopyranoside
xmi-er1	*Xenopus* mesoderm induction early-response 1

LIST OF APPENDICES

SECTION I

CHAPTER 1 GENERAL INTRODUCTION

1.1 Development of cancer

In mammals and other multicellular organisms, normal cells are controlled by their surroundings through signals such as growth factors, extracellular matrix components or ligands on neighbouring cells that dictate their behaviour: to grow, to stop growing, to die and so on. A normal cell also has its own signal transduction network to integrate and respond to the external messages so that the cell is in harmony with its environment (reviewed in Hanahan and Weinberg, 2000). Misregulated signal transduction pathways can result in the development of cancer, which is characterized by the uncontrolled growth of cells (reviewed in Hanahan and Weinberg, 2000).

Cancer is a genetic disease caused by accumulation of mutations. Genetic mutation leading to cancer development comes from DNA damage or errors made by the DNA replication machinery (reviewed in Hoeijmakers, 2001). A mutation(s) in a single cell may cause the cell to divide abnormally, forming a mass of localized benign tumor cells. Subsequent mutations in a subset of the benign tumor cells may generate a malignant tumor cell, and this cell continues to divide and the progeny may invade the basal lamina that surrounds the tissue. Some of these tumor cells spread into blood vessels that will distribute them to other sites in the body. Some of the tumor cells may exit from the blood vessels and grow at distant sites; a patient with such a tumor is said to

have cancer. It is thought that cancer development in humans is a multi-step process caused by the progressive accumulation of lesions in multiple genes governing vital cellular functions (reviewed in Hanahan and Weinberg, 2000; Vogelstein and Kinzler, 1993).

Normal cells have the ability not only to identify and repair DNA damage (reviewed in Hoeijmakers, 2001), but also to prevent the expansion of mutation-laden daughter cells by cell growth arrest (see section 1.2) or apoptosis (see section 1.3). As such, cells maintain genomic stability and prevent mutation accumulation. Evidence suggests that the acceleration of the development of cancer lies in the disruption of those genes maintaining genomic stability (reviewed in Hoeijmakers, 2001; Shiloh, 2003). Once these genes are inactivated, further mutations would easily follow (Loeb, 2001). Indeed, genetic defects that disrupt DNA-damage response and repair systems, have been shown to cause severe syndromes that are characterized by the predisposition to cancer (reviewed in Hoeijmakers, 2001). Disruption of the DNA-damage response and repair systems result in genomic instability, which may also develop during tumour progression. For instance, the ability of tumour cells to evade cell grow arrest (see section 1.2) or apoptosis (see section 1.3) is associated with genomic instability. The increasing frequency of mutational events accompanying tumour progression provides a driving force for several of the properties of malignancy, such as self-sufficient growth signals (see section 1.1), limitless replicative potential (see section 1.4), sustained angiogenesis (see section 1.5), tissue invasion and metastasis (see section 1.6).

1.1.1 Self-sufficient growth signals

It is thought that a normal somatic cell proliferates only in response to growth signals such as growth factors. In a cancer cell, the requirement for growth signals from their surroundings can be circumvented by three common molecular strategies: (1) enabling the cell to produce its own mitogenic signals; (2) through mutations activating growth factor receptors which lead to signalling in the absence of ligand; or, (3) through mutations in genes encoding components of the intracellular signalling network that translate those signals into action. The mutant forms of *ras* gene which encodes a constitutive activated protein, for example, is found in perhaps one quarter of all tumours in humans (reviewed in McCormick, 1991). These mutations release cells from dependence on exogenously derived signals and thus disrupt a critically important homeostatic mechanism that normally operates to ensure the appropriate behaviour of the various cell types within a tissue.

1.1.2 Disrupted growth regulatory circuits

Normal somatic cell proliferation not only depends on mitogenic signals, but is also usually limited by growth regulatory networks, such as the $p16^{INK4A}$-cyclin D-RB pathway and the $p19^{ARF}$-MDM2-p53 pathway (reviewed in Harbour and Dean, 2000; Sharpless and DePinho, 2002).

At the heart of the $p16^{INK4A}$-cyclin D-RB pathway is transcription factor E2F. E2F transcription factors regulate the expression of a number of genes that act in the G1 and S phases of the cell cycle. E2F activity is negatively regulated by RB through

multiple mechanisms (reviewed in Stevens and La Thangue, 2003). In turn, the ability of RB to inhibit the E2F activity is regulated through phosphorylation of RB by complexes consisting of cyclin dependent kinases (CDK) -4 and -6 and D type cyclins. Furthermore, the activity of CDK4 and 6 is negatively regulated by the CDK inhibitor $p16^{INK4A}$. Disruption of the RB pathway results in inactivating the RB protein that releases E2Fs and thus allows cell proliferation (reviewed in Harbour and Dean, 2000). Inactivation of RB is associated with a significant proportion of human cancers including familial retinoblastoma, cervical carcinomas, prostate carcinomas, and breast carcinomas (reviewed in Sellers and Kaelin, Jr., 1997). Moreover, in a number of cancers, functional inhibition of RB is achieved through inactivation of the *p16*^{INK4A} gene, by genetic lesions or by *p16*^{INK4A} promoter methylation (Sharpless and DePinho, 1999; Sherr, 2001). Overexpression of cyclin D1 has also been shown to inhibit RB (reviewed in Sicinski and Weinberg, 1997). Thus, the $p16^{INK4A}$-cyclin D-RB pathway is a central growth regulatory pathway. When this pathway is inactivated, cells acquire a selective growth advantage during oncogenic transformation.

p53 can arrest cell proliferation or induce apoptosis in response to DNA damage or inappropriate cell growth signals (reviewed in Sherr, 1998). p53 functions as a transcription factor that can activate the expression of a group of cellular genes involved in cell cycle arrest and induction of apoptosis. The importance of p53 in tumour suppression is underscored by the fact that p53 is probably the most frequently mutated gene in human cancer (reviewed in Sharpless and DePinho, 2002). Activity of p53 is negatively regulated by MDM2, a protein that has been detected in various human

tumours (reviewed in Momand and Zambetti, 1997). The $p19^{ARF}$ protein can bind MDM2, thereby preventing MDM2 from acting as a negative regulator of p53. Mutations involving the *p19*^{ARF} gene are again common events in cancer development. Impairment of p53 function by mutation, $p19^{ARF}$ inactivation, or by MDM2 overexpression is associated with increased tumour incidence (reviewed in Sharpless and DePinho, 2002). Thus, the $p19^{ARF}$-MDM2-p53 pathway is another central growth regulatory pathway whose inactivation is strongly selected for during oncogenic transformation.

1.1.3 Evasion of apoptosis

Apoptotic cell death can be characterized morphologically by blebbing of the plasma membrane, chromatin condensation, and phagocytosis by neighbouring cells (reviewed in Kerr *et al.*, 1972; Savill and Fadok, 2000). Along with the obvious morphological changes, distinct biochemical alterations are also associated with apoptosis. The most prominent is the induced activation of a family of cysteine proteases called caspases which cleave a variety of specific protein substrates (reviewed in Shi, 2002; Thornberry and Lazebnik, 1998). Caspases are crucial components of apoptosis pathways. They are normally present in the cell as zymogens that require proteolysis for activation of enzymatic activity. The mammalian caspases have been divided into upstream initiator caspases such as caspase-8 and caspase-9, and downstream effector caspases such caspase-3, based on their sites of action in the proteolytic caspase cascade (reviewed in Shi, 2002; Thornberry and Lazebnik, 1998). Binding of initiator caspase

precursors to activator molecules appears to promote procaspase oligomerization and autoactivation by enzymatic cleavage of the procaspase into fragments. Enzymatic activation of initiator caspases leads to proteolytic activation of downstream effector caspases and then cleavage of a set of proteins, resulting in destruction of the cell (reviewed in Shi, 2002; Thornberry and Lazebnik, 1998).

Two main apoptotic pathways leading to the execution of apoptosis have been identified: a mitochondria-dependent pathway and a parallel pathway involving activation of cell death receptors, such as those involved in Fas and tumour necrosis factor (TNF) receptor signalling (reviewed in Reed, 2000b). Some cytokines, such as TNF-α can bind to their receptors on the plasma membrane, causing trimerization of the receptors. Trimerized cell death receptors interact with death adaptor proteins such as TNF receptor-1 (TNFR1)-associated death domain protein (TRADD), which can cause activation of initiator caspases such as caspase-8 (Baker and Reddy, 1998; Nunez *et al.*, 1998).

The mitochondria-dependent pathway includes the release of the intermembrane space proteins such as cytochrome c (reviewed in Gross *et al.*, 1999). Once binding to cytochrome c, the apoptotic proteinase activating factor-1 (APAF-1) can form a complex with and then activate initiator caspase-9 (Li *et al.*, 1997; Zou *et al.*, 1997). The mitochondrial pathway is governed by bcl-2 family proteins (reviewed in Coultas and Strasser, 2003).

In multicellular organisms, homeostasis is maintained through a balance between cell proliferation and cell death (reviewed in Thompson, 1995). Mounting evidence

suggests that decreased cell death, especially by apoptosis, may result in an abnormal accumulation of cells during some of the tumour progression (reviewed in Reed, 2000a; Thompson, 1995). For example, studies in transgenic mice have confirmed the notion that the tumour suppressor role of p53 *in vivo* is closely linked to its ability to induce apoptosis by regulating the expression of members of the Bcl-2 family (Donehower *et al.*, 1995; Symonds *et al.*, 1994). A number of viral oncoproteins have also been shown to play roles in regulating apoptosis. Examples are the E1B of adenovirus and E6 of human papillomavirus (HPV). E1B and E6 can disable the p53 pathway in apoptosis, interfering with the RB-mediated cell death response to E1A or E7, respectively (Debbas and White, 1993; White, 1996).

1.1.4 Limitless replicative potential

The end of the human chromosome, the telomere, comprises an array of tandem repeats of the hexanucleotide 5 -TTAGGG-3 and its binding proteins (reviewed in Blackburn, 2001). These nucleoprotein structures function to anchor chromosomes within the nucleus, to protect chromosomes against exonucleases and ligases, and to prevent the activation of DNA-damage checkpoints (reviewed in DePinho and Wong, 2003; Maser and DePinho, 2002). Moreover, telomeres assist the replication of chromosomes. Conventional DNA replication machinery utilizes an RNA primer to initiate DNA synthesis. The removal of the terminal RNA primer after DNA synthesis is complete, results in the absence of the 5 end of the daughter DNA strand. The incompletely replicated telomeres will then be inherited by the daughter cells. In the

absence of a molecular mechanism to overcome this end replication problem , the telomeres progressively shorten as cells proceed through successive cell divisions. Eventually, the telomeres will be too short to continue to function as protective caps of the chromosomes, which are necessary for maintaining genomic stability (reviewed in DePinho and Wong, 2003; Maser and DePinho, 2002). Telomere shortening has been proposed as the mitotic clock marking the progress of a cell toward the end of its replicative life span, a process is evident by the consistent shortening of telomeres with aging of human tissues (Allsopp *et al.*, 1992). However, telomere maintenance is evident in virtually all types of cancer cells (Shay and Bacchetti, 1997). Ongoing maintenance of telomeres is thought to be a prerequisite for the indefinite proliferation of cells - the phenotype of cell immortalization - and is thus clearly an intrinsic part of the tumorigenesis (Sharpless and DePinho, 2004). The majority of cancers succeed in maintaining telomeres by upregulating expression of the telomerase enzyme, which adds the telomeric TTAGGG hexanucleotide repeats *de novo* onto the ends of telomeric DNA (Artandi and DePinho, 2000; Masutomi and Hahn, 2003). The remainder of cancer cells maintain telomeres though an alternative telomere (ALT) mechanism *via* recombination-based interchromosomal exchanges of sequence information, which in turn permits unlimited multiplication of descendant cells (Artandi and DePinho, 2000; Masutomi and Hahn, 2003).

1.1.5 Sustained angiogenesis

The process of new blood vessel formation is called angiogenesis. Angiogenesis is a complex multi-step process, which begins with local degradation of the basement membrane surrounding capillaries and then is followed by invasion of the surrounding stroma by the underlying endothelial cells in the direction of the angiogenic signals. Endothelial cell migration is accompanied by the proliferation of endothelial cells and their organization into three-dimensional structures that join with other structures to form a new blood vessel in tissues (Auerbach and Auerbach, 1994; Bussolino *et al.*, 1997).

The oxygen and nutrients in blood supplied by the vasculature are crucial for the survival of both normal and tumour cells. Tumour cells induce angiogenesis through their ability to release angiogenic signals which attract and stimulate endothelial cells. Tumours appear to induce angiogenesis by upregulating pro-angiogenic factors such as vascular endothelial growth factor (VEGF) and fibroblast growth factors (FGFs), while downregulating anti-angiogenic factors such as thrombospondin-1 (Fidler *et al.*, 2000; Hanahan and Folkman, 1996). These signals stimulate the endothelia cells which, in turn, construct capillaries within the tumour.

1.1.6 Tissue invasion and metastasis

Metastases are the major causes of human cancer deaths (Sporn, 1996). Metastasis is defined as the formation of secondary tumour foci at sites discontinuous from the primary lesion. The metastatic process is a complex composed of many steps in which tumour cells will: (1) detach from the primary tumour; (2) invade through the

ECM; (3) invade into the bloodstream; (4) migrate to the target site; (5) attach to target endothelium; (6) invade into the target tissue; and, (7) progressively growing secondary tumour. To invade and metastasize to other parts of the body, the cancer cell must have acquired altered expression of genes involved in the regulation of cell-cell and cell-matrix interactions. For example, E-cadherin function is lost in a majority of epithelial cancers (Christofori and Semb, 1999), while ECM degrading proteases are upregulated in tumour cells (Coussens and Werb, 1996).

In summary, cancer development is a multiple step process, and the transformation of a normal cell into a cancerous one requires deregulation of multiple distinct control mechanisms and the acquisition of multiple properties, which results from changes in cellular gene expression pattern. Gene expression is mainly controlled by transcription regulation (see below).

1.2 Transcription regulation

There are major differences in the gene expression patterns between cell types, despite the fact that all have the same set of inherited genes. The temporal and spatial control of gene expression encompasses many layers of complex mechanisms. The genetic information for proteins is encoded in DNA sequences. In order to produce proteins, the cell needs first to transcribe the DNA code into a sequence of messenger RNA (mRNA) molecules. This process is called transcription. Transcription is followed by translation, the synthesis of proteins according to the mRNA sequence. Eukaryotic cells utilize a many intricate mechanisms, which is mostly centered at the transcriptional

regulation, to ensure that gene expression is properly regulated (reviewed in Alberts *et al.*, 2002). In eukaryotic cells, transcription is controlled by interactions among the general transcriptional machinery, sequence-specific transcriptional regulators, mediators, DNA *cis*-acting regulatory elements, and unique local chromatin structure at *cis*-acting regulatory elements (see below).

1.2.1 General transcription machinery and gene transcription

Transcription involves several separate stages, including: (1) template recognition; (2) initiation of transcription; (3) elongation of transcription; and (4) termination of transcription. Transcriptional regulation is centered at promoter recognition and initiation (reviewed in Alberts *et al.*, 2002).

Eukaryotic genes are transcribed by three distinct RNA polymerases, namely RNA polymerase I, RNA polymerase II and RNA polymerase III (reviewed in Alberts *et al.*, 2002). RNA polymerase I transcribes genes encoding the ribosomal RNAs, while RNA polymerase III transcribes genes encoding the transfer RNAs and small nuclear RNAs (reviewed in Alberts *et al.*, 2002).

RNA polymerase II is the core protein of the enzyme complex responsible for all mRNA synthesis (reviewed in Alberts *et al.*, 2002). The core eukaryotic RNA polymerase II enzyme requires a set of proteins called general transcription factors, which must be assembled at the promoter before transcription can begin. The term general refers to the fact that these proteins assemble on all promoters transcribed by RNA polymerase II. Therefore, they differ from sequence-specific transcriptional

regulatory proteins (section 1.2.2), which act only at particular gene promoters and/or enhancers (Alberts *et al.*, 2002). The general transcription factors together with RNA polymerase constitute the basal transcription machinery, which is sufficient to mediate basal transcription. The general transcription factors for RNA polymerase II are called TFII (for Transcription Factor for RNA polymerase II), including TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH (Alberts *et al.*, 2002). They are required for the initiation and elongation of mRNA (Alberts *et al.*, 2002). The *in vitro* assembly scheme was shown in the Figure 1.1. Firstly, the core TATA box is recognized by the components of the TFIID complex (Figure 1.1A and B), including TATA-binding protein (TBP), and TBP-associated factors (TAFs) (Alberts *et al.*, 2002). TBP binds TFIIA and TFIIB (Figure 1.1C), which further recruits the RNA polymerase II-TFIIF complex to the promoter (Figure 1.1D) (Alberts *et al.*, 2002). Initiation of transcription additionally requires the binding of TFIIE and TFIIH (Alberts *et al.*, 2002).

Although the general transcription factors and RNA polymerase II assemble in a stepwise order *in vitro* (see above), there are cases in living cells where some of them are brought to the promoter as a large pre-assembled complex that is called the RNA polymerase II holoenzyme (Figure 1.2A and B) (reviewed in Alberts *et al.*, 2002). In addition to some of the general transcription factors and RNA polymerase II, the holoenzyme typically contains a multi-subunit protein complex called the mediator (see section 1.2.3), which was first identified as being required to activate transcription by sequence-specific transcription regulators (see section 1.2.2).

Figure 1.1 Initiation of transcription of a eukaryotic gene by RNA polymerase II *in vitro* (adapted from Alberts *et al.*, 2002).

(A) To begin transcription, RNA polymerase requires a number of general transcription factors. First, TBP binds the TATA box. **(B)** The TATA box is recognized and bound by transcription factor TFIID, which then enables the adjacent binding of TFIIB **(C).** For simplicity, the DNA distortion produced by the binding of TFIID is not shown. **(D)** The rest of the general transcription factors including TFIIE, TFIIH and TFIIF, as well as the RNA polymerase itself, assemble at the promoter. TFIIH then uses ATP to pry apart the DNA double helix at the transcription start point, allowing transcription to begin. The assembly scheme shown in the figure was deduced from experiments performed *in vitro*.

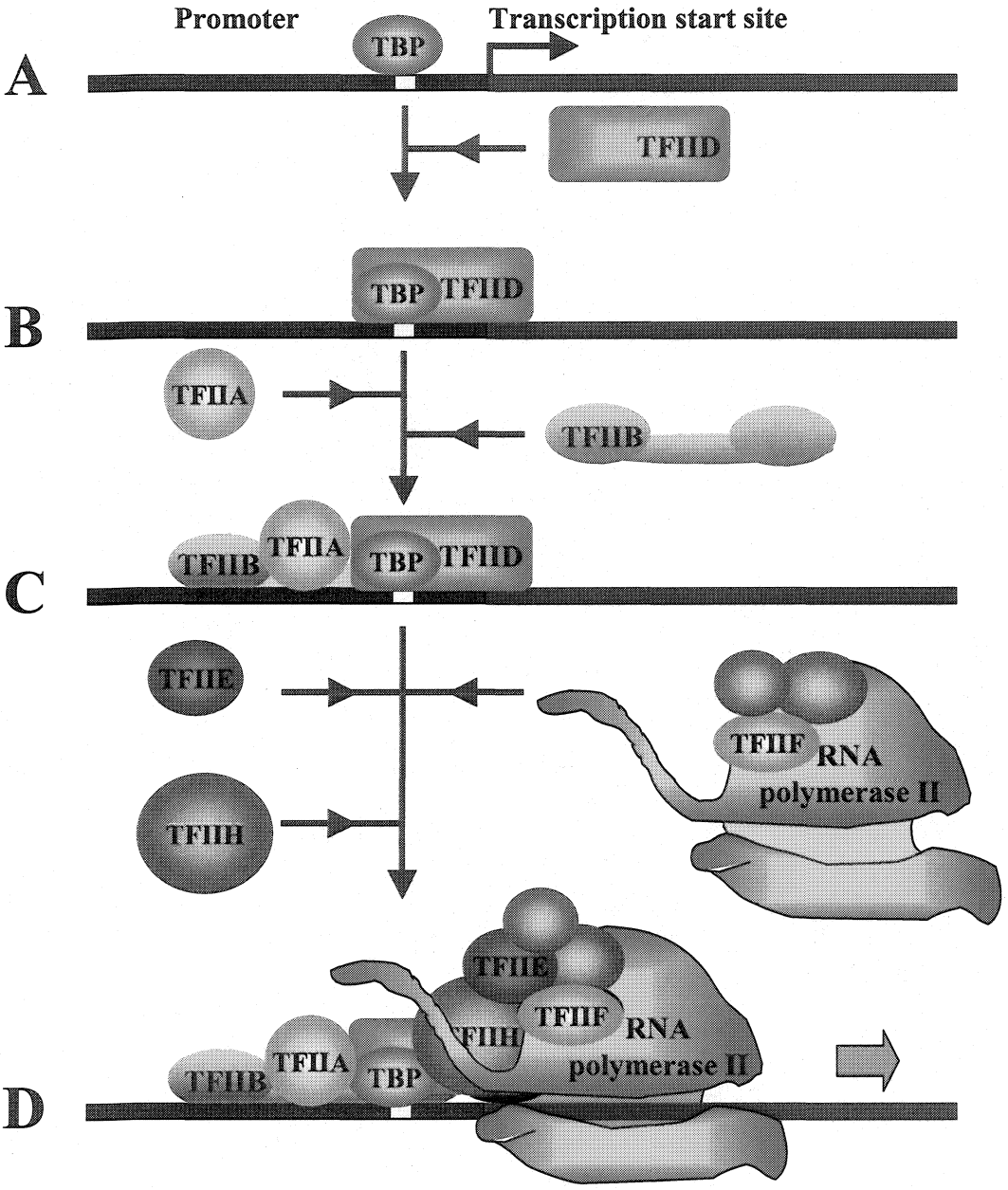
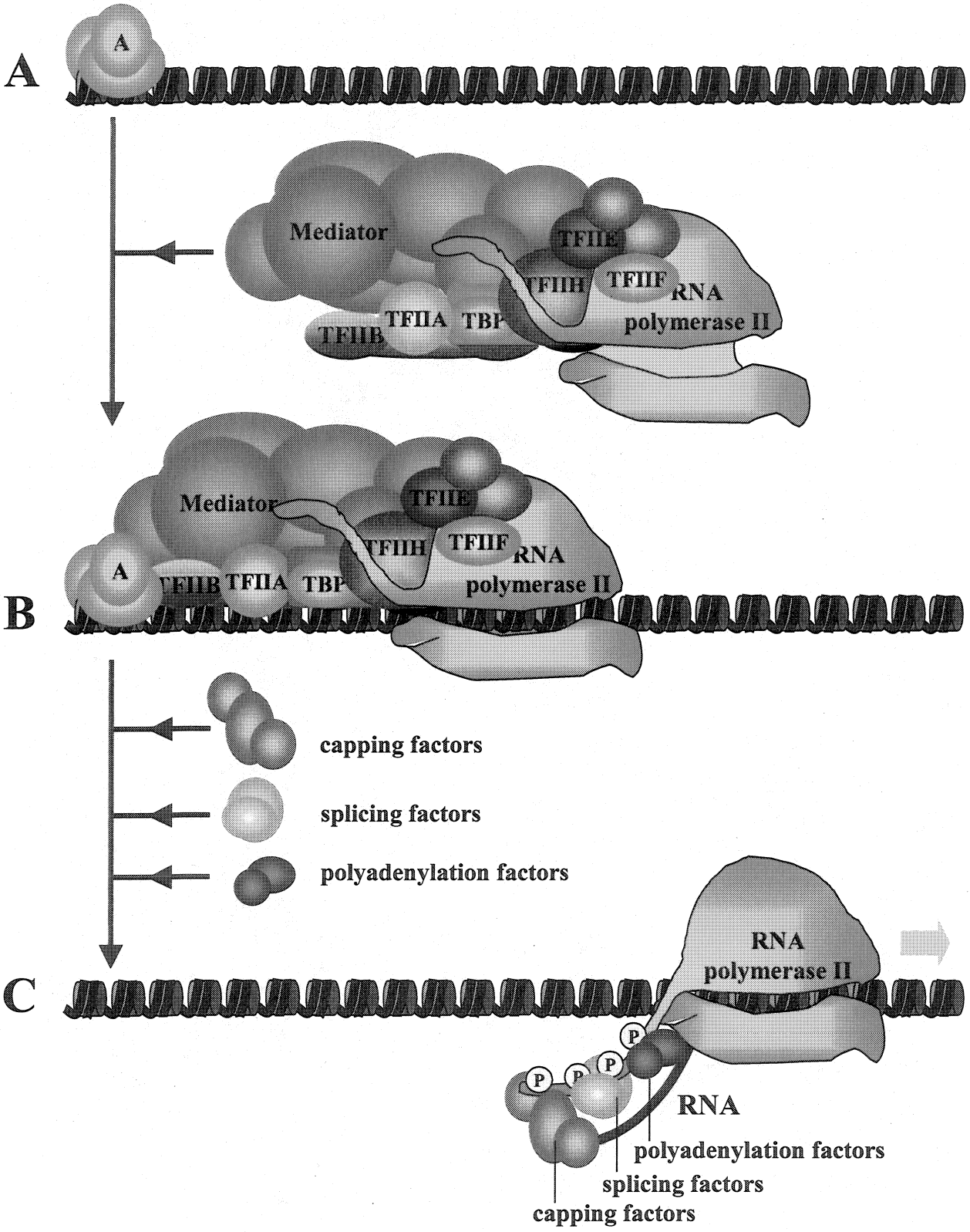

Promoter
Transcription start site
A
TBP
TFIID
B
TBP
TFIID
TFIIA
TFIIB
C
TFIIB
TFIIA
TBP
TFIID
TFIIE
TFIIH
TFIIF
RNA
polymerase II
D
TFIIB
TFIIA
TBP
TFIIE
TFIIH
TFIIF
RNA
polymerase II

Figure 1.2 Transcription of a eukaryotic gene by RNA polymerase II in cells.

(A) To begin transcription, transcriptional activators bind to *cis*-regulatory elements in promoters. **(B)** RNA polymerase II is brought to the promoter as a large pre-assembled complex that is called the RNA polymerase II holoenzyme, which also contains transcriptional regulator called mediator. For simplicity, the alterations of the chromatin structure are not shown. **(C)** Not only does the polymerase transcribe DNA into RNA, but it also carries a few critical components of pre-mRNA-processing proteins on its tail when it is phosphorylated late in the processing of transcription initiation. These components of pre-mRNA-processing proteins then transferred to the nascent RNA at the appropriate time. Once transferred to an RNA molecule, they serve as a nucleation site for the remaining components.



Transcription *in vivo*

Elongation of transcription is a stage during which the polymerase continues its movement along the DNA template as the precursor-mRNA chain (pre-mRNA) grows. It is regulated by specific elongation factors as well as by phosphorylation of the RNA polymerase II tail, the C-terminal domain (CTD) of the largest subunit of RNA polymerase II (Figure 1.2C) (Alberts *et al.*, 2002). The phosphorylation of the RNA polymerase II tail dissociates the RNA polymerase II from other proteins present at the start point of transcription, and recruits a few critical components of pre-mRNA-processing proteins (Figure 1.2C). These components of pre-mRNA-processing proteins are then transferred to the nascent RNA at the appropriate time. Once transferred to an RNA molecule, they serve as a nucleation site for other related components.

The RNA processing events, including 5 end capping, splicing, and 3 end formation, are tightly coupled to transcription elongation and termination (Alberts *et al.*, 2002; Proudfoot *et al.*, 2002). The long C-terminal tail of the RNA polymerase coordinates these processes by transferring pre-mRNA-processing factors directly to the nascent RNA as the RNA emerges from the enzyme.

As soon as RNA polymerase II has produced about 25 nucleotides of RNA, the 5 end of the new pre-mRNA molecule is modified by addition of a cap structure that consists of a modified guanine nucleotide. The capping reaction is performed by a phosphatase, a guanyl transferase and a methyl transferase acting in succession (Alberts *et al.*, 2002; Proudfoot *et al.*, 2002).

A typical euykaryotic gene in the genome consists of short blocks of protein-coding sequence (exons) separated by long non-protein-coding sequence (introns). Both

intron and exon sequences are transcribed into pre-mRNA. As the polymerase continues its traverses the DNA template, the spliceosome components assemble on the pre-mRNA and the intron sequences are removed through the process that is called RNA splicing (Alberts *et al.*, 2002; Proudfoot *et al.*, 2002).

Termination of transcription is a complex process that is regulated by mRNA 3 end processing. The specific signals in the pre-RNA are recognized by mRNA 3 end processing factors. A multisubunit protein complex called cleavage and polyadenylation specificity factor (CPSF) recognizes the highly conserved AAUAAA hexamer sequence in the pre-mRNA, while another protein complex called cleavage stimulation factor F (CstF) recognizes a downstream sequence element (DSE) which is a U- or GU-rich motif (Proudfoot *et al.*, 2002). Both of CPSF and CstF protein complexes travel with the RNA polymerase tail and are transferred to the 3 end of pre-mRNA sequence as it emerges from the RNA polymerase II. Once CstF and CPSF bind to specific nucleotide sequences on an emerging RNA molecule, additional proteins assemble with them to perform the processing that creates the 3 end of the mRNA. The pre-mRNA is cleaved between the AAUAAA hexamer and the DSE and a polyadenosine [poly(A)] tail is added to the 3 end of RNA upstream of the cleavage site by an poly-A polymerase enzyme, a process called as polyadenylation (reviewed in Alberts *et al.*, 2002; Proudfoot *et al.*, 2002). The mRNA is then transported out to the cytoplasm, where it serves as a template to translate information into protein. After the 3 end of a pre-mRNA molecule has been cleaved, the RNA polymerase II continues to transcribe, but soon releases from the template and transcription terminates. The piece of RNA downstream of the cleavage site is then

degraded in the cell nucleus. The phosphates on the RNA polymerase II tail are removed by soluble phosphatases, and it can reinitiate transcription.

1.2.2 Sequence-specific transcriptional regulators

RNA polymerase II and the general transcription factors are sufficient for the basal transcription (Alberts *et al.*, 2002). However, gene selective transcriptional activation requires the action of sequence-specific transcription activators that bind to the *cis*-acting regulatory elements. Conversely, transcription can be inhibited by sequence-specific transcription repressors which, when bound to their *cis*-acting regulatory DNA elements, prevent the transcriptional machinery from transcribing a gene. It is thought that transcription activators and repressors serve as critical components of the transcription regulatory network that ensures accurate transcription of a given gene according to the needs of a cell at the right time.

The majority of sequence-specific transcription regulators recognize and bind to *cis*-acting regulatory elements on the DNA. These *cis*-acting regulatory elements are often parts of larger regulatory entities called promoters and/or enhancers (see section 1.2.5). Transcriptional regulators can bind to the *cis*-acting regulatory elements in the core promoter region or in a distant enhancer region, but their role in the regulation of the accessibility and activity of transcription machinery is similar (Alberts *et al.*, 2002). DNA can bend and form loops, and thus even those transcription factors located at a distance from the promoter can still make physical contacts with the basal transcription machinery. Sequence-specific transcription factors may interact with the basal

transcription machinery by recruiting mediator complexes or modify the chromatin structure by recruiting chromatin modification complexes (reviewed in Alberts *et al.*, 2002).

Eukaryotic sequence-specific transcription regulators typically possess two characteristic domains: a DNA binding domain (DBD) that binds sequence-specific regulatory sites directly, and a second domain that exhibits transcriptional activation or repression potential. In some cases this dual requirement is shared by separate proteins in a complex, so that the sequence-specific binding domain and transcription activation domain and/or repression domain occur on separate proteins. In this case, transcription regulators without a DNA binding domain assemble on other DNA-bound protein complexes through protein-protein interactions. For example, RB has no DNA binding domain but can form complexes with members of E2F family on the E2F target gene promoters to repress gene expression from the E2F-target gene such as cyclin E (reviewed in Stevaux and Dyson, 2002).

It is generally thought that sequence-specific transcription regulators serve as protein-protein interaction interfaces which recruit other transcription regulators, including mediator (section 1.2.3) and/or chromatin modifiers (section 1.3), to regulate transcription (Alberts *et al.*, 2002). For example, RB can function as a transcriptional repressor by recruiting multiple chromatin remodelling and/or chromatin modifier complexes to the E2F target gene promoters (reviewed in Ferreira *et al.*, 2001). Moreover, multiple transcription activator and/or repressor complexes may act simultaneously on a specific promoter/enhancer. If the activity of activator complexes

exceeds the activity of repressor complexes, promoter-proximal nucleosomes sequentially adopt an activation-specific modification profile, and *vice versa* (Berger, 2002; Daujat *et al.*, 2002; Zhang and Reinberg, 2001). Therefore, multiple sequence-specific transcription regulatory protein complexes work together allow individual genes to be turned on or off to carry out a desired cellular response. Transcriptional regulatory proteins may be expressed differently in different cell types and thereby results in distinct gene expression profiles that give each cell type its unique characteristics.

1.2.3 Mediator complexes

Mediator complexes are general transcription machinery-interacting coactivator complexes, which transmit regulatory signals from sequence-specific transcriptional regulators to the general transcriptional machinery (Naar *et al.*, 2001). It has been suggested that mediator complexes bridge between transcription regulators and the general transcription machinery and promote the formation of a stable pre-initiation complex at the promoter (Naar *et al.*, 2001). Several mammalian mediator complexes, such as cofactor required for Sp1 activation (CRSP), vitamin D receptor-interacting protein (DRIP), thyroid hormone receptor-associated protein (TRAP) and activator-recruited cofactor (ARC) mediator complexes have been identified (Naar *et al.*, 2001). These mediator complexes share a many subunits (Naar *et al.*, 2001). It is thought, therefore, that there may be both core protein components which are shared by different mediator complexes and numerous variable components which are dependent on particular sequence-specific transcriptional regulators, and the composition and activity

of a mediator complex on a particular promoter may be regulated by specific signals and may also be dependent on the distinct cell type.

1.2.4 Regulating the transcriptional regulators

The transcriptional regulators themselves may also be subjected to multiple levels of regulation, in which their expression and/or activity is affected. The expressions of many of transcriptional regulators are regulated by specific signals (Brivanlou and Darnell, Jr., 2002). For example, the early-response gene *Xbra* encodes an embryonic mesoderm-specific transcriptional regulator whose expression can be induced by FGF (Smith *et al.*, 1991). Transcriptional regulators may be also subjected to post-translational modifications such as phosphorylation (reviewed in Freiman and Tjian, 2003). Phosphorylation of RB, for example, is mediated by cyclin-dependent kinases (CDKs) at the restriction site at later G1 of the cell cycle. Release of E2F following phosphorylation of RB allows for the transcription of E2F target genes (reviewed in Harbour and Dean, 2000). In addition, it has been recently shown that transcriptional regulator activities may also be regulated by acetylation (Gu and Roeder, 1997; Imhof *et al.*, 1997; Martinez-Balbas *et al.*, 2000; Sartorelli *et al.*, 1999). For example, acetylation of several lysine residues at the C terminus of p53 by p300 protein regulates its transcriptional activity (Barlev *et al.*, 2001; Prives and Manley, 2001). Other potential post-translational modifications, such as ubiquitination and sumoylation, can also regulate the activity of transcriptional regulators (Freiman and Tjian, 2003).

The activity of transcriptional regulators may also be subjected to regulation by subcellular localization. For example, eukaryotic cells are characterized by distinct nuclear and cytoplasmic compartments separated by the nuclear envelope. Active nuclear transport of transcriptional regulators is based on recognition of specific signals in the molecule that is being transported. Proteins carrying a nuclear localization signal (NLS) can be transported into the nucleus, while proteins carrying a nuclear export signal (NES) can be transported out of the nucleus (Moroianu, 1999). Moreover, the subcellular localization of proteins may also be regulated through other mechanisms, including piggy back transportation through interaction with other proteins, masking of transport signals, and subcellular department sequestration (Macara, 2001). For example, cytoplasmic retention has been described for some hormone nuclear receptors (Jans and Hubner, 1996; Vandromme *et al.*, 1996). This form of regulation leads to the transportation and/or immobilization of a protein within a specific cellular compartment, thus ensuring that a precise presence and function of a protein in a given compartment at a given time.

1.2.5 *Cis*-acting regulatory elements

Eukaryotic promoters for RNA polymerase II contain a variety of core *cis*-acting regulatory elements required for promoter function. Some examples of these elements include but are not limited to the initiator (Inr), TATA box sequence, downstream promoter element (DPE) and the GC box (Alberts *et al.*, 2002). Each of those elements is

found in only a subset of promoters, and any specific promoter may contain some, all, or none of them (Butler and Kadonaga, 2002).

The efficiency and specificity of transcriptional activity from a promoter depends upon multiple *cis*-acting regulatory elements. The *cis*-acting regulatory elements found in any individual promoter differ in number, location, and orientation. In some cases, the activity of a promoter is also affected by the presence of *cis*-regulatory elements called enhancers located at a variable distance from the promoter. The enhancers are distinct from that of the promoters and have two distinguishing characteristics: (1) the position of the enhancers relative to the promoter can vary substantially; (2) they can function in either orientation. The DNA *cis*-acting regulatory elements can be recognized by a corresponding transcriptional regulator or by a number of a family of factors. Binding of the transcription regulators at these *cis*-acting regulatory elements may exert influence on the formation of the transcription initiation complex. There are usually multiple *cis*-acting regulatory elements clustered in the promoter/enhancer, where protein-protein interactions between transcription regulators may play an important role in regulating the promoter activity. The activities of many transcription factors have shown to be regulated by other regulators bound nearby (reviewed in Alberts *et al.*, 2002). Thus, a single transcription factor can play distinct regulatory roles in different gene promoters. This context-dependent regulation of transcription allows cells to respond to a surprisingly diverse array of stimuli using the same factors (reviewed in Alberts *et al.*, 2002). The *cis*-regulatory sequences usually contain binding sites for multiple transcription factors and thus allow each gene promoter to respond to multiple signalling

pathways and facilitate the precise control of gene transcription (reviewed in Alberts *et al.*, 2002).

1.2.6 DNA methylation and transcriptional regulation

Even though each individual has the same set of genes inherited in their somatic cells, the structure and function of these cells may still differ from each other. Tissue-specific gene expression patterns are established and maintained by certain epigenetic 'marks', such as DNA methylation, within the genome to provide an important mechanism for distinguishing genes that are active from those that are not (Li, 2002). DNA methylation is the enzymatic addition of methyl groups to cytosine nucleotides that is mediated by DNA methyltransferases (DNMTs) (Bestor, 1988). DNMTs, such as DNMT3a, DNMT3b, are enzymes that are responsible for the *de novo* methylation. Other DNMTs, such as DNMT1, are responsible for the maintenance of methylation (Newell-Price *et al.*, 2000). The methylated form of cysteine, 5-methylcytosine (5-methyl C), has no effect on base-pairing. The methylation in DNA is restricted to cytosine (C) nucleotides in the sequence CpG, which is base-paired to exactly the same sequence in the opposite orientation on the other strand of the DNA helix. The existing pattern of DNA methylation can be inherited directly by the daughter DNA strands.

These CpG sequences are found to be clustered as dense regions of CpGs, called CpG islands (Antequera and Bird, 1993). They are mainly in the 5'-regions of genes (Antequera and Bird, 1993). The state of gene promoter methylation is associated with transcriptional repression (Bird and Wolffe, 1999). The biological importance of CpG

methylation is directly demonstrated by the fact that mice lacking the *dnmt1* gene exhibit defects in embryogenesis at mid-gestation (Li *et al.*, 1992). The transcriptional repression effects of CpG methylation are mainly mediated by methyl-CpG-binding domain (MBD)-containing proteins (reviewed in Ballestar and Wolffe, 2001). Much evidence has accumulated to indicate a close association between DNA methylation-dependent repression and chromatin modification complexes (see section 1.3.6).

1.3 Chromatin and transcription regulation

Additional levels gene transcription regulatory complexity in eukaryotic cells are achieved by the folding of the DNA template into chromatin (Pazin and Kadonaga, 1997). Thus, transcriptional regulation is not only dictated by the DNA sequence, but is also affected by the structure of chromatin, which controls the access of general transcription machinery to the promoter. The unique local chromatin structure at promoter and enhancer regions allows the cell to tightly control specific gene expression (Ashraf and Ip, 1998; Gregory and Horz, 1998; Torchia *et al.*, 1998).

1.3.1 Chromatin structure

The basic building block of chromatin is the nucleosome. It consists of a short segment of DNA wrapped around an octamer of histones represented by two copies each of histone H2A, H2B, H3, and H4 (Alberts *et al.*, 2002). . Each nucleosome is connected to its neighbours by a short segment of linker DNA. This polynucleosome string is folded into a compact fibre which is stabilized by the binding H1 histone, to each

nucleosome and to its adjacent linker (Alberts *et al.*, 2002). The DNA wound around the surface of the histone octamer is only partially accessible to other proteins and therefore provides a poor template for biochemical reactions such as transcription.

The N-terminal tails of histones that protrude from the nucleosome (Davie and Chadee, 1998). The histone tails are unstructured and serve as targets for variable covalent post-translational modifications, including acetylation, phosphorylation, methylation, sumoylation, and ubiquitination (reviewed in Spencer and Davie, 1999; Strahl and Allis, 2000). This has led to the histone code hypothesis that the modification state of the histone termini make up a histone code , which can be read by other ptoeins proteins that modulate transitions between the different chromatin states (Jenuwein and Allis, 2001; Orphanides and Reinberg, 2002; Strahl and Allis, 2000). This histone code hypothesis suggests that the presence of a given modification on histone tails may dictate or prevent the occurrence of a second modification elsewhere on the chromatin. This histone code hypothesis thus predicts that histone modifications create binding sites for accessory proteins. In this way, these modifications serve as marks for the recruitment of different proteins or protein complexes to regulate gene expression (Jenuwein and Allis, 2001; Strahl and Allis, 2000; Thiagalingam *et al.*, 2003). For instance, heterochromatin protein 1 (HP1) uses its chromodomain to recognize the histone H3 lysine 9 modification found in heterochromatin (Richards and Elgin, 2002). A number of studies have highlighted the important role of histone tails in chromatin folding and gene transcriptional regulation (Hayes and Hansen, 2001). The two most

important ways of locally altering chromatin structure are through nucleosome remodelling and covalent histone modification such as histone acetylation.

1.3.2 Chromatin remodelling complexes and transcription regulation

Chromatin remodelling factors use the energy derived from ATP hydrolysis to catalyze nucleosome mobilization, which involves a change in the position of the histone octamer on DNA. This change is believed to facilitate the access and function of key components of the transcriptional machinery to regulate gene transcription (Imbalzano *et al.*, 1994).

Chromatin remodelling complexes may be directed by transcription regulators to specific sites of a gene promoter to affect local chromatin structure (Peterson and Workman, 2000). For example, chromatin remodelling complexes can physically associate with nuclear receptors (Wallberg *et al.*, 2000; Yoshinaga *et al.*, 1992). The transcription factor c-Myc was also shown to bind directly to subunits of mammalian SWI/SNF complexes (Cheng *et al.*, 1999; Kowenz-Leutz and Leutz, 1999). Interestingly, chromatin remodelling complexes, such as the NURD complex, appear to be important not only for transcriptional activation, but also for transcriptional repression (Bird and Wolffe, 1999; Knoepfler and Eisenman, 1999). Moreover, the recruitment of a remodelling complex to the promoter may also facilitate binding of other transcriptional regulators whose binding sites were not previously exposed in the nucleosomes (Narlikar *et al.*, 2002).

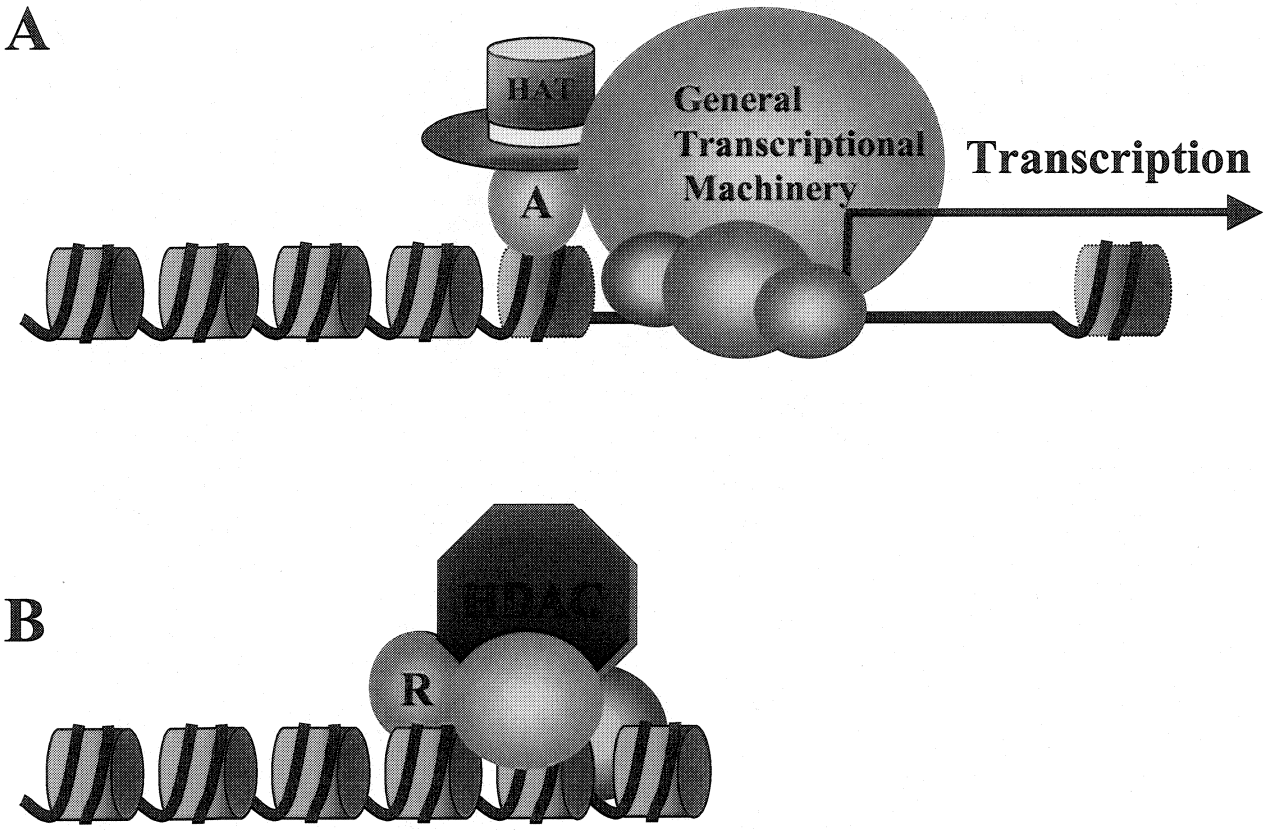
1.3.3 Histone acetyltransferase (HAT) complexes and transcription regulation

Histone acetylation is a process in which acetyl groups are added to the ε-groups at specific lysine residues in the N-terminal regions of the core histone proteins. Histone acetylation is believed to destabilize nucleosomes, thereby causing the opening of the chromatin structure. With the opening of the chromatin structure, the transcription activators and the basal transcription machinery have better access to their specific binding sites in the gene promoter and/or enhancer sites (Figure 1.3A) (Ayer, 1999; Grunstein, 1997). Indeed, actively transcribed genes have been correlated with increased levels of histone acetylation (hyperacetylation), whereas silenced genes are generally associated with decreased levels of histone acetylation (hypoacetylation) (Roth *et al.*, 2001).

Histone acetylation is catalyzed by HAT enzymes. Many studies have demonstrated the importance of HAT activity in the activator-dependent transcriptional activation (Roth *et al.*, 2001). There is a direct functional link between the transcriptional machinery and HATs. Some HATs are actually tightly associated with the RNA polymerase II transcription machinery itself. For example, the TBP-associated factor, RNA polymerase II 251-KD (TAFII251) subunit of TFIID, was found to contain intrinsic HAT activity (Mizzen *et al.*, 1996). Moreover, a number of mammalian transcriptional regulators, primarily identified by interactions with activation domains of sequence-specific transcription activators, have been found to exhibit HAT activity. Examples are the cAMP response element (CRE) binding protein (CREB)-binding protein (CBP) and the related adenovirus E1A-interacting protein p300 (Goodman and Smolik, 2000).

Figure 1.3 HAT and HDAC-dependent mechanisms of gene transcription regulation.

Local alterations of chromatin structure on promoters are emerging as a mechanism of precise regulation of transcription. Numerous studies have linked post-translational histone modifications, particularly the acetylation and deacetylation of lysine residues in histone N-terminal tails, with the transcriptional capacity of chromatin. Acetylation and deacetylation are catalyzed by HATs and histone deacetylases (HDACs), respectively. **(A)** Acetylation is believed to destabilize nucleosomes, at least in part, and thus facilitate access of the general transcription machinery, thereby promoting transcription. **(B)** Deacetylation is likely to stabilize nucleosomes and/or the higher order structure of chromatin further and thus reinforces the inhibitory effect of chromatin.


A
HAT
A
General
Transcriptional
Machinery
Transcription
B
HDAC
R

CBP/p300 is recruited by many transcription factors (Kornberg, 1999), and the action of CBP/p300 appears to require its HAT activity (Glass and Rosenfeld, 2000; McKenna *et al.*, 1999). Different HATs were shown to exist in multisubunit complexes, and may have different subunit compositions (reviewed in Roth *et al.*, 2001). For example, yeast GCN5p has been found in two distinct complexes, the SPT-ADA-GCN5-acetyltransferase (SAGA) complexes and the adaptor (ADA) complexes (Grant *et al.*, 1997). Two mammalian homologs of GCN5p, hGCN5 and PCAF have been identified (Candau *et al.*, 1996; Wang *et al.*, 1997; Yang *et al.*, 1996). Both hGCN5 and PCAF has been shown to be present in multisubunit complexes (Brand *et al.*, 1999; Martinez *et al.*, 1998; Ogryzko *et al.*, 1998; Wieczorek *et al.*, 1998).

1.3.4 Histone deacetylase complexes and transcription regulation

The degree of acetylation of core histone tails has been shown to be dependent on the opposing activities of two types of enzymes, HATs and histone deacetylases (HDACs). Histone deacetylases (HDACs) are enzymes that remove the acetyl groups from the lysine residues of core histone proteins, influencing nucleosome structure in the opposite way to histone acetylation. Histone deacetylation, and hence histone hypoacetylation, is likely to stabilize nucleosomes and/or the higher order structure of chromatin, further reinforcing the inhibitory effect of chromatin structure on gene transcription (Figure 1.3B) (Braunstein *et al.*, 1996; Lee *et al.*, 1993a).

HDACs are subunits of large protein complexes such as mSin3A, NURD, and CoREST, that play important roles in transcription repression (reviewed in Thiagalingam *et al.*, 2003). Many sequence-specific transcription factors form a complex with mSin3A, NURD and CoREST. For example, the transcription factor Snail recruit mSin3A-HDAC1/2 complex to repress E-cadherin expression (Peinado *et al.*, 2004).

1.3.5 Interplay of distinct chromatin modification complexes and transcription regulation

The modification of chromatin by acetylation/deacetylation to allow transcription is more complex than simple addition or removal of acetyl groups to histones. Recently, it has been suggested that there is interplay of histone modification with chromatin-remodelling as well as with DNA methylation (see below).

(a) Complexes containing chromatin remodelling enzymes and HDACs

The chromatin-remodelling complexes play a major role in chromatin remodelling and are not only associated with activators but also found to form complexes with corepressors such as HDACs. Indeed, chromatin-remodelling complexes plays important roles in both regulation of transcription and gene repression (Neely and Workman, 2002). For example, the NURD complex from HeLa human cervical carcinoma cells contains Mi-2, HDAC1/2 and the RBAP46/48, MTA2 and MBD3 proteins (Guschin *et al.*, 2000; Knoepfler and Eisenman, 1999). The ATPase activity of Mi-2 can increase the efficiency of histone deacetylation by NURD complexes in cell-free systems (Guschin *et al.*, 2000;

Tong *et al.*, 1998; Xue *et al.*, 1998b). The nucleosome remodelling activity of NURD is required for the deacetylation activity of NURD, which suggests that chromatin remodelling facilitates the access of the HDACs in NURD to the histone tails.

(b) HAT-HDAC complex

It is possible that some complexes possess both HAT and HDAC activity and therefore may be recruited for both transcriptional activation and repression, depending on the cellular context and under the appropriate stimuli. A recent study, for instance, showed that endogenous HDACs are associated with PCAF and GCN5 HATs, in HeLa cells (Yamagoe *et al.*, 2003). These large multiprotein HAT-HDAC complexes were found to be distinct from the previously described mSin3A, CoREST and NURD complexes (Yamagoe *et al.*, 2003). The dynamic association with HAT-HDAC complexes may be regulated by distinct signals. For example, coactivator CBP and corepressor HDACs complexes have been shown to bind the homeodomain heterodimer PBX-HOX simultaneously, while protein kinase A (PKA) stimulation of CBP has been found to facilitate the switch from transcriptional repression to activation in this system (Asahara *et al.*, 1999; Saleh *et al.*, 2000).

(c) Interplay of DNA methylation and HDAC complexes

DNA methylation correlates with gene repression (section 1.2.2). A number of studies lead to the suggestion that DNA methyltransferases might be link to the histone deacetylation and gene transcriptional repression. The two repression mechanisms,

histone deacetylation and DNA methylation, are apparently connected by the methyl-CpG binding proteins (MBPs), such as MeCP2, MBD1, MBD2, MBD3, and MBD4. MBPs are believed to recruit chromatin remodelling and modification complexes, including the NURD complex (Ng *et al.*, 1999; Tyler and Kadonaga, 1999; Zhang *et al.*, 1999). It has been proposed that DNA methylation and histone deacetylation might work together in some gene promoters to establish a repressive chromatin environment and silence gene expression (Cameron *et al.*, 1999).

1.4 *Xenopus* mesoderm induction early-response gene (*xmi-er1*)

1.4.1 FGF signalling transduction

FGFs comprise a family of 22 distinct proteins, numbered consecutively from 1 to 22 (reviewed in Ornitz and Itoh, 2001). FGF signalling plays important roles in regulate developmental processes and adult physiology, and deregulated FGF signalling pathways have been shown to be involved in tumorigenesis (Powers *et al.*, 2000). FGFs signal through FGF receptors (FGFRs), which are transmembrane tyrosine kinase receptors. There are four FGFRs, designated FGFR1, FGFR2, FGFR3 and FGFR4 (Johnson and Williams, 1993). Following ligand binding and dimerization, the receptors phosphorylate specific tyrosine residues on their own and each others cytoplamic domains (Lemmon and Schlessinger, 1994). The signal cascade is propagated by the recruitment of other signalling molecules onto the phosphorylated tyrosine residues of the activated receptors resulting in the activation of downstream transduction pathways that eventually lead to initiation of a specific set of gene expression (Pawson, 1995). The biological outcome of

FGF stimulation depends on a number of factors, including receptors and signalling network present in a particular cell. The effects of FGFs on different cellular functions are mediated by distinct combinations of signalling pathways including the MAPK pathway, phosphoinositide 3 kinase pathway (PI3K), and the phospholipase C-γ1 (PLCγ1) pathway (Hawkins *et al.*, 1997; Kamat and Carpenter, 1997; Szebenyi and Fallon, 1999). These transduction cascades are initiated by the autophosphorylation of the FGFRs upon ligand binding.

The MAPK pathway is propagated through the recruitment of SH2 domain and/or PTB domain docking proteins to the activated FGF receptor. Docking molecules, such as FRS2, bind to the phosphorylated receptor and recruit the GRB2-son of sevenless (SOS) complex. This puts SOS in close vicinity with RAS, which is also membrane bound. SOS promotes the dissociation of GDP from RAS, allowing RAS to acquire GTP molecules and to become activated (Ferrell, Jr., 1996). The activated membrane-associated RAS then recruits and activates RAF-1, a serine-threonine MAPK kinase kinase (MAPKKK). In turn, RAF-1 activates MEK, a MAPK kinase (MAPKK), which activates MAPK also known as ERK1 (p44) and ERK2 (p42). MAPK can signal directly to the nucleus by phosphorylating transcriptional regulators such as Elk1 (Buchwalter et al., 2004).. In addition, besides nuclear substrates, MAPK has been found to phosphorylate cytosketal proteins, phospholipase, and protein kinases (Ferrell, Jr., 1996).

Previous reports have established that activation of PI3K, which is a lipid kinase, can be achieved either by binding of PI3K to the receptor complex or by direct binding of activated RAS to the catalytic subunit of PI3K (Cantrell, 2001; Carballada *et al.*, 2001;

Ryan *et al.*, 1998). Activated PI3K phosphorylates the 3 -OH position of the inositol ring of membrane phosphatidylinositols (Ptdlns) (reviewed in Lawlor and Alessi, 2001). 3 - phosphorylated Ptdlns in turn act as second messagers to recruit AKT and its regulators PDK1 and PDK2 to the cell membrane, where PDK1 and PDK2 phosphorylate AKT (Scheid and Woodgett, 2003). Activated AKT in turn phosphorylates a number of targets, including transcription factors such as transcription factor FOXO (Birkenkamp and Coffer, 2003).

PLCγ has been shown to be associated with FGFR and is phosphorylated following ligand-dependent activation (Burgess *et al.*, 1990; Gillespie *et al.*, 1992; Mohammadi *et al.*, 1991; Ryan *et al.*, 1998; Ryan and Gillespie, 1994). Activated PLCγ then initiates hydrolysis of phosphatidylinositol 4,5-biphosphate to diacylglcerol (DAG) and inositol 1,4,5-trisphosphate (IP_3). DAG is an activator of protein kinase C (PKC), a serine-threonine kinase, and IP_3 initiates Ca^{2+} release from intracellular stores. Second messengers, like PKC and Ca^{2+}, activate several molecules including transcription factors such as c-FOS and thereby bring about desired cellular responses (Kamat and Carpenter, 1997; Slootweg *et al.*, 1991).

1.4.2 Early-response genes

In a signal transduction cascade, such as that induced by FGFs, communication from the extracellular environment to the nucleus results in rapid changes in the steady-state levels of mRNAs of many genes. These genes are the first genes to be transcribed in signal transduction pathways and known as early-response genes or immediate-early

genes (reviewed in Thomson *et al.*, 1999). For example, *Xbra is a* well-characterized FGF early-response gene in *Xenopus*. Expression of *Xbra* is an early response to mesoderm induction by FGFs (Smith *et al.*, 1991). The expression of the early-response genes is not dependent on *de novo* protein synthesis and is normally rapid. The majority of early response genes are transcription factors. Their expression can therefore lead to the initiation or termination of transcription of other genes that ultimately carry out the requirements of functions relayed by the original signal. Therefore, the early response genes are considered to be master regulators in signalling transduction, such as FGF signalling transduction. Thus, isolation and characterization of early response genes may lead to the identification of pivotal points in the signal transduction cascade which determine the response of a target cell.

1.4.3 Isolation of *xmi-er1*

Embryos from the amphibian *Xenopus laevis*, have been used as a model system for studying early response genes in FGF signalling transduction in Paterno and Gillespie laboratories (Paterno *et al.*, 1997; Ryan *et al.*, 1998; Ryan and Gillespie, 1994; Teplitsky *et al.*, 2003). Most studies were conducted by animal cap experiments in which an explant is taken from the prospective ectoderm of a blastula stage embryo and incubated with FGF-2 in an attempt to identify the particular early response genes that are active in the FGF signalling cascade. The discovery of a novel, developmentally regulated gene that was activated by FGF-2 in *Xenopus* was a result from this search (Paterno *et al.*, 1997). This gene was designated mesoderm induction early response 1 (*mi-er1*), since it

expresses during mesoderm induction and may function as a negative regulator of mesoderm induction (Teplitsky *et al.*, 2003). *Xenopus mi-er1* was identified through the PCR-based differential display method and demonstrated increased expression levels in animal cap explants in response to treatment by FGF-2 (Paterno *et al.*, 1997). The cDNA sequence isolated from the differential display was then used to isolate a full-length 2.3 kilobase-pair (kb) cDNA from a *Xenopus* blastula library (Paterno *et al.*, 1997). The *Xenopus mi-er1* cDNA cloned consists of a 1.479 kb single open reading frame (ORF) predicted to encode a protein of 493 amino acid residues (Paterno *et al.*, 1997).

The expression levels of *xmi-er1* were found to increase in *Xenopus* embryo animal cap explants during mesoderm induction by FGF-2. The steady-state levels of *xmi-er1* were shown to increase 3-4 fold upon treatment with FGF-2. However, the increase in *xmi-er1* levels was not due to *de novo* protein synthesis as cycloheximide, which is an inhibitor of protein synthesis, did not affect the FGF-induced increase in *xmi-er1* levels, thus demonstrated that *xmi-er1* is an early-response gene (Paterno *et al.*, 1997).

A database homology search revealed similarity in the predicted xMI-ER1 amino acid sequence to rat and human proteins encoded by the metastasis-associated gene 1 (*mta1*), a gene that was isolated by differential cDNA library screening and whose expression was associated with a metastatic phenotype (Nicolson *et al.*, 2003; Toh *et al.*, 1994). However, *xmi-er1* is not the *Xenopus* homolog of *mta1*, since a human homolog of *xmi-er1, hmi-er1,* has been found and is distinct from human *mta1* (Paterno *et al.*, 1997; Paterno *et al.*, 1998).

1.4.4 Expression and subcellular localization of xMI-ER1

Northern blot analysis of RNA extracted from *Xenopus laevis* embryos showed the presence of *xmi-er1* during initial cleavage stages, suggesting that most *xmi-er1* transcripts may be initially maternally derived. The levels of *xmi-er1* mRNA expression were constant during early cleavage, increased slightly at blastula stage and decreased 6-fold during gastrula, neurula and tailbud stage. The levels then remained below detectable levels during subsequent development (Paterno *et al.*, 1997).

The subcellular localization of endogenous xMI-ER1 protein during *Xenopus laevis* development has been investigated using immunohistochemistry with anti-MI-ER1 antibody (Luchman *et al.*, 1999). xMI-ER1 is found to be expressed during early stages of development, but only localizes to the nucleus around the mid-blastula transition (MBT). xMI-ER1 is found exclusively in the cytoplasm during early stages but then is found to progressively accumulate in the nuclei of marginal zone cells of stage 8 blastulae. xMI-ER1 is found exclusively in all nuclei in the animal hemisphere by late blastula stage and ubiquitous nuclear localization is observed by early gastrula. Nuclear xMI-ER1 staining decreases during tailbud stages (Luchman *et al.*, 1999). Further investigation revealed that there is a domain responsible for the cytoplasmic retention of xMI-ER1 before MBT, and this retention domain is located between the ELM2 and SANT domain (Post *et al.*, unpublished data).

With forced expression in NIH 3T3 cells, xMI-ER1 protein is found to be constitutively targeted to the nucleus in these cells (Paterno *et al.*, 1997). Analysis of the

basic amino acids and comparison of groups of basic amino acids to NLS described in the literature showed that xMI-ER1 has 4 putative NLSs, which were designated as NLS1, NLS2, NLS3, and NLS4 (Paterno *et al.*, 1997; Post *et al.*, 2001). Further investigation revealed that NLS4 on the C-terminal region is a functional NLS, and NLS1 may function as a weak NLS for xMI-ER1 (Post *et al.*, 2001).

1.4.5 xMI-ER1 functions as a transcriptional regulator

As mentioned above, xMI-ER1 was identified as an early response gene induced by FGF-2 (see 1.4.3). xMI-ER1 is a nuclear protein (see 1.4.4) (Luchman *et al.*, 1999; Post *et al.*, 2001). Structurally, xMI-ER1 has conserved domains found in a number of transcriptional regulators, including an acid activation domain (Paterno *et al.*, 1997), an ELM2 domain (Solari *et al.*, 1999) and a signature SANT domain (Aasland *et al.*, 1996). These observations suggested that xMI-ER1 may function as a transcriptional regulator. The possibility that xMI-ER1 is a transcriptional regulator was investigated by testing the transactivation potential of various regions of the xMI-ER1 protein. Constructs containing different regions of xMI-ER1 fused to the GAL4 DNA binding domain were used along with a chloramphenicol acetyltransferase (CAT) reporter plasmid in transient transfections (Paterno *et al.*, 1997). Although full-length xMI-ER1 did not activate transcription in NIH 3T3 cells, the region consisting of the N-terminal 98 amino acids of xMI-ER1 stimulated transcription 80-fold. Fusion with other parts of xMI-ER1 did not have any transcriptional activity. These results indicated that the N-terminal region

contains a transcription activation domain, and xMI-ER1 has the potential to function as a transcription activator (Paterno *et al.*, 1997).

A recent study revealed that overexpression of xMI-ER1 in *Xenopus* embryos resulted in a dramatic reduction of *Xbra* expression, suggesting that xMI-ER1 may function as a transcription repressor as well (Teplitsky *et al.*, 2003). *Xbra* is a mesoderm induction regulator which is able to induce mesoderm (Cunliffe and Smith, 1994). In agreement with the reduction in *Xbra* expression, overexpression of xMI-ER1 in *Xenopus* embryos significantly inhibited mesoderm induction induced by FGF-2, and resulted in truncations of the anteroposterior axis (Teplitsky *et al.*, 2003). These observations indicated that xMI-ER1 may be a negative regulator of the FGF signalling pathway, perhaps by acting as both a transcriptional repressor on target genes such as *Xbra*, and a transcription activator on other genes.

1.5 Human homologue of xMI-ER1, hMI-ER1

1.5.1 *hmi-er1* isolation and genomic structure

A human orthologue of *xmi-er1* was previously cloned from human testis cDNA library in the laboratory, and designated as hMI-ER1α (Paterno *et al.*, 1998; Paterno *et al.*, 2002). Extensive 5 and 3 rapid amplification of cDNA ends (RACE) was performed using forward and reverse primers designed against various regions of the original published hMI-ER1α sequence in combination with cDNAs from a number of human tissues and cell lines (Paterno *et al.*, 2002). Sequence analysis of these RACE clones and BLAST comparison to the human genome revealed the structure of the *hmi-*

er1 gene and the presence of 12 distinct transcripts. Searches of the human EST database revealed that all reported sequences are represented in the cDNA clones obtained in the laboratory (Paterno *et al.*, 2002).

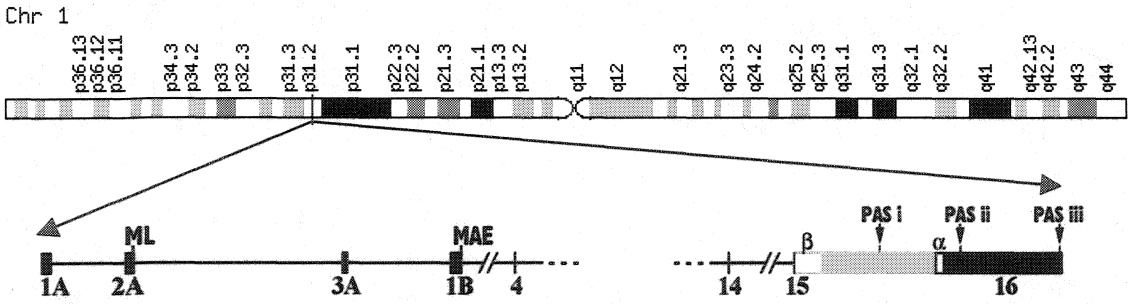
hmi-er1 is a single copy gene located at 1p31.2 and spanning 63 kb (Figure 1.4A). This gene consists of 17 exons. Most exons are smaller than 160 bp (range 20 bp to 2.5 kb), while most introns are several kb in size (range 630 bp to 11.7 kb). As shown in Figure 1.4B, there are a total of 12 distinct transcripts (sequence data in the GenBank Database under the accession numbers AF515446, AY124191, AY124194,; AF515447, AY124190, AY124193; AF515448, AY124189, AY124192, AY124186, AY124187, AY124188) generated by alternative splicing, alternate promoter usage and/or polyadenylation signal (PAS) usage (Paterno *et al.*, 2002). The alternate 5 ends result from alternate promoter usage or alternate inclusion of exon 3A to generate three distinct 5 ends encoding three amino terminal regions, N1, N2 and N3. The four variant 3 ends, a, bi, bii and biii, result from alternative splicing or alternate PAS usage. The a 3 end portion of *hmi-er1* transcripts (Figure 1.4) is predicted to encode the 23 aa α C-terminus domain of hMI-ER1; while the bi, bii and biii 3 portions of *hmi-er1* transcripts (Figure 1.4) are all predicted to encode the same 102 aa β C-terminal domain of hMI-ER1 (Paterno *et al.*, 2002). The production of two distinct C-terminal regions was verified using polyclonal antiserum generated against peptides corresponding to hMI-ER1α C-terminus, the β C-terminus or the common internal sequence (Paterno *et al.*, 2002).

As shown in Figure 1.4B, alternate promoter usage and splicing at the 5 end results in the generation of mRNAs predicted to encode three distinct N-terminal

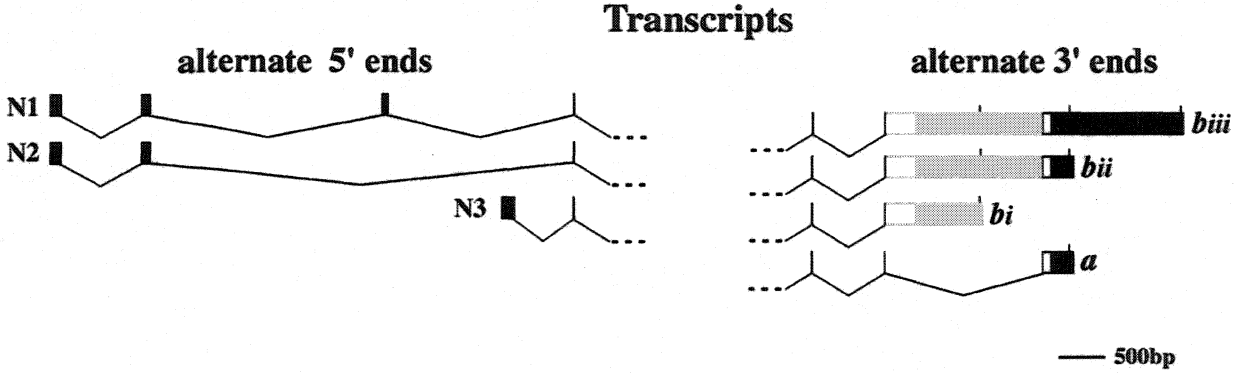
Figure 1.4 Structure of the human *hmi-er1* gene and splice variants (adapted from Paterno *et al.*, 2002).

Schematics illustrating the organization of the *hmi-er1* gene and the various *hmi-er1* transcripts. **(A)** Exon/intron organization of the *hmi-er1* gene. *hmi-er1* is a single copy gene located at 1p31.2 (adapted from http://genome-www.stanford.edu/cgi-bin/genecards/carddisp?MI-ER1). The two alternate starts of translation, ML- and MAE- are indicated. Exon numbers are indicated below each schematic. **(B)** Schematics illustrating the variant 5' and 3' ends of *hmi-er1* transcripts. Alternate 5' ends result from alternate promoter usage or alternate inclusion of exon 3A to generate three distinct amino terminal regions, N1, N2 and N3. The four variant 3' ends, a, bi, bii and biii, result from alternative splicing or alternate PAS usage. Scale bar, 500 bp.

A



B



■ **Exon**

— **Intron**

Facultative Intron

□ **α or β C-terminal coding region**

ML, MAE- alternate AUG starts

PASi, ii, iii- alternate poly A signals

N1, N2, N3- alternate N-termini

domains: N1, N2 and N3. N1 has the same translational start as N2 but includes exon 3A sequences, which is inserted after the first 2 amino acid residues (aa). This additional sequence is predicted to encode a 25 aa cysteine-rich domain that remains in frame with the rest of the hMI-ER1 protein. The three distinct N-terminal domains, in combination with the two possible C-terminal regions, are predicted to produce six distinct hMI-ER1 proteins: N1α (457 aa), N1β (536 aa), N2α (433 aa), N2β (511 aa), N3α (433 aa), and N3β (512 aa). The predicted sequences were verified using an anti-hMI-ER1 antibody to immunoprecipitate *in vitro* translated products (Paterno *et al.*, 2002).

1.5.2 Protein domains and motifs in hMI-ER1 isoforms and their possible function

Figure 1.5 shows a schematic of the putative functional domains and motifs in hMI-ER1 isoforms. Putative domains and motifs found in hMI-ER1 isoforms were identified by computer analysis using software MOTIF (http://www.Motif.genome.ad.jp), PSORT (http://www.psort.nibb.ac.jp), Domain Architecture Retrieval Tool (DART; http://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi), Protein Families Data Base of Alignments (PFAM) at http://pfam.wustl.edu/hmmsearch.shtml. Some of the motifs and domains found are indicated in Figure 1.5 and will be described.

(a) The acidic activation domain

As shown in Figure 1.5, hMI-ER1 isoforms have an acidic activation domain.

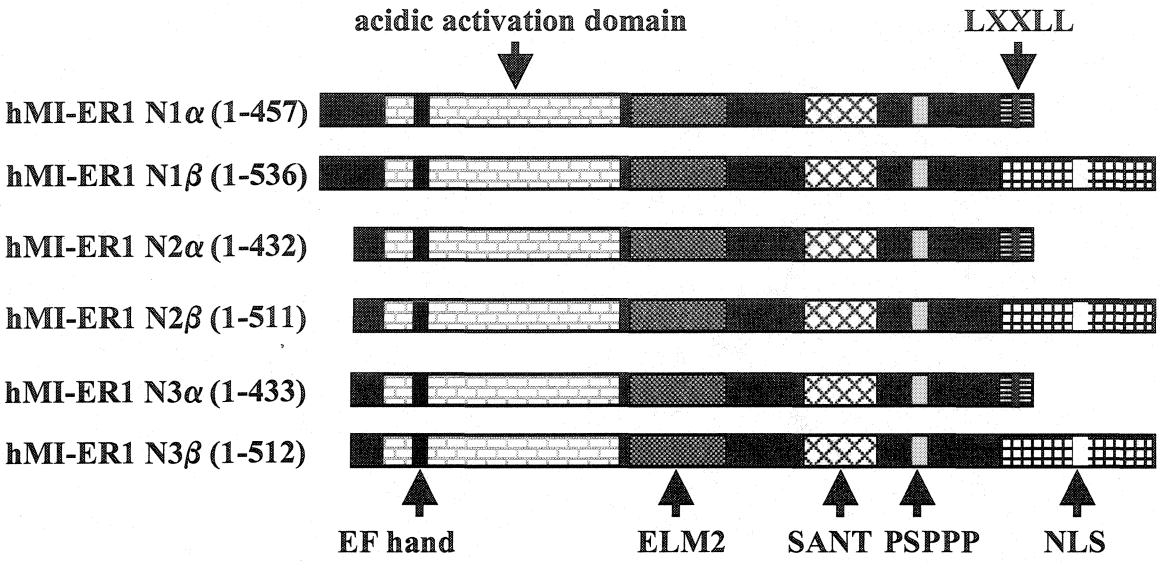
Acidic activation domains were first described in the yeast GAL4 activator protein (Ma and Ptashne, 1987). In the GAL4 protein, transcriptional activation is mediated by acidic amino acid residues and transcriptional activation closely correlates with the net negative charge. This indicated that acidic amino acids are crucial for transactivation of GAL4. Moreover, several studies with acidic activation domains also indicated the importance of hydrophobic amino acids for activation (Blair *et al.*, 1994; Cress and Triezenberg, 1991; Lin *et al.*, 1994). The N-terminus of xMI-ER1 also has regions rich in acidic residues which are important for transcriptional activation function (Paterno *et al.*, 1997). Although full-length xMI-ER1 did not activate transcription in NIH 3T3 cells, the region consisting of the N-terminal 98 amino acids of xMI-ER1 stimulated transcription 80-fold. Fusion with other parts of xMI-ER1 did not have any transcriptional activity. These results indicated that the N-terminal region of MI-ER1 contains an acidic activation domain, and xMI-ER1 has the potential to function as a transcription activator (Paterno *et al.*, 1997). The N-terminal acidic sequences are conserved from *Xenopus* MI-ER1 to its human homologue (Paterno *et al.*, 1998), and suggested a transcriptional activation function for this domain in hMI-ER1.

(b) The EF-hand motif

As shown in Figure 1.5, hMI-ER1 isoforms have a predicted EF-hand motif. The EF-hand motif is involved in binding intracellular calcium (Julenius *et al.*, 2002). The basic EF-hand consists of two 10 to 12 residue alpha helices with a 12-residue loop

Figure 1.5 Protein domains and motifs in hMI-ER1 isoforms.

Schematics illustrating the protein motifs common and unique to the hMI-ER1 isoforms. The acidic activation domain, the EF-hand motif, the ELM2 domain, the SANT domain, and the proline rich motif PSPPP are common to the hMI-ER1 isoforms. The nuclear localization signal (NLS) is localized on the C-terminal region of hMI-ER1α. The LXXLL motif is localized on the C-terminal region of hMI-ER1β.


acidic activation domain
LXXLL
hMI-ER1 N1α (1-457)
hMI-ER1 N1β (1-536)
hMI-ER1 N2α (1-432)
hMI-ER1 N2β (1-511)
hMI-ER1 N3α (1-433)
hMI-ER1 N3β (1-512)
EF hand
ELM2
SANT
PSPPP
NLS

region between, thereby forming a single calcium-binding site (helix-loop-helix). While calcium ions interact with residues contained within the alpha helices, each of the 12 residues in the loop region is important for calcium coordination. Binding of Ca^{2+} to the EF-hand domains may trigger a conformational change in calcium-binding proteins and this change may underlie the Ca^{2+}-dependent biological effect. One example of Ca^{2+}-dependent conformational change by EF-hand domain is the downstream regulatory element antagonist modulator (DREAM) protein (Craig *et al.*, 2002). DREAM acts as a direct transcriptional repressor of c-*fos* genes (Carrion *et al.*, 1999). The EF-hands domain of the DREAM protein senses the intracellular concentration of Ca^{2+}. EF-hand occupancy by Ca^{2+} blocks binding of DREAM to its DNA binding sites. This results in relief of the transcriptional repression by DREAM on its target gene such as c-fos (Carrion *et al.*, 1999). Moreover, the conformational change of DREAM triggered by Ca^{2+} creates a protein-protein interaction surface for CREB. As a result of the interaction, DREAM prevents the recruitment of CBP by CREB and represses CRE-dependent transcription (Ledo *et al.*, 2002).

Activation of most growth factor tyrosine kinase receptors, as well as stimulation of G-protein coupled receptors, results in the release of Ca^{2+} stores from the endoplasmic reticulum (reviewed in Soderling, 1999). Ca^{2+} release is controlled by phosphoinositide-specific phospholipase C (PLC) isozymes. An increase in intracellular Ca^{2+} induces Ca^{2+}-responsive signalling cascades *via* the EF-hand domain-containing proteins,

resulting in activation or repression a specific set of genes, thus triggering numerous biological outcomes.

(c) The ELM2 domain

As shown in Figure 1.5, each of the hMI-ER1 isoforms has an ELM2 domain. ELM2 consists of approximately 60 amino acid residues and stands for EGL-27 and MTA1 homology domain 2 (Solari *et al.*, 1999). The ELM2 domain was initially identified in the EGL-27 protein as well as in MTA1 (Solari *et al.*, 1999). EGL-27 is a *C. elegans* protein that is involved in embryonic patterning and plays a key role in Wnt signalling possibly by regulating HOX gene expression (Herman *et al.*, 1999; Solari *et al.*, 1999). MTA1 is known to be part of a protein complex possessing HDAC activities (Xue *et al.*, 1998a). However, no function has been ascribed to the ELM2 domain before the studies described in this thesis (see Chapter 3).

Many of ELM2 domain-containing proteins also contain one or more putative SANT domains, as predicted by BLAST search (RPS-BLAST 2.2.1, NCBI; http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Interestingly, the ELM2 domain is usually found in a position N-terminal to a conserved SANT domain. It is thought that there might be some important cooperative function between the two motifs (Solari *et al.*, 1999).

(d) The SANT domain

As shown in Figure 1.5, each of the hMI-ER1 isoforms has a SANT domain C-terminal to the ELM2 domain. The SANT domain was named for the four transcription regulators in which it was first discovered: SWI3, ADA2, nuclear receptor corepressor (N-COR), and TFIIIB (Aasland *et al.*, 1996). The SANT domain consists of approximately 50 amino acid residues and is highly related to the DNA binding domain (DBD) of the oncoprotein MYB (Aasland *et al.*, 1996). Like the DBD of MYB, the SANT domain has two or three repeated subdomains each resembling the helix-turn-helix design in secondary structure.

Although the function of the SANT domain has not been fully understood, it is thought that the SANT domain acts in transcriptional regulation either through DNA binding or protein-protein interactions (Aasland *et al.*, 1996). Recent studies with SANT domain proteins have suggested that SANT domains are more likely to function in protein-protein interaction. For example, the yeast ADA2 SANT domain has been shown to be essential for HAT activity in GCN5-containing complexes and was also found to assist GCN5 to bind to histones (Boyer *et al.*, 2002; Sterner *et al.*, 2002). Another example is CoREST, which is a transcriptional corepressor containing two SANT domains (You *et al.*, 2001). The first SANT domain is important for the co-repression function of the protein, by virtue of its interaction with HDAC1/2 (You *et al.*, 2001). Similarly, the first SANT domain of N-COR and silencing mediator of retinoic acid and thyroid hormone receptor (SMRT) was shown to bind and activate HDAC3 (Guenther *et al.*, 2001). The second SANT domain of SMRT is essential for binding to

hypoacetylated histone and therefore may play a role in interpretation of the histone code (Yu *et al.*, 2003). Most recently, overexpression of xMI-ER1, which is a SANT domain containing molecule, in *Xenopus* embryos was found to inhibit mesoderm induction and result in truncations of the anteroposterior axis (Teplitsky *et al.*, 2003). However, a mutagenesis analysis of the SANT domain revealed that the SANT domain is not required for the effects of xMI-ER1 on embryonic development and mesoderm induction (Teplitsky *et al.*, 2003).

(e) The proline rich motif

As shown in Figure 1.5, each of hMI-ER1 isoforms also has a proline rich motif C-terminal to the SANT domain. Proline rich motifs are characterized by the presence of the consensus PXXP tetrapeptide, in which the P represents proline and X represents any amino acid residue (Alexandropoulos *et al.*, 1995). The proline-rich region of MI-ER1 fits the consensus for an SH3 recognition motif (Cesareni *et al.*, 2002). As well, it also conforms to the consensus S/T-P for a proline-directed phosphorylation site (Teplitsky *et al.*, 2003). Since MI-ER1 was found not to be phosphorylated in cells (Paterno *et al.*, unpublished data), its proline-rich region most likely functions as a SH3 recognition motif. In agreement with this, mutagenesis analysis of the proline-rich region of xMI-ER1 (^{365}P^{366}SPPP) revealed that only ^{365}P in the proline-rich region is required for the overexpression effects of xMI-ER1 on embryonic development and mesoderm induction (Teplitsky *et al.*, 2003).

It is known that SH3 domains recognize proline rich motifs (Mayer, 2001). A variety of proteins involved in intracellular signal transduction pathways contain SH3 domains or proline rich motifs. Proline rich motifs are commonly found in situations requiring the rapid recruitment of proteins through protein-protein interactions, such as in signalling cascades (Kay *et al.*, 2000). For example, following receptor activation, the proline rich region on SOS1 binds to the two SH3 domains of Grb2, thereby bringing SOS1 to the cell membrane, where SOS1 in turn activates the RAS pathway (reviewed in Williamson, 1994).

(f) The nuclear localization signals

Eukaryotic cells are characterized by distinct nuclear and cytoplasmic compartments separated by the nuclear envelope, which is penetrated by nuclear pore complexes (NPCs) (Alberts *et al.*, 2002). Nuclear transport is proving to be a fundamental mechanism for regulating protein localization, and protein function (Hood and Silver, 2000). Protein nuclear transport in either direction across the nuclear envelope through NPCs involves sequential steps mediating the intricate interplay of multiple protein components starting from recognition of the protein import/export signal by an import or export receptor (Gorlich, 1998; Nigg, 1997). Proteins carrying a NLS can be transported into the nucleus. There are several classes of NLS (Post *et al.*, 2001). Proteins carrying a NES can be transported out of the nucleus (Moroianu, 1999).

As shown in Figure 1.5, hMI-ER1β contains a functional NLS in the C-terminal region (Paterno *et al.*, 2002). The hMI-ER1β was found localized in the nucleus, while

hMI-ER1α shuttles between the nucleus and the cytoplasm (Paterno *et al.*, 2002; Paterno *et al.*, unpublished data).

(g) The LXXLL motif

The LXXLL motif is characterized by the presence of a consensus LXXLL peptide, in which the L represents leucine and X represents any amino acid residue. The LXXLL motif is a protein-protein interaction motif that has been implicated in the interaction with nuclear hormone receptors. It was also reported to be present in other transcription regulators, including the short form of MTA1 (MTA1s) (Kumar *et al.*, 2002). This core sequence also has very strong similarity to the Sin3A interaction domain (SID) present in the family of MAD (MAD1-4) transcriptional repressors (Brubaker *et al.*, 2000). The selectivity of the LXXLL motif may depend upon flanking amino acid sequences. Many peptides that bind strongly to nuclear hormone receptors possess a hydrophobic amino acid in the -1 position relative to the core and a non-hydrophobic amino acid at the +2 position (Heery *et al.*, 2001).

The LXXLL domain has been identified in the hMI-ER1α C-terminus and contains a core LXXLL motif embedded in an amphipathic helix that is not present in the hMI-ER1β isoform. This domain within hMI-ER1α has a hydrophobic amino acid at both the -1 and +2 positions.

1.5.3 Expression of hMI-ER1

The tissue distribution of the various *hmi-er1* transcripts was examined first by Northern blot analysis and then by RT-PCR analysis. A probe hybridizing to the common internal region revealed that *hmi-er1* transcripts could be resolved into only four bands by Northern blot analysis (Paterno *et al.*, 2002). This was expected since the three 5 ends differ little in size. The estimated molecular sizes of the bands, 1.7, 2.5, 3.4 and 4.8 kb, were consistent with the predicted size of transcripts containing a, bi, bii and biii 3 ends, respectively.

Of the 23 human tissues examined by Northern blot analysis, *hmi-er1* transcripts were barely detectable or absent in most. Of the five tissues (heart, testis, ovary, thyroid and adrenal gland) that showed readily detectable levels of *hmi-er1*, testis expressed the highest level and was the only tissue in which all four bands were visible. The 4.8 kb band was the most widely expressed and could be detected, albeit at low levels, in most tissues. The 2.5 kb band was the least abundant and was only visible in a few tissues, while expression of the 3.4 kb band was highly variable. The 1.7 kb band was only detectable in a few endocrine tissues such as testis, ovary and thyroid (Paterno *et al.*, 2002). RT-PCR analysis using distinct primers for different isoforms revealed that the *hmi-er1* expression pattern is complex as different transcripts were shown to have cell type specific expression (Paterno *et al.*, 2002). *hmi-er1* was found to have ubiquitous, but low level, expression in normal human tissues (Paterno *et al.*, 1998; Paterno *et al.*, 2002). Breast carcinoma cell lines and tumours, on the other hand, showed elevated

levels (Paterno *et al.*, 1998), suggesting that *hmi-er1* expression may be associated with the neoplastic state in human breast carcinoma.

1.5.4 Subcellular localization of hMI-ER1α and β

The high degree of similarity (91% overall similarity) shared between the N-terminal domains of hMI-ER1α protein and xMI-ER1 suggests that they are evolutionarily conserved. The C-terminal regions of xMI-ER1 and hMI-ER1α proteins have highly divergent sequences (Paterno *et al.*, 1997; Paterno *et al.*, 2002). However, the amino acid sequence of the hMI-ER1β C-terminus displayed a high degree of similarity (66%) to that of the xMI-ER1 C-terminal domain, suggesting that the published xMI-ER1 sequence encodes the orthologue of the hMI-ER1β isoform (Paterno *et al.*, 2002). Interestingly, although xMI-ER1 contains multiple putative NLSs, it was revealed that this C-terminal domain of xMI-ER1 contains the only functional NLS (Post *et al.*, 2001). Similar to xMI-ER1, hMI-ER1β also contains a functional NLS in the C-terminal region. The subcellular localization of hMI-ER1 was examined by transfecting NIH 3T3 cells with plasmids expressing Myc-tagged hMI-ER1α or β. Immunohistochemical staining with the anti-MYC antibody 9E10 showed that hMI-ER1β was targeted exclusively to the nucleus, while hMI-ER1α shuttled between the nucleus and cytoplasm (Paterno *et al.*, 2002; Paterno *et al.*, unpublished data). Moreover, hMI-ER1 displayed less nuclear localization in breast cancer samples, when compared with the normal breast tissue (Paterno *et al.*, unpublished data).

1.6 Purpose of this study

hMI-ER1 isoforms localized in the nucleus or shuttled between the nucleus and cytoplasm (Paterno *et al.*, 2002; Paterno *et al.*, unpublished data). Structurally, hMI-ER1 has conserved domains found in a number of transcriptional regulators, including an acidic activation domain (Paterno *et al.*, 1997), an ELM2 domain (Solari *et al.*, 1999) and a signature SANT domain (Aasland *et al.*, 1996). hMI-ER1 was found to have ubiquitous, but low level, expression in normal human tissues (Paterno *et al.*, 1998; Paterno *et al.*, 2002). Breast carcinoma cell lines and tumours, on the other hand, showed elevated levels (Paterno 1998), suggesting that *hmi-er1* expression is associated with the neoplastic state in human breast carcinoma. Moreover, overexpression of hMI-ER1 in NIH 3T3 cells suppressed colony formation efficiency (Paterno *et al.*, unpublished data). Depletion of hMI-ER1 in multiple breast cancer cell lines by *hmi-er1* antisense plasmid transfection also results in suppression of colony formation efficiency (Huang *et al.*, unpublished data). These results suggested that hMI-ER1 plays an important role in cell growth regulation. Therefore, my hypothesis is that hMI-ER1 may be a potent transcriptional regulator, and deregulation of its expression and/or functions may contribute to tumorigenesis. At the time this study was undertaken, the N3a (the GenBank Database accession number AY124188) and N3bii (the GenBank Database accession number AY124192) *hmi-er1* cDNA were the only two cloned isoforms in the laboratory. The purpose of this study was to isolate and characterize the promoters of *hmi-er1*, and investigate the role of hMI-ER1α (N3a) and hMI-ER1β (N3bii) in

transcriptional regulation. To obtain a better understanding of the function of hMI-ER1, different approaches have been used to identify and characterize the hMI-ER1α- and hMI-ER1β-interacting proteins, which may be critical for regulating specific cellular functions with hMI-ER1α and hMI-ER1β.

Objective 1: Molecular cloning and characterization of the human of *hmi-er1* promoters

hmi-er1 was found to have ubiquitous, but low level, expression in normal human tissues (Paterno *et al.*, 1998; Paterno *et al.*, 2002). Breast carcinoma cell lines and tumours, on the other hand, showed elevated levels (Paterno *et al.*, 1998), suggesting that *hmi-er1* expression is associated with the neoplastic state in human breast carcinoma. My goal was to study the molecular mechanism for the transcriptional regulation of *hmi-er1*, through cloning of the 5 promoter region flanking the transcription start sites, and identifying regulatory regions important for transcriptional regulation.

Objective 2: Characterization of the role of hMI-ER1 in transcriptional regulation

Structurally, hMI-ER1 shares features with other transcriptional regulators. It possesses a putative acidic activation domain, conserved ELM2 domain and conserved SANT domain. Although the precise function of ELM2 and SANT domains remain unclear, proteins containing these domains were recently found to be present in multisubunit complexes with HAT or HDAC enzymatic activity and implicated in transcriptional regulation. Examples of these transcription regulators are CoREST,

MTA1, MTA2, MTA3, ADA2, N-COR and SMRT. There is evidence that *Xenopus* xMI-ER1 can function as a transcriptional regulator (Paterno *et al.*, 1997). Thus, my goal was to investigate whether human hMI-ER1 functions as a transcriptional regulator implicated by these conserved domains in other proteins, and to investigate the molecular mechanisms it utilizes to regulate this function.

Objective 3: Identify and characterize the hMI-ER1-interacting proteins

To obtain a full picture of hMI-ER1 function, it is necessary to perform complementary approaches in protein-protein interaction analysis to obtain more specific hMI-ER1 interacting proteins. Pioneering work toward this goal in this study was undertaken using two different approaches. Immunoprecipitation and Western blot analysis was used to test the possible interaction between hMI-ER1 and predicted interacting proteins. I also established the yeast two-hybrid system to screen hMI-ER1-interacting proteins from a testis cDNA library. My goal was to identify new hMI-ER1-inteacting proteins that would assist in the elucidation of the regulatory networks in which hMI-ER1 plays a role in the cell.

SECTION II ISOLATION AND CHARACTERIZATION OF THE PROMOTERS OF *hMI-ER1*

CHAPTER 2 MOLECULAR CLONING AND CHARACTERIZATION OF THE hMI-ER1 P2 PROMOTER: THE ROLE OF SP1 IN PROMOTER ACTIVITY REGULATION*

2.1 Introduction

It has been revealed that *hmi-er1* transcripts have ubiquitous, but low level, expression of *hmi-er1* in all normal human tissues (Paterno *et al.*, 1998; Paterno *et al.*, 2002). Breast carcinoma cell lines and tumours, on the other hand, expressed significantly elevated levels (Paterno *et al.*, 1998), suggesting that *hmi-er1* expression is associated with the neoplastic state in some human breast carcinomas. *Cis*-regulatory elements in the *hmi-er1* promoter(s) would represent an important element for the specific expression of this gene. Isolation and analysis of the *hmi-er1* promoter(s) may thus shed light to a better understanding of the molecular mechanism for the transcriptional regulation of *hmi-er1*. To locate the putative promoter, extensive CapSelecting rapid amplification of cDNA ends (RACE) was performed, and two

---

* **Part of this chapter has been published in:** Paterno, G.D., Ding, Z., Lew, Y.Y., Nash, G.W., Mercer, F.C., and Gillespie, L.L. (2002) Genomic organization of the human mi-er1 gene and characterization of alternatively spliced isoforms: regulated use of a facultative intron determines subcellular localization. *Gene*. 295:79-88.

Part of this chapter has been published in: The SANT domain of human MI-ER1 interacts with Sp1 to interfere with GC box recognition and repress transcription from its own promoter. J Biol Chem. 2004 Apr 26 [Epub ahead of print].

transcriptional start sites were found in the exons A1 and B1, respectively (section 2.3.1 and Paterno *et al.*, 2002). The genomic DNA sequences encompassing the two transcriptional start sites were then cloned. Two alternative promoters, designated as P1 and P2, were found (section 2.3.1 and Paterno *et al.*, 2002). Further characterization of the promoter focused on the P2 promoter of *hmi-er1*.

2.2 Materials and Methods

2.2.1 Cell culture

Human cell lines, HeLa cervical carcinoma cells, C33A cervical carcinoma cells, MCF-7 breast carcinoma cells, BT-20 breast carcinoma, SK-OV-3 ovarian carcinoma, HEK 293 transformed embryonic kidney cells and U87 glioblastoma cells, were obtained from the American Tissue Culture Collection and cultured at 37°C in 5% CO_2 in DMEM containing 10% FCS.

2.2.2 CapSelecting RACE for *hmi-er1 5'* transcriptional start sites

In order to produce clones with complete 5 ends, cDNA preparations were enriched for full-length 5 ends using the CapSelect method (Schmidt and Mueller, 1999). Multiple clones were obtained by 5 RACE and sequenced.

2.2.3 Plasmids and constructs

A 3045 bp genomic sequence, containing 1264 bp of sequence downstream of the first transcriptional start site of *hmi-er1* and 1781 bp of genomic sequence upstream of the first transcription start site of *hmi-er1* (section 2.3.1 and Paterno *et al.*, 2002) was generated by PCR [the PCR product was designated as (-1781)] from human endocervical primary cell genomic DNA, using the primer pairs listed in Table 2.1. A 1460 bp sequence, containing 144 bp of genomic sequence downstream of the second transcriptional start site of *hmi-er1* and 1316 bp of genomic sequence upstream of the second transcription start site of *hmi-er1* (section 2.3.1 and Paterno *et al.*, 2002) was generated by PCR [the PCR product was designated as (-1316)] from human endocervical primary cells, using the primer pairs listed in Table 2.1. The PCR product (-1781) and (-1316) was cloned into the pCR2.1 vector [pCR(-1781) and pCR(-1316)], using the TOPO-TA cloning kit (Invitrogen, Inc.) and sequenced on both strands. The sequences were verified by comparison to the previously reported *hmi-er1* sequences (Paterno *et al.*, 2002) and to the human genome sequence data (Sanger Centre). The pGL3(-1781) plasmid was generated by digestion of pCR(-1781) with *Kpn*I/*Eco*RI and subcloning into *Kpn*I/*Sma*I sites of the promoterless pGL3-Basic vector (Promega Corporation). The pGL3(-1316) plasmid was generated by digestion of pCR(-1316) and subcloning into the *Xho*I/*Hind*III sites of the promoterless pGL3-Basic vector.

The pGL3(-133) construct was generated by *Sma*I digestion of pGL3(-1316) followed by re-ligation. The pGL3(-657) construct was generated by *Sma*I/*Bgl*II

Table 1 Table 2.1 PCR primer pairs used for preparing *hmi-er1* promoter constructs.

Construct	Forward primer	Reverse primer
(-1781)	5'-TGCAGGTTGGTAGCCTAG AAGCAACA-3'	5'-TCCGTCTTGTCTGCATTG AACC-3'
(-1316)	5'-GACTGTCTGTAGACTCTT TTCC-3'	5'-CGTACTGCCGGGTCACAT CTCC-3'
(-945)	5'-TGAAGATTAGGAAAAAAA TCCCAGTC-3'	5'-CGTACTGCCGGGTCACAT CTCC-3'
(-469)	5'-GATATAGAATTTTACATT TCCTGTCG -3'	5'-CGTACTGCCGGGTCACAT CTCC-3'
(-312)	5'-ACGTATTTTTCCTCTGCT GTGTCA -3'	5'-CGTACTGCCGGGTCACAT CTC C-3'
(-68)	5'-TTTCCCTCCAGTCCAGCC CAGCCG-3'	5'-CGTACTGCCGGGTCACAT CTCC-3'
(+28)	5'-AGTGGCGGCGGGAGCGG CAGAGA-3'	5'-CGTACTGCCGGGTCACAT CTCC-3'

digestion of pCR(-1316) and subcloning into the *SmaI/BamHI* sites of pGL3-Basic. The other deletion mutants were constructed by PCR, using the primer pairs listed in Table 2.1; the PCR products were cloned into pCR2.1, then subcloned into pGL3-Basic, using either the *XhoI/Hind*III sites [pGL3(-469), (-312), and (+28)] or the *KpnI/Xho*I sites [pGL3(-945) and (-68)].

Full-length human Sp1 cDNA was obtained from Dr. R. Tjian, University of California. The Sp1 expressing construct was generated by PCR from a testis library using 5 -CAAGATCACTCCATGGATGAAATGACAG-3 and 5 -TGCCTGATCTCA GAAGC CATTGCCA-3 as forward and reverse primers, respectively, and cloned into pCR3.1. The GST-Sp1 fusion was constructed by subcloning the *EcoR*I fragments from Sp1-pCR3.1 into the *EcoR*I sites of pGEX-4T-2. All plasmids were sequenced to verify the junctions and the Sp1 sequence.

2.2.4 Computer analysis of the *hmi-er1* promoter region

To identify regulatory DNA elements in the 1460 bp sequence of P2 promoter (-1316), computer-assisted analysis was performed using the following programs: (1) for promoter prediction: Promoter Scan (PROSCAN) at http://bimas.dcrt.nih.gov /molbio/proscan/ and Neural Network Promoter Prediction (NNPP) at http://www.fruitfly.org/seq_tools/promoter.html; (2) for CpG islands: Webgene at http://www.itba.mi.cnr.it/webgene; (3) for transcription factor binding sites: Transcription Factor Binding Site (TFSEARCH) at http://www.cbrc.jp/research

/db/TFSEARCHJ.html; Transcription Element Search System (TESS) at http://www.cbil.upenn.edu/tess/; and PROSCAN.

2.2.5 Transfection and reporter assays

All transfections were performed as previously described (Paterno *et al.*, 1997) in duplicate in 6-well plates. 1.5×10^5 cells/well was seeded 18 h prior to transfection and cells were harvested after 48 h in culture. Luciferase assays were performed on cell lysates using a Monolight 2010 Luminometer (Analytical Luminescence Laboratory) and a luciferase assay reagent (Promega), according to the manufacturer s directions. The values obtained, in relative luciferase units (RLU), were normalized to the level of cellular protein in each sample. Each experiment was repeated three times.

2.2.6 Western blot analysis

Western blot analysis was performed as described previously (Ryan and Gillespie, 1994) with anti-Sp1 monoclonal antibody (dilution 1:100) (Catalog No. 1C6: sc-420, Santa Cruz Biotechnology, Inc.).

2.2.7 GST-fusion protein production

GST fusion proteins were expressed in *E. coli* BL21 and purified according to the instructions supplied with the pGEX-4T vector. GST fusion protein level and purity were determined by SDS-PAGE.

2.2.7 Electophoretic mobility shift assays (EMSAs)

EMSAs were performed as (Kinzler and Vogelstein, 1990). Briefly, the *hmi-er1* P2 minimal functional promoter (-68) was labelled with γ-[^{32}P]ATP and T4 polynucleotide kinase, and purified on NucTrap Probe Purification Columns (Stratagene, Inc.). The labelled probe was incubated with 2 µl of HeLa nuclear extract (Promega) or 1 µg GST-Sp1 protein at room temperature for 20 min in 20 µl containing 5% glycerol, 5 mM $MgCl_2$, 1 mM dithiothreitol, 50 mM KCl, 10 µM $ZnSO_4$, 85 µg/ml bovine serum albumin (BSA), and 50 mM HEPES (pH 7.5). Poly(dI-dC) was used as heterologous competitor in the reaction (2 µg/reaction). Where indicated, a 20 molar excess of unlabelled probe (competitor) was included in the binding reaction to demonstrate specificity for the probe. For antibody supershift assays, the extract was incubated for 30 min at room temperature with 1 µl anti-Sp1 polyclonal antiserum (Catalog No. PEP 2 X: sc-59, Santa Cruz Biotechnology, Inc.). Bound and free probes were resolved by non-denaturing electrophoresis on 4% polyacrylamide gels and analyzed by autoradiography.

2.3 Results

2.3.1 Cloning of the *hmi-er1* 5 regulatory sequences

In order to obtain cDNA clones with complete ends, cDNA preparations were enriched for full-length 5 ends using the CapSelect method (Schmidt and Mueller, 1999). Multiple cDNA clones were obtained by extensive 5 RACE and sequenced. Sequence analysis of these clones and BLAST comparison to the human genome (Sanger

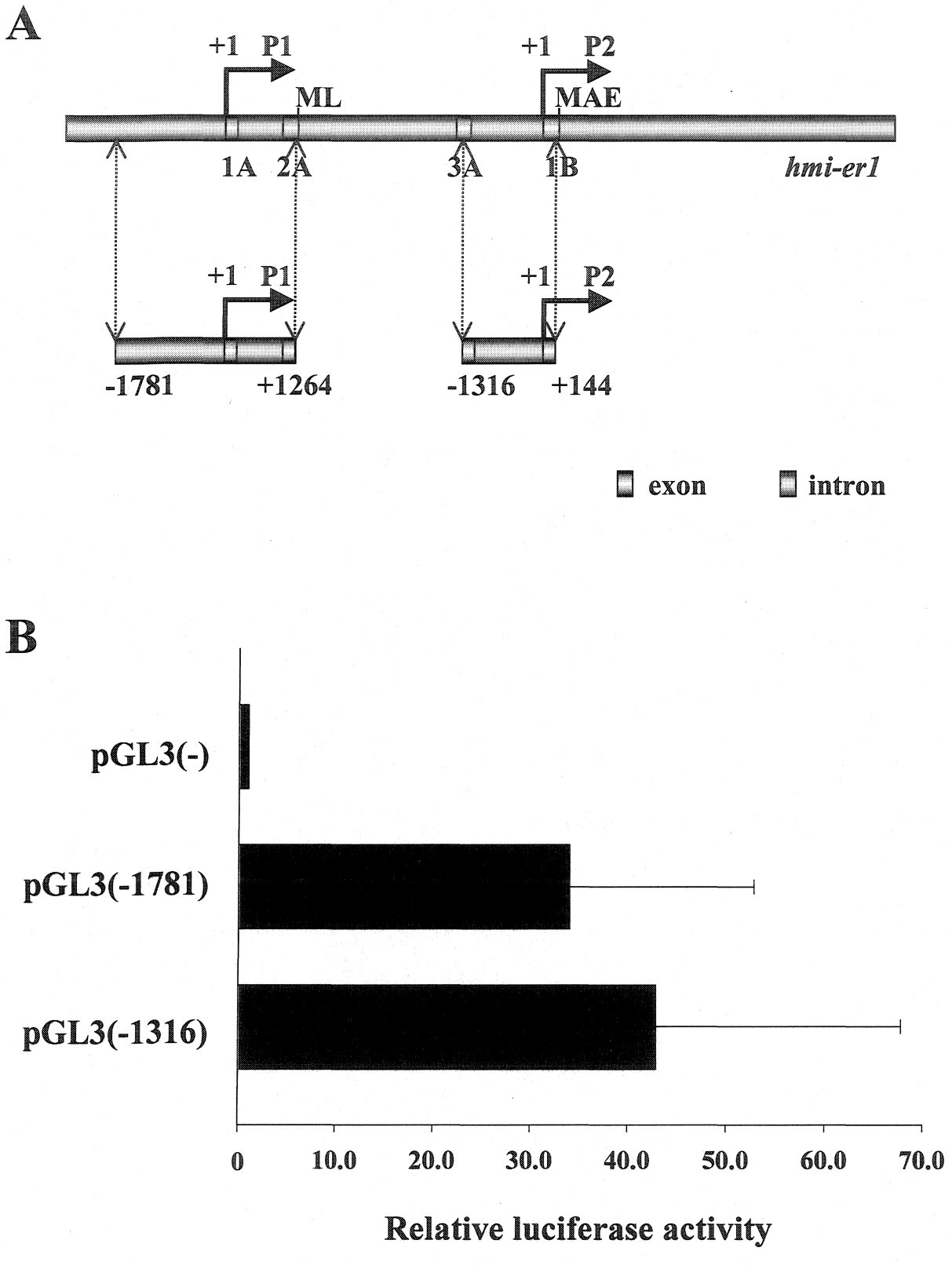
Centre) revealed that there are two *hmi-er1* transcription start sites, which designated as P1 and P2, in exons A1 and B1, respectively (Figure 2.1A, and Paterno *et al.*, 2002). Similarity searches of the human EST database revealed that all reported sequences are represented in the cDNA clones obtained. Organization of these two transcriptional starts suggested the presence of two distinct promoters of *hmi-er1*.

To provide insight into the regulation of *hmi-er1* expression, two genomic DNA fragments encompassing the two transcriptional start sites were isolated by PCR from human endocervical tissue and were subcloned into the promoterless luciferase reporter vector, pGL3-Basic. As shown in Figure 2.1A, the genomic DNA fragment encompassing the first transcriptional start site contains 3045 bp upstream of the ML translational start site. It contains 1264 bp downstream of transcriptional start site (+1 to +1264) and 1781 bp (-1 to -1781) upstream of transcriptional start site. The genomic DNA fragment encompassing the second transcriptional start site in exon B1 contains 1460 bp of sequence upstream of the MAE translational start site. It contains 144 bp of sequence downstream of the transcriptional start site (+1 to +144) and 1316bp (-1 to -1316) of sequence upstream of the transcriptional start site.

The promoter activities of these DNA fragments were analyzed *in vivo* using luciferase reporter assays in HeLa cervical carcinoma cells. As shown in Figure 2.1B, the luciferase activity of luciferase reporter construct driven by the genomic DNA fragment encompassing the P1 transcription start site [pGL3(-1781)] and the genomic DNA fragment encompassing the P2 transcription start site [pGL3(-1316)] was significantly higher, compared to the pGL3-Basic vector [pGL3(-)]. These results

Figure 2.1 Promoter activities of the DNA fragments encompassing either P1 or P2 5' flanking sequence.

(A) Diagrammatic representation to show the two transcriptional start sites, P1 and P2, respectively. The two alternate starts of translation, ML- and MAE- are indicated. Exon numbers are indicated below each schematic. Two genomic DNA fragments encompassing either P1 or P2 isolated by PCR are shown below. **(B)** Promoter activity of the *hmi-er1* P1 and P2 promoters in HeLa cells. HeLa cells were transiently transfected with the luciferase reporter vector pGL3-Basic [pGL3(-)] or pGL3-Basic containing the DNA fragments of *hmi-er1* 5 -flanking the two transcriptional start sites, for P1: 3045 bp (-1781 to -1 and +1 to +1264) [(pGL3(-1781)] and for P2: 1460 bp (-1316 to -1 and +1 to +144) [pGL3(-1316)]. Following transfection, the cells were cultured for an additional 48 h and then lysed in luciferase lysis buffer, and the relative luciferase units (RLU) and β-galactosidase activity were determined. The luciferase activity per μg protein was then normalized to β-galactosidase activity per μg protein. Luciferase activity is presented as the fold increase of luciferase activity from pGL3(-1316) relative to that from the pGL3-Basic vector control. Shown are the average values and standard deviation for three independent experiments.


A
+1 P1
+1 P2
ML
MAE
1A 2A
3A 1B
hmi-er1
+1 P1
+1 P2
-1781
+1264
-1316
+144
exon
intron
B
pGL3(-)
pGL3(-1781)
pGL3(-1316)
0
10.0
20.0
30.0
40.0
50.0
60.0
70.0
Relative luciferase activity

indicated that these two 5 flanking genomic DNA sequences function as transcription promoters, which were designated as *hmi-er1* P1 and P2 promoters.

Further studies were focused on the characterization of the *hmi-er1* P2 promoter (-1316), which was located in a relatively shorter DNA fragment. Further experiments were performed to test the luciferase activity of *hmi-er1* P2 promoter-driven luciferase reporter construct pGL3(-1316) in six different cell lines: C33A, HEK293, BT-20, MCF-7, SK-OV-3 and U87. Luciferase activity was high in all pGL3(-1316)-transfected cells and ranged from 10-fold to 34-fold higher than pGL3-Basic controls (Figure 2.2), demonstrating that the *hmi-er1* P2 promoter (-1316) can function in a variety of cell types.

2.3.2 Location of the *hmi-er1* P2 minimal promoter

Computer-assisted analysis of the 1460 bp P2 DNA fragment using Webgene at http://www.itba.mi.cnr.it/webgene to identify potential *cis*-regulatory elements predicted a CpG island located between nucleotide position -389 and the start of translation (Figure 2.3A). Further analysis revealed that this GC-rich region does not contain a TATA box; nor does it contain an Inr or a DPE sequence, as is often found in TATA-less promoters (reviewed in Butler and Kadonaga, 2002). A number of potential transcription factor binding sites were identified using TFSEARCH, PROSCAN and TESS, including 2 TEF-2 sites, an HT4F site, an estrogen response element (ERE) half-site as well as multiple binding sites for the Sp1 family (Figure 2.3B).

Figure 2.2 *hmi-er1* P2 promoter activities in multiple cancer cell lines.

hmi-er1 P2 promoter-driven luciferase reporter construct pGL3(-1316) was transfected in C33A, MCF-7, BT-20, SK-OV-3, U87, and 293K human cell lines. Following transfection, the cells were cultured for an additional 48 h and then lysed in luciferase lysis buffer, and the level of relative luciferase units (RLU) and β-galactosidase activity were determined. The luciferase activity per μg protein was then normalized to β-galactosidase activity per μg protein. Luciferase activity is presented as the fold increase of luciferase activity from pGL3(-1316) relative to that from the pGL3-Basic vector control. Shown are the average values and standard deviation for three independent experiments.

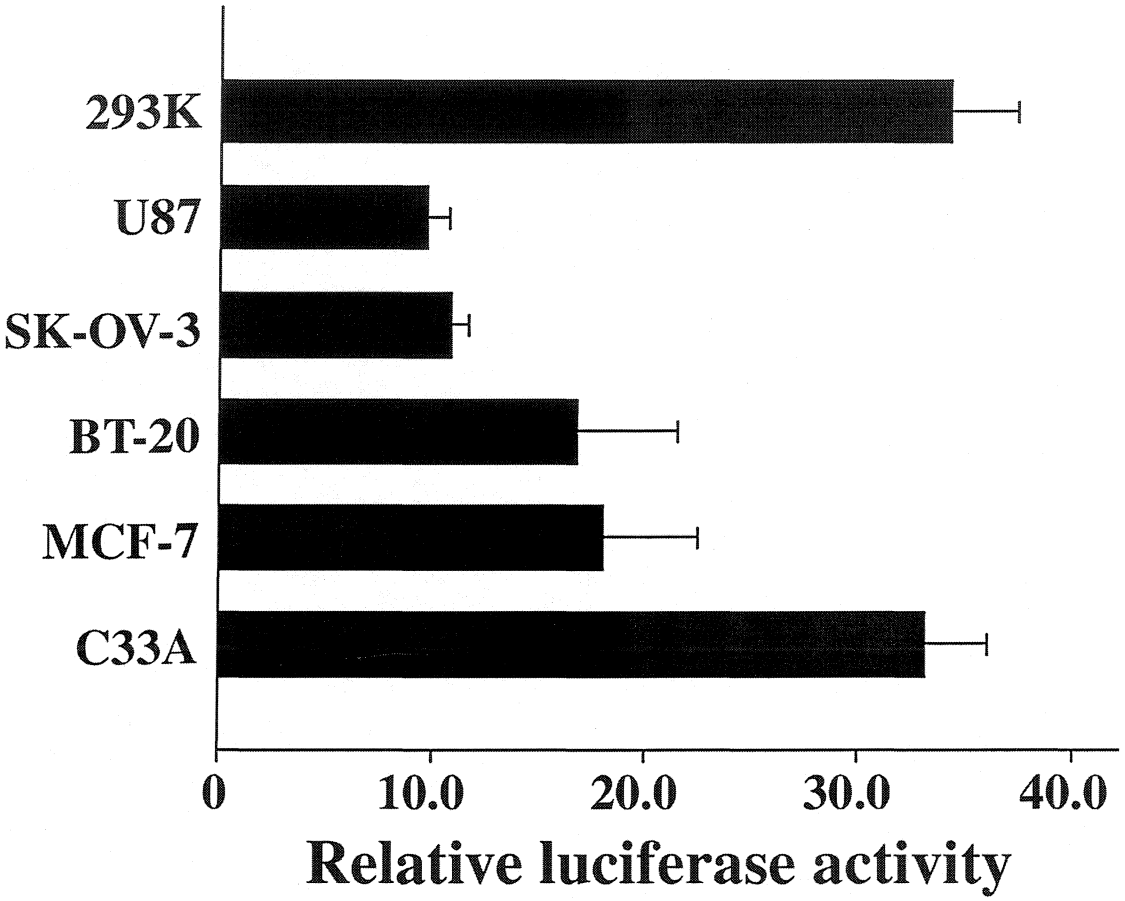
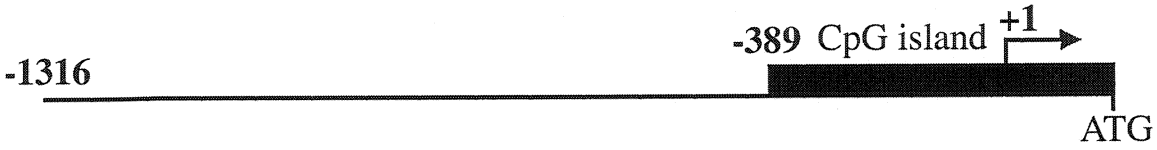

293K
U87
SK-OV-3
BT-20
MCF-7
C33A
0
10.0
20.0
30.0
40.0
Relative luciferase activity

Figure 2.3 Nucleotide sequence of *hmi-er1* P2 promoter, putative transcription factor binding sites and CpG island.

(A) Diagrammatic representation of *hmi-er1* P2 promoter (-1316). The first nucleotide of the *hmi-er1* translation initiation site ATG (+1) is indicated by an arrow. The position of a CpG island is shown as a black rectangle. **(B)** Nucleotide sequence of GC-rich region of the *hmi-er1* P2 promoter (-1316). The number of nucleotides upstream of the first nucleotide A (+1) of *hmi-er1* P2 transcriptional start site are shown on the left of the sequence. The promoter region is designated by the gray region. The GC-rich Sp1 binding sites are shown as white nucleotides in black rectangles. Sequence motifs for other potential transcription factor binding are underlined and designated.

A



B

-389 TTCTTGCCGCCGTTGCATCACAGCAACATCCGT

-356 TTTTCAGCAAATGCATTTCAAAAACGACCTACATGTAAAATATTACGTATTTTTCCTCTG
Elf-1

-296 CTGTGTCAATGCTGGATGTGACTCCTTTGGCACTGTTCCTTGACCTCTCTGATCCGTAGA
1/2 ERE

-236 CAACCCCCTCCGCCTGGAGTGGCGGATCAGCTGGAGCCAGCGAAGCGCCCCGCGCGGTTG
-133

-176 CCCACCTCCTCCCACACCCACCTTGACTCCGCCCCCTCCGCTCTTCCCGGGGAGGGCTGG
TEF-2 -68 H4TF-2

-116 CCGCGGGGCCGCGCGCGCCCCTGCTCCGGCGCGTGCTCGCTGGTCTTTTCCCTCCAGT

-56 CCAGCCCAGCCGGGGCGCCGCGAGGGGGCGGAGTGGGGTGTGGTGGGCGCGCTCGGGCGG
TEF-2 +1

+28

+5 CTCCTGCGCGTTCCCGCCGAGGCAGTGGCGGCGGGAGCGGCAGAGACGGCAGCGGCCGGA

+65 GTCCCGTTGCTGAGTCTCACATCCGGGTTCTGGCCGTGACCCAGCTGCGGCCGCCGCGGA

+125 GATGTGACCCGGCAGTACGGCAAATATG

To identify the *hmi-er1* P2 minimal functional region, a series of 5 deletion fragments was constructed in pGL3 and each construct was transiently transfected into HeLa cells. Deletion of nucleotides -1316 to -133 did not result in a significant change in luciferase activity. Further deletion to nucleotide -68 reduced, but did not abolish luciferase activity, while deletion to nucleotide +28 completely abolished luciferase activity (Figure 2.4). These results indicated that the *hmi-er1* P2 minimal functional promoter is located within the sequence -68/+144, contained in the (-68) construct.

2.3.3 Sp1 binds to the *hmi-er1* P2 minimal promoter and activates transcription

Since the *hmi-er1* P2 minimal promoter (-68) is predicted to contain four Sp1 binding sites, a number of experiments were performed to investigate the possibility that Sp1 would bind to this sequence. EMSAs were performed using a HeLa cell nuclear extract and ^{32}P-labelled (-68) as a probe. Two bands representing DNA-protein complexes appeared in samples containing nuclear extract (Figure 2.5A, lane 2 band b and c), however only the higher band (band b) was specific, as revealed by competition with excess nonlabelled (competitor) probe (Figure 2.5A, lane 3). Supershift assays, using a Sp1 antibody that does not cross-react with SP2, SP3 or SP4, resulted in the appearance of an additional high mobility band (band a) (Figure 2.5A, lane 4) that disappeared in the presence of excess nonlabelled (competitor) probe (Figure 2.5A, lane 5), thus confirming the presence of Sp1 in the DNA-protein complexes. The binding of Sp1 protein to the *hmi-er1* P2 minimal promoter (-68) was confirmed by the specific

Figure 2.4 Localization of the *hmi-er1* P2 minimal promoter.

The diagram on the left represents the luciferase (Luc) constructs made by a series of deletions extending from +144 bp downstream of the P2 transcriptional start site. The 5 end of the nt of each construct is listed. HeLa cells were transfected with the constructs indicated on the left. Following transfection, the cells were cultured for an additional 48 h and then lysed in luciferase lysis buffer, and the level of relative luciferase units (RLU) and β-galactosidase activity were determined. The luciferase activity per μg protein was then normalized to β-galactosidase activity per μg protein. Luciferase activity is presented as the fold increase of normalized luciferase activity relative to that from the pGL3(-) vector control. Shown are the average values and standard deviation for three independent experiments.

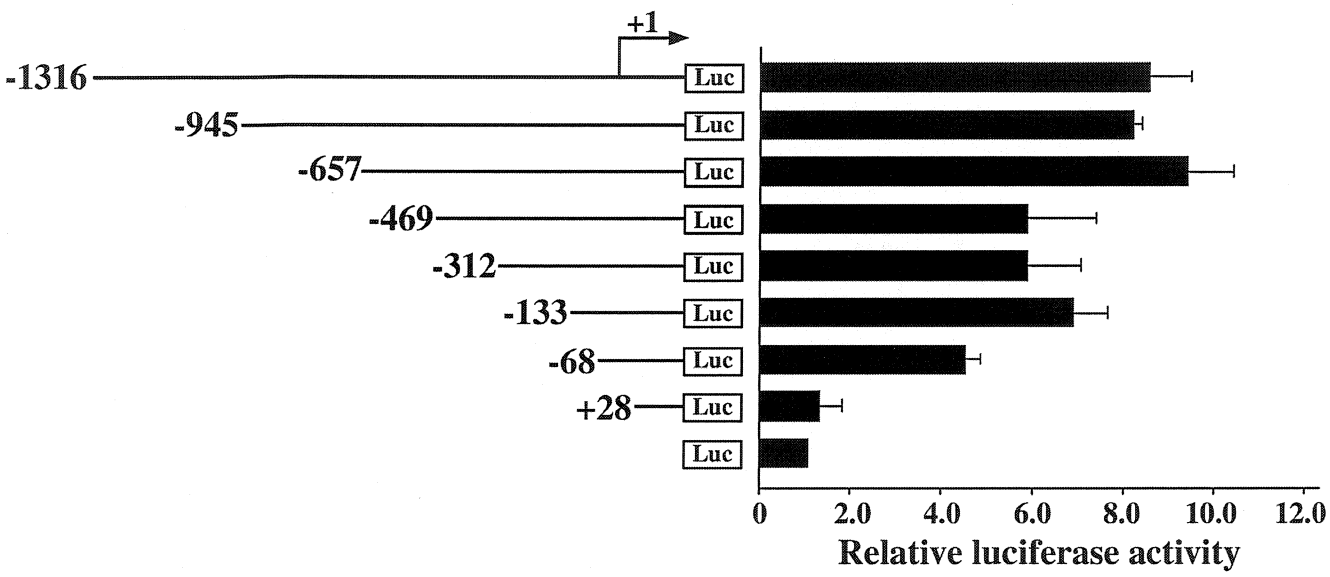
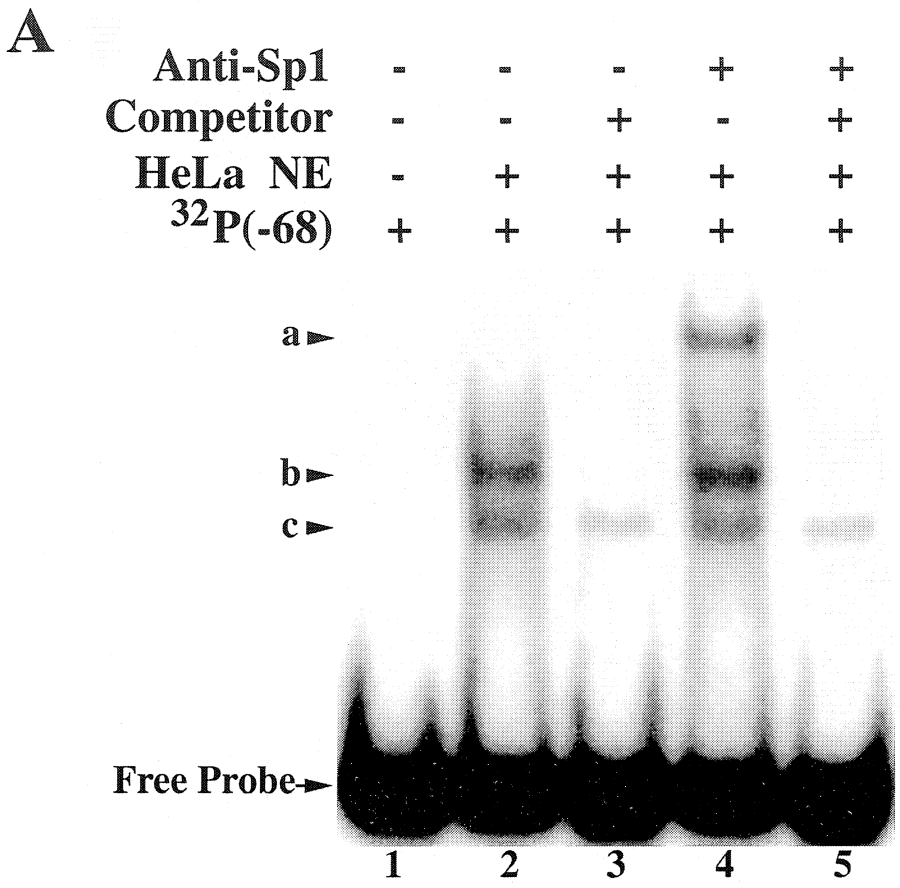
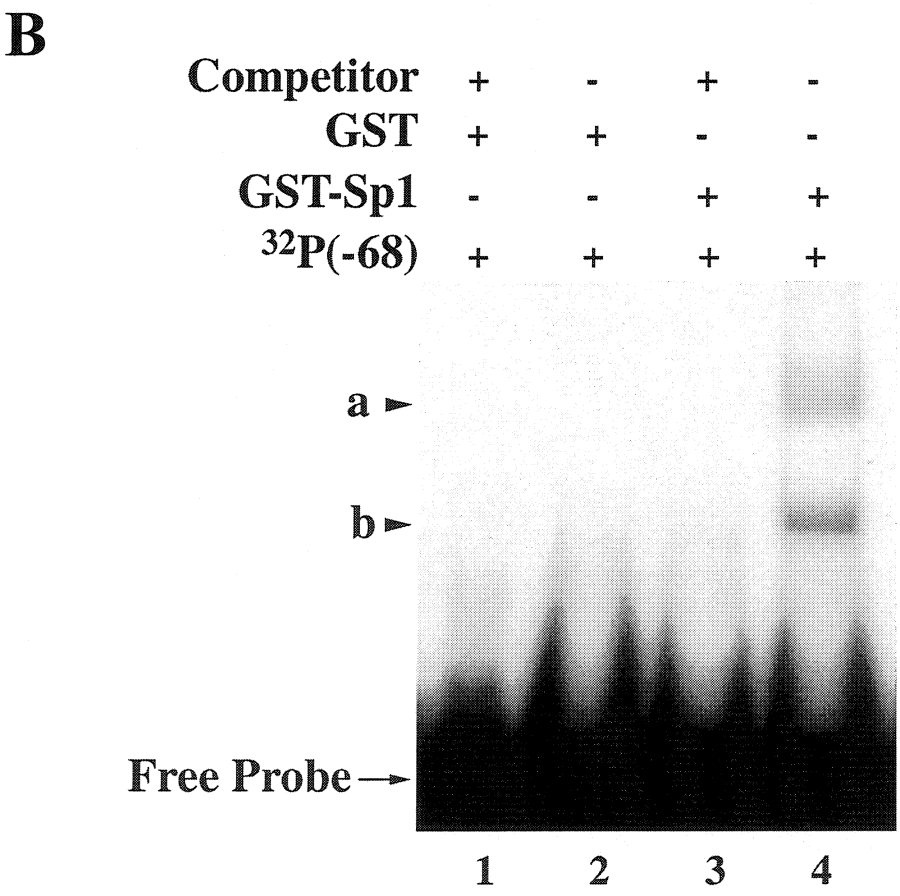

+1
-1316
-945
-657
-469
-312
-133
-68
+28
Luc
0
2.0
4.0
6.0
8.0
10.0
12.0
Relative luciferase activity

Figure 2.5 Sp1 binds to the *hmi-er1* P2 minimal promoter.

(A) Shown is a representative EMSA performed using HeLa cell nuclear extracts and a ^{32}P-labelled the *hmi-er1* P2 minimal promoter (-68) as a probe [^{32}P(-68)]. In lane 1, the labelled probe was incubated without HeLa nuclear extract. In lanes 2-5, HeLa nuclear extracts were incubated with (lane 3, 5) or without (lane 2, 4) 20-fold molar excess of nonlabelled (competitor) probe and with (lanes 4, 5) or without (lanes 2, 3) 1 μl anti-Sp1 antibody. The positions of DNA-protein complexes (band a, b and c) are indicated by arrowheads. The position of the free probe is indicated by arrow. **(B)** Shown is a representative EMSA performed using ^{32}P(-68) and the indicated GST-fusion protein, in the presence (lanes 2) or absence (lanes 1) of 20-fold molar excess of nonlabelled (competitor) probe. The positions of DNA-protein complexes (band a, b) are indicated by arrowheads. The position of the free probe is indicated by arrow.


A
Anti-Sp1 - - - + +
Competitor - - + - +
HeLa NE - + + + +
32P(-68) + + + + +
a
b
c
Free Probe
1 2 3 4 5


B
Competitor + - + -
GST + + - -
GST-Sp1 - - + +
32P(-68) + + + +
a
b
Free Probe
1 2 3 4

retarded bands a and b in EMSAs using purified GST-Sp1 fusion protein (Figure 2.5B, lane 4), which was significantly competed out with nonlabelled (competitor) probe (Figure 2.5B, lane 3). In contrast, GST control did not bind to the *hmi-er1* P2 minimal promoter (-68) (Figure 2.5B, lane 1 and 2). Thus, these experiments demonstrated that Sp1 binds to the *hmi-er1* P2 minimal promoter (-68).

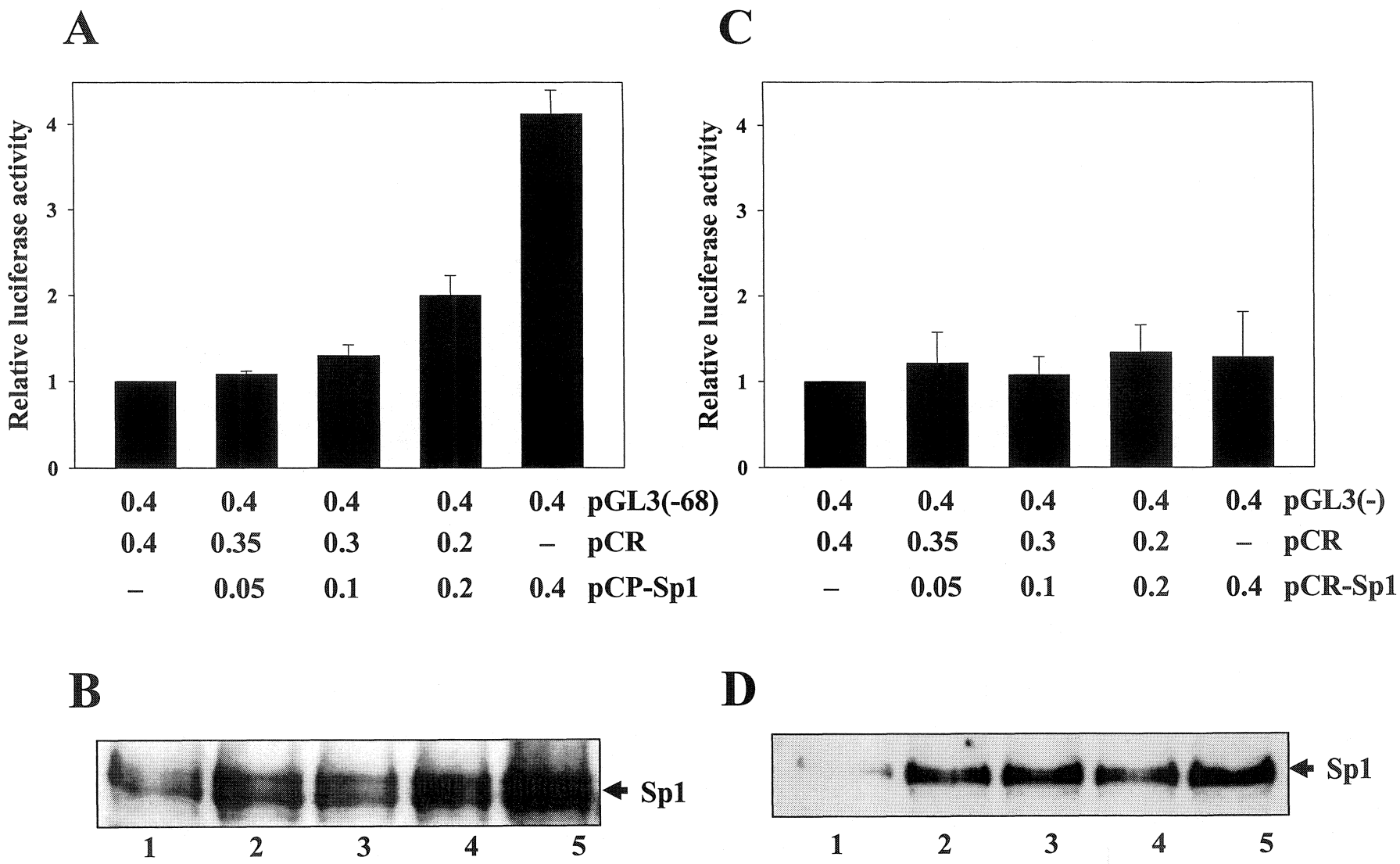
The ability of Sp1 to regulate transcription from the *hmi-er1* P2 minimal promoter (-68) was investigated *in vivo* using luciferase reporter assays. HeLa cells were co-transfected with pGL3(-68), along with a Sp1 expression vector (pCR-Sp1) or control empty vector (pCR). Overexpression of Sp1 (Figure 2.6B) resulted in a dose-dependent increase in luciferase activity of pGL3(-68), compared to pCR empty vector control (Figure 2.6A). However, overexpression of Sp1 (Figure 2.6C) did not significantly alter the luciferase activity of pGL3(-) (Figure 2.6D). These results indicated that Sp1 can activate transcription from the *hmi-er1* P2 minimal promoter (-68).

2.4 Discussion

The transcription of a eukaryotic gene is regulated by the combined action of general transcription factors, mediators, multiple sequence-specific transcription factors, and histone modifiers including histone acetyltransferase and histone deacetylase that regulate the activities of the transcription factors and the chromatin structure (reviewed in Dynan and Tjian, 1985; Naar *et al.*, 2001; Pazin and Kadonaga, 1997). *hmi-er1* expression was found to be associated with the neoplastic state in human breast

Figure 2.6 Overexpression of Sp1 enhances *hmi-er1* P2 promoter activity.

HeLa cells were co-transfected with the indicated promoter-luciferase reporter constructs with the pCR3.1 (pCR) empty vector, pCR3.1-Sp1 (pCR-Sp1). The amount of plasmid (μg) used for transfection is indicated below each bar. Cells were harvested 48h after transfection and the level of relative luciferase units (RLU) was determined, and this value was normalized to the level of cellular protein in each sample. For each sample, the value is presented as the fold increase over that obtained with the pCR empty vector. **(A)** Shown are the average values and standard deviations from three independent experiments with *hmi-er1* P2 minimal promoter-driven luciferase reporter construct pGL3(-68). **(B)** The level of Sp1 protein expressed in each sample described in panel A was determined by Western blot, using anti-Sp1 antibody. Shown is a representative Western blot, and the position of the Sp1 is indicated by arrow. **(C)** Shown are the average values and standard deviations from three independent experiments with luciferase reporter construct pGL3-Basic [pGL3(-)]. **(D)** The level of Sp1 protein expressed in each sample described in panel C was determined by Western blot analysis, using anti-Sp1 antibody. Shown is a representative Western blot, and the position of the Sp1 is indicated by arrow.


A
Relative luciferase activity
0
1
2
3
4
0.4 0.4 0.4 0.4 0.4 pGL3(-68)
0.4 0.35 0.3 0.2 – pCR
– 0.05 0.1 0.2 0.4 pCP-Sp1
C
Relative luciferase activity
0
1
2
3
4
0.4 0.4 0.4 0.4 0.4 pGL3(-)
0.4 0.35 0.3 0.2 – pCR
– 0.05 0.1 0.2 0.4 pCR-Sp1
B
Sp1
1 2 3 4 5
D
Sp1
1 2 3 4 5

carcinoma (Paterno *et al.*, 2002), and plays a role in the regulation of cell growth (Paterno *et al.*, 1998; Paterno *et al.*, unpublished data). As a first step toward understanding the molecular mechanism of regulation of hMI-ER1 expression, the *hmi-er1* P1 promoter (-1781) and P2 promoter (-1316) were cloned (Figure 2.1). Analysis of the P2 promoter (-1316) sequence revealed that this sequence contains a CpG island within which a number of putative transcription factor binding sites were identified. These included 9 Sp1 sites, 2 TEF-2 sites, an HT4F site, an ERE half-site (Figure 2.3). The ERE half-site is of particular interest, as many estrogen responsive genes containing 1/2ERE motifs also possess GC boxes in close proximity such that estrogen receptor (ER) protein co-activator complexes co-operate with Sp1 to regulate transcription (Sanchez *et al.*, 2002). This suggests that *hmi-er1* may be an estrogen responsive gene, and experiments are currently being carried out to test this hypothesis (Paterno *et al.*, unpublished data).

Although transcription regulators, such as Sp3, Sp4 and BTEB3, may compete with Sp1 for binding to GC boxes and repress transcription (Kaczynski *et al.*, 2001), my study showed that overexpression of Sp1 binds to, and enhances transcriptional activity from the *hmi-er1* P2 minimal promoter (-68) (Figure 2.5 and 2.6). Sp1 is a sequence-specific transcription factor that binds GC and GT boxes to activate a wide range of viral and cellular genes (reviewed in Philipsen and Suske, 1999). Overexpression of Sp1 enhances the transcriptional activity from many gene promoters such as that associated with Interleukin-1β (IL-1β) (Chadjichristos *et al.*, 2003). Sp1 interacts with $dTAF_{II}110$ (Gill *et al.*, 1994; Hoey *et al.*, 1993) and $hTAF_{II}130$ (Tanese *et al.*, 1996) through its

activation domain, and with $hTAF_{II}55$ through its DNA-binding domain (Chiang and Roeder, 1995), suggesting that Sp1 may facilitate the assembly of the general transcription machinery *via* multiple protein-protein interactions with components of TFIID. Sp1 also recruits an additional cofactor complex, CRSP, to mediate synergistic activation with SREBP-1a on chromatin templates (Naar *et al.*, 1999; Ryu *et al.*, 1999). Sp1 and the transcriptional coactivators CBP/p300 are associated in complexes that regulate NGF- and progesterone-mediated induction of p21Waf/Cip1 (Billon *et al.*, 1999; Owen *et al.*, 1998). In addition to its role as a transcriptional activator, Sp1 has also been reported to repress transcription of murine TK promoter by recruiting histone deacetylase activity (Doetzlhofer *et al.*, 1999). Therefore, Sp1 is important both in transcription activation and repression in a promoter context-specific and a cell-type specific manner (Bouwman and Philipsen, 2002).

The presence of a CpG island in the *hmi-er1* P2 promoter (-1316) indicates that promoter activity may also be regulated by DNA methylation. Methyl-CpG binding proteins such as MeCP2 and MBD3 recruit chromatin remodelling and modification complexes, including the NURD complex (Ng *et al.*, 1999; Tyler and Kadonaga, 1999; Zhang *et al.*, 1999). It has been proposed that DNA methylation and histone deacetylation might work together to establish a repressive chromatin environment and silence gene expression (Cameron *et al.*, 1999). Moreover, Sp1 has also been linked to the maintenance of methylation-free CpG islands (Macleod *et al.*, 1994) and hypermethylation around Sp1 binding sites reduces Sp1 binding, thereby decreasing transcription (Zhu *et al.*, 2003). Aberrant methylation on the promoter CpG islands of

genes required for cell growth regulation such as $p16^{INK4A}$ has been found in certain cancers (reviewed in Esteller and Herman, 2002; Nephew and Huang, 2003). For example, aberrant methylation of the promoter of the tumour suppressor genes $p16^{INK4A}$ is closely related to tumour growth characteristics (Bender *et al.*, 1998; Cameron *et al.*, 1999; Issa *et al.*, 1994). Deregulated expression of hMI-ER1 in some breast carcinomas (Paterno *et al.*, 1998; and data at http://genome-www5.stanford.edu/cgi-bin/SMD/source/expressionSearch?option=cluster&criteria=Hs.222746&dataset=2&organism=Hs) thus may be due to aberrant methylation on the promoter CpG island of *hmi-er1*.

SECTION III INVESTIGATION OF hMI-ER1α AND β INTERACTING PROTEINS AND hMI-ER1α AND β FUNCTION

CHAPTER 3 TRANSCRIPTIONAL REPRESSION BY hMI-ER1 INVOLVES RECRUITMENT OF HISTONE DEACETYLASE *VIA* THE ELM2 DOMAIN

3.1 Introduction

Structurally, hMI-ER1α and β isoforms share a number of features with other transcriptional regulators. hMI-ER1α and β contain an acidic transcription activation domain, an ELM2 domain, and a signature SANT domain. The latter was first identified in the transcription factors SWI3, ADA2, N-COR, and TFIIIB, from which the acronym SANT is derived (Aasland *et al.*, 1996). Other SANT domain-containing proteins include transcriptional regulatory molecules involved in nuclear hormone activity, such as N-COR (Aasland *et al.*, 1996) and SMRT (Ordentlich *et al.*, 1999), molecules that are components of transcription and chromatin-regulatory complexes, such as MTA1 (Tong *et al.*, 2002), MTA2 (Wade *et al.*, 1999), MTA3 (Fujita *et al.*, 2003), and CoREST (You *et al.*, 2001); and molecules that are important for regulating developmental events, such as EGL-27 (Solari *et al.*, 1999). The SANT domain has been implicated in protein-

---

This chapter has been published in: Ding, Z., Gillespie, L.L., and Paterno, G.D. (2003) Human MI-ER1 alpha and beta function as transcriptional repressors by recruitment of HDAC1 to their conserved ELM2 domain. *Molecular and Cellular Biology,* 23: 250-258.

protein interactions (Aasland *et al.*, 1996), including interactions with complexes containing HDAC enzymatic activity (Guenther *et al.*, 2001; You *et al.*, 2001) and HAT activity (Sterner *et al.*, 2002). The ELM2 (EGL-27 and MTA1 homology 2) domain was first described in EGL-27, a *Caenorhabditis elegans* protein that plays a fundamental role in patterning during embryonic development (Solari *et al.*, 1999). The ELM2 domain is conserved throughout evolution, but to date no function has been ascribed to this motif before this study. More interestingly, ELM2 domain-containing proteins, including xMI-ER1, hMI-ER1, MTA1, MTA2, MTA3 and CoREST, also contain one or more putative SANT domains, suggesting that there might be some important cooperative function between the two motifs (Solari *et al.*, 1999).

Given the potential role of hMI-ER1α and β as novel transcription regulators and their association with tumorigenesis, the role of the hMI-ER1α and β isoforms in transcriptional regulation was investigated.

3.2 Materials and Methods

3.2.1 Cell culture

HeLa human cervical carcinoma, C33A human cervical carcinoma cells, HEK 293 human transformed embryonic kidney cells, and mouse NIH 3T3 fibroblast cells were obtained from the American Type Culture Collection and cultured at 37°C in 5% CO_2 in DMEM containing 10% FCS.

3.2.2 Plasmids and constructs

The plasmids used in this study include the G5tkCAT reporter plasmid (a kind gift from Diane Hayward, Johns Hopkins School of Medicine), which contains chloramphenicol acetyltransferase (CAT) linked to five GAL4 DNA binding sites and the herpes simplex virus minimal thymidine kinase (tk) promoter (Hsieh *et al.*, 1999; Teodoro and Branton, 1997; Yew *et al.*, 1994); the pM plasmid (Clontech), which contains the GAL4 DBD and an NLS; CS3+MT, containing the Myc epitope tag (a kind gift from David Turner, University of Michigan); the pGBKT7 plasmid (Clontech), HA-HDAC3 (a kind gift from Mark Featherstone, McGill University); and a luciferase T7 control plasmid encoding full-length luciferase protein (Promega).

Expression vectors were engineered to contain full-length or deletion mutants of hMI-ER1α or β fused to the GAL4 DBD of the pM plasmid or to the Myc epitope tag of the CS3+MT plasmid. Specific primers incorporating 5 and 3 *BamH*I sites were used to amplify the entire coding sequence of either hMI-ER1α or β, and the digested PCR fragments were inserted into the *Bgl*II site of either the pM or the CS3+MT plasmid. Fragments encoding the appropriate amino acid residues of hMI-ER1α or β deletion mutants were amplified by PCR using the primer pairs listed in Table 3.1. PCR products were cloned into pCR3.1 using the TA cloning kit (Invitrogen, Inc.), and *Eco*RI fragments were then inserted into the complementary sites of the pM or pGBKT7 plasmid. The deletion constructs were named according to the encoded amino acid residues of the hMI-ER1α or β protein. The GAL4-hMI-ER1(164-283) ^{214}W→A and GAL4-hMI-ER1(164-283) ^{227}FL→AA mutant constructs were generated by site-directed

Table 2 Table 3.1 PCR primer pairs used for constructing hMI-ER1 and mutating hMI-ER1 plasmids in PM vector

Construct[a]	Forward primer	Reverse primer
hmi-er1α aa 1-433	5 -CGGGATCCATATGGCGG AGCCATCTGTTG-3	5 -CGGGATCCAAAACAAGAC CACAGAAGC-3
hmi-er1β aa 1-512	5 -CGGGATCCATATGGCGG AGCCATCTGTTG-3	5 -CGGGATCCTTAGTCATCT GTGTTTTCAAG-3
aa 1-283	5 -CACCATGGCGACATCTG TTGAATC-3	5 -ATCCTCTCTAGCTGCTTTT ACA-3
aa 1-155	5 -CACCATGGCGACATCTG TTGAATC-3	5 -CAAAATATTTACATCGACG TGGGCG-3
aa 1-83	5 -CACCATGGCGACATCTG TTGAATC-3	5 -TCTTCAGGTAGTCGAACAG TAC-3
aa 83-155	5 -CACCATGGAAGAAGATG AGGAAGAGGAAGAAGAG-3	5 -CAAAATATTTACATCGACG TGGGCG-3
aa 83-433	5 -CACCATGGAAGAAGATG AGGAAGAGGAAGAAGAG-3	5 -CGGGATCCAAAACAAGAC CACAGAAGC-3
aa 164-433	5 -CACCATGGAAGAATCTG AAGAAGATGAAGATT-3	5 -CGGGATCCAAAACAAGAC CACAGAAGC-3
aa 287-433	5 -CACCATGGTTTGGACAG AGGAAGAGTGTA-3	5 -CGGGATCCAAAACAAGAC CACAGAAGC-3
aa 287-512	5 -CACCATGGTTTGGACAG AGGAAGAGTGTA-3	5 -CGGGATCCTTAGTCATCT GTGTTTTCAAG-3
aa 164-283	5 -CACCATGGAAGAATCTG AAGAAGATGAAGATT-3	5 -ATCCTCTCTAGCTGCTTTT ACA-3

Construct[a]	Forward primer	Reverse primer
aa 180-283	5 -AAGGAGATTATGGTGGG CTCCATGTTTCAA-3	5 -ATCCTCTCTAGCTGCTTTT ACA-3
aa 164-239	5 -CACCATGGAAGAATCTG AAGAAGATGAAGATT-3	5 -CTTCTCATCACCTGTTCTT CTAGATGCATC-3
aa 240-283	5 -GGTGTAGAAGCAATTCC TGAAGGATCTCAC-3	5 -ATCCTCTCTAGCTGCTTTT ACA-3
aa 164-273	5 -CACCATGGAAGAATCTG AAGAAGATGAAGATT-3	5 -TCATCTTCTCAATGCTTCT TCTGTATCAAAATTGCA-3
aa 164-262	5 -CACCATGGAAGAATCTG AAGAAGATGAAGATT-3	5 -TCATTTAACCAATTCATATA AAGCCTGTTCATTGTC-3
aa 164-251	5 -CACCATGGAAGAATCTG AAGAAGATGAAGATT-3	5 -TCATTTTATGTGAGATCCT TCAGGAATTGCTTC-3
^{214}W→A	5 -GATCAGCTCCTGGCGGA CCCTGAGTACTTACC-3	5 -GGTAAGTACTCAGGGTCC GCCAGGAGCTGATC-3
^{227}FL→AA	5 -GTGATTATAGCTGCTAAA GATGCATCTAGAAGAACAG GTGATGAGAAGG-3	5 -CCTTCTCATCACCTGTTCT TCTAGATGCATCTTTAGCAGC TATAATCAC-3

[a] Deletion constructs were named according to the encoded amino acid residues of the hMI-ER1α or β isoforms

mutagenesis mutagenesis using two complementary primers (Table 3.1) (Oligos, Etc.) designed to produce the mutation and a QuikChange site-directed mutagenesis kit (Stratagene Inc.). All plasmids were sequenced to verify the junctions and the hMI-ER1 sequence.

Full-length human *hdac1* cDNA was amplified by PCR from a testis cDNA library by using 5 -ACGGGAGGCGAGCAAGATGGCG-3 and 5 -TCAGGCCAACTT GACCTCCTCCTTGAC-3 as forward and reverse primers, respectively, and then cloned into pCR3.1 as described above. Plasmid was sequenced to verify the junctions and the *hdac1* sequence (GenBank Database accession number NM_004964).

3.2.3 Transfection and reporter assays

All transfections were performed in duplicate in six-well plates with plasmid DNA as previously described (Paterno *et al.*, 1997). A total of 1.5×10^5 cells/well were seeded 18 h prior to transfection, and cells were harvested after 48 h in culture. For trichostatin A (TSA) treatment, the medium was replaced 24 h after transfection with fresh medium with or without the indicated concentration of TSA, and the cells were cultured for an additional 24 h. Cell lysates were prepared and assayed for CAT protein by using a CAT enzyme-linked immunosorbent assay kit (Boehringer Mannheim) as described previously (Paterno *et al.*, 1997). The level of CAT protein expressed in each sample was determined using a CAT standard curve supplied by the manufacturer, and this value was normalized to the level of cellular protein in each sample.

3.2.4 Co-immunoprecipitation (co-IP) and Western blot analysis

In vitro coupled transcription-translation reactions (TNTs; Promega) were performed as described previously (Ryan and Gillespie, 1994). For *in vitro* co-IP assays, 50,000 cpm each of ^{35}S-labelled and unlabelled TNT mixtures programmed with the appropriate cDNAs were combined and incubated for 3 h at 4°C. Immunoprecipitation was performed as described previously (Ryan and Gillespie, 1994) with 10 µl anti-HDAC1 polyclonal antiserum (Catalog No. H51: sc-7872, Santa Cruz Biotechnology, Inc.) or 10 µl anti-Myc monoclonal antibody 9E10. For all assays, 1/20 volume of the indicated TNT was loaded into the input lanes. All samples were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) followed by autoradiography.

For *in vivo* co-IP assays, either nontransfected HeLa cells or cells transfected with Myc-tagged plamid pCS3+MT (pMyc), pCS3+MT-*hmi-er1*α (pMyc-α), or pCS3+MT-*hmi-er1*β (pMyc-β) were used. Cells were lysed in the immunoprecipitation buffer [50 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol, 1% Triton X-100, 40 mM Sodium Pyrophosphate, 50 mM Sodium Fluoride, 1mM Sodium Vanadate], and the insoluble material was removed by centrifugation at 12,000× *g* for 10 min. For input lanes, one-third of the total amount of protein in the soluble fraction was used. Supernatants were subjected to immunoprecipitation with either 1 µl anti-hMI-ER1 antibody produced in our laboratory (24).

Western blot analysis was performed as described in section 3.2.4 with anti-Myc monoclonal antibody 9E10 (1:100 dilution), anti-HDAC1 monoclonal antibody (1:100

dilution) (Catalog No. H-11: sc-8410, Santa Cruz Biotechnology, Inc.), or anti-GAL4 monoclonal antibody dilution (Catalog No. RK5C1: sc-510, Santa Cruz Biotechnology, Inc.).

3.2.5 HDAC assays

The [^{3}H]acetate-labelled histone substrate for the HDAC enzyme assays was prepared from HeLa cells labelled with [^{3}H]acetate in the presence of 300 nM TSA as described previously (Yoshida *et al.*, 1990). For each sample, 1.5×10^5 HeLa cells were transfected with the appropriate construct and, after 48 h, the cells were lysed as described previously (Hu *et al.*, 2000) and insoluble material was removed by centrifugation at 12,000 × *g* for 10 min. Supernatants were subjected to immunoprecipitation using 1 µl anti-GAL4 polyclonal antiserum (Catalog No. 06-262, Upstate Biotechnology, Inc.), 10 µl anti-Myc monoclonal antibody 9E10, or 10 µl anti-HDAC1 polyclonal antiserum (Santa Cruz Biotechnology, Inc.). The immunoprecipitates were analyzed for HDAC enzymatic activity as described previously (Hu *et al.*, 2000) with 5,000 cpm of [^{3}H]acetate-labelled histone substrate.

3.3 Results

3.3.1 hMI-ER1α and β function as transcriptional repressors

In this study, the transcriptional regulatory activity of the hMI-ER1α and β isoforms was investigated using the G5tkCAT reporter plasmid. This plasmid contains

CAT reporter gene linked to five GAL4 DNA binding sites along with the herpes simplex virus (HSV) minimal tk promoter (Hsieh *et al.*, 1999; Teodoro and Branton, 1997; Yew *et al.*, 1994) to provide a constitutive level of CAT expression; and thereby allows for a measure of both activation and repression. HeLa cells were transfected with the reporter plasmid along with a plasmid expressing the GAL4 DBD alone or fused to hMI-ER1α (GAL4-α) or hMI-ER1β (GAL4-β). Transcriptional regulation was analyzed by determining the ability of such fusion proteins to alter the level of CAT expression. Overexpression of hMI-ER1α and β (Figure 3.1B) significantly repressed transcription of G5tkCAT in a dose-dependent manner ($P < 0.05$) (Figure 3.1A). Maximum repression resulted in a reduction in CAT expression to 10% and 8% of control levels by GAL4-α and GAL4-β, respectively. This repression was not unique to HeLa cells, as similar results were also obtained with C33A human cervical carcinoma cells, HEK 293 human transformed embryonic kidney cells, and NIH 3T3 mouse fibroblasts. As shown in Figure 3.2, maximum repression resulted in a reduction in CAT expression to 16% and 26% of control levels by GAL4-α and GAL4-β respectively in C33A cells; to 28% and 24% of control levels respectively in 293K cells; to 15% and 39% of control levels respectively in NIH 3T3 cells.

3.3.2 Transcriptional repression by hMI-ER1α and β involves recruitment of HDAC enzymatic activity

Recent studies have shown that transcriptional repression is often associated with recruitment of HDACs and/or chromatin-regulatory complexes containing HDAC

Figure 3.1 hMI-ER1α and β function as transcriptional repressors on the G5TKCAT promoter in HeLa cells.

(A) HeLa cells were transfected with the G5tkCAT reporter plasmid alone or with the pM plasmid containing an NLS and the GAL4 DBD (GAL4) alone or fused to *hmi-er1*α (GAL4-α) or *hmi-er1*β (GAL4-β); the amount of plasmid (in μg) used for transfection is indicated below each bar. Cells were harvested 48 h after transfection, and the level of CAT protein was determined, and this value was normalized to the level of cellular protein in each sample. For each sample, the value is presented as a proportion of the value obtained for control G5tkCAT alone (relative CAT expression). Shown are the average values and standard deviations from three independent experiments. **(B)** The level of GAL4-α or GAL4-β protein expressed in each sample was determined by Western blot analysis using an anti-GAL4 antibody. Shown is a representative Western blot, and the position of the GAL4-α or GAL4-β is indicated by arrows.

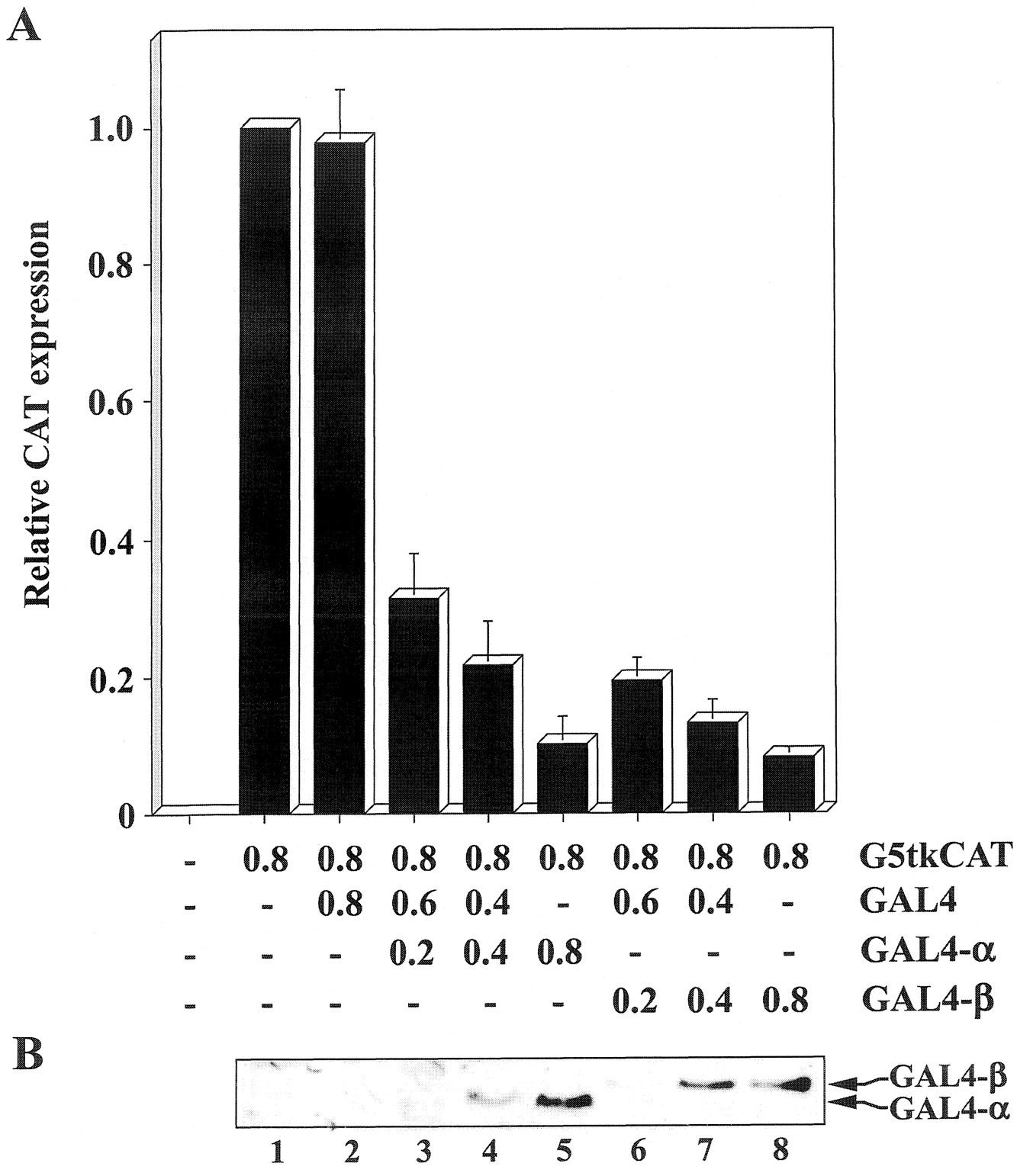
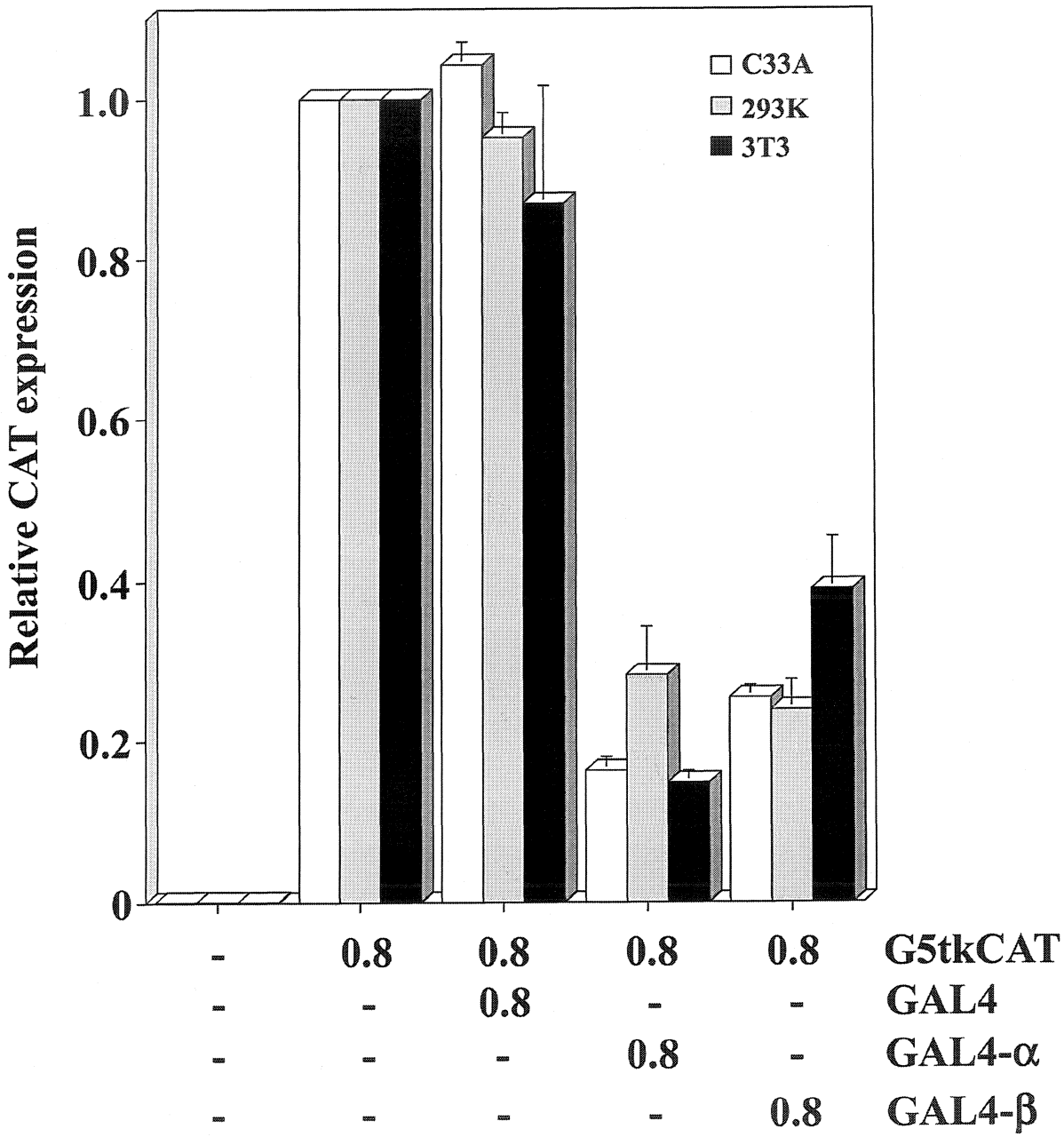

A
Relative CAT expression
1.0
0.8
0.6
0.4
0.2
0
- 0.8 0.8 0.8 0.8 0.8 0.8 0.8 0.8 G5tkCAT
- - 0.8 0.6 0.4 - 0.6 0.4 - GAL4
- - - 0.2 0.4 0.8 - - - GAL4-α
- - - - - - 0.2 0.4 0.8 GAL4-β
B
GAL4-β
GAL4-α
1 2 3 4 5 6 7 8

Figure 3.2 hMI-ER1α and β function as transcriptional repressors on the G5TKCAT promoter in C33A, HEK 293 and NIH 3T3 cells.

C33A, HEK 293 (293K), and NIH 3T3 cells were transfected with the G5tkCAT reporter plasmid alone or with the pM plasmid (GAL4) alone or fused to *hmi-er1*α (GAL4-α) or *hmi-er1*β (GAL4-β); the amount of plasmid (in μg) used for transfection is indicated below each bar. Cells were harvested 48 h after transfection, and the level of CAT protein was determined, and this value was normalized to the level of cellular protein in each sample. For each sample, CAT expression value is presented as a proportion of the value obtained for control G5tkCAT alone (relative CAT expression). Shown are the average values and standard deviations from three independent experiments.


Relative CAT expression
1.0
0.8
0.6
0.4
0.2
0
C33A
293K
3T3
- 0.8 0.8 0.8 0.8 G5tkCAT
- - 0.8 - - GAL4
- - - 0.8 - GAL4-α
- - - - 0.8 GAL4-β

enzymatic activity (reviewed in Ayer, 1999). To test whether hMI-ER1α and β repress transcription through such a mechanism, HeLa cells were treated with TSA, a specific inhibitor of class I and class II HDACs (Yoshida *et al.*, 1990), and the effect was examined on hMI-ER1α and β-mediated repression. Addition of TSA to the culture medium partially relieved repression of the tk promoter by hMI-ER1α and β (Figure 3.3), suggesting that transcriptional repression by hMI-ER1α and β involves recruitment of HDAC enzymatic activity.

To further investigate the nature of the TSA-sensitive repression, co-IP analysis was utilized to examine the ability of hMI-ER1α and β to physically associate with HDACs both *in vitro* and *in vivo*. ^{35}S-labelled Myc-tagged hMI-ER1α (Myc-α) and hMI-ER1β (Myc-β) were synthesized *in vitro* and mixed with unlabelled *in vitro*-translated HDAC1 and then subjected to immunoprecipitation with anti-HDAC1 antibody. Both hMI-ER1α and hMI-ER1β were detected in HDAC1 immunoprecipitates, while luciferase, a control protein, was not (Figure 3.4A). Reciprocal experiments using anti-Myc antibody 9E10 for immunoprecipitation confirmed that HDAC1 could associate with either hMI-ER1 isoform (Figure 3.4B). However, there were no interactions between either hMI-ER1 isoform with HDAC3 in the similar *in vitro* assay condition (Figure 3.4C).

In vivo analysis involved transient expression of the Myc-tag empty vector, Myc-α or Myc-β in HeLa cells. Expression of Myc-α and Myc-β was verified by Western blot analysis with anti-Myc antibody 9E10 (Figure 3.5A, lane 2 and 3). Cell extracts were

Figure 3.3 Transcription repression by hMI-ER1α and β occurs through a mechanism involving HDAC activity.

HeLa cells were co-transfected with 0.8 μg of the G5tkCAT reporter plasmid and 0.8 μg of the GAL4, GAL4-α, or GAL4-β plasmid and cultured in the presence or absence of TSA, a specific inhibitor of class I and class II HDACs. Cells were harvested 48 h after transfection, and the level of CAT protein was determined, and this value was normalized to the level of cellular protein in each sample. CAT expression values for GAL4-α- and GAL4-β-transfected cells are presented as a proportion of the value obtained with the GAL4 empty vector (relative CAT expression). Shown are the average values and standard deviations from three independent experiments.

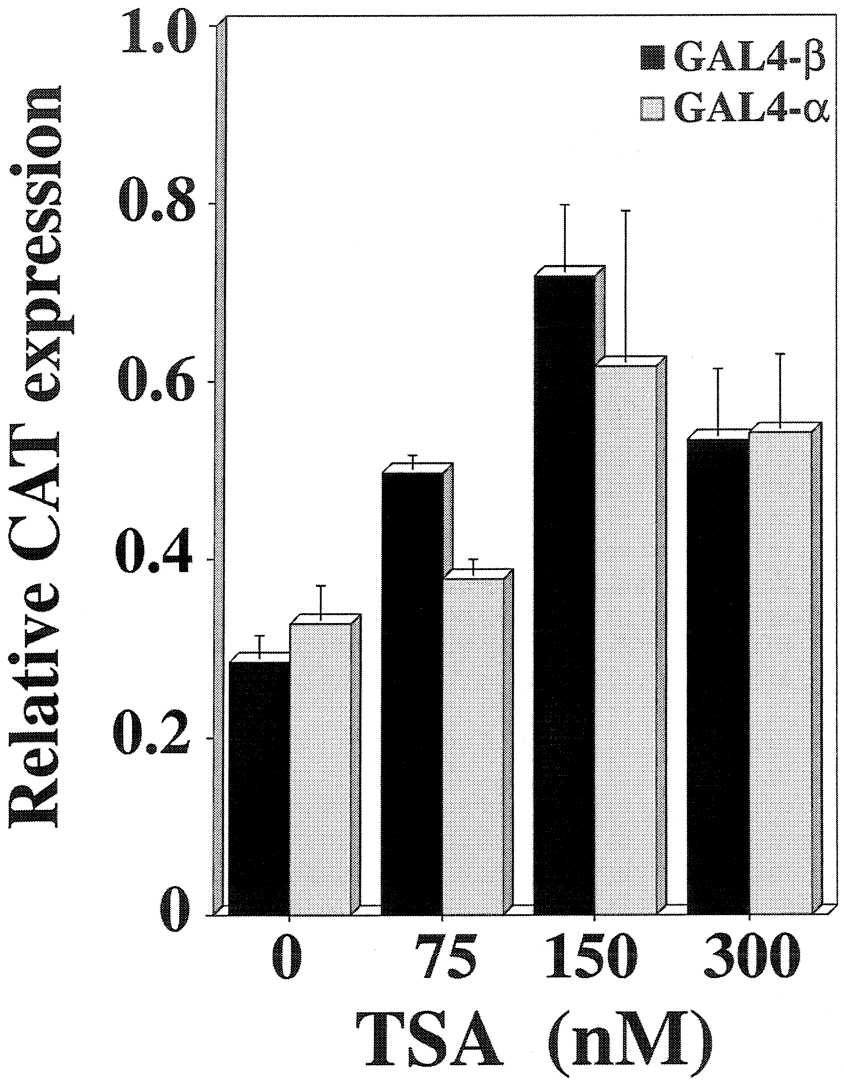

Relative CAT expression
1.0
0.8
0.6
0.4
0.2
0
GAL4-β
GAL4-α
0
75
150
300
TSA (nM)

Figure 3.4 hMI-ER1α and β physically associate with HDAC1 *in vitro*.

(A) ^{35}S-labelled TNT mixtures programmed with cDNA encoding luciferase (Luc), Myc tagged hMI-ER1α [Myc-α (α)], or Myc tagged hMI-ER1β [Myc-β (β)] were loaded directly on the gel (1/20 of input) (lanes 1 to 3) or incubated with unlabelled TNT mixtures programmed with *hdac1* cDNA and subjected to immunoprecipitation (IP) with anti-HDAC1 antibody (lanes 4 to 6). Proteins were resolved by SDS-PAGE and visualized by autoradiography. Shown is a representative autoradiograph, and the positions of the Myc-α, Myc-β, and luciferase are indicated by arrowheads. **(B)** ^{35}S-labelled TNT mixtures programmed with cDNA encoding luciferase or HDAC1 were loaded directly on the gel (1/20 of input) (lanes 1 and 2) or incubated with unlabelled TNT mixtures programmed with cDNA encoding Myc-α (lanes 3 and 4) or Myc-β (lanes 5 and 6) and subjected to IP with anti-Myc antibody 9E10; proteins were resolved by SDS-PAGE and visualized by autoradiography. Shown is a representative autoradiograph, and the positions of the HDAC1 and luciferase are indicated by bars. **(C)** ^{35}S-labelled TNT mixtures programmed with cDNA encoding luciferase or HDAC3 were loaded directly on the gel (1/20 of input) (lanes 1 and 2) or incubated with unlabelled TNT mixtures programmed with cDNA encoding Myc-α (lanes 3 and 4) or Myc-β (lanes 5 and 6) and subjected to IP with anti-Myc antibody 9E10; proteins were resolved by SDS-PAGE and visualized by autoradiography. Shown is a representative autoradiograph, and the positions of the HDAC3 and luciferase is indicated by arrows.

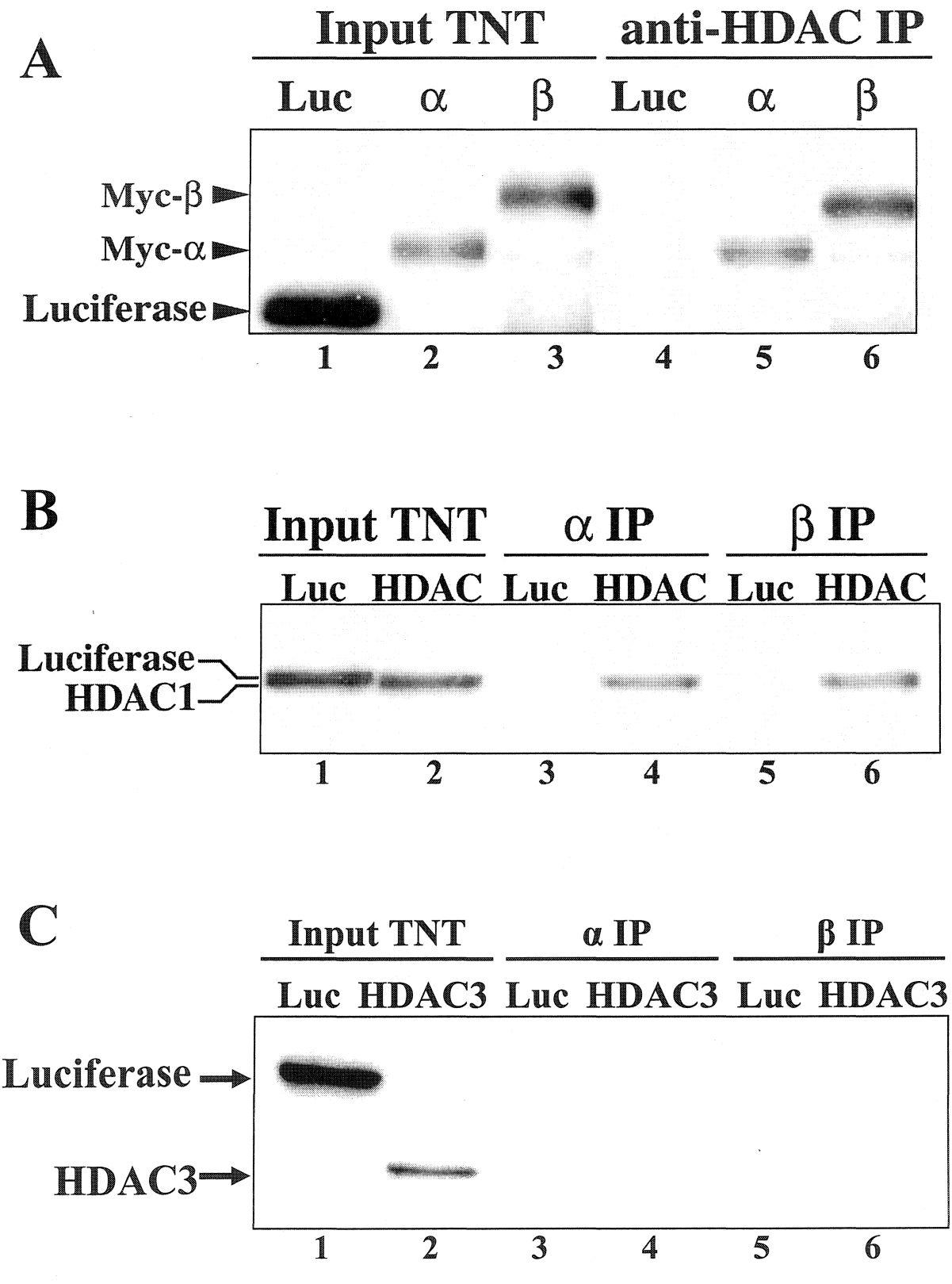
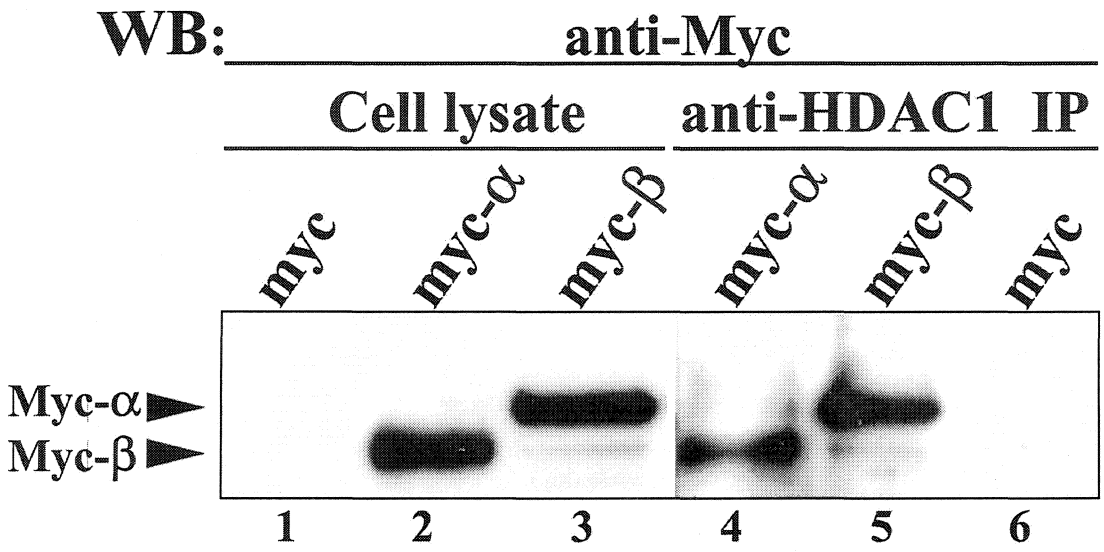

A
Input TNT
anti-HDAC IP
Luc
α
β
Luc
α
β
Myc-β
Myc-α
Luciferase
1
2
3
4
5
6
B
Input TNT
α IP
β IP
Luc HDAC Luc HDAC Luc HDAC
Luciferase
HDAC1
1
2
3
4
5
6
C
Input TNT
α IP
β IP
Luc HDAC3 Luc HDAC3 Luc HDAC3
Luciferase
HDAC3
1
2
3
4
5
6

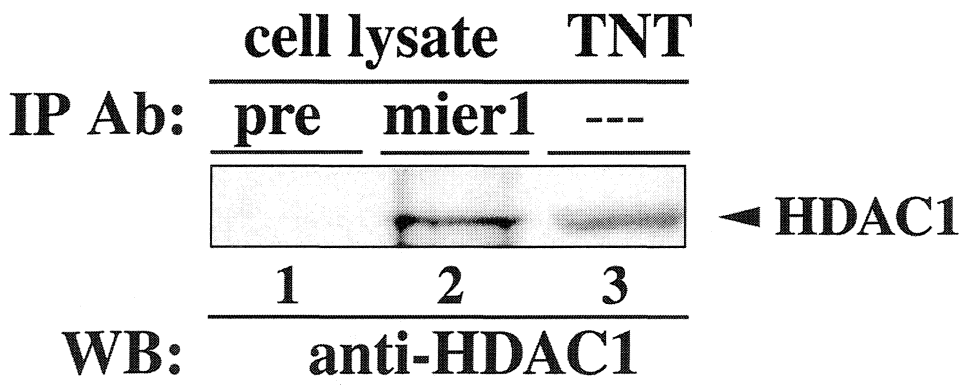
Figure 3.5 hMI-ER1α and β physically associate with HDAC1 *in vivo*.

(A) HeLa cell lysates from HeLa cells transiently transfected with Myc tag empty vector or Myc tagged hMI-ER1α (pMyc-α), or Myc tagged hMI-ER1β (pMyc-β) plasmid were prepared, and either added directly to sample buffer (1/3 of input) (lanes 1 to 3) or subjected to immunoprecipitation (IP) with anti-HDAC1 antibody (lanes 4 to 6); Western blot (WB) analysis was performed using anti-Myc antibody 9E10. Shown is a representative Western blot. The positions of the Myc-α and Myc-β proteins are indicated by arrowheads. **(B)** Untransfected HeLa cell lysates were subjected to IP with pre-immune serum (lane 1) or anti-hMI-ER1 antiserum (lane 2). HDAC1 protein from an *in vitro* TNT mixture was loaded in lane 3. Western blot (WB) analysis was performed using anti-HDAC1 antibody. Shown is a representative Western blot. The position of the HDAC1 protein is indicated by arrowheads.

A


WB:
anti-Myc
Cell lysate
anti-HDAC1 IP
myc
myc-α
myc-β
myc-α
myc-β
myc
Myc-α
Myc-β
1
2
3
4
5
6

B


cell lysate
TNT
IP Ab:
pre
mier1

HDAC1
1
2
3
WB:
anti-HDAC1

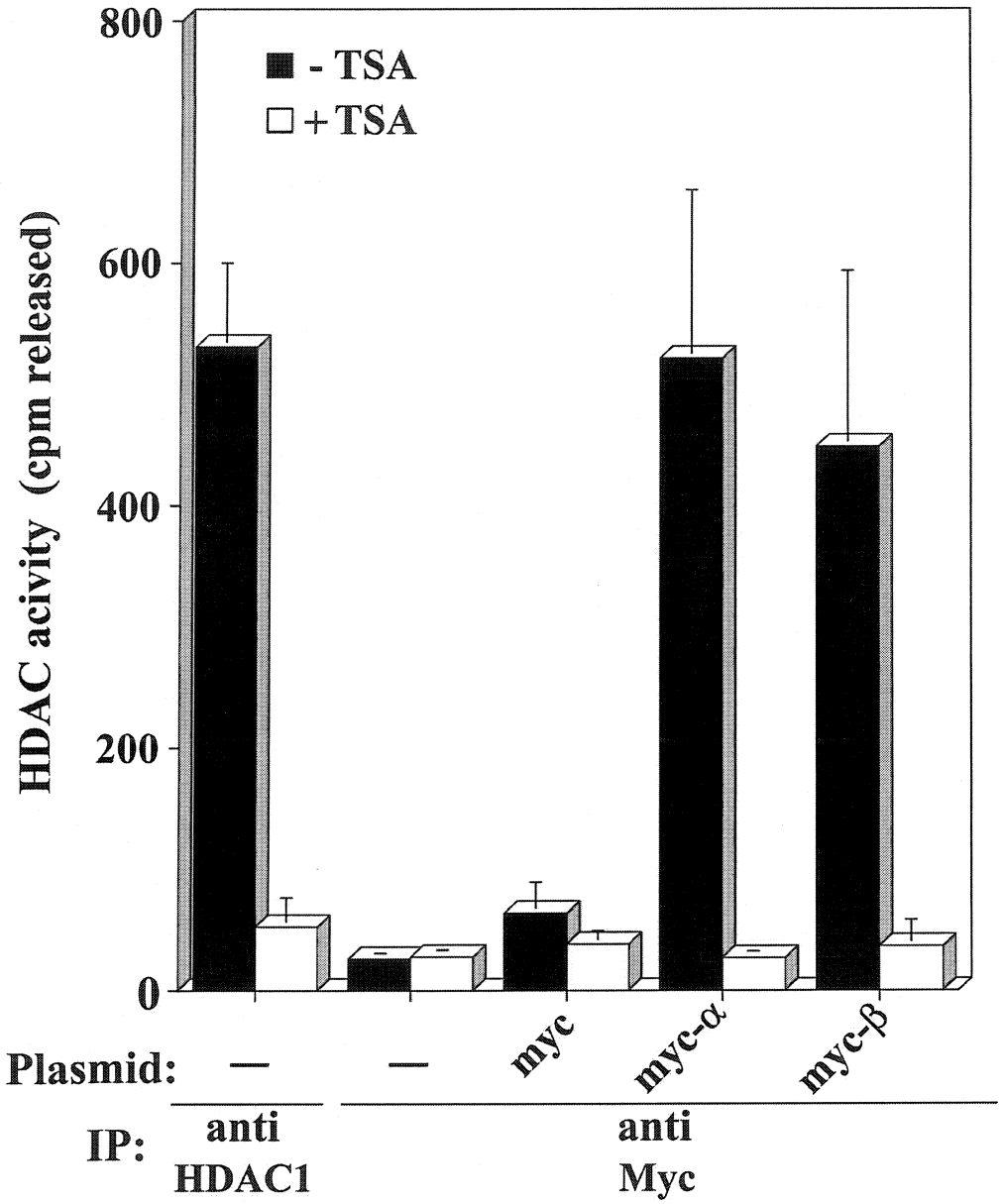
subjected to immunoprecipitation with anti-HDAC1 antibody, followed by Western blot analysis with anti-Myc antibody 9E10. Both hMI-ER1α and β isoforms were co-immunoprecipitated with endogenous HDAC1 (Figure 3.5A, lane 4 and 5).

The next set of experiments was designed to investigate whether endogenous complexes containing hMI-ER1 and HDAC1 exist in the cell and thus to exclude the possibility that the observed interaction was an artifact of overexpression. Co-immunoprecipitation analysis of extracts from nontransfected HeLa cells was performed using anti-hMI-ER1 antiserum or pre-immune serum (Figure 3.5B, lane 2). Endogenous HDAC1 protein was detected in hMI-ER1 immunoprecipitate but not in the pre-immune serum control (Figure 3.5B, lane 1), demonstrating that hMI-ER1 can physically associate with endogenous HDAC1 *in vivo*.

Next, HDAC enzymatic activity was analyzed from the hMI-ER1 immunoprecipitates. Cell extracts from HeLa cells transfected with the Myc tag empty vector, Myc-α, or Myc-β were incubated with anti-Myc antibody 9E10, and the immunoprecipitates were assayed for HDAC enzymatic activity. As shown in Figure 3.6, both Myc-α and Myc-β immunoprecipitates with anti-Myc antibody 9E10 contained significant levels of HDAC enzymatic activity and this activity was inhibited by TSA. On the other hand, immunoprecipitates with anti-Myc antibody 9E10 from mock-transfected cells or cells transfected with the Myc tag empty vector did not contain significant HDAC enzymatic activity (Figure 3.6). Positive control consisted of extracts

Figure 3.6 HDAC activities in hMI-ER1α and β immunoprecipitates from HeLa cells.

HeLa cells were transiently transfected with Myc tag empty vector or Myc tagged hMI-ER1α (pMyc-α), or Myc tagged hMI-ER1β (pMyc-β). Cell extracts were prepared 48 h after transfection and subjected to immunoprecipitation with anti-HDAC1 antibody or anti-Myc antibody 9E10. Immunoprecipitates were assayed for HDAC enzymatic activity in the presence or absence of 300 nM TSA. The histogram shows the average values and standard deviations from three independent experiments.


800
600
400
200
0
■ - TSA
□ + TSA
HDAC acivity (cpm released)
Plasmid:
—
—
myc
myc-α
myc-β
IP:
anti
HDAC1
anti
Myc

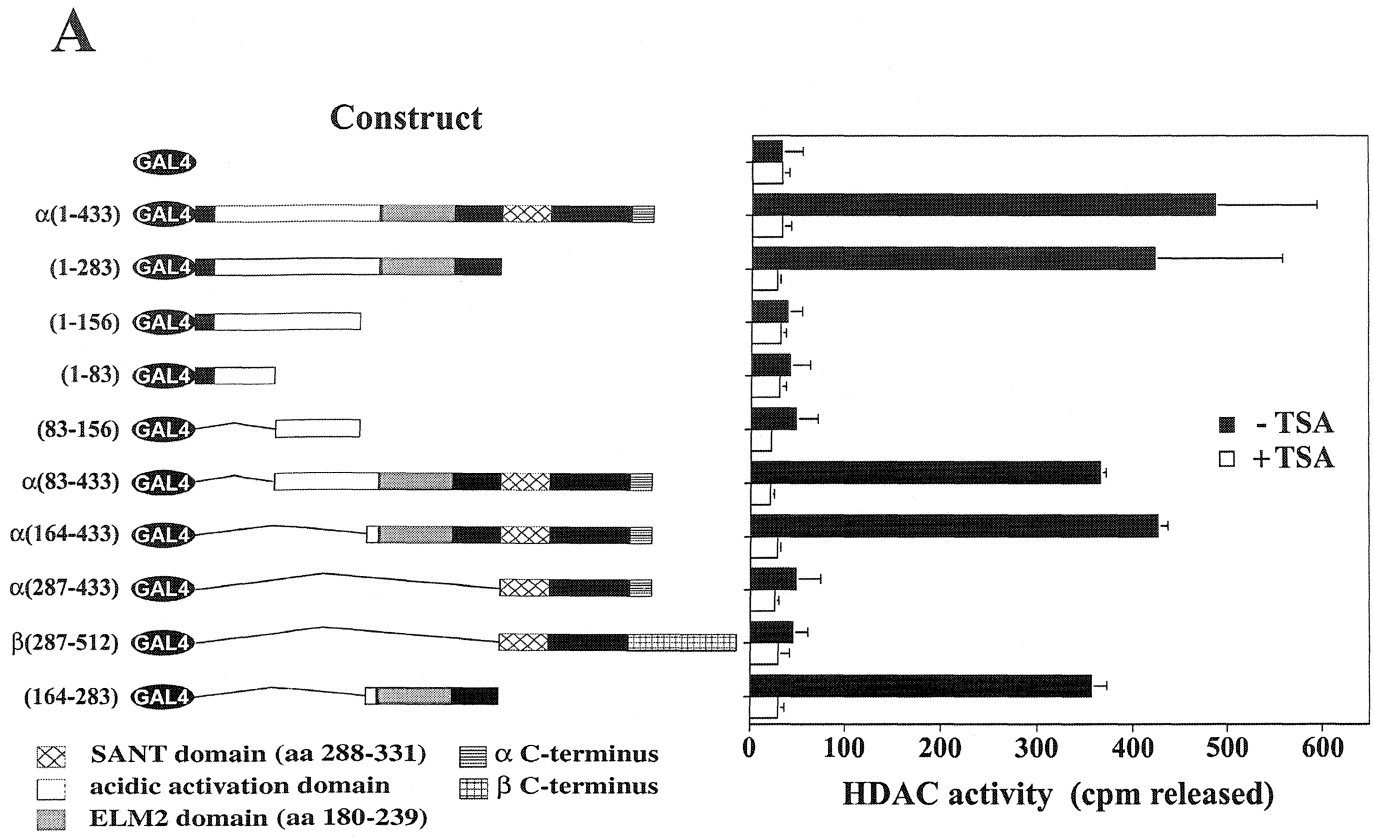
from mock-transfected cells immunoprecipitated with anti-HDAC1 antibody and contained significant levels of HDAC enzymatic activity (Figure 3.6). Taken together, these results demonstrate that hMI-ER1α and β isoforms are physically associated with a functional HDAC1 protein.

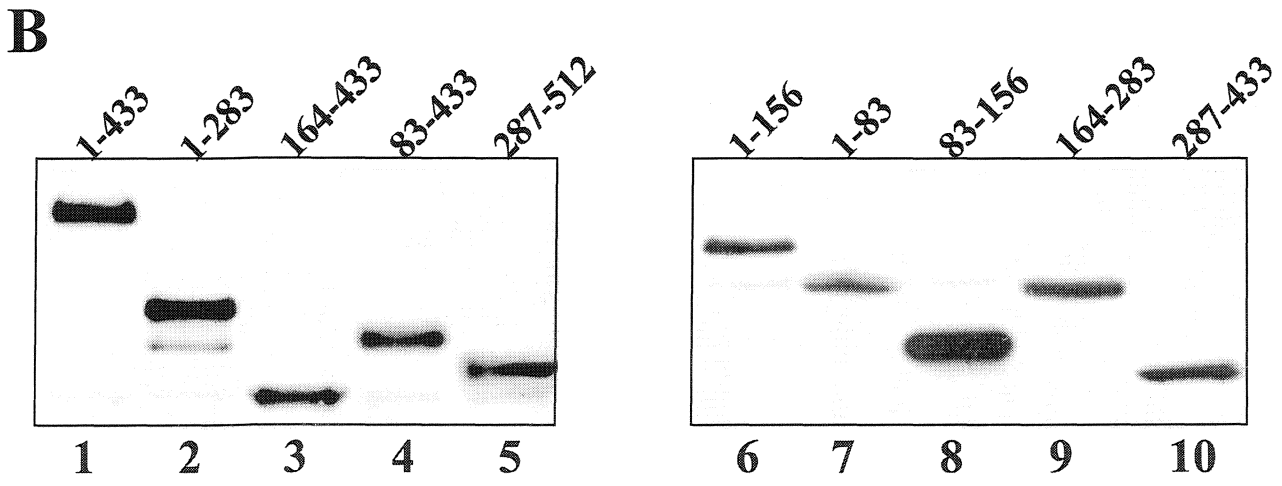
3.3.3 hMI-ER1α and β recruits HDAC enzymatic activity through a region containing the ELM2 domain

To determine which region of hMI-ER1 protein was responsible for recruitment of HDAC enzymatic activity, a series of GAL4 DBD-hMI-ER1α and GAL4 DBD-hMI-ER1β deletion mutants were constructed (Figure 3.7A, left panel) and transiently expressed in HeLa cells. Expression of all constructs was confirmed by Western blot analysis (Figure 3.7B), and immunoprecipitates with anti-GAL4 antibody were tested for HDAC enzymatic activity in the presence or absence of TSA (Figure 3.7A, right panel). This deletion analysis revealed that a single 120-aa region common to both hMI-ER1α and β was sufficient for recruitment of HDAC enzymatic activity. This region is located between aa 164 and 283. Sequence analysis of this region using PFAM at http://pfam.wustl.edu/hmmsearch.shtml revealed a conserved ELM2 domain located between aa 180 and 239 (Appendices 1).

Figure 3.7 The region of hMI-ER1 containing the ELM2 domain is associated with HDAC enzymatic activity.

(A) The diagram on the left illustrates the deletion mutants of hMI-ER1 fused to GAL4 DBD in pM vector. The individual domains are identified in the legend below the diagram, and the hMI-ER1 amino acid residues encoded by each construct are listed. HeLa cells were transfected with constructs indicated on the left. Cell extracts were prepared 48 h after transfection and subjected to immunoprecipitation with anti-GAL4. Immunoprecipitates were assayed for HDAC enzymatic activity in the presence or absence of 300 nM TSA. The histogram shown on the right is the average values and standard deviations from three independent experiments. **(B)** The expression of the GAL4-hMI-ER1 fusion protein in each sample used in panel A was examined by Western blot analysis using an anti-GAL4 antibody. Shown is a representative Western blot. Indicated above each lane (lane 1 to 10) are the hMI-ER1 residues encoded by the construct as described in panel A.


A
Construct
GAL4
α(1-433)
(1-283)
(1-156)
(1-83)
(83-156)
α(83-433)
α(164-433)
α(287-433)
β(287-512)
(164-283)
SANT domain (aa 288-331)
acidic activation domain
ELM2 domain (aa 180-239)
α C-terminus
β C-terminus
- TSA
+TSA
0
100
200
300
400
500
600
HDAC activity (cpm released)


B
1-433
1-283
164-433
83-433
287-512
1-156
1-83
83-156
164-283
287-433
1
2
3
4
5
6
7
8
9
10

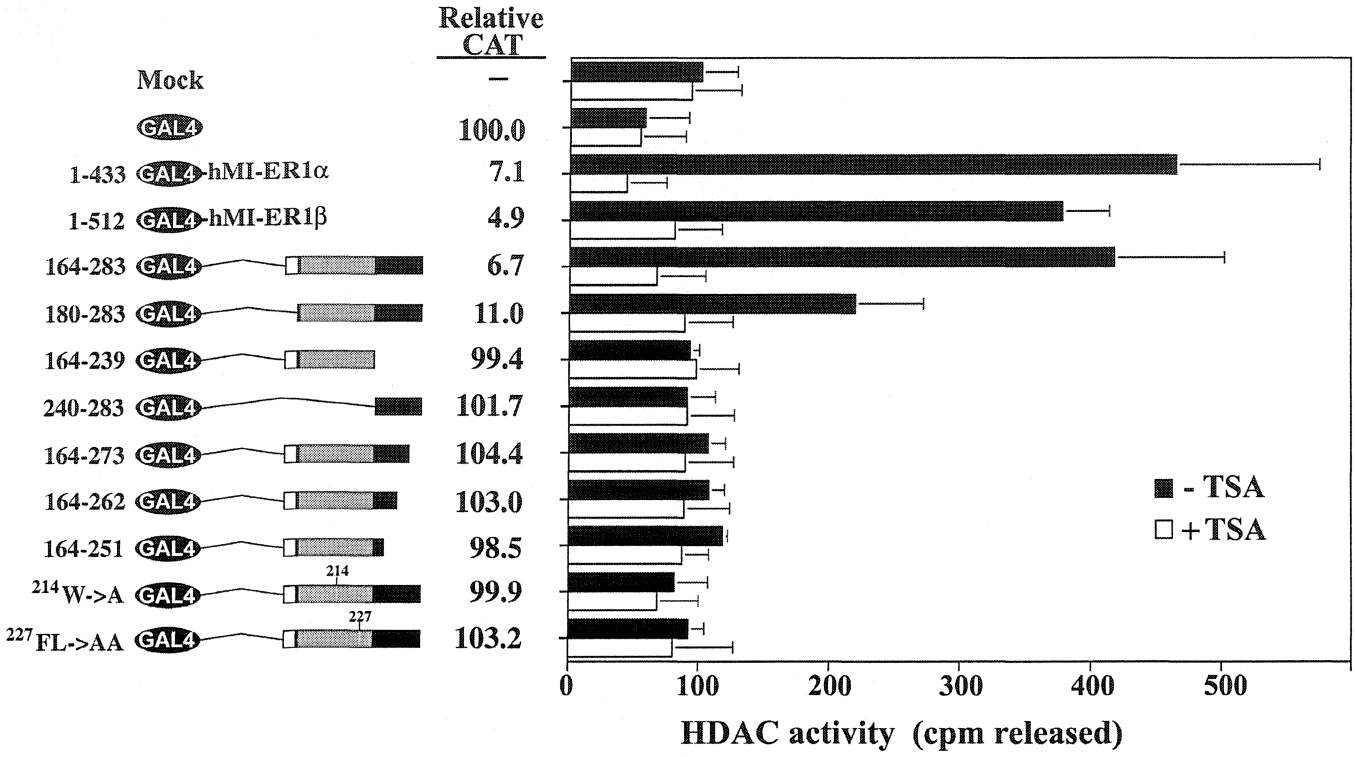
3.3.4 The ELM2 domain functions as a transcriptional repression domain through recruitment of HDAC1

The ELM2 domain-containing region is located between aa 164 and 283 and contains 16 aa N-terminal sequences and 44 aa C-terminal sequences to the ELM2 domain (aa 180-239). Further analysis of the region containing aa residues 164 to 283 was performed to determine the minimum sequence required for recruitment of HDAC1 and to determine whether this region was important for the transcriptional repression activity of hMI-ER1. Comparison of aa in the ELM2 domain by PFAM (Appendices 1) revealed that tryptophan (W) at position 214, and phenylalanine-leucine (FL) at positions 227 and 228 are highly conserved among the ELM2 domains. Therefore, a site-directed mutagenesis approach was employed to specifically examine the role of the ELM2 domain. For this purpose, ^{214}W and ^{227}FL were changed to alanine (A) to produce two mutant GAL4-hMI-ER1(164-283) constructs, ^{214}W→A and ^{227}FL→AA, respectively (Figure 3.8A, left panel). Moreover, a series of GAL4-hMI-ER1(164-283) deletion mutants were constructed (Figure 3.8A, left panel). All constructs were transfected into HeLa cells, and expression of the GAL4 fusion proteins was verified by Western blot analysis (Figure 3.8B). Extracts from these cells were assayed for CAT expression and HDAC enzymatic activity (Figure 3.8A, right panel). Analysis revealed that mutant proteins ^{214}W→A and ^{227}FL→AA, were unable to recruit HDAC enzymatic activity or to function in transcriptional repression (Figure 3.8A, right panel), confirming the role of the ELM2 domain in mediating these activities.

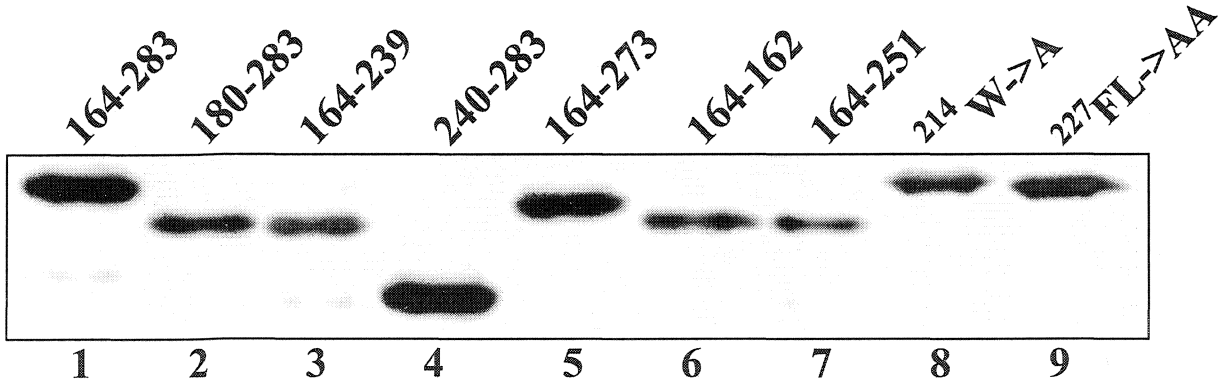
Figure 3.8 The ELM2 domain and an additional 45 aa C-terminal sequence of hMI-ER1α and β recruit HDAC enzymatic activity and repress transcription *in vivo*.

(A) The diagram on the left illustrates the deletion mutants of hMI-ER1 fused to GAL4 DBD in pM vector. The hMI-ER1 amino acid residues encoded by each construct are listed. ^{214}W→A and ^{227}FL→AA constructs contain residues 164 to 283 with alanine substitutions at ^{214}W and ^{227}FL, respectively. For HDAC enzymatic activity measurements, HeLa cells were transfected with constructs indicated on the left. Cell extracts were prepared 48 h after transfection and subjected to immunoprecipitation with anti-GAL4. Immunoprecipitates were assayed for HDAC enzymatic activity in the presence or absence of 300 nM TSA. The histogram shown on the right is the average values and standard deviations from three independent experiments. For transcriptional repression assays, HeLa cells were transfected with the G5tkCAT reporter plasmid alone or with constructs indicated on the left. Cells were harvested 48 h after transfection, and the level of CAT protein was determined, and this value was normalized to the level of cellular protein in each sample. CAT expression values for all constructs are presented in the middle as a proportion of the value obtained with the GAL4 empty vector (relative CAT). **(B)** The expression of the GAL4-hMI-ER1 fusion protein in each sample used in panel A was examined by Western blot analysis using an anti-GAL4 antibody. Indicated above each lane (lane 1 to 9) are the hMI-ER1 residues encoded by each construct.

A


Relative CAT
Mock
GAL4
1-433 GAL4-hMI-ER1α
1-512 GAL4-hMI-ER1β
164-283 GAL4
180-283 GAL4
164-239 GAL4
240-283 GAL4
164-273 GAL4
164-262 GAL4
164-251 GAL4
214W->A GAL4
227FL->AA GAL4
214
227
–
100.0
7.1
4.9
6.7
11.0
99.4
101.7
104.4
103.0
98.5
99.9
103.2
- TSA
+TSA
0
100
200
300
400
500
HDAC activity (cpm released)

B


164-283
180-283
164-239
240-283
164-273
164-162
164-251
214 W->A
227FL->AA
1
2
3
4
5
6
7
8
9

Moreover, deletion of the first 16 aa (aa 164 to 179) of the N-terminus [GAL4-hMI-ER1(180-283)] resulted in a reduction in the associated HDAC enzymatic activity to 53% of GAL4-hMI-ER1(164-283), but had no significant effect on transcription repression. Further deletion of 60 aa (aa 180-239) of N-terminal residues reduced HDAC enzymatic activity and transcriptional repression to control levels (Figure 3.8A, right panel). At the C-terminal end, all deletions, even the one as small as 10 aa from aa 274 to 283 [GAL4-hMI-ER1(164-273)], completely eliminated HDAC enzymatic activity and transcriptional repression. Thus, only GAL4-hMI-ER1(164-283) and GAL4-hMI-ER1(180-283) showed significant levels of associated HDAC enzymatic activity and transcriptional repression. Moreover, these were the only two constructs that were able to co-immunoprecipitate HDAC1 (Figure 3.9). These results demonstrate that the minimum sequence for recruitment of HDAC1 activity and repression of transcription is contained in aa 180 to 283.

While the ELM2 domain (aa 180 to 239) was shown to be required for HDAC1 binding and transcriptional repression, it was clear that an additional sequence C terminal to this domain was also essential. The ELM2 domain was originally defined on the basis of sequence conservation in a small number of available protein sequences (Solari *et al.*, 1999). A re-examination of the alignment, by using a larger number of ELM2-containing proteins, revealed conservation of an additional sequence C terminal to the defined ELM2 domain (Figure 3.10A). This highly conserved sequence encompasses aa 256 to 274 of hMI-ER1 and contains the consensus $ALXXLX_5DX_3ALXXL$, in which the second and last leucines, ^{260}L and ^{274}L, are invariant. Secondary-structure analysis of this

Figure 3.9 The ELM2 domain and an additional 45 aa C-terminal sequence of hMI-ER1α and β are essential for recruitment of HDAC1 *in vitro*.

^{35}S-labelled TNT mixture programmed with cDNA encoding HDAC1 was loaded directly on the gel (lanes 1) or incubated with unlabelled TNT mixtures programmed with Myc tagged hMI-ER1 mutants in pGBKT7 plasmid and subjected to immunoprecipitation (IP) with anti-Myc antibody 9E10. Proteins were resolved by SDS-PAGE and visualized by autoradiography. Shown is a representative autoradiograph, and the position of the HDAC1 protein is indicated by arrow.

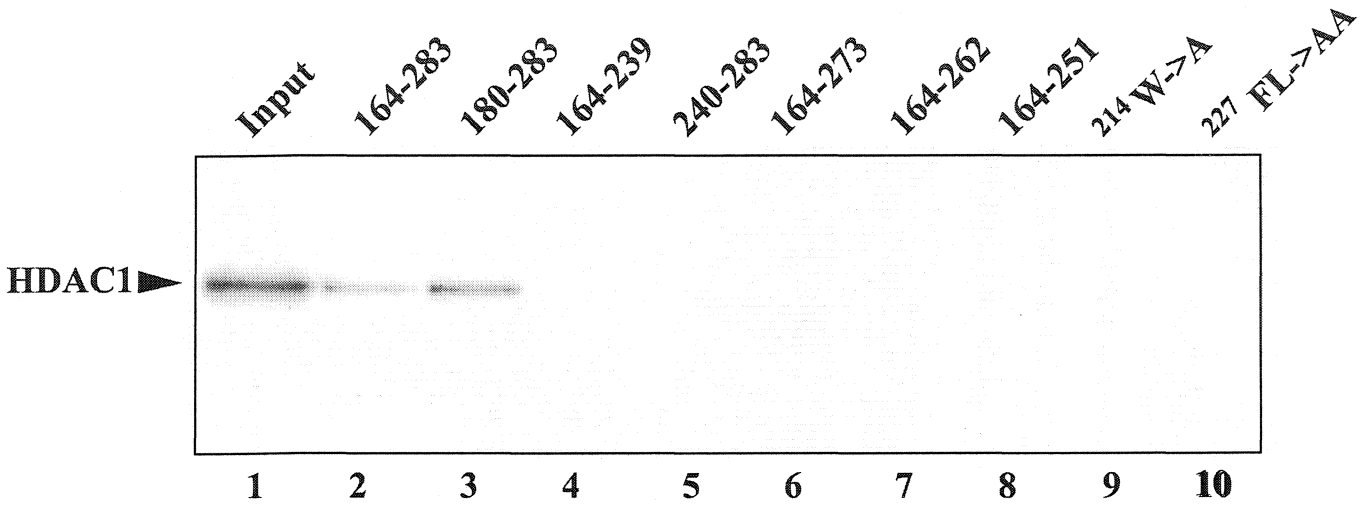
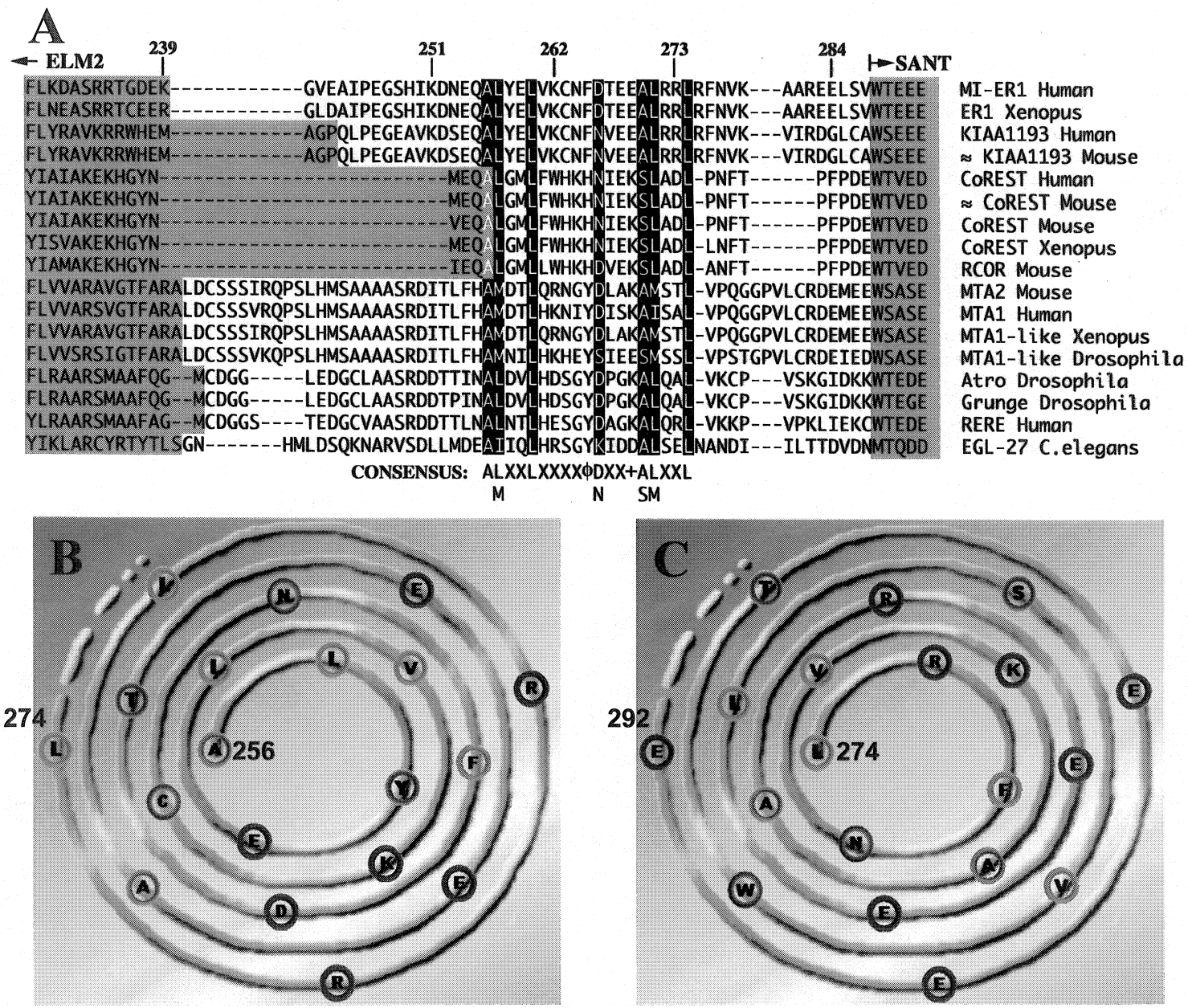

Input
164-283
180-283
164-239
240-283
164-273
164-262
164-251
214 W->A
227 FL->AA
HDAC1
1
2
3
4
5
6
7
8
9
10

Figure 3.10 Alignment of ELM2 domains reveals additional conserved sequence C-terminal to the ELM2 domain.

(A) The ELM2 regions of proteins from the PFAM and GenBank databases were aligned using ClustalW. Shown is the amino acid sequence from the C-terminal end of the ELM2 domain to the beginning of the SANT domain in each protein. Residues belonging to these two domains are shaded. Highly conserved residues in the region C terminal to the ELM2 domain are shown with white lettering and highlighted in black. The consensus sequence is listed below the alignment; X represents any amino acid; Φ represents Y, F, or H; and + represents a charged residue. The numbers listed above the alignment correspond to amino acid positions in the hMI-ER1 protein sequence. The accession numbers for the sequences used in this alignment (from top to bottom) are as follows: AF515447, O42194, AB033019, XM_125783, Q9UKL0, XM_127140, XM_127140, AJ311849, Q9JMK4, Q9R190, Q13330, O94776, Q9VNF6, Q9VNF6, Q9NHX6, Q9P2R6, and Q09228. **(B)** The amino acid sequence from the C-terminal end of the ELM2 domain to the beginning of the SANT domain in hMI-ER1 was analyzed using HelixDraw v1.00. Shown is the helical wheel of aa from ^{256}A to ^{274}E. **(C)** Shown is the helical wheel of aa from ^{274}L to ^{292}E, as described in panel B.


A
ELM2
SANT
239
251
262
273
284
FLKDASRRTGDEK----------GVEAIPEGSHIKDNEQALYELVKCNFDTEEALRRLRFNVK---AAREELSVWTEEE MI-ER1 Human
FLNEASRRTCEER----------GLDAIPEGSHIKDNEQALYELVKCNFDTEEALRRLRFNVK---AAREELSVWTEEE ER1 Xenopus
FLYRAVKRRWHEM----------AGPQLPEGEAVKDSEQALYELVKCNFNVEEALRRLRFNVK---VIRDGLCAWSEEE KIAA1193 Human
FLYRAVKRRWHEM----------AGPQLPEGEAVKDSEQALYELVKCNFNVEEALRRLRFNVK---VIRDGLCAWSEEE ≈ KIAA1193 Mouse
YIAIAKEKHGYN----------MEQALGMLFWHKHNIEKSLADL-PNFT------PFPDEWTVED CoREST Human
YIAIAKEKHGYN----------MEQALGMLFWHKHNIEKSLADL-PNFT------PFPDEWTVED ≈ CoREST Mouse
YIAIAKEKHGYN----------VEQALGMLFWHKHNIEKSLADL-PNFT------PFPDEWTVED CoREST Mouse
YISVAKEKHGYN----------MEQALGMLFWHKHNIEKSLADL-LNFT------PFPDEWTVED CoREST Xenopus
YIAMAKEKHGYN----------IEQALGMLWHKHDVEKSLADL-ANFT------PFPDEWTVED RCOR Mouse
FLVVARAVGTFARALDCSSSIRQPSLHMSAAAASRDITLFHAMDTLQRNGYDLAKAMSTL-VPQGGPVLCRDEMEEWSASE MTA2 Mouse
FLVVARSVGTFARALDCSSSVRQPSLHMSAAAASRDITLFHAMDTLHKNIYDISKAISAL-VPQGGPVLCRDEMEEWSASE MTA1 Human
FLVVARAVGTFARALDCSSSIRQPSLHMSAAAASRDITLFHAMDTLQRNGYDLAKAMSTL-VPQGGPVLCRDEMEEWSASE MTA1-like Xenopus
FLVVSRSIGTFARALDCSSSVKQPSLHMSAAAASRDITLFHAMNILHKHEYSIEESMSSL-VPSTGPVLCRDEIEDWSASE MTA1-like Drosophila
FLRAARSMAAFQG--MCDGG-----LEDGCLAASRDDTTINALDVLHDSGYDPGKALQAL-VKCP---VSKGIDKKWTEDE Atro Drosophila
FLRAARSMAAFQG--MCDGG-----LEDGCLAASRDDTPINALDVLHDSGYDPGKALQAL-VKCP---VSKGIDKKWTEGE Grunge Drosophila
YLRAARSMAAFAG--MCDGGS----TEDGCVAASRDDTTLNALNTLHESGYDAGKALQRL-VKKP---VPKLIEKCWTEDE RERE Human
YIKLARCYRTYTLSGN-------HMLDSQKNARVSDLLMDEAIQLHRSGYKIDDALSELNANDI---ILTTDVDNMTQDD EGL-27 C.elegans
CONSENSUS: ALXXLXXXXΦDXX+ALXXL
M N SM
B
274
256
C
292
274

region with the SOPMA at http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html revealed that this region has a propensity to form α-helical structures. This region was further analyzed by HelixDraw v1.00 at http://bioinf.man.ac.uk/~gibson/HelixDraw/helixdraw.html. The charged amino residues are clustered on one side, while hydrophobic leucine residues are clustered on the opposite face (Figure 3.10B and C). This sequence is predicted to form an alpha-helical coiled-coil by NPS@:COILED-COILS PREDICTION at http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_lupas.html which may be important for the recruitment of HDAC1 activity and transcriptional repression.

The results presented here illustrate that the previously delineated ELM2 domain is not sufficient for transcriptional repression and recruitment of HDAC enzymatic activity, but requires an additional sequence further downstream. This additional sequence, which is required for recruitment of HDAC1 activity and transcription repression, was found to be highly conserved among ELM2-containing proteins, indicating that the functional ELM2 domain extends further downstream than previously described and includes aa 180 to 283 in hMI-ER1.

3.4 Discussion

In this study, the transcriptional regulatory function of the hMI-ER1α and β isoforms was investigated. Both hMI-ER1α and β can function equally well as transcriptional repressors through recruitment of HDAC1 and the evolutionarily

conserved ELM2 domain (aa 180 to 283), which is common to both, is required for these activities. Thus, the two alternate C-terminal domains appear not to be involved in HDAC1 activity recruitment and repression.

Interestingly, the hMI-ER1α isoform was found to shuttle between the nucleus and cytoplasm in NIH 3T3 cells (Paterno *et al.*, 2002; Paterno *et al.*, unpublished data). Thus, controlling the subcellular localization of hMI-ER1 by alternate splicing provides a mechanism for regulation of its nuclear activities. It is also possible that the hMI-ER1α isoform may be transported to the nucleus through regulated interactions with other nuclear protein(s), such as HDAC1. Co-transport to the nucleus through such a piggy back mechanism has been reported for a number of proteins, including RB and HSP90 (Kang *et al.*, 1994; Zacksenhaus *et al.*, 1999). The C terminus of the hMI-ER1α isoform possesses a class III LXXLL motif (NR box), a nuclear hormone receptor interaction domain found in a number of transcriptional coactivators and corepressors (Aranda and Pascual, 2001; Chang *et al.*, 1999; Heery *et al.*, 1997; Yu, 2002). Thus the C-terminus of the hMI-ER1α may be important in regulating the differential subcellular localizations of hMI-ER1α in steroid hormone responsive tissues. In support of this speculation, a variant of MTA1, MTA1s, was recently found to contain the LXXLL motif and localize in the cytoplasm. MTA1s inhibits nuclear signalling by sequestering ER in the cytoplasm, which is distinct from effects of the nuclear form of MTA1 (Kumar *et al.*, 2002).

To date, 18 HDACs have been identified in humans, and their activities have been implicated in multiple cellular functions (reviewed in Thiagalingam *et al.*, 2003). Three classes of HDACs have been described based on homology to *Saccharomyces cerevisiae*

RPD3, HDA1 and SIR2 proteins (reviewed in Thiagalingam *et al.*, 2003). It is believed that the different HDACs may have unique temporal and spatial expression patterns that contribute to tissue-specific regulation of chromatin- and transcription-regulatory complexes. It has been shown that both hMI-ER1α and β interact with HDAC1 and this is a native interaction (Figure 3.5), not just an artifact of overexpression by transfection. However, it is possible that hMI-ER1 recruits HDAC1 activity either directly or indirectly through other proteins in the HDAC1-containing complexes. Since there are multiple endogenous proteins in cell lysates and in the reticulocyte TNT lysates, I cannot, at present, exclude this possibility. Alternative protein-protein interaction assays using both purified proteins and then Western blot analysis may address this question.

There were no interactions between either hMI-ER1 isoform with HDAC3 *in vitro* assay (Figure 3.4C), suggesting some specificity in the interaction of hMI-ER1 with HDAC family members. Specific association with different types of HDACs has been reported for other transcription regulators. For example, NCOR and SMRT specifically associate with HDAC3 (Guenther *et al.*, 2001; Yu *et al.*, 2003), while RB interacts with both HDAC1/2 and HDAC3 (Ferreira *et al.*, 1998; Lai *et al.*, 1999). Investigations are currently ongoing in the laboratory to determine whether other HDAC family members can interact specifically with hMI-ER1 to regulate transcription in the same and/or distinct cell types.

A region from aa 180 to 283, which is common to all hMI-ER1 isoforms and containing the ELM2 domain, was found to be responsible for recruiting HDAC1 activity (Figure 3.7). Other ELM2-containing proteins, including CoREST, MTA1, MTA2, and

MTA3 are present in HDAC-containing complexes (Fujita *et al.*, 2003). While all of these are known components of characterized chromatin- and transcription-regulatory complexes containing one or more HDACs, most of the interaction domain(s) for recruitment of HDACs has not been determined. You *et al.* concluded that the SANT domain of the corepressor CoREST was required for association with HDAC1 (You *et al.*, 2001); however, the construct used in that study also contained the ELM2 domain. The deletion mutant constructs of hMI-ER1 used in the present study were designed to separate the two domains and clearly showed that the SANT domain of hMI-ER1 does not recruit HDAC1 activity (Figure 3.7). It is possible that different molecules may utilize these domains differently and/or may cooperate in binding HDAC-containing complexes. Indeed, the SANT domain of the related corepressor SMRT, which does not contain an ELM2 domain, has been implicated in HDAC3 binding (Guenther *et al.*, 2001).

The ELM2 containing region (aa 164-283) containing point mutations of the highly conserved ^{214}W or ^{227}FL residues on the ELM2 domain were unable to recruit HDAC1 enzymatic activity or to function in transcriptional repression (Figure 3.8). These results suggested that the conserved aa in the ELM2 domain may play a key role in directing the recruitment of HDAC1 activity. Secondary-structure analysis of this region with SOPMA revealed that the ELM2 domain has a propensity to form α-helical structures. Although the substitution of ^{214}W or ^{227}FL residues on the ELM2 domain with alanine did not result in a alteration in the propensity to form α-helical structures, it is possible that ^{214}W or ^{227}FL play a key role in forming an interacting surface for HDAC1

or other protein(s) to bridge HDAC1. These mutations may result in alteration of the structural conformation of the protein. These results indicated that the structural integrity of the ELM2 containing region (aa 164-283) is critical for recruitment of HDAC1 activity and transcription repression.

The deletion mutational analysis of the region from aa 180 to 283 revealed that an intact ELM2 domain is required for these activities but that additional sequences, not previously included in the ELM2 domain as revealed by sequence comparisons (Solari *et al.*, 1999) or PFAM alignments, are also critical for recruiting HDAC1 activity and transcriptional repression activities (Figure 3.8). A more detailed analysis in the present studies of the sequences C-terminal to the ELM2 domain, in hMI-ER1 and other ELM2 domain-containing proteins, revealed conservation of additional sequences, including the consensus sequence $ALXXLX_5DX_3ALXXL$. Secondary-structure analysis of this region with SOPMA revealed that this region has a propensity to form α-helical structures. Analysis of this region by HelixDraw v1.00 revealed that the charged amino residues are clustered on one side, while hydrophobic residues such as leucine are clustered on the opposite face (Figure 3.10B and C). Secondary-structure analysis of this region with the COILS algorithms at http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_lupas.html revealed that residues from 263 to 290 has a high probability of forming a coiled-coil. Coiled-coil structures are involved in protein-protein interaction (reviewed in Lupas, 1996). For example, the coiled-coil structure in PML is responsible for homodimerization (Kastner *et al.*, 1992) and for the specific interaction with Nur77 (Wu *et al.*, 2002) and the coiled-coil structure of c-SKI was shown to recruit mSin3A-HDAC

complexes (Nomura *et al.*, 1999). Thus, the structural integrity of this ELM2 domain containing region of hMI-ER1 may be critical for the recruitment of HDAC1 and transcriptional repression.

Although the ELM2 containing region (aa 164-283) recruited HDAC1 activity and was involved in transcription repression by hMI-ER1, treatment of TSA can only partially relieve the transcription repression by hMI-ER1 (Figure 3.3). These results suggested that transcriptional repression by hMI-ER1α and β not only may be due to the recruitment of HDAC enzymatic activity, but that other mechanisms independent of HDAC may also be involved (see Chapter 4).

CHAPTER 4 THE SANT DOMAIN CONTAINING REGION OF hMI-ER1 INTERACTS WITH SP1 TO INTERFERE WITH GC BOX RECOGNITION AND REPRESS TRANSCRIPTION FROM ITS OWN P2 PROMTER

4.1 Introduction

hMI-ER1 has conserved domains found in a number of transcriptional regulators, including an acid activation domain (Paterno *et al.*, 1997), an ELM2 domain (Solari *et al.*, 1999) and a signature SANT domain (Aasland *et al.*, 1996). In a previous study, I found that the ELM2 domain functions in the recruitment of HDAC1 and transcriptional repression (Chapter 3 and Ding *et al.*, 2003). The SANT domain is located immediately downstream of the ELM2 domain and, in other proteins, has been implicated in protein-protein interactions (Aasland *et al.*, 1996), including interactions with HDAC1 and HDAC3 (Guenther *et al.*, 2001; You *et al.*, 2001) and HAT (Boyer *et al.*, 2002; Sterner *et al.*, 2002) containing complexes. To date, no function has been ascribed to the SANT domain of hMI-ER1.

Studies in our laboratory revealed that overexpression of exogenous hMI-ER1α and β repress the transcription of endogeneous hMI-ER1 (Paterno *et al.*, unpublished data). However, the molecular mechanisms of this transcriptional auto-regulation remain

---

This chapter has been published in: The SANT domain of human MI-ER1 interacts with Sp1 to interfere with GC box recognition and repress transcription from its own promoter. J Biol Chem. 2004 Apr 26 [Epub ahead of print].

unknown. I hypothesized that hMI-ER1α and β may function as transcriptional repressors on their own promoter, thus forming a negative feedback loop to tightly control the activity of hMI-ER1. This would ensure that hMI-ER1 regulates the expression of other genes at appropriate times.

In a previous study of the molecular mechanism for the transcriptional regulation of *hmi-er1*, the P1 promoter (-1781) and the P2 promoter (-1316) of *hmi-er1* have been cloned (Figure 2.1, Chapter 2). This chapter describes experiments to investigate whether hMI-ER1α and β repress *hmi-er1* P2 promoter activity. The molecular mechanism of the auto-repression by hMI-ER1α and β from the *hmi-er1* P2 promoter was further analyzed.

4.2 Materials and Methods

4.2.1 Cell culture

The human cervical carcinoma HeLa cell line was obtained from the American Tissue Culture Collection and cultured at 37°C in 5% CO_2 in DMEM containing 10% FCS.

4.2.2 Plasmids and constructs

The pGL3(-1316) and pGL3(-68) plasmids have been previously described (section 2.2.3). Myc-tagged plasmids (CS3+MT) containing either full-length *hmi-er1α* (pMyc-α) or β (pMyc-β) have been previously described in section 3.2.2.

To obtain GST-hmi-er1α and β fusion constructs, cDNA representing the appropriate isoform was subcloned in-frame into the pGEX-4T-1 vector (Pharmacia, Biotech). A series of hMI-ER1 deletion mutations was generated by amplifying fragments encoding the appropriate amino acid residues of hMI-ER1α or β, using the primer pairs listed in Table 4.1. PCR products were cloned into pCR3.1 and *EcoR*I fragments were then inserted into the *EcoR*I sites of the pGEX-4T-1 plasmid. Deletion constructs were named according to the encoded amino acid residues of the hMI-ER1α and β proteins. The GST-Sp1 fusion was constructed by subcloning the *EcoR*I fragments from Sp1-pCR3.1 into the *EcoR*I sites of pGEX-4T-2. All plasmids were sequenced to verify the junctions and the hMI-ER1 or Sp1 sequence.

4.2.3 Transfection and reporter assays

All transfection reporter assays were performed as described in section 2.2.5.

4.2.4 GST-fusion protein production

GST fusion proteins were expressed as described in section 2.2.6.

4.2.5 EMSAs

EMSAs were performed as described in section 2.2.7.

Table 3 Table 4.1 PCR primer pairs used for constructing hMI-ER1 and mutating hMI-ER1 plasmids in GST-4T1-1 vector

Construct[a]	Forward primer	Reverse primer
hmi-er1α aa 1-433	5 -CGGGATCCATATGGCGG AGCCATCTGTTG-3	5 -CGGGATCCAAAACAAGAC CACAGAAGC-3
hmi-er1β aa 1-512	5 -CGGGATCCATATGGCGG AGCCATCTGTTG-3	5 -CGGGATCCTTAGTCATCT GTGTTTTCAAG-3
aa 1-283	5 -CACCATGGCGACATCTG TTGAATC-3	5 -ATCCTCTCTAGCTGCTTTT ACA-3
aa 287-433	5 -CACCATGGTTTGGACAG AGGAAGAGTGTA-3	5 -CGGGATCCAAAACAAGAC CACAGAAGC-3
aa 325-433	5 -CACCATGGCATTCTATTA CATGTGGAAAAAATCT-3	5 -CGGGATCCAAAACAAGAC CACAGAAGC-3
aa 287-410	5 -CACCATGGTTTGGACAG AGGAAGAGTGTA-3	5 -CTGGTCCATTAGATGACA CTCCA-3
aa 287-357	5 -CACCATGGTTTGGACAG AGGAAGAGTGTA-3	5 -CCGTTACACCAGGATGAA GATT-3
aa 287-330	5 -CACCATGGTTTGGACAG AGGAAGAGTGTA-3	5 -CCACATGTAATAGAATGC TACA-3
aa 287-512	5 -CACCATGGTTTGGACAG AGGAAGAGTGTA-3	5 -CGGGATCCTTAGTCATCT GTGTTTTCAAG-3

[a] Deletion constructs were named according to the encoded amino acid residues of the hMI-ER1α or β protein

4.2.6 GST pull-down assays

GST pull-down assays were performed as previously described (Routledge *et al.*, 2000). Briefly, 1 μg of GST fusion protein was incubated with 50 μl of GST-Sepharose beads for 1 h at 4°C in 500 μl binding buffer (10 mM HEPES pH 7.2, 140 mM NaCl, 5 mM $MgCl_2$, 1 mM EDTA, 0.5% BSA and 0.25 % Nonidet P-40). Unbound proteins were removed by washing with 1000 μl of binding buffer three times. 50,000 cpm of [^{35}S]-labelled TNT product was incubated with the GST fusion protein in 500 μl binding buffer, for 3 h at 4°C. The beads were washed with 1000 μl of binding buffer three times, 1000 μl modified binding buffer (without Nonidet P-40) two times, and 1000 μl 150 mM NaCl two times, then boiled 3 min in 40 μl SDS sample buffer. The supernatants were subjected to SDS-PAGE followed by autoradiography. For all assays, 1/20 of the volume of the indicated TNT was loaded into the input lanes.

4.2.7 Co-IP and Western blot analysis

In vitro co-IP assays were performed as described in section 3.2.4 with either 10 μl anti-Sp1 polyclonal antiserum (Catalog No. PEP 2 X: sc-59, Santa Cruz Biotechnology, Inc.), 10 μl anti-Myc monoclonal antibody 9E10.

In vivo co-IP assays were performed as described in section 3.2.4 with 1 μl anti-hMI-ER1 antibody produced in our laboratory (24), or 10 μl anti-Sp1 polyclonal antiserum (Catalog No. PEP 2 X: sc-59, Santa Cruz Biotechnology, Inc.). Western blot analysis was performed as described in section 3.2.4 with anti-Myc monoclonal antibody 9E10 (1:100 dilution).

4.3 Results

4.3.1 hMI-ER1α and β repress *hmi-er1* P2 promoter activity through a HDAC-independent mechanism

Studies in our laboratory revealed that overexpression of exogenous hMI-ER1α and β represses the transcription of endogeneous hMI-ER1 (Paterno *et al.*, unpublished data). However, the molecular mechanisms of this transcriptional auto-regulation remain unknown. Since hMI-ER1α and β were shown to interact with HDAC1 and mediate transcriptional repression through recruitment of HDAC enzymatic activity, experiments were designed to investigate whether hMI-ER1α and β are involved in directing HDAC enzymatic activity to *hmi-er1* P2 promoter (-1316). HeLa cells were co-transfected with the *hmi-er1* P2 promoter-driven luciferase reporter construct pGL3(-1316) with Myc tag empty vector alone or fused to hMI-ERα (pMyc-α) or hMI-ER1β (pMyc-β). Transcriptional regulation was analyzed by determining the ability of such fusion proteins to alter the level of luciferase activity, compared to the Myc tag empty vector. Overexpression of both hMI-ER1α and β isoforms (Figure 4.1B) significantly repressed luciferase activity of pGL3(-1316), in a dose-dependent manner (Figure 4.1A). Maximum repression resulted in a reduction in luciferase activity to 36% and 29% of control levels by Myc-α and Myc-β, respectively. Likewise, overexpression of Myc-α and Myc-β (Figure 4.2B) repressed the activity of pGL3(-68) to 40% and 33% of control levels, respectively (Figure 4.2A), while overexpression of hMI-ER1α and β (Figure 4.2D) had no significant affect on pGL3-Basic [pGL3(-)] (Figure 4.2C). These results

Figure 4.1 hMI-ER1α and β repress *hmi-er1* P2 promoter activity.

HeLa cells were co-transfected with the *hmi-er1* P2 promoter-driven luciferase reporter construct pGL3(-1316) and with the Myc-tag empty vector, pMyc-α or pMyc-β. Cells were harvested 48 h after transfection and the level of relative luciferase units (RLU) was determined. For each sample, the value is presented as a proportion of the value obtained for Myc tag empty vector transfected cells (relative luciferase activity). **(A)** Shown are the average values and standard deviations from three independent experiments. **(B)** The level of Myc-α or Myc-β protein expressed in each sample as described above in the panel A was determined by Western blot, using anti-Myc antibody 9E10. Shown is a representative Western blot, and positions of the Myc-α and Myc-β proteins are indicated by arrowheads.

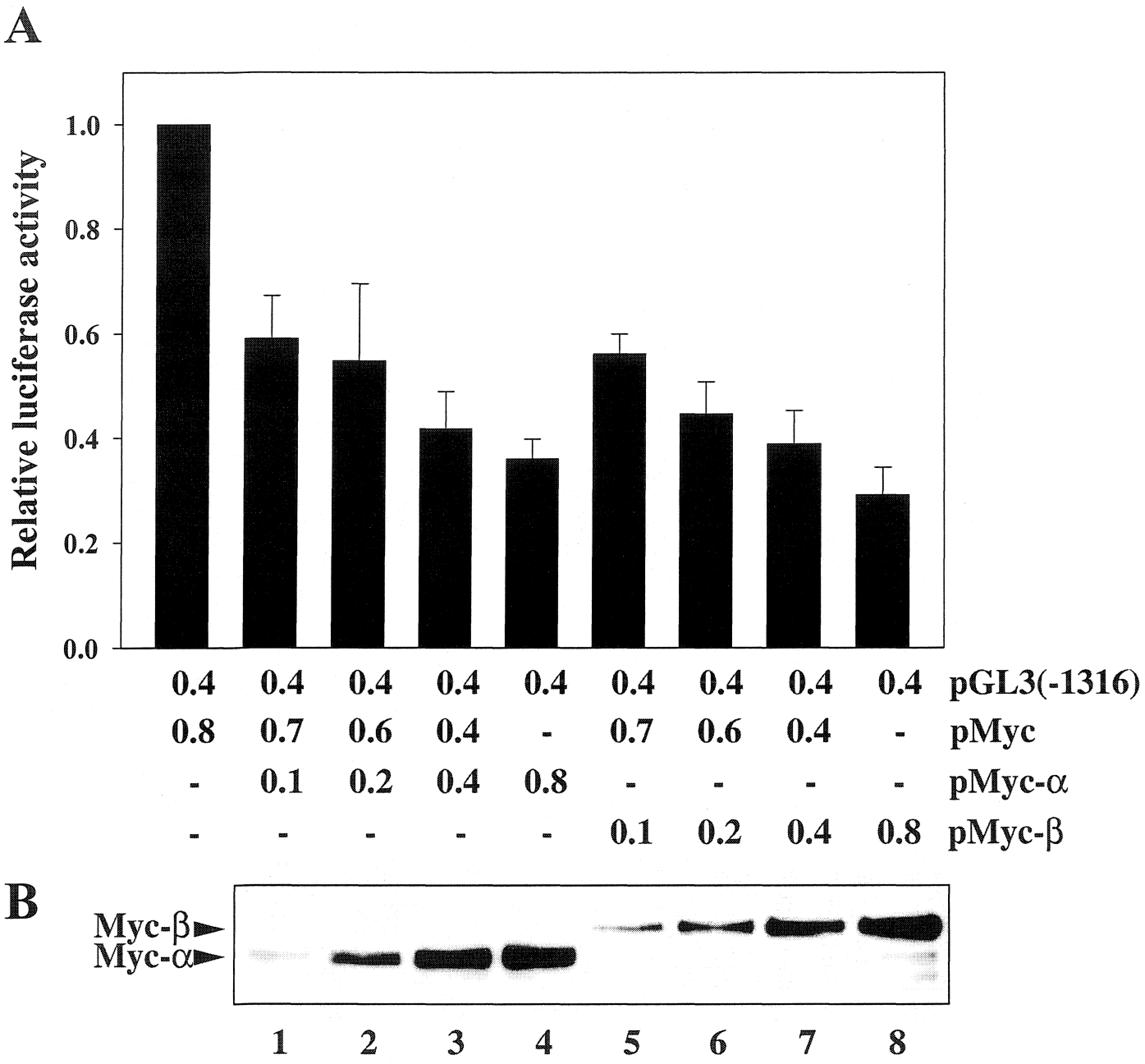
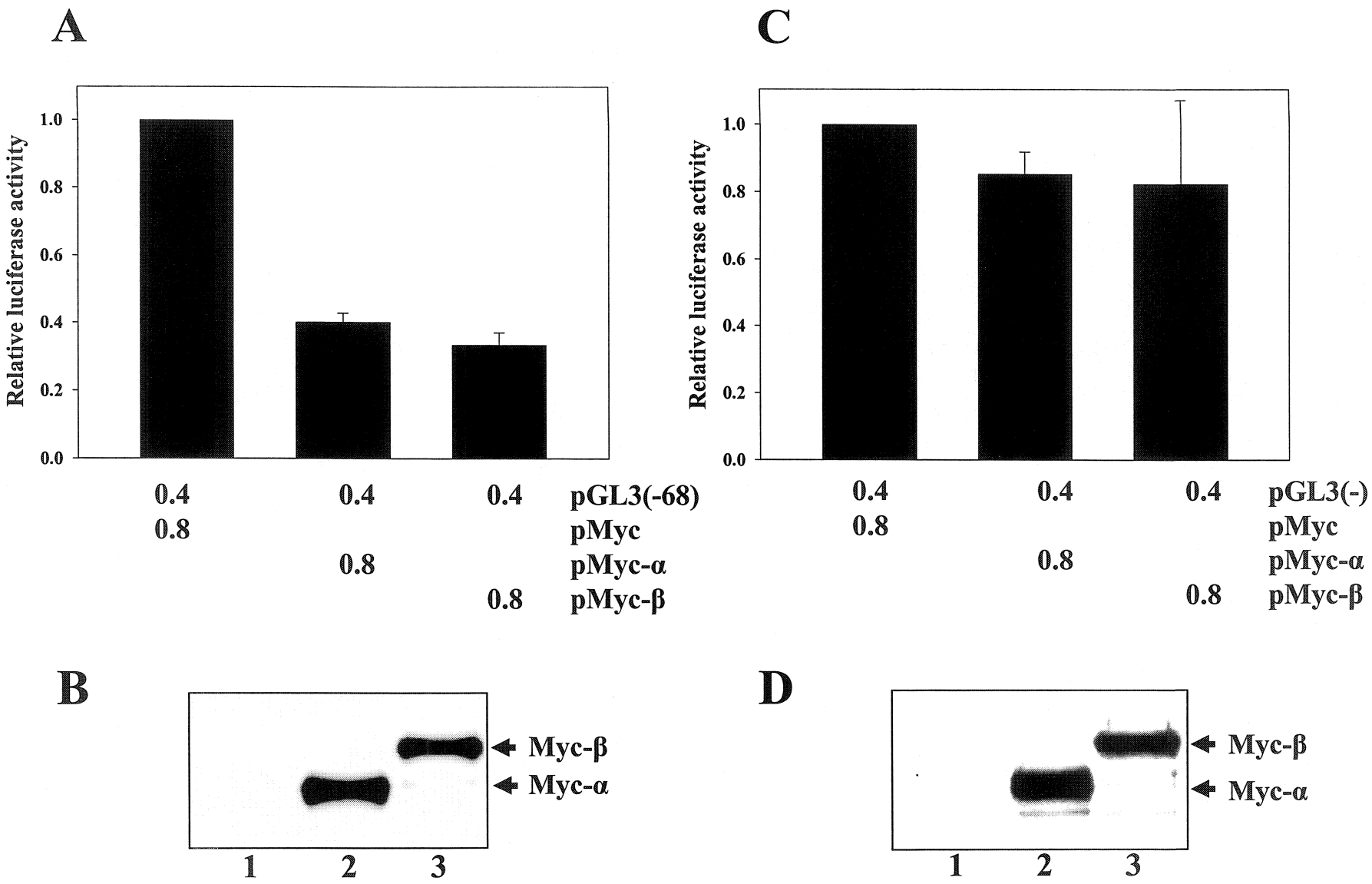

A
Relative luciferase activity
1.0
0.8
0.6
0.4
0.2
0.0
0.4 0.4 0.4 0.4 0.4 0.4 0.4 0.4 0.4 pGL3(-1316)
0.8 0.7 0.6 0.4 - 0.7 0.6 0.4 - pMyc
- 0.1 0.2 0.4 0.8 - - - - pMyc-α
- - - - - 0.1 0.2 0.4 0.8 pMyc-β
B
Myc-β
Myc-α
1 2 3 4 5 6 7 8

Figure 4.2 hMI-ER1α and β repress the *hmi-er1* P2 minimal promoter activity.

HeLa cells were co-transfected with the *hmi-er1* P2 mimimal promoter-driven luciferase reporter construct pGL3(-68) or the pGL3-Basic control vector pGL3(-) and with the Myc-tag empty vector, pMyc-α or pMyc-β. Cells were harvested 48 h after transfection and the level of relative luciferase units (RLU) was determined. For each sample, the value is presented as a proportion of the value obtained for Myc tag empty vector transfected cells (relative luciferase activity). **(A)** Shown are the average values and standard deviations from three independent experiments with the reporter construct pGL3(-68). **(B)** The level of Myc-α or Myc-β protein expressed in each sample with pGL3(-68) as described above in the panel A was determined by Western blot, using anti-Myc antibody 9E10. Shown is a representative Western blot, and the positions of the Myc-α and Myc-β proteins are indicated by arrows. **(C)** Shown are the average values and standard deviations from three independent experiments with the reporter construct pGL3(-). **(D)** The level of Myc-α or Myc-β protein expressed in each sample as described above in the panel C was determined by Western blot, using anti-Myc antibody 9E10. Shown is a representative Western blot, and the positions of the Myc-α and Myc-β proteins are indicated by arrows.


A
Relative luciferase activity
1.0
0.8
0.6
0.4
0.2
0.0
0.4 0.4 0.4 pGL3(-68)
0.8 pMyc
0.8 pMyc-α
0.8 pMyc-β
C
Relative luciferase activity
1.0
0.8
0.6
0.4
0.2
0.0
0.4 0.4 0.4 pGL3(-)
0.8 pMyc
0.8 pMyc-α
0.8 pMyc-β
B
Myc-β
Myc-α
1 2 3
D
Myc-β
Myc-α
1 2 3

demonstrated that the DNA sequence required for repression by hMI-ER1 is contained within the *hmi-er1* P2 minimal promoter (-68).

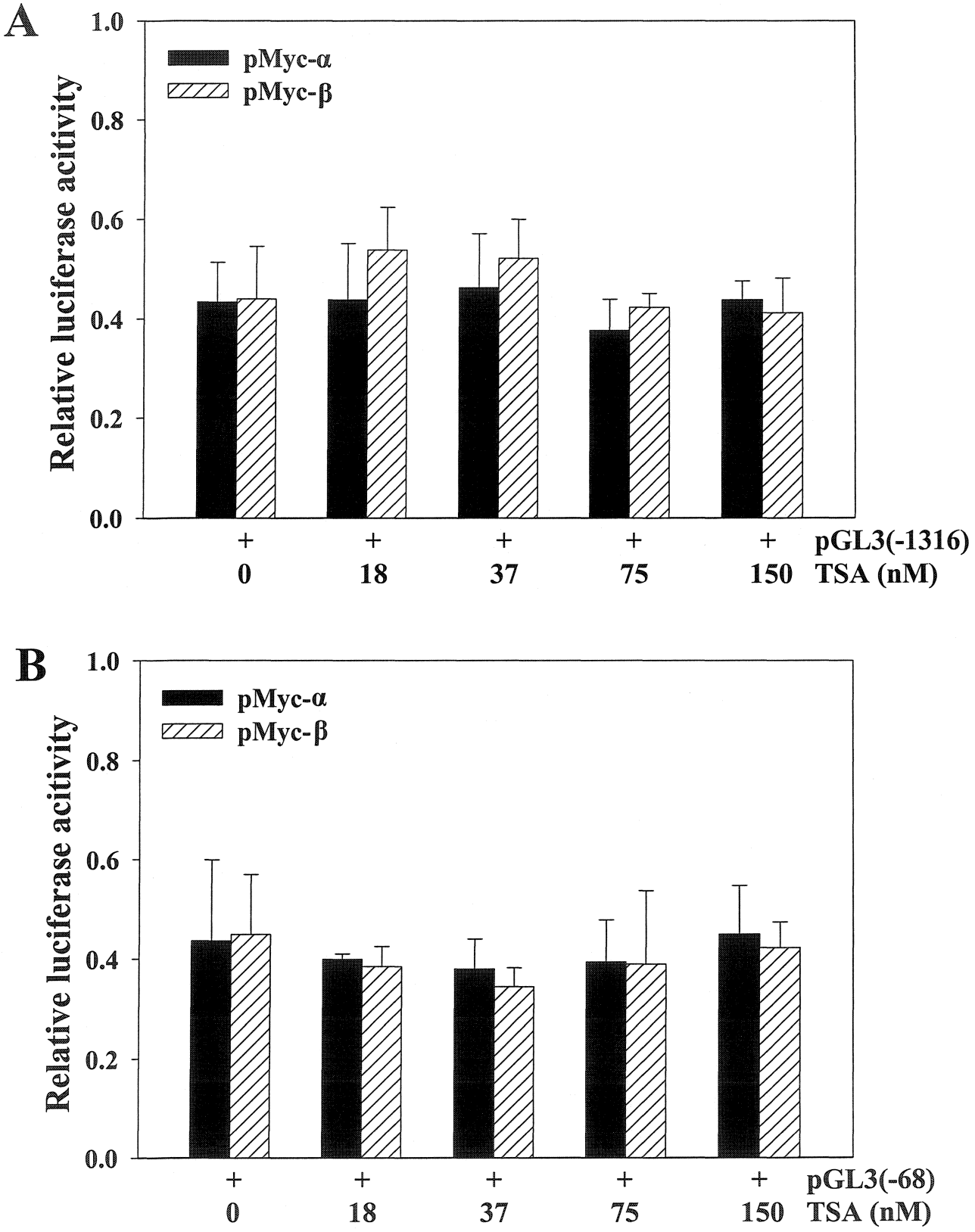
Since hMI-ER1 was shown to recruit HDAC1 enzymatic activity, one might then speculate that hMI-ER1 represses promoter activity by targeting HDAC1 enzymatic activity to the promoter. However, other mechanisms may also contribute to repression activity, such as competition for DNA binding sites on the *hmi-er1* P2 minimal promoter DNA sequence with other transcription regulators or interacting with other transcriptional regulators that can bind to the promoter and interfere with their functions. To test whether hMI-ER1α and β repress transcription through recruitment of HDAC1 enzymatic activity, HeLa cells were treated with TSA, and the effect on hMI-ER1α and β-mediated repression was examined. However, addition of TSA to the culture medium did not relieve the repression by hMI-ER1α and β on the luciferase activity of pGL3(-1316) (Figure 4.3A), or on the luciferase activity of pGL3(-68) (Figure 4.3B). These results suggested that transcriptional repression of the luciferase activity of pGL3(-1316) by hMI-ER1α and β involves a predominantly HDAC-independent mechanism, unlike what was observed with the TK promoter of G5TKCAT (Chapter 3).

4.3.2 hMI-ER1α and β do not bind to their own P2 minimal promoter

To investigate whether hMI-ER1 could bind to its P2 minimal promoter (-68) and compete for DNA binding sites with other transcription regulators, EMSAs were performed using purified GST fusion proteins and ^{32}P-labelled (-68) as a probe. These

Figure 4.3 Repression of the *hmi-er1* P2 promoter by hMI-ER1α and β is not relieved by treatment with TSA.

HeLa cells were co-transfected with 0.4 μg of the *hmi-er1* P2 promoter-driven luciferase reporter construct pGL3(-1316) or the P2 minimal promoter-driven luciferase reporter construct pGL3(-68) with the Myc-tag empty vector, pMyc-α or pMyc-β and cultured in the presence or absence of TSA. Cells were harvested 48 h after transfection, and the level of relative luciferase units (RLU) was determined. For each sample, the value is presented as a proportion of the value obtained for Myc tag empty vector-transfected cells (relative luciferase activity). **(A)** Shown are the average values and standard deviations from three independent experiments with the reporter construct pGL3(-1316). **(B)** Shown are the average values and standard deviations from three independent experiments with the reporter construct pGL3(-68).


A
1.0
0.8
0.6
0.4
0.2
0.0
Relative luciferase acitivity
pMyc-α
pMyc-β
+ + + + + pGL3(-1316)
0 18 37 75 150 TSA (nM)
B
1.0
0.8
0.6
0.4
0.2
0.0
Relative luciferase acitivity
pMyc-α
pMyc-β
+ + + + + pGL3(-68)
0 18 37 75 150 TSA (nM)

assays revealed that hMI-ER1α and β do not bind to the ^{32}P-labelled (-68) (Figure 4.4, lane 6 and 8), while Sp1 does (band a and b) (Figure 4.4, lane 4). These results are in agreement with the previous observation that MI-ER has no DNA binding activity (Paterno *et al.*, unpublished data), which was determined from an unbiased set of degenerate oligonucleotides using CASTing methods (Wright *et al.*, 1991). These results indicated that repression of transcriptional activation on the P2 minimal promoter (-68) is not the result of competition for binding to the promoter sequence with other transcription factors, such as Sp1.

4.3.3 hMI-ER1α and β physically interact with Sp1 *in vitro* and *in vivo*

Since overexpression of Sp1 was found to activate the luciferase activity of *hmi-er1* P2 minimal promoter-driven luciferase reporter construct pGL3(-68) (Figure 2.6, Chapter 2), experiments were performed to investigate whether overexpression of hMI-ER1α and β alter the luciferase activity of pGL3(-68) stimulated by Sp1. HeLa cells were co-transfected with pGL3(-68), pCR-Sp1 and either pMyc, pMyc-α, or pMyc-β. Luciferase reporter assays revealed that overexpression of both hMI-ER1α and β (Figure 4.5B) repressed promoter activity stimulated by Sp1 in a dose-dependent manner (Figure 4.5A).

Four Sp1 binding sites were predicted in the *hmi-er1* P2 minimal functional promoter and Sp1 was also found to bind to this region (Chapter 2). Since hMI-ER1α and β do not bind to (-68) (Figure 4.4) but repress the luciferase activity of pGL3(-68)

Figure 4.4 hMI-ER1α and β do not bind to *hmi-er1* P2 minimal promoter.

Shown is a representative EMSA performed using GST fusion proteins and a ^{32}P-labelled the *hmi-er1* P2 minimal promoter as a probe [^{32}P(-68)]. Shown is an EMSA using labelled minimal promoter region as a probe. In lanes 1 and 2, the ^{32}p(-68) probe was incubated with GST in the presence or absence, respectively, of 20-fold molar excess of nonlabelled (competitor) probe. In lanes 3 and 4, the labelled probe was incubated with GST-Sp1 in the presence or absence of 20-fold molar excesses of nonlabelled (competitor) probe, respectively. In lanes 5 and 6, the labelled probe was incubated with GST-hMI-ER1α (GST-α) in the presence or absence of 20-fold molar excesses of nonlabelled (competitor) probe, respectively. In lanes 7 and 8, the labelled probe was incubated with GST-hMI-ER1β (GST-β) in the presence or absence of 20-fold molar excesses of nonlabelled (competitor) probe, respectively. The positions of DNA-protein complexes (band a and b) are indicated by arrowheads. The position of the free probe is indicated by arrow.

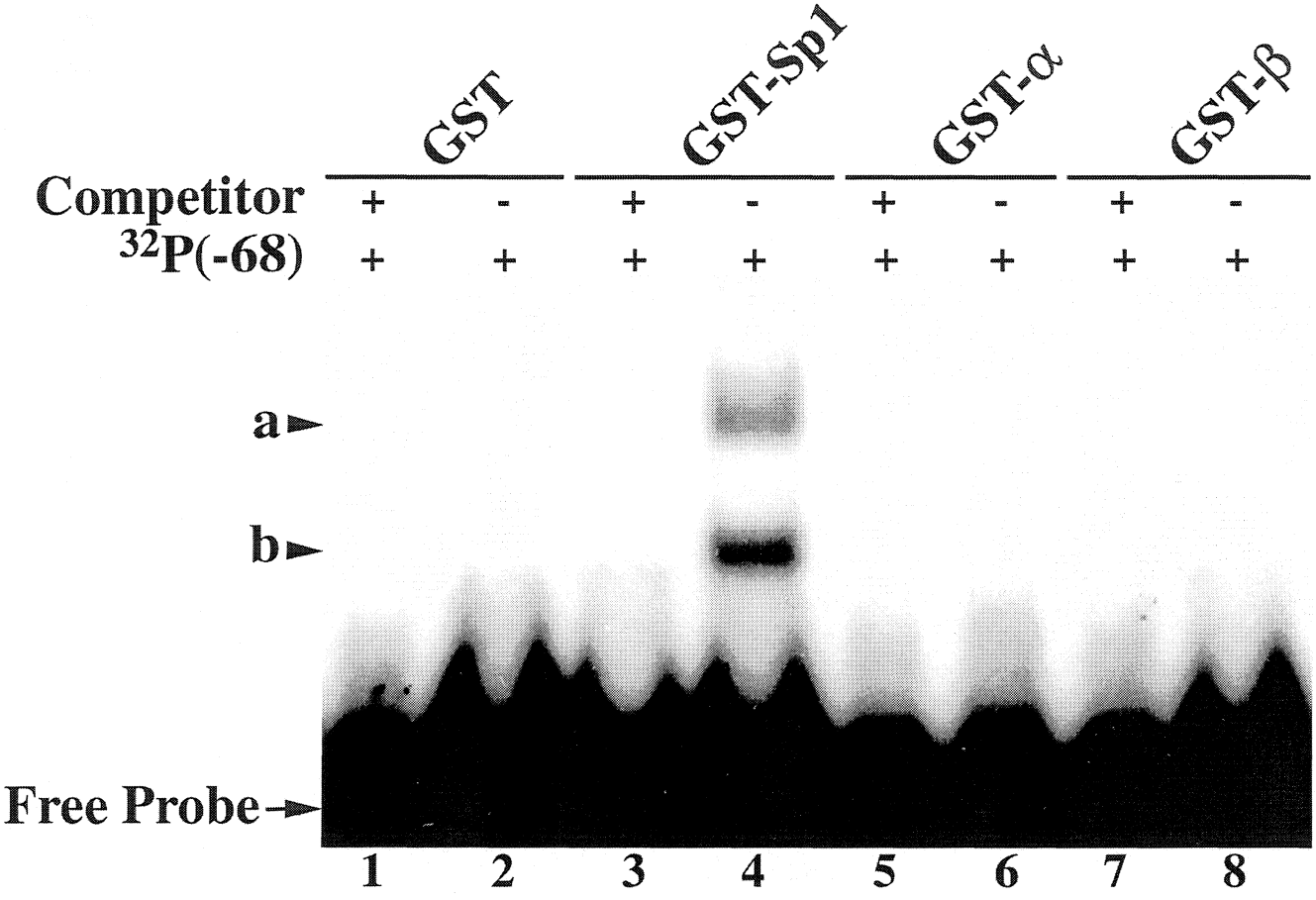
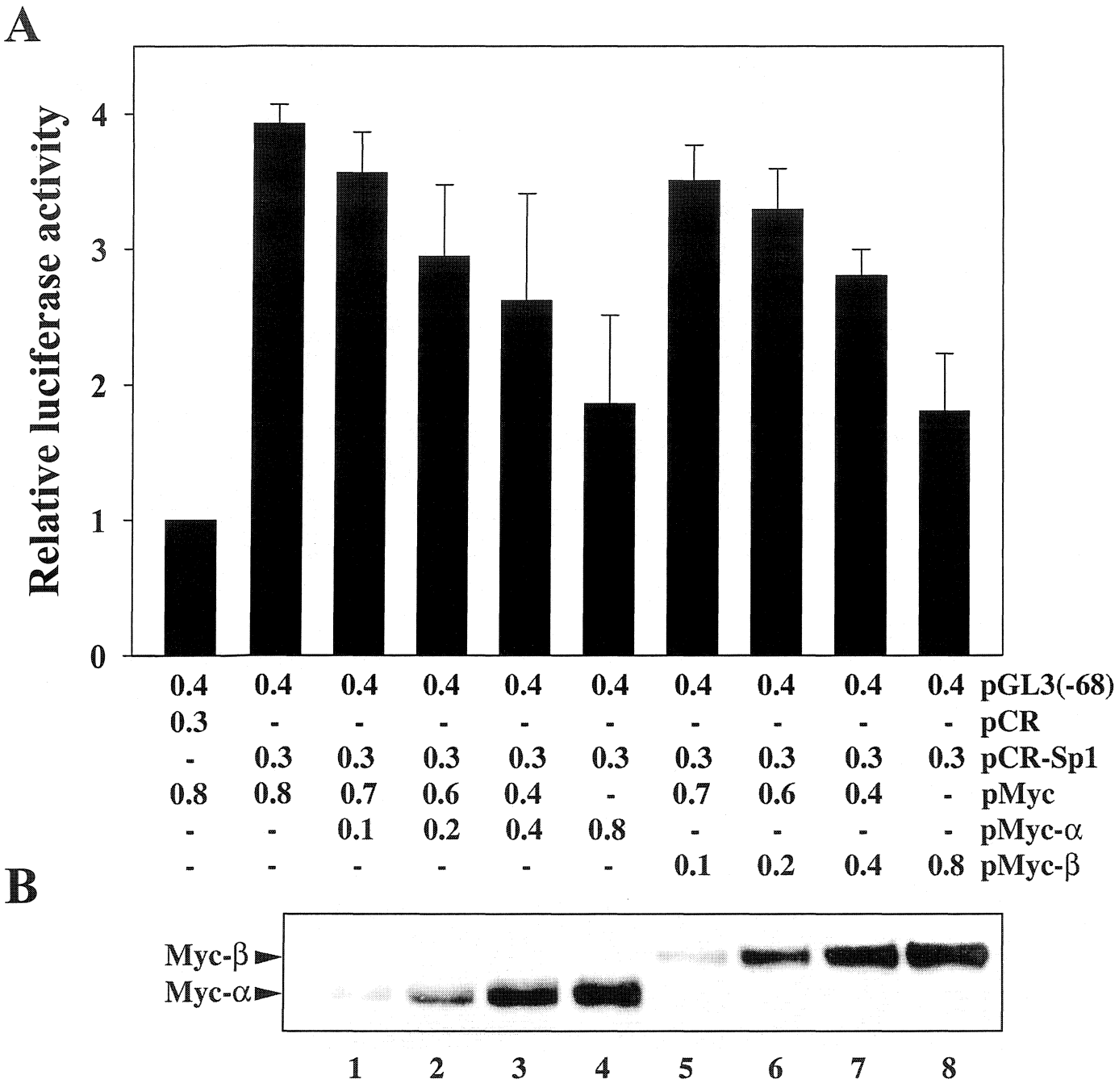

GST
GST-Sp1
GST-α
GST-β
Competitor
+ - + - + - + -
32P(-68)
+ + + + + + + +
a
b
Free Probe
1 2 3 4 5 6 7 8

Figure 4.5 Overexpression of hMI-ER1α and β disrupt *hmi-er1* P2 promoter activity stimulated by Sp1.

(A) HeLa cells were co-transfected with P2 minimal promoter-driven luciferase reporter construct pGL3(-68) and pCR3.1-Sp1 (pCR-Sp1) construct with the Myc-tag empty vector, pMyc-α or pMyc-β; the amount of plasmid (μg) used for transfection is indicated below each bar. Cells were harvested 48h after transfection and the relative luciferase units (RLU) was determined. For each sample, the value is presented as a proportion of the value obtained for Myc tag empty vector and empty pCR3.1 vector transfected cells (relative luciferase activity). **(B)** The level of Myc-α or Myc-β protein expressed in each sample was determined by Western blot, using anti-Myc antibody 9E10. Shown is a representative Western blot, and the positions of the Myc-α and Myc-β proteins are indicated by arrowheads.


A
Relative luciferase activity
0
1
2
3
4
0.4 0.4 0.4 0.4 0.4 0.4 0.4 0.4 0.4 0.4 pGL3(-68)
0.3 - - - - - - - - - pCR
- 0.3 0.3 0.3 0.3 0.3 0.3 0.3 0.3 0.3 pCR-Sp1
0.8 0.8 0.7 0.6 0.4 - 0.7 0.6 0.4 - pMyc
- - 0.1 0.2 0.4 0.8 - - - - pMyc-α
- - - - - - 0.1 0.2 0.4 0.8 pMyc-β
B
Myc-β
Myc-α
1 2 3 4 5 6 7 8

stimulated by Sp1 (Figure 4.5), it is possible that hMI-ER1α and β may repress the activity from this promoter region through interaction with the transcriptional regulator Sp1. Thus, experiments were performed to see if an interaction between hMI-ER1 and Sp1 was possible. Co-immunoprecipitation analysis was utilized to examine the ability of hMI-ER1α and β isoforms to physically associate with Sp1 both *in vitro* and *in vivo*. ^{35}S-labelled Myc-tagged hMI-ER1α (Myc-α) and hMI-ER1β (Myc-β) were synthesized *in vitro* and mixed with unlabelled *in vitro*-translated *in vitro* translated Sp1, then the mixture was subjected to immunoprecipitation with anti-Sp1 antibody. Both hMI-ER1α and β isoforms were detected in Sp1 immunoprecipitates while luciferase, a control protein, was not (Figure 4.6A). Reciprocal experiments, using anti-Myc antibody 9E10 for immunoprecipitation, confirmed that Sp1 could associate with either hMI-ER1 isoform (Figure 4.6B). GST pull-down assays again confirmed that *in vitro* translated Sp1 specifically associated with GST-hMI-ER1α or hMI-ER1β, but not with the GST control (Figure 4.6C).

In vivo analysis involved transient expression of the Myc-tag empty vector, Myc-α or Myc-β in HeLa cells. Expression of Myc-α and Myc-β was verified by Western blot analysis with anti-Myc antibody 9E10 (Figure 4.7, lane 2 and 3). Cell extracts were subjected to immunoprecipitation with anti-Sp1 antibody, followed by Western blot analysis with anti-Myc antibody 9E10. Both hMI-ER1α and β isoforms were co-immunoprecipitated with endogenous Sp1 by anti-Sp1 antibody (Figure 4.7A, lane 7 and 9), but not in immunoprecipitates with non-immune serum control (Figure 4.7A, lane 6 and 8).

Figure 4.6 hMI-ER1α and β interact physically with Sp1 *in vitro*.

(A) ^{35}S-labelled TNT mixtures programmed with cDNA encoding luciferase (Luc), Myc tagged hMI-ER1α [Myc-α (α)], or Myc tagged hMI-ER1β [Myc-β (β)] were loaded directly on the gel (1/20 of input) (lanes 1 to 3) or incubated with unlabelled TNT mixtures programmed with *Sp1* cDNA and subjected to immunoprecipitation (IP) with anti-Sp1 antibody (lances 4-6). Proteins were resolved by SDS-PAGE and visualized by autoradiography. Shown is a representative autoradiograph, and the positions of the Myc-α, Myc-β and luciferase proteins are indicated by arrowheads. **(B)** ^{35}S-labelled TNTs programmed with cDNA encoding luciferase (Luc), or Sp1 were loaded directly on the gel (1/20 of input) (lanes 1, 2) or incubated with unlabelled TNTs programmed with cDNA encoding Myc-α (lanes 3, 4) or Myc-β (lanes 5, 6) and subjected to immunoprecipitation (IP) with anti-Myc antibody 9E10. Proteins were resolved by SDS-PAGE and visualized by autoradiography. Shown is a representative autoradiograph, and the positions of the Sp1 and luciferase proteins are indicated by arrowheads. **(C)** ^{35}S-labelled TNTs programmed with cDNA encoding Sp1 was loaded directly on the gel (1/20 of input) (lane 1) or incubated with GST (lane 2), GST-hMI-ER1α (GST-α) (lane 3), GST-hMI-ER1β (GST-β) (lane 4). Proteins were resolved by SDS-PAGE and visualized by autoradiography. Shown is a representative autoradiograph, and the position of the Sp1 protein is indicated by arrowheads.

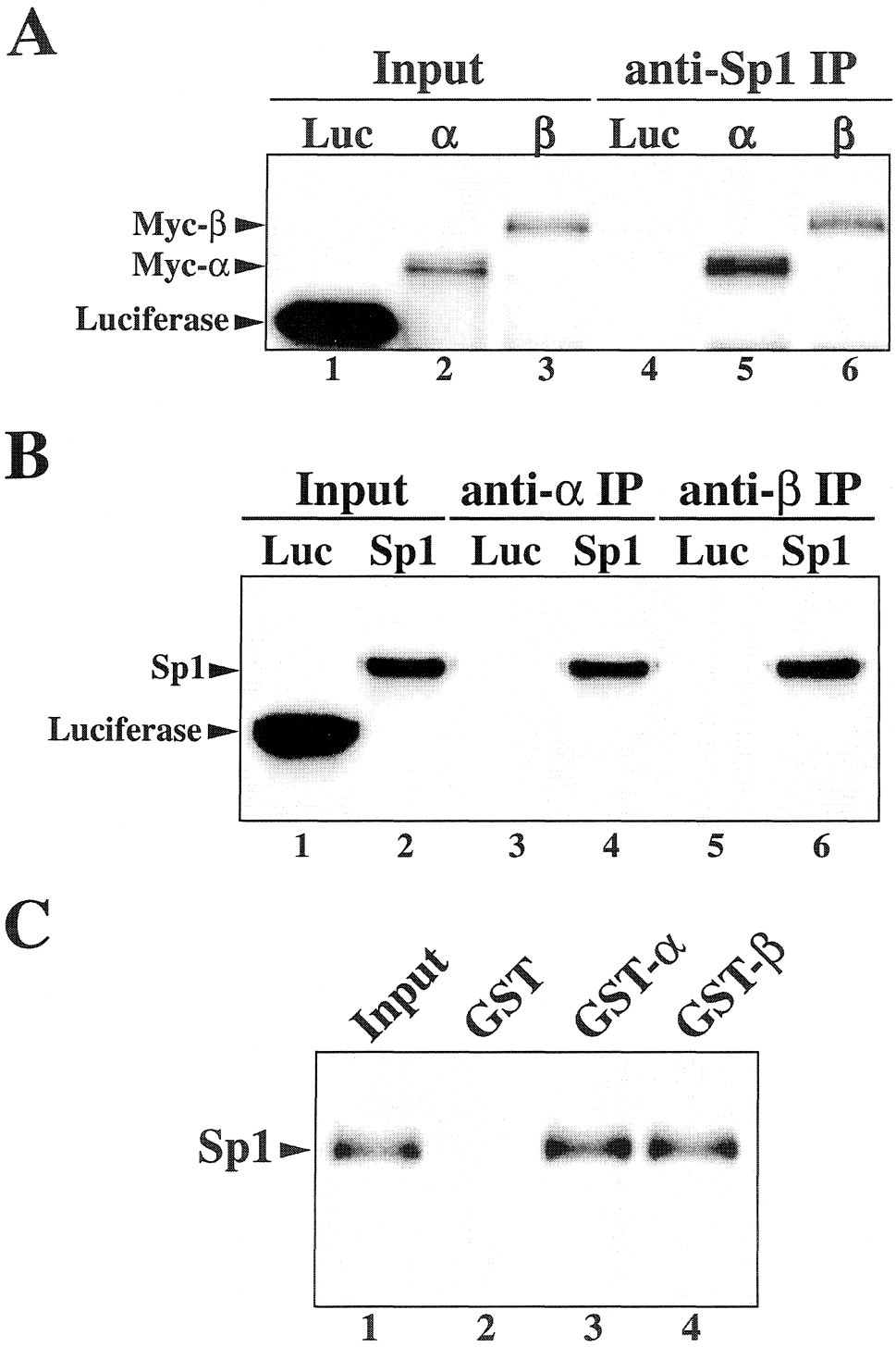
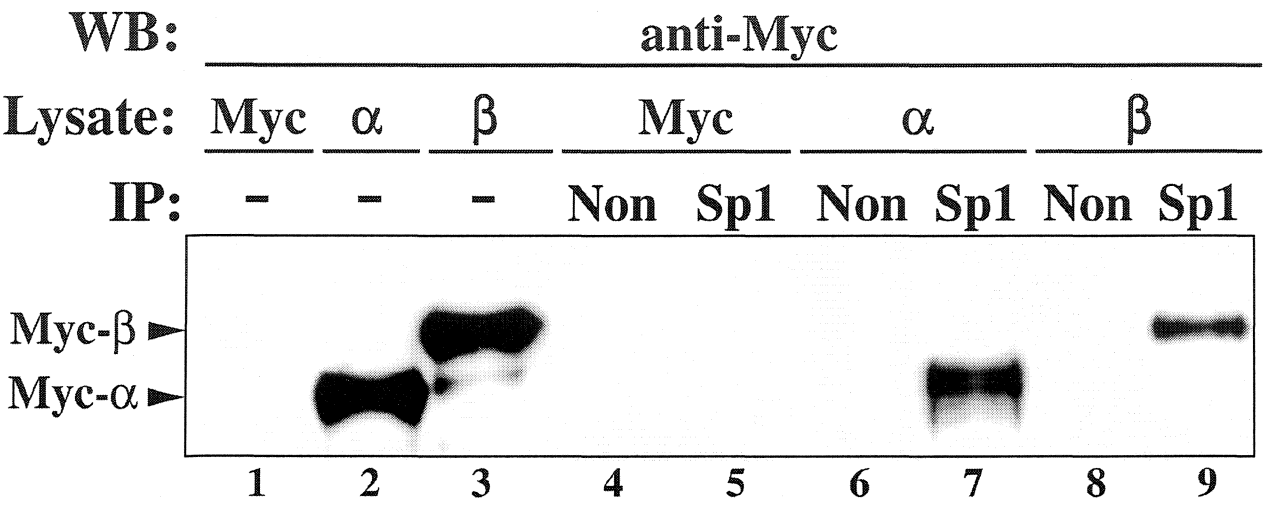

A
Input
anti-Sp1 IP
Luc
α
β
Luc
α
β
Myc-β
Myc-α
Luciferase
1
2
3
4
5
6
B
Input
anti-α IP
anti-β IP
Luc
Sp1
Luc
Sp1
Luc
Sp1
Sp1
Luciferase
1
2
3
4
5
6
C
Input
GST
GST-α
GST-β
Sp1
1
2
3
4

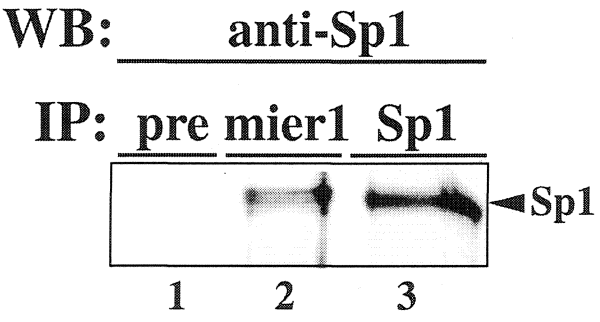
Figure 4.7 hMI-ER1α and β interact physically with Sp1 *in vivo*.

(A) Cell lysates from HeLa cells, transiently transfected with Myc-tag empty vector, Myc-α or Myc-β, were prepared and equivalent amounts of protein from each sample were either added directly to sample buffer (1/3 of input) (lanes 1-3) or subjected to IP with anti-Sp1 antibody (lanes 5, 7, 9), or with non-immune serum (lanes 4, 6, 8); Western blot (WB) analysis was performed using anti-Myc antibody 9E10. Shown is a representative Western blot, and the positions of the Myc-α and Myc-β proteins are indicated by arrowheads. **(B)** Untransfected HeLa cell lysates were subjected to IP with pre-immune serum (lane 1) or anti-hMI-ER1 (lane 2), or anti-Sp1 antiserum (lane 3). Western blot (WB) analysis was performed using anti-Sp1 antibody. Shown is a representative Western blot, and the position of the Sp1 protein is indicated by arrowheads.

A


WB:
anti-Myc
Lysate:
Myc
α
β
Myc
α
β
IP:
-
-
-
Non
Sp1
Non
Sp1
Non
Sp1
Myc-β
Myc-α
1
2
3
4
5
6
7
8
9

B

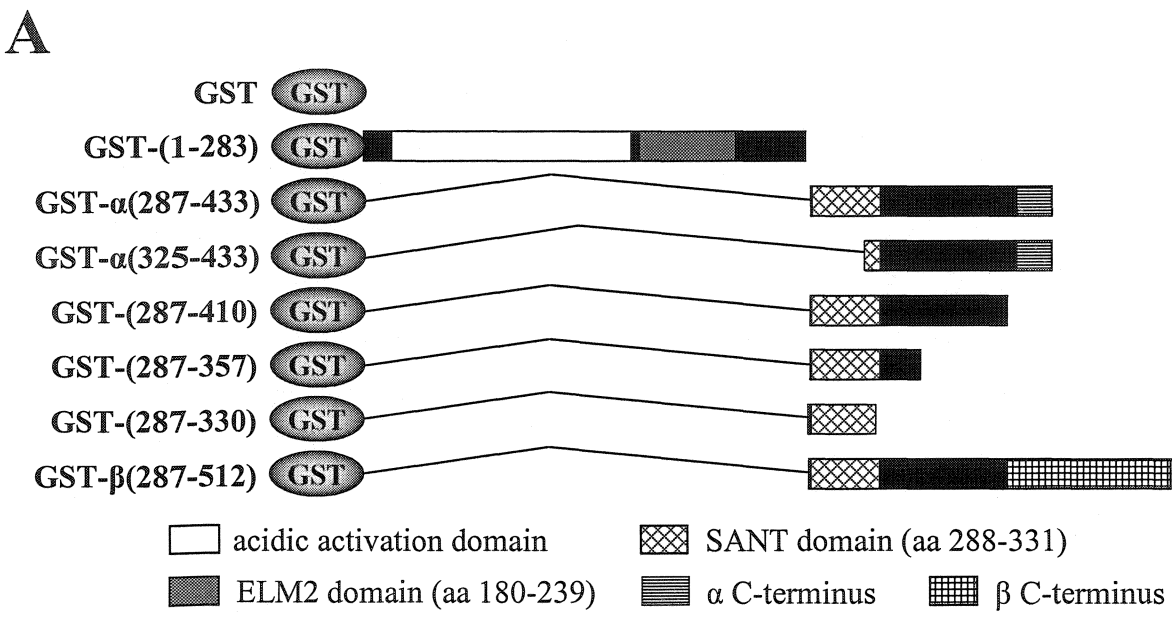

WB:
anti-Sp1
IP:
pre mier1
Sp1
Sp1
1
2
3

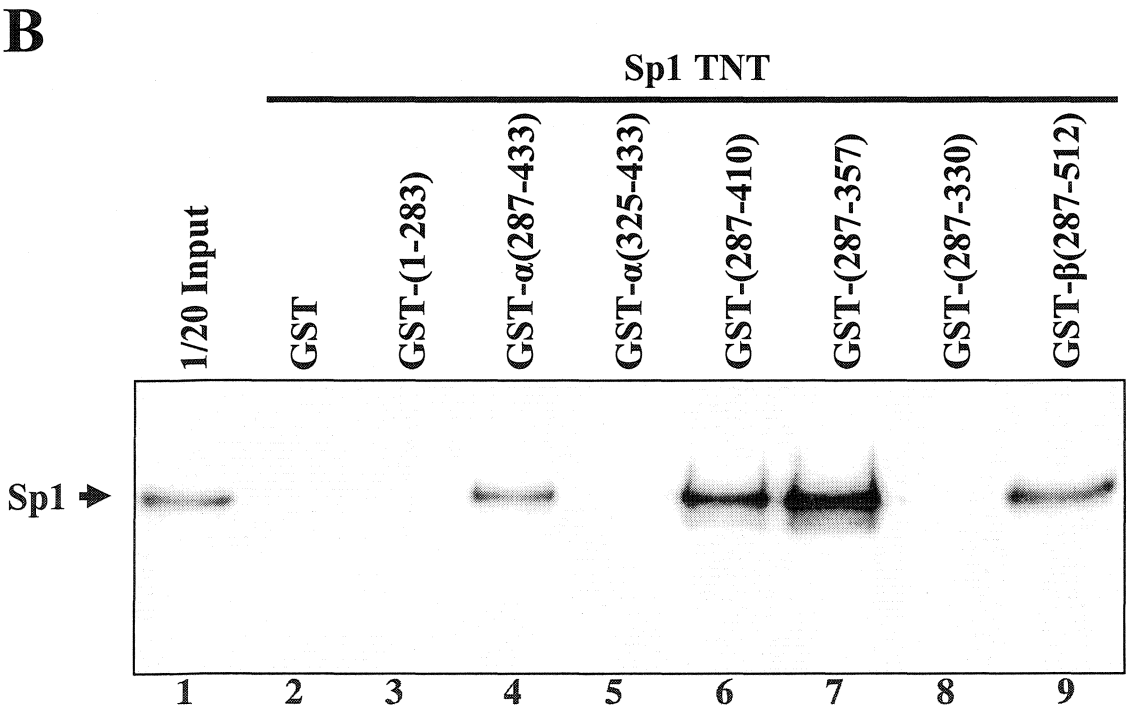
The next set of experiments was designed to investigate whether endogenous complexes containing hMI-ER1 and Sp1 exist in the cell to exclude the possibility that the observed interaction was an artifact of overexpression of transfected constructs. Co-immunoprecipitation analysis of extracts from non-transfected HeLa cells was performed, using pre-immune serum or an anti-hMI-ER1 antibody. As shown in Figure 4.7B, Sp1 protein was detected in the hMI-ER1 immunoprecipitate (lane 2), but not in the control (lane 1), demonstrating that hMI-ER1 can physically associate with endogenous Sp1 *in vivo*.

To determine which region of the hMI-ER1 protein is required for interaction with Sp1, a series of GST-hMI-ER1α and β deletions were constructed (Figure 4.8A) and pull-down assays were performed with ^{35}S-labelled Sp1. The deletion analysis revealed that the sequence required for interaction with Sp1 maps to residues 287-357 (Figure 4.8B, lane 4, 6, 7 and 9). Sequence analysis of this region using PFAM at http://pfam.wustl.edu/cgi-bin/getalignment revealed a conserved SANT domain was located between aa 288 and 331 (Appendices 2). Further deletion of this region at the C-terminal end to remove residues from aa 331 to 357 (^{331}K to ^{357}T), which includes a highly conserved lysine (K) of SANT domain at position 331 and the C-terminal 26 aa, completely abolished interaction with Sp1 (Figure 4.8B, lane 8). Sp1 did not interact with GST control (Figure 4.8B, lane 6), the N-terminal aa 1-283 deletion mutant of hMI-ER1 (Figure 4.8B, lane 3), or the C-terminal aa 325-433 deletion mutant of hMI-ER1 (Figure 4.8B, lane 5). These results demonstrate the importance of an intact SANT domain and the C-terminal 26 aa for the interaction with Sp1.

Figure 4.8 The region of hMI-ER1 containing SANT domain interacts with Sp1 *in vitro*.

(A) The diagram illustrates the deletion mutants of hMI-ER1 fused to GST in pGEX-4T-1 vector. The individual domains are identified in the legend below the diagram, and the hMI-ER1 amino acid residues encoded by each construct are listed. **(B)** Shown is GST pull-down assays of hMI-ER1 deletion mutants with Sp1. ^{35}S-labelled TNTs programmed with cDNA encoding *Sp1* was loaded directly on the gel (lane 1) or incubated with GST (lane 2), GST-(1-283) (lane 3), GST-(287-433) (lance 4), GST-(325-433) (lane 5), GST-(287-410) (lane 6), GST-(287-357) (lane 7), GST-(287-330) (lane 8), and GST-(287-512) (lane 9). Proteins were resolved by SDS-PAGE and visualized by autoradiography. The position of the Sp1 protein is indicated by arrowheads.


A
GST
GST-(1-283)
GST-α(287-433)
GST-α(325-433)
GST-(287-410)
GST-(287-357)
GST-(287-330)
GST-β(287-512)
acidic activation domain
SANT domain (aa 288-331)
ELM2 domain (aa 180-239)
α C-terminus
β C-terminus


B
Sp1 TNT
1/20 Input
GST
GST-(1-283)
GST-α(287-433)
GST-α(325-433)
GST-(287-410)
GST-(287-357)
GST-(287-330)
GST-β(287-512)
Sp1
1
2
3
4
5
6
7
8
9

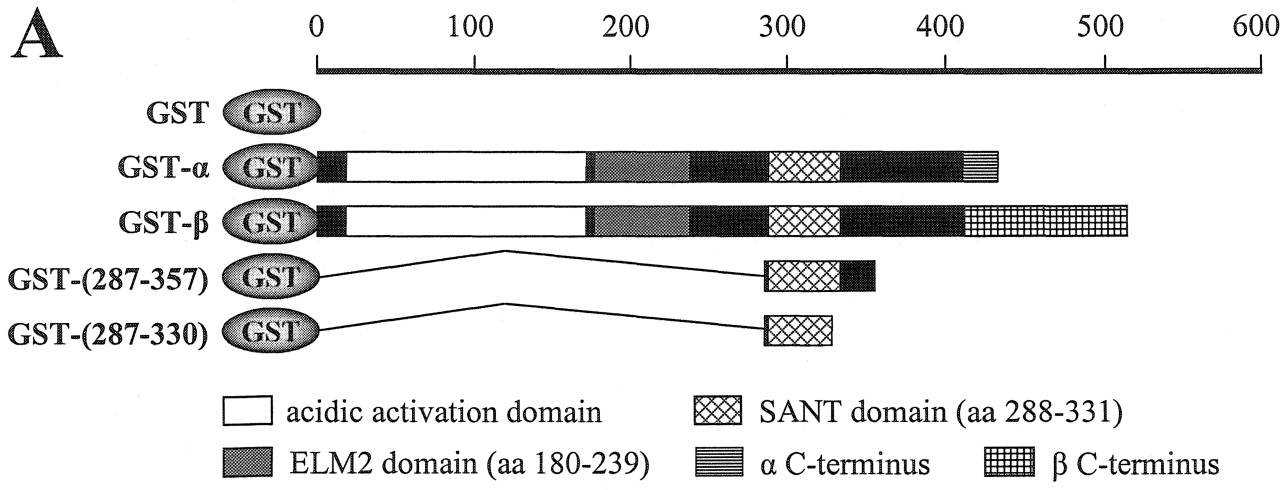
4.3.4 hMI-ER1α and β interfere with GC box recognition by Sp1

The finding that GST-hMI-ER1α and β isoforms can associate with Sp1 raises the possibility that interaction with Sp1 may affect the ability of Sp1 to bind DNA, thus providing a mechanism for the regulation of the *hmi-er1* promoter by hMI-ER1. To examine the effects of GST-hMI-ER1α (GST-1α) and GST-hMI-ER1β (GST-β) on Sp1 DNA-binding activity, EMSAs were performed using GST-Sp1 with the *hmi-er1* minimal promoter sequence, in the presence or absence of various GST-hMI-ER1 fusion proteins (Figure 4.9A). As shown in Figure 4.9B, addition of GST-1α or GST-β to the EMSA reaction mixtures resulted in the disruption of the Sp1-DNA complexes (band a and b) (lanes 3, 4), while addition of GST alone had no effect (lane 2). Furthermore, addition of GST-(287-357), containing an intact SANT domain, disrupted the Sp1-DNA complexes (band a and b) (lane 5), while addition of GST-(287-330) construct, in which the last amino acid residue of the SANT domain and C-terminal 26 aa were removed, did not (lane 6). An uncharacterized third band (c) appeared in the samples with addition of GST-1α, GST-β, or GST-(287-357), but not in the samples with GST alone or GST-(287-330).

Next, experiments were performed to examine whether hMI-ER1 could affect the ability of Sp1 to interact with its consensus binding site. EMSAs were performed using GST-Sp1 and ^{32}P-labelled double-stranded oligonucleotide consisting of the Sp1 consensus sequence (Sp1CS), in the presence or absence of various GST-hMI-ER1 fusion proteins (Figure 4.10A). As seen with the *hmi-er1* minimal promoter, addition of GST-α, GST-β or GST-(287-357) to the EMSA reaction mixtures resulted in the disruption of

Figure 4.9 hMI-ER1α and β interfere with GC box recognition by Sp1 to the *hmi-er1* P2 minimal promoter.

(A) The diagram illustrates the deletion mutants of hMI-ER1 fused to GST in pGEX-4T-1 vector. The individual domains are identified in the legend below the diagram, and the hMI-ER1 amino acid residues encoded by each construct are listed. **(B)** Shown is a representative EMSA performed using ^{32}P-labelled the *hmi-er1* P2 minimal promoter (-68) as a probe [^{32}P(-68)]. The ^{32}p(-68) probe was incubated with GST-Sp1 in the presence of GST elution buffer (lane 1), GST (lane 2), GST-hMI-ER1α (GST-α) (lane 3), GST-hMI-ER1β (GST-β) (lane 4), GST-(287-357) (lane 5), or GST-(287-330) (lane 6). The positions of DNA-protein complexes (band a, b and c) are indicated by arrowheads. The position of the free probe is indicated by arrow.


A
0
100
200
300
400
500
600
GST
GST-α
GST-β
GST-(287-357)
GST-(287-330)
acidic activation domain
SANT domain (aa 288-331)
ELM2 domain (aa 180-239)
α C-terminus
β C-terminus

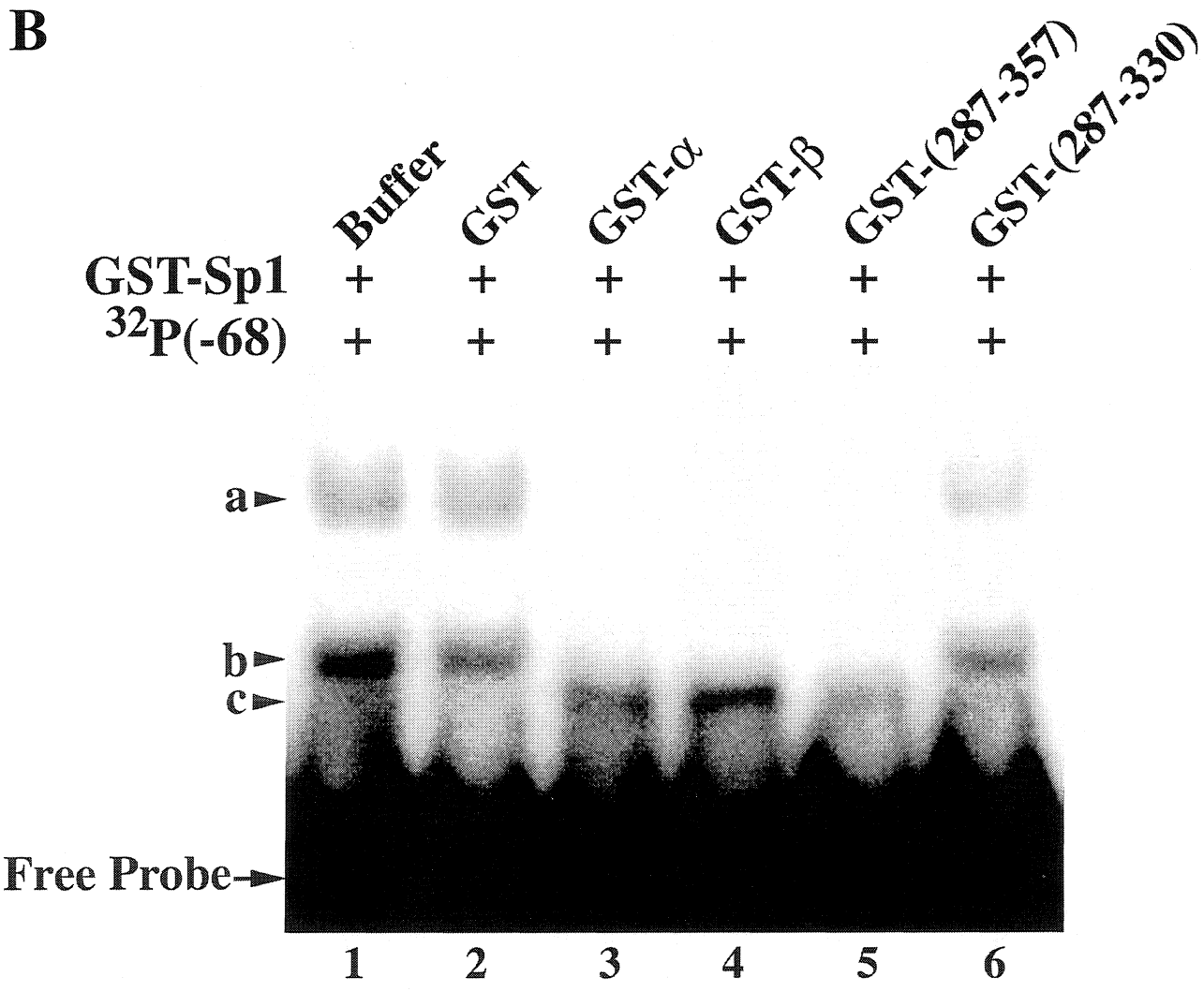
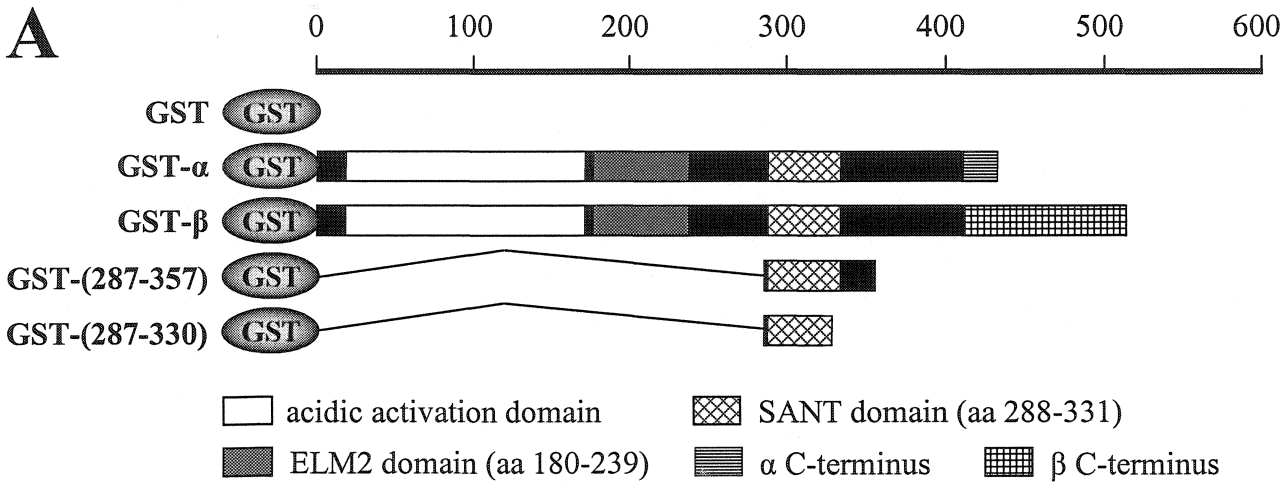
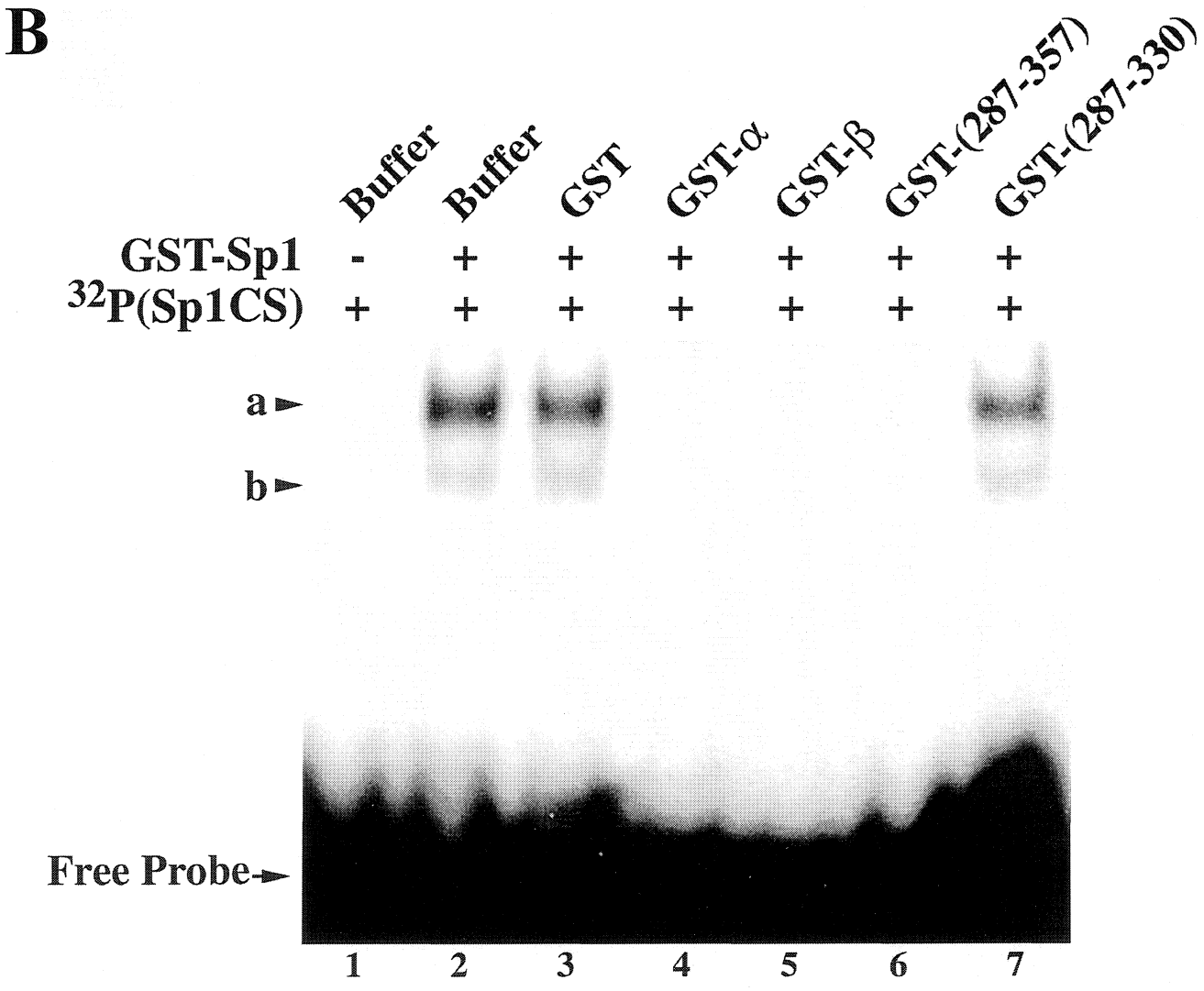

B
Buffer
GST
GST-α
GST-β
GST-(287-357)
GST-(287-330)
GST-Sp1 + + + + + +
32P(-68) + + + + + +
a
b
c
Free Probe
1
2
3
4
5
6

Figure 4.10 hMI-ER1α and β interfere with DNA binding by Sp1 to the Sp1 consensus oligonucleotide.

(A) The diagram illustrates the deletion mutants of hMI-ER1 fused to GST in pGEX-4T-1 vector. The individual domains are identified in the legend below the diagram, and the hMI-ER1 amino acid residues encoded by each construct are listed. **(B)** Shown is a representative EMSA performed using labelled consensus Sp1 oligo [^{32}p(Sp1CS)] as a probes. The ^{32}p(Sp1CS) probe was incubated with GST-Sp1 in the presence of elution buffer (lane 1), GST (lane 2), GST-hMI-ER1α (GST-α) (lane 3), GST-hMI-ER1β (GST-β) (lane 4), GST-(287-357) (lane 5), or GST-(287-330) (lane 6). Proteins were resolved by SDS-PAGE and visualized by autoradiography. The positions of DNA-protein complexes (band a, b) are indicated by arrowheads. The position of the free probe is indicated by arrow.


A
0
100
200
300
400
500
600
GST
GST-α
GST-β
GST-(287-357)
GST-(287-330)
acidic activation domain
SANT domain (aa 288-331)
ELM2 domain (aa 180-239)
α C-terminus
β C-terminus


B
Buffer
Buffer
GST
GST-α
GST-β
GST-(287-357)
GST-(287-330)
GST-Sp1
32P(Sp1CS)
a
b
Free Probe
1
2
3
4
5
6
7

Sp1-DNA complexes (band a and b) (Figure 4.10B, lane 4, 5 and 6), while addition of GST and GST-(287-330) had no effect (Figure 4.10B, lane 3 and 7). These results confirm that hMI-ER1α or β interferes with the ability for Sp1 to bind to its cognate site on DNA. Furthermore, this interference was not due to competition with Sp1 for binding to the Sp1 consensus sequence, as neither hMI-ER1 isoform could form a DNA-protein complex with the Sp1 consensus oligonucleotide (Figure 4.11, lane 6 and 8).

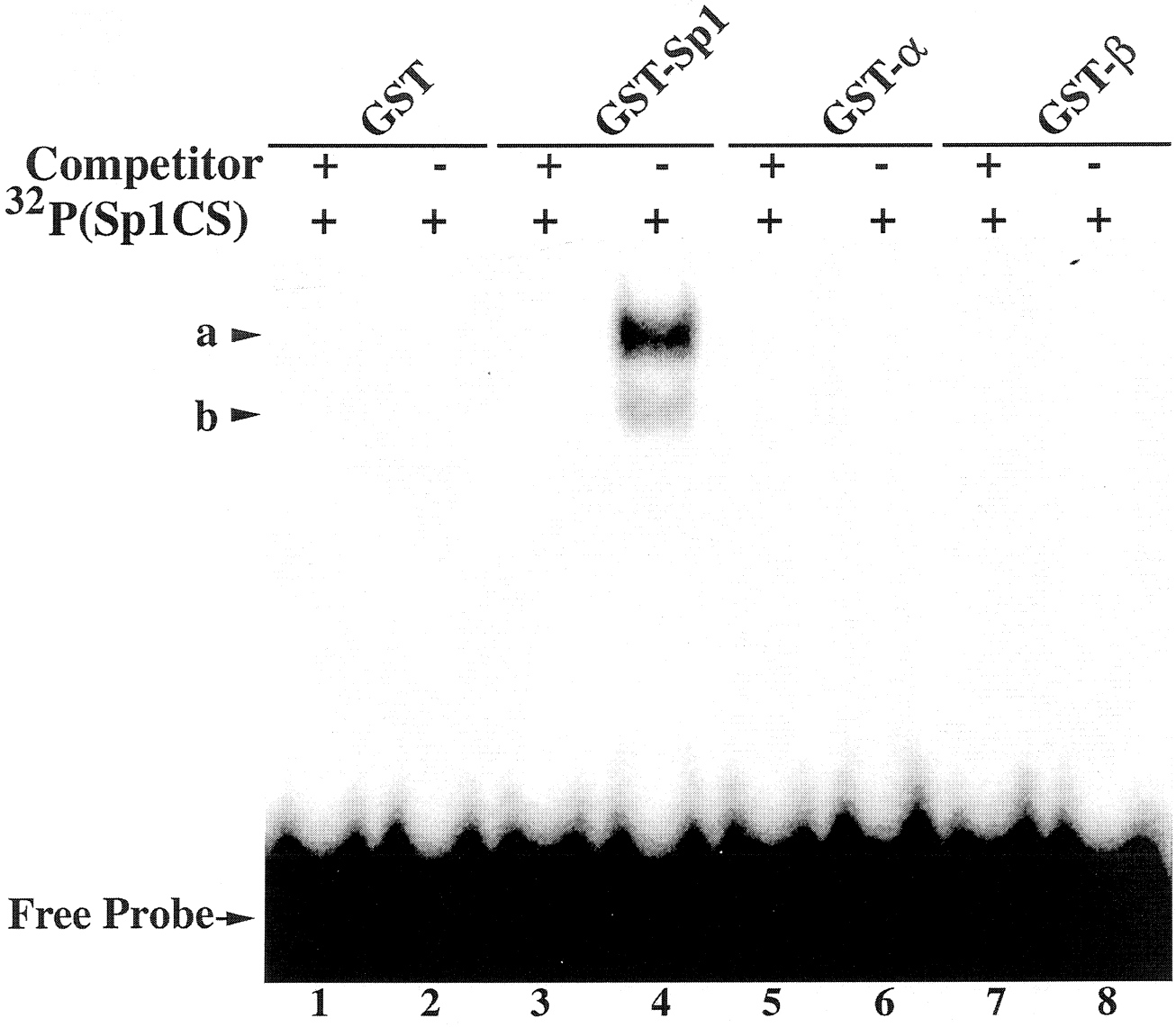
4.4 Discussion

Our previous studies have shown that overexpression of both GAL4 tagged hMI-ER1α and hMI-ER1β significantly repressed transcription from the TK promoter of the G5TKCAT reporter plasmid (Chapter 3). Maximum repression resulted in a reduction in CAT expression to 10% and 8% of control levels by hMI-ER1α and hMI-ER1β, respectively (Figure 3.1). In the studies described in this chapter, overexpression of both Myc tagged hMI-ER1α and hMI-ER1β also significantly repressed luciferase activity of pGL3(-1316), in a dose-dependent manner (Figure 4.1). Maximum repression resulted in a reduction in reporter gene expression to 36% and 29% of control levels by hMI-ER1α and hMI-ER1β, respectively (Figure 4.1).

The DNA sequence required for repression by hMI-ER1 on the luciferase activity of pGL3(-1316) is contained within the P2 minimal promoter (-68) (Figure 4.2). There are four putative Sp1 and one TEF-2 binding sites in this region (Figure 2.3). One might speculate that hMI-ER1 represses promoter activity on this region by: (1) targeting HDAC1 activity to the promoter; (2) competing for DNA binding sites with other

Figure 4.11 hMI-ER1α and β do not bind to the Sp1 consensus oligonucleotide.

Shown is a representative EMSA performed using labelled Sp1 consensus oligonucleotide [^{32}p(Sp1CS)] as a probes. In lanes 1 and 2, ^{32}p(Sp1CS) was incubated with 1 μg GST fusion protein in the presence or absence, respectively, of 20-fold molar excess of nonlabelled (competitor) probe. In lanes 3 and 4, the labelled probe was incubated with GST-Sp1 in the presence or absence, respectively, of 20-fold molar excesses of nonlabelled (competitor) probe. In lanes 5 and 6, the labelled probe was incubated with GST-hMI-ER1α (GST-α) in the presence or absence of 20-fold molar excesses of nonlabelled (competitor) probe, respectively. In lanes 7 and 8, the labelled probe was incubated with GST-hMI-ER1β (GST-β) in the presence or absence of 20-fold molar excesses of nonlabelled (competitor) probe, respectively. The position of the free probe is indicated by arrow.


GST
GST-Sp1
GST-α
GST-β
Competitor + - + - + - + -
32P(Sp1CS) + + + + + + + +
a
b
Free Probe
1 2 3 4 5 6 7 8

transcription regulators such as Sp1 and/or TEF-2; (3) interacting and interfering with other transcriptional regulators such as Sp1 and/or TEF-2 that can bind to the promoter; or (4) changing endogenous Sp1 concentration and/or activity. The present study showed that transcriptional repression of P2 *hmi-er1* promoter activity by hMI-ER1α and β is not relieved by treatment with TSA (Figure 4.3), unlike what was observed with the TK promoter of G5TKCAT (Chapter 3). EMSAs also revealed that hMI-ER1α and β can not bind to the *hmi-er1* P2 minimal functional promoter (Figure 4.4), suggesting that hMI-ER1α and β may not compete for the DNA binding sites for other transcription factors such as Sp1 and TEF-2. Since our previous study showed that Sp1 binds to the *hmi-er1* P2 minimal promoter (-68) and enhances the luciferase activity of pGL3(-68) (Figure 2.5 and Figure 2.6 in Chapter 2), the study described in this chapter investigated the possibility that hMI-ER1 interferes with the luciferase activity of pGL3(-68) stimulated by Sp1. Overexpression of hMI-ER1 was shown to repress the luciferase activity of pGL3(-68) stimulated by Sp1 (Figure 4.5). *In vitro* and *in vivo* protein-protein interaction assays revealed that hMI-ER1 indeed interacts with Sp1 (Figure 4.7). Since the cell lysates and the TNT lysate may contain other factors that interact with Sp1 or hMI-ER1, it remains unknown as to whether the interaction of hMI-ER1α and β with Sp1 is direct or indirect through other proteins. To investigate whether they interact directly, purified hMI-ER1 and Sp1 protein from bacterial GST recombinant products will be required for pull-down assays, followed by Western blot analysis.

The interaction between hMI-ER1 and Sp1 interfered with GC box recognition by Sp1 to the *hmi-er1* P2 minimal promoter (-68) *in vitro* (band a and b) (Figure 4.9). This

result was further confirmed by EMSAs using the consensus Sp1 binding sequence as a probe (Figure 4.10). Interestingly, in the presence of ^{32}P-labelled *hmi-er1* P2 minimal promoter (-68) as a probe, addition of purified hMI-ER1α or β GST fusion protein to the EMSA mixture resulted in the presence of an extra band c (Figure 4.9 lane 3, 4, and 5), which is less likely due to the binding of hMI-ER1α or β themselves, since: (1) hMI-ER1α or β cannot bind to the *hmi-er1* P2 minimal promoter (-68) (Figure 4.4); (2) band c is on the same position in all the samples with GST-1α, GST-β, or GST-(287-357), which have differing molecular weights. Interestingly, this extra band (band c) did not appear when the consensus Sp1 was used as a probe (Figure 4.10). Thus this extra band may be a specific form of DNA-Sp1 complexes that bind to the *hmi-er1* P2 minimal promoter (-68), but which does not bind to the consensus Sp1 sequence. Further studies such as using anti-Sp1 antibody and/or consensus Sp1 sequence as competitor are required to understand the composition of this extra band.

The present study suggested that the hMI-ER1 interacts with Sp1 and interferes with GC box recognition by Sp1 through the SANT domain-containing region (aa 287-357). These findings support the possibility of a novel mechanism for negative regulation of expression from genes containing Sp1 target promoters, which includes the promoter of *hmi-er1*. However, it is still possible that overexpression of hMI-ER1 may also affect Sp1 concentration in the cells, which in turn affects activity from the hmi-er1 promoter. Further studies to monitor the cellular level of Sp1 are required to support or refute this scenario. Moreover, it is also possible that overexpression of hMI-ER1 may affect Sp1 activity indirectly through other factors. There are multiple factors that have

been reported to regulate Sp1 activities *via* several mechanisms: (1) Many transcription regulators, such as Sp3, Sp4 and BTEB3 compete with Sp1 for the core *cis*-acting elements and repress promoter activity (Kaczynski *et al.*, 2001; Kwon *et al.*, 1999). (2) Sp1 activity was also found to be regulated through protein-protein interactions. Factors that interact with Sp1 include E2F1, GATA1, and YY1, all of which act synergistically with Sp1 on DNA to increase transcriptional activity (Gregory *et al.*, 1996; Karlseder *et al.*, 1996; Lee *et al.*, 1993b; Lin *et al.*, 1996b). (3) Another set of Sp1-interacting transcription factors that impair Sp1-mediated transcriptional activity includes p107, PML, FBI-1, TAF-1 and MDM2 (Datta *et al.*, 1995; Johnson-Pais *et al.*, 2001; Lee *et al.*, 2002; Suzuki *et al.*, 2003; Vallian *et al.*, 1998). (4) Furthermore, Sp1 and the transcriptional coactivator CBP/p300 have been shown to be associated in complexes and to upregulate the promoter activity of p21Waf/Cip1 (Billon *et al.*, 1999; Owen *et al.*, 1998). (5) HDAC1-containing complexes can be recruited directly by Sp1 for transcriptional repression (Doetzlhofer *et al.*, 1999). (6) While RB has not been shown to bind Sp1 directly, it can increase Sp1 activity by releasing Sp1 from MDM2-Sp1 complexes (Johnson-Pais *et al.*, 2001). Thus several different kinds of mechanisms of regulating Sp1 transcriptional activity appear to exist, and the change of the concentration and/or activity of those proteins may also affect the Sp1-targeted promoter activity.

The SANT domain is present exclusively in proteins that have an important role in the regulation of transcription (Aasland *et al.*, 1996). This domain is related in primary and secondary structure to the DNA binding motif of the MYB oncoprotein and other MYB-like domains (Aasland *et al.*, 1996). A subset of SANT-containing proteins

possesses two or three related repeats of this domain, each of which serves distinct functions. For example, the nuclear hormone co-repressor, N-CoR, contains two SANT domains. Its N-terminal SANT1 domain serves as a deacetylase activation domain (DAD) that binds and activates HDAC3 while the C-terminal SANT2 domain functions as a histone interaction domain (HID), preferentially targeting non-acetylated histones and inhibiting HAT activity (Yu *et al.*, 2003). Thus the two domains function synergistically to repress specific regions of chromatin by targeting HDACs and inhibiting HAT activity.

The SANT domains of single SANT proteins, like ADA2, are also involved in protein-protein interactions. Interestingly, the single SANT domain of ADA2 has been shown to have two different functions in yeast. The N-terminal portion of the ADA2 SANT domain is critical for efficient acetylation of histone targets by GCN5, while the C-terminal portion of ADA2 SANT domain functions as an interaction domain for GCN5 (Sterner *et al.*, 2002). In this chapter, I demonstrated that the SANT domain-containing region (aa 287-357) of hMI-ER1 interacted with Sp1, and the C-terminal deletion from ^{331}K to ^{331}T eliminated both hMI-ER1 interactions with Sp1 (Figure 4.8) and interfered with the GC box recognition by Sp1 (Figure 4.10). These results suggest that the SANT domain-containing region spanning residues aa 287-357 is an important interaction region for hMI-ER1 to interact with Sp1 and interfere with the GC box recognition by Sp1. However, whether an intact SANT domain or the region spanning between the SANT domain and its C-terminal 26 aa is critical for the interaction with Sp1 remains

unclear. Site-directed mutagenesis and/or more deletion mutagenesis to this SANT domain-containing region (aa 287-357) could be employed to address this question.

The transcription repression mechanism mediated by the SANT domain of hMI-ER1 is distinct from transcription repression mediated by the ELM2 domain-containing region (aa 164-283) which recruits HDAC1 activity to deacetylate the chromatin (Chapter 3). Results suggested that hMI-ER1 can function as a transcription repressor via both HDAC-dependent and -independent mechanisms through ELM2 and SANT domains, respectively. Both HDAC-dependent and -independent transcriptional repression mechanisms have been reported for other transcription regulators, such as RB (Lai *et al.*, 1999), Stra13 (Sun and Taneja, 2000), LCoR (Fernandes *et al.*, 2003) and MTA1 (Yan *et al.*, 2003).

In summary, hMI-ER1 interferes with DNA recognition by Sp1 on its own promoter and represses its own gene promoter activity, thus forming a negative feedback loop to tightly control the activity of hMI-ER1. This ensures that hMI-ER1 regulates the expression of other genes at appropriate times.

CHAPTER 5 A YEAST TWO-HYBRID SYSTEM cDNA LIBRARY SCREENING FOR HMI-ER1-INTERACTING PROTEINS: HSP40 AND TRABID WERE FOUND TO INTERACT PHYSICALLY WITH HMI-ER1

5.1 Introduction

Proteins rarely function by themselves. They normally interact with other proteins to execute their functions. Networks of such protein-protein interactions constitute the basis for regulation of all cellular functions, including cell growth and cell death (reviewed in Figeys, 2002; Tucker *et al.*, 2001). Thus deciphering of protein interaction networks is vital to our understanding of cell growth and cell death as an interacting system of molecules. Understanding these protein interaction networks and their roles in diseases is critical for successful therapeutic strategies (reviewed in Archakov *et al.*, 2003).

Our previous studies using immunoprecipitation, GST-pull down assays, and Western blot analysis have identified multiple interacting partners for hMI-ER1, based mainly on the observation that hMI-ER1 functions as a transcription regulator (Chapter 3 and 4). These hMI-ER1-interacting proteins include HDAC1 (Chapter 3), Sp1 (Chapter 4), ERα (Savicky *et al.*, unpublished data), CBP (Blackmore *et al.*, unpublished data), and histone protein (Paterno *et al.*, unpublished data).

To obtain a full picture of hMI-ER1 function, it is also necessary to perform a complementary approach in protein-protein interaction analysis to obtain more unbiased representation of hMI-ER1 interaction proteins. Such an analysis would expand our

knowledge on the interacting partners of hMI-ER1. Pioneering work toward this goal in this study was undertaken using the yeast two-hybrid system cDNA library screening for hMI-ER1-interacting proteins. Because of the nature of the reporter system (see section 5.2.3), positive signals are amplified many-fold, thus making the yeast two-hybrid system a very sensitive method for detecting weak and transient protein interactions that are not detected by other methods. Since it is performed in yeast, the proteins produced in this eukaryotic system are more likely folded properly than those produced in bacterial systems and the *in vitro* reticulocyte TNT lysate system. Thus, the yeast two-hybrid cDNA library screening system has advantages over other cDNA library screening approaches for studying protein-protein interaction.

5.2 Materials and Methods

5.2.1 Plasmids and constructs

Expression vectors for the yeast two-hybrid system were engineered to contain full-length or deletion mutants of hMI-ER1α fused to the GAL4 DBD of the pAS2-1 plasmid (Clontech). Fragments encoding the appropriate amino acid residues of hMI-ER1α deletion mutants were amplified by PCR using the primer pairs listed in Table 3.1 (Chapter 3). PCR products were cloned into pCR3.1 using the TA cloning kit (Invitrogen, Inc.), and *EcoR*I fragments were then inserted into the *EcoR*I sites of the pAS2-1 plasmid. The deletion constructs were named according to the encoded amino acid residues of the hMI-ER1α protein. All plasmids were sequenced to verify the

junctions and the hMI-ER1 sequence. The human testis MATCHMAKER cDNA pACT2 AD Library of plasmids inserted into the yeast two-hybrid system transcriptional activation domain (AD) vector pACT2 was purchased from Clontech. pCMV-Tag2C-HSP40(43-340) was generated by subcloning the *Bgl*II fragment from the positive clone No. 3764 obtained from the yeast two-hybrid system cDNA library screening (see section 6.3.2) to the *BamH*I site of pCMV-Tag2C vector (Stratagene). pCMV-Tag2C-TRABID was generated by subcloning the *Bgl*II fragment from the positive clone No. 3801 obtained from yeast two-hybrid system cDNA library screening (see section 6.3.3) to the *BamH*I site of pCMV-Tag2C vector. Full-length human *HSP40* cDNA was amplified by PCR from a human keratinocyte MATCHMAKER cDNA Library (Clontech) by using 5 -GGG CCG CAG GAG GGG GTC ATG-3 and 5 -CTA TAT TGG AAG AAC CTG CTC-3 as forward and reverse primers, respectively, and then cloned into pCR3.1 using the TA cloning kit (Invitrogen, Inc.). All plasmids were sequenced to verify the junctions and the TRABID or HSP40 sequence.

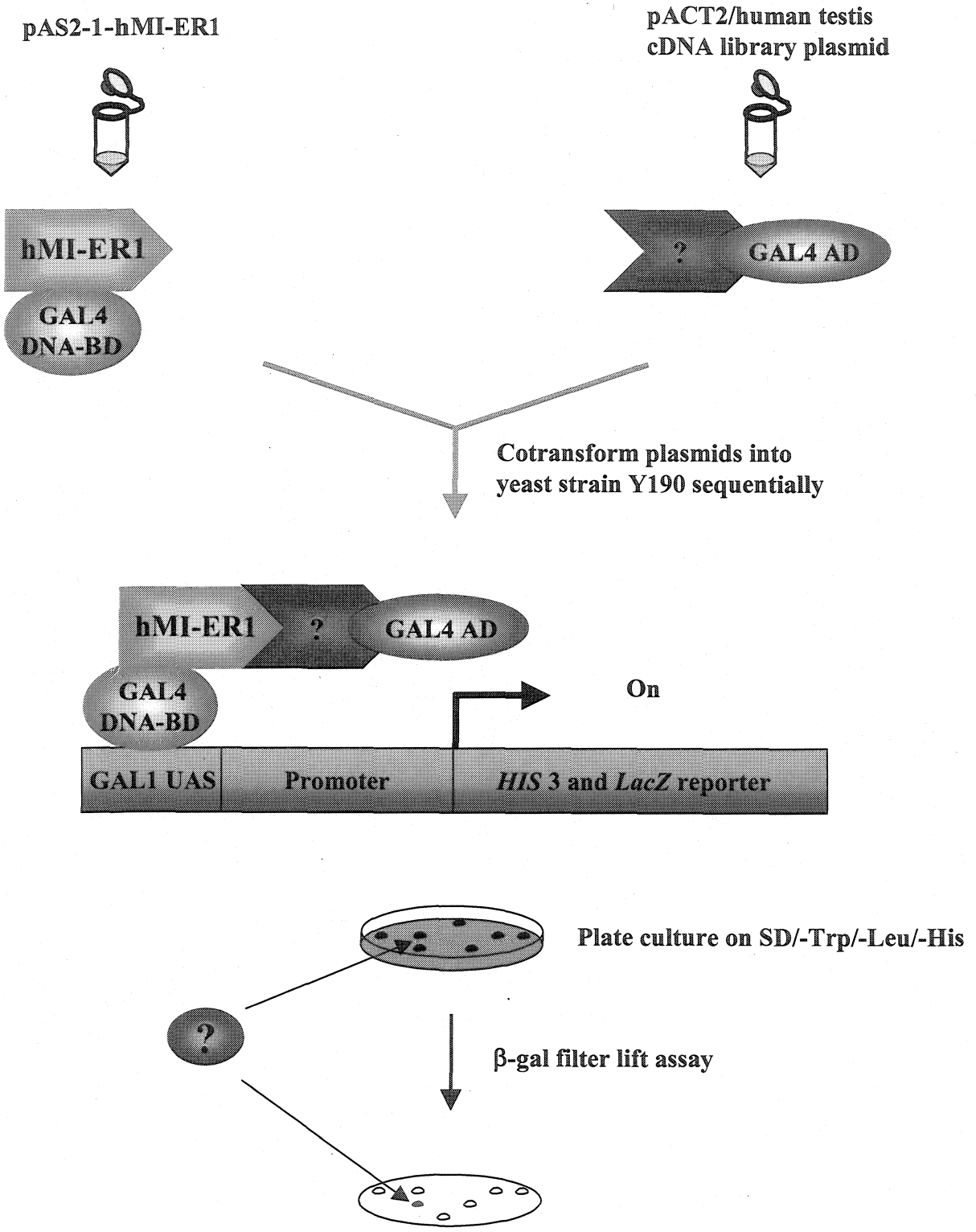
5.2.3 Yeast two-hybrid system cDNA library screening

The yeast two-hybrid system (Fields and Song, 1989; Fields and Sternglanz, 1994) was used as a genetic system to isolate hMI-ER1-interacting proteins *in vivo*. It uses the restoration of transcriptional activation to assay the interaction between two proteins. It relies on the modular nature of many site-specific transcriptional activators such as GAL4, consisting of a DNA-binding domain and a transcriptional activation domain (AD). The yeast two-hybrid system is based on the observation that the two domains of

the activator need not be a single polypeptide and can be brought together by any two interacting fusion proteins, one of which contains the GAL4 DBD (in pAS2-1 vector) while the other has the GAL4 AD (in pACT2 vector) (Figure 5.1). The two hybrid proteins, often coined as "bait" and "prey," respectively, are assembled onto GAL4 binding sites [GAL1 upstream activating sequence (UAS)] in the yeast genome. The assembly functionally reconstitutes the GAL4 transcription factor and induces the expression of yeast *HIS*3 and the bacterial *lac*Z reporter genes integrated in the region downstream of the GAL4 binding sites (Figure 5.1). The *S. cerivisiae* strain Y190, which is Trp^-, Leu^- and His^-, was used for the yeast two-hybrid system (Clontech). This strain harbours the yeast *HIS*3 and the bacterial *lac*Z reporter genes, which contain an upstream GAL4 binding site. For cDNA library screening, Y190 was co-transformed with pAS2-1-(287-433) and the human testis MATCHMAKER cDNA pACT2 AD Library, using the lithium acetate procedure as described by the manufacturer (Clontech). The transformation mixture was then plated on 150-mm petri dishes containing synthetic dropout yeast culture medium (SD) lacking tryptophan, leucine, and histidine (SD/-Trp/-Leu/-His) but including 25 mM 3-amino-1,2,4-triazole (3-AT), and incubated at 30 °C for 5-8 days. The transformants were screened for β-gal activity using a filter lift assay, according to the protocol recommended by the manufacturer (Clontech). Briefly, colonies were transferred to Whatman filters and cells were permeabilized by freezing for 10 seconds in liquid nitrogen, and thawing at room temperature. Filters were then overlaid onto another Whatman filter saturated with Z buffer/X-gal solution [16.1 mg/ml $Na_2HPO_4.7H_2O$, 5.5 mg/ml $NaH_2PO_4.H_2O$, 0.75 mg/ml KCl, 0.246 mg/ml $MgCl_2.7H_2O$,

Figure 5.1 Mechanism of the method of the yeast two-hybrid system cDNA library screening for hMI-ER1-interacting proteins.

The application of this system requires that two hybrid fusion plasmids be constructed for expressing DBD and AD fusion proteins: a DBD fused to the bait protein (in this case DBD-hMI-ER1 from plasmid pAS2-1-hMI-ER1), and an AD fused to the prey proteins which may interact with the bait protein (in this case the human testis MATCHMAKER cDNA pACT2 AD Library proteins). For cDNA library screening, the two hybrid plasmids are co-transformed into a yeast *S. cervisiae* host strain Y190 harbouring the yeast *HIS*3 and the bacterial *lac*Z reporter genes, which downstream the GAL4 binding site (GAL1 UAS). The interaction of hMI-ER1 with a novel library protein (marked as ?) will activate the *HIS*3 and the *lacZ* reporter genes, while library proteins not interacting with hMI-ER1 will not activate the reporter genes. The transformation mixture was then plated on 150-mm petri dishes containing synthetic dropout yeast culture medium (SD) lacking tryptophan, leucine, and histidine (SD/-Trp/-Leu/-His) but including 25 mM 3-amino-1,2,4-triazole (3-AT), and incubated at 30 ^{o}C for 5-8 days. The transformants were screened for β-galatosidase (β-gal) activity using a filter lift assay.


pAS2-1-hMI-ER1
pACT2/human testis cDNA library plasmid
hMI-ER1
GAL4 DNA-BD
?
GAL4 AD
Cotransform plasmids into yeast strain Y190 sequentially
hMI-ER1
?
GAL4 AD
GAL4 DNA-BD
On
GAL1 UAS
Promoter
HIS 3 and LacZ reporter
Plate culture on SD/-Trp/-Leu/-His
?
β-gal filter lift assay

0.327 mg/ml X-gal, 0.3% (v/v) 2-mercaptoethanol], and incubated at room temperature 3 h for color development.

5.2.4 Purification and confirmation of the yeast two-hybrid system positive colonies

Each of the initial yeast $LacZ^+$ positive colonies was streaked out one to five times to segregate multiple pACT2-library plasmids within each single colony and yeast two-hybrid β-gal filter lift assays were repeated on well-isolated colonies. The plasmids were then isolated from yeast, transfected into *E. coli*, and further amplified. These pACT2-library plasmids isolated from *E. coli* were then individually re-transformed into yeast strain Y190 to test the specific interaction of the candidate library clones with the bait pAS2-1-(287-433), non-related protein pAS2-1-LAM5, and empty pAS2-1 vector. Nonspecific interactions (those conferring His^+, $LacZ^+$ when paired with pAS2-1 and pAS2-1-LAM5) were considered false positives and eliminated, while clones specifically interacting with pAS2-1-hMI-ER1 bait fusion were retained. The cDNA obtained were sequenced and analyzed for sequence homology using the National Center for Biotechnology Information sequence databases through the Basic Alignment Search Tool (BLAST) program at http://www.ncbi.nlm.nih.gov/BLAST/.

5.2.5 GST-fusion protein production

GST fusion proteins were expressed as described in section 2.2.6.

5.2.6 GST pull-down

GST pull-down assays were performed as described in section 4.2.6.

5.3 Results

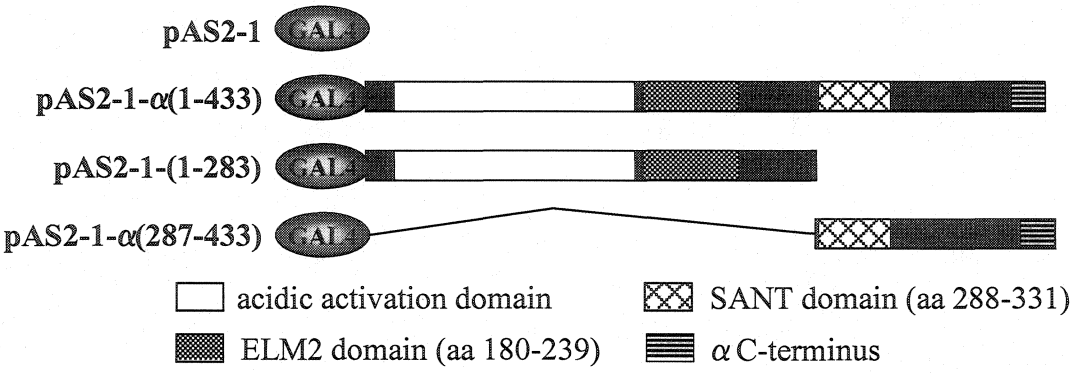
5.3.1 Yeast two-hybrid system cDNA library screening for hMI-ER1-interacting positive colonies

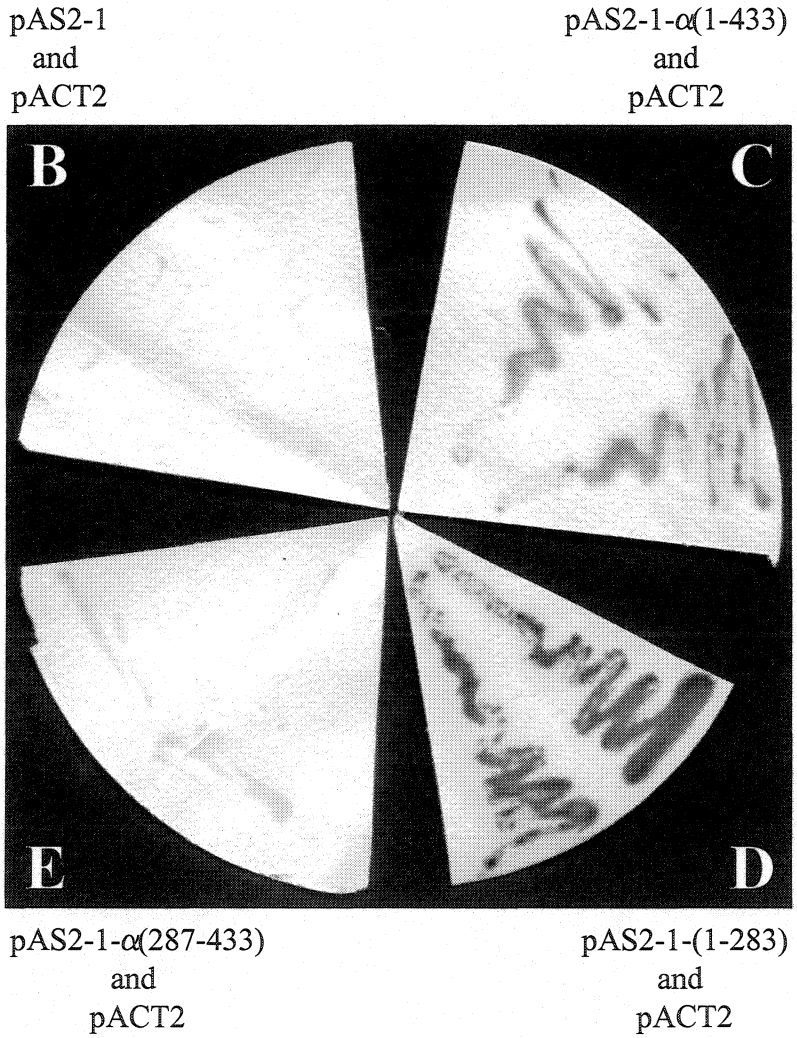
The yeast two-hybrid β-gal filter lift assays of the transformants with pAS2-1-hMI-ER1α and pACT2 empty vector revealed expression of the *LacZ* reporter gene (Figure 5.2C), while no expression of the *LacZ* reporter gene was detected in the transformants with pAS2-1 vector and pACT2 vector (Figure 5.2B). The N-terminal (1-283 aa) and C-terminal (287-433 aa) deletion mutants of hMI-ER1α were generated and fused to DBD in pAS2-1 vector. β-gal filter lift assays of the yeast transformants with pAS2-1-hMI-ER1(1-283) and empty vector pACT2 revealed the expression of the *LacZ* reporter gene (Figure 5.2D), while pAS2-1-hMI-ER1α(287-433) and AD empty vector did not activate the reporter gene (Figure 5.2E). These results indicate that the N-terminal (1-283 aa) region of hMI-ER1 may interact with the GAL4 AD alone and/or recruit the transcriptional machinery of yeast cells to activate the transcription of the reporter gene. These results suggest that full-length hMI-ER1α and the N-terminal (1-283 aa) region of hMI-ER1 can not be used as bait in the yeast two-hybrid screening. Thus pAS2-1-hMI-ER1α(287-433) was used for the yeast two-hybrid cDNA library screening. The yeast was transformed with pAS2-1-hMI-ER1α(287-433) and the human testis

Figure 5.2 Yeast two-hybrid β-gal filter-lift assays of tranformants of empty pACT2 AD vector with full-length and deletion mutants of hMI-ER1α.

(A) The diagram illustrates the deletion mutants of hMI-ER1 fused to GAL4 DBD in pAS2-1 vector. The individual domains are identified in the legend below the diagram, and the hMI-ER1 amino acid residues encoded by each construct are listed. Yeast two-hybrid β-gal filter-lift assays of the transformant with pAS2-1 DBD empty vector and pACT2 AD empty vector **(B)**; hMI-ERα full-length (1-433 aa) in pAS2-1 DBD vector and pACT2 AD empty vector **(C)**; hMI-ER1 N-terminal 1-283 aa in pAS2-1 DBD vector and pACT2 AD empty vector **(D)**; or hMI-ER1 C-terminal 287-433 aa in pAS2-1 DBD vector and pACT2 AD empty vector **(E)**.

A


pAS2-1
GAL4
pAS2-1-α(1-433)
GAL4
pAS2-1-(1-283)
GAL4
pAS2-1-α(287-433)
GAL4
acidic activation domain
SANT domain (aa 288-331)
ELM2 domain (aa 180-239)
α C-terminus


pAS2-1
and
pACT2
pAS2-1-α(1-433)
and
pACT2
B
C
E
D
pAS2-1-α(287-433)
and
pACT2
pAS2-1-(1-283)
and
pACT2

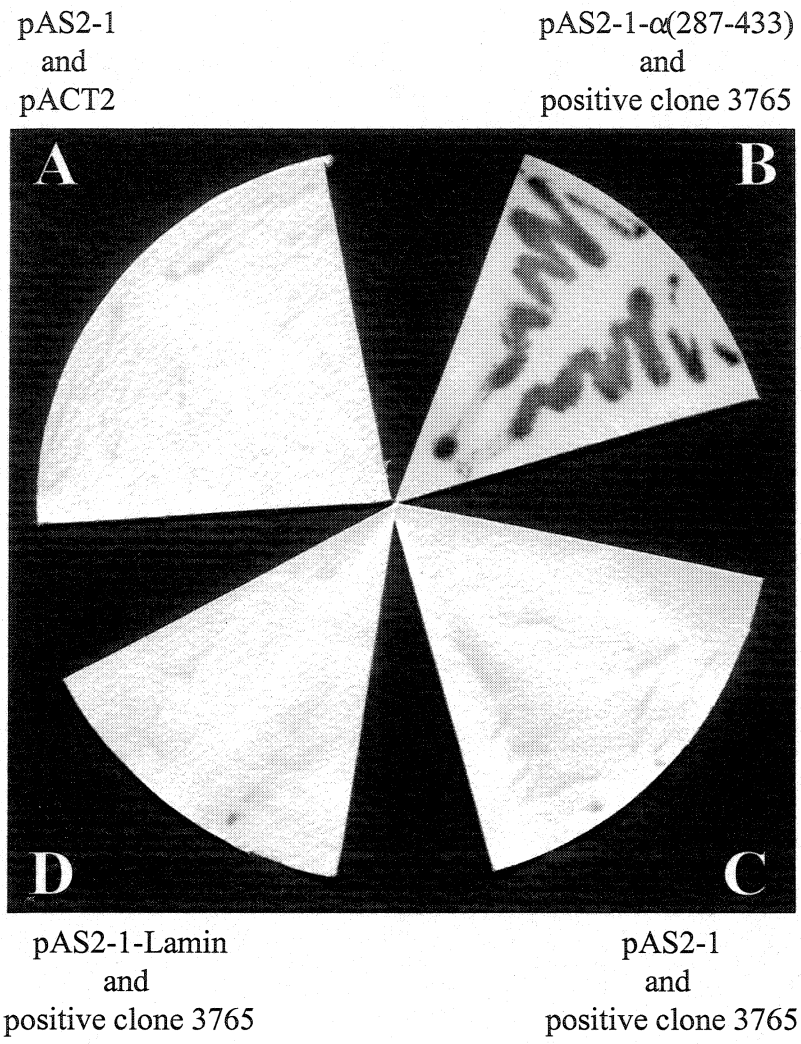
MATCHMAKER cDNA pACT2 AD Library. Approximately 6.6×10^6 yeast transformants were screened for pAS2-1-hMI-ER1α(287-433)-interacting proteins. $Trp^+/Leu^+/His^+$ colonies were then screened for β-gal activity using β-gal filter lift assays, and 108 colonies were $LacZ^+$. These positive colonies were then subjected to further verification by isolating the pACT2 plasmids from these clones and then individually reintroducing them into yeast strain Y190 to test for interaction of the candidate library clone against pAS2-1-hMI-ER1α(287-433), empty pAS2-1 vector, or non-related protein pAS2-1-Lamin. Two positive clones No. 3764 (section 5.3.3) and No. 3801 (section 5.3.4) were shown to be specific interactions and sequenced.

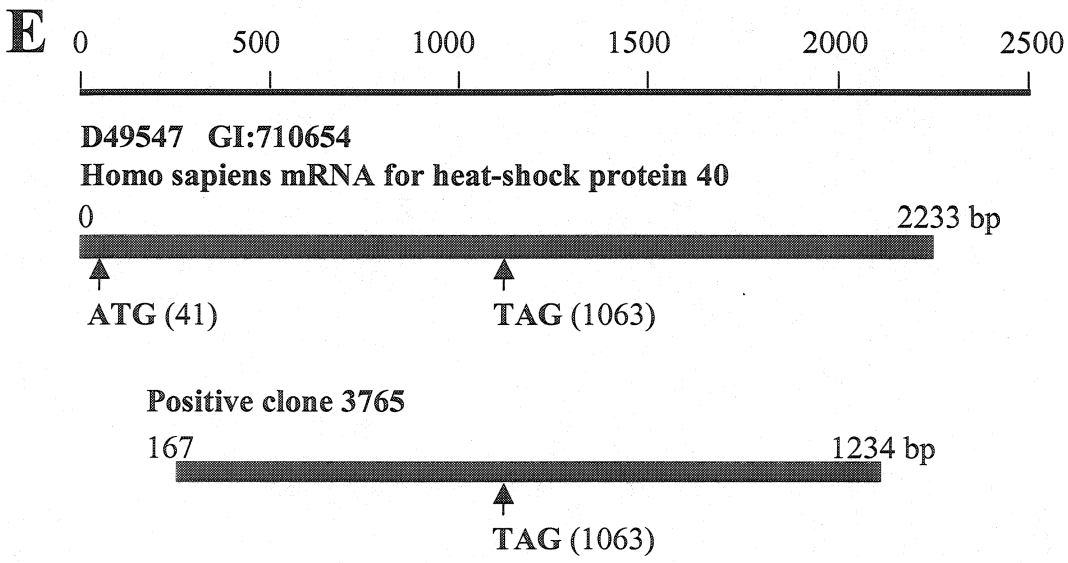
5.3.2 HSP40 was found to interact with hMI-ER1 in yeast and *in vitro*

Yeast two-hybrid β-gal filter lift assays revealed that the positive clone No. 3764 specifically interacts with pAS2-1-hMI-ER1α(287-433) (Figure 5.3B), but not with empty vector (Figure 5.3C), or a non-related protein pAS2-1-Lamin (Figure 5.3D). The cDNA insert of positive clone (No. 3764) was found to contain 1067 bp sequence. Subsequent DNA sequencing analysis followed by BLAST at http://www.ncbi.nlm.nih.gov/BLAST/ revealed that the sequence of the positive clone No. 3764 was identical to the C-terminal of *HSP40* encoding a truncated HSP40(43-340) (Figure 5.3E, lower blue bar).

Figure 5.3 Yeast two-hybrid β-gal filter-lift assays revealed that pAS2-1-hMI-ER1α(287-433) interacts specifically with the positive clone No. 3765 (human *HSP40*).

pAS2-1-hMI-ER1α(287-433) was the bait and pAS2-1 empty vector and pAS2-1-Lamin were negative controls to verify the specificity of the interaction between pAS2-1-hMI-ER1α(287-433) and positive clone No. 3765. Yeast two-hybrid β-gal filter-lift assays of the transformant with pAS2-1 empty vector and pACT2 empty vector **(A)**; pAS2-1-hMI-ER1α(287-433) and positive clone No. 3765 **(B)**; pAS2-1 empty vector and positive clone No. 3765 **(C)**; pAS2-1-Lamin and positive clone No. 3765 **(D)**; **(E)** positive clone No. 3765 (lower blue bar) is identical to previously known sequences of the C-terminal portion of *HSP40* (higher blue bar). Translation start codon ATG and stop codon TAG are indicated below each bar, and corresponding nt is given.


pAS2-1
and
pACT2
pAS2-1-α(287-433)
and
positive clone 3765
A
B
D
C
pAS2-1-Lamin
and
positive clone 3765
pAS2-1
and
positive clone 3765


E
0
500
1000
1500
2000
2500
D49547 GI:710654
Homo sapiens mRNA for heat-shock protein 40
0
2233 bp
ATG (41)
TAG (1063)
Positive clone 3765
167
1234 bp
TAG (1063)

To determine which region of hMI-ER1α interacts with HSP40, four deletion mutants were made in frame with GAL4 DBD (Figure 5.4A). Yeast two-hybrid β-gal filter lift assays revealed that pACT2-HSP40(43-340) did not interact with the C-terminal aa 325-433 deletion mutant of hMI-ER1α (Figure 5.4B). The sequence required for interaction with pACT2-HSP40(43-340) maps to hMI-ER1 aa 287-357, which includes the SANT domain (Figure 5.4C and D). Further deletion of this region at the C-terminal end to remove residues from aa 331 to 357 (^{331}K to ^{357}T), which includes a highly conserved lysine (K) of SANT domain at position 331 and the C-terminal 26 aa, completely abolished interaction with pACT2-HSP40(43-340) in yeast (Figure 5.4E).

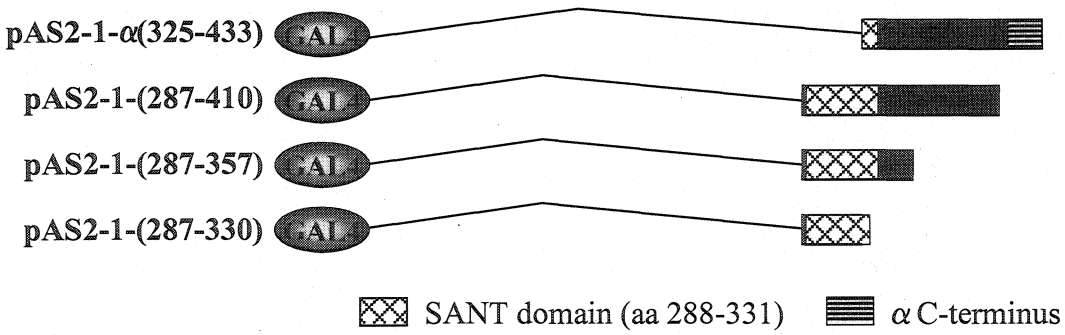
To confirm the interaction between hMI-ER1 with HSP40(43-340) *in vitro*, the *HSP40* cDNA portion isolated from yeast two-hybrid library screening was then subcloned into pCMV-Tag2C vector [pCMV-Tag2C-HSP40(43-340)] (see section 5.2.2). *In vitro* translated HSP40(43-340) specifically associated with GST-α (Figure 5.5B, lane 3) or GST-β (Figure 5.5B, lane 4), but not with GST control (Figure 5.5B, lane 2). To investigate whether the full-length HSP40 interacted with hMI-ER1 *in vitro*, the *HSP40* cDNA encoding full-length HSP40 was amplified from a cDNA library by PCR and cloned into pCR3.1. *In vitro* translated full-length HSP40 (Figure 5.5B, lane 5) specifically interacted with GST-α (Figure 5.5B, lane 7) or GST-β (Figure 5.5B, lane 8), but not with GST control (Figure 5.5B, lane 6).

Further GST pull-down assays with deletion mutants of hMI-ER1 (Figure 5.5A) and full-length HSP40 confirmed that the sequences required for interaction with HSP40 are located in residues spanning aa 287-357, which includes the SANT domain of

Figure 5.4 Yeast two-hybrid β-gal filter-lift assays revealed that a region containing SANT domain of hMI-ER1 interacts specifically with HSP40(43-340).

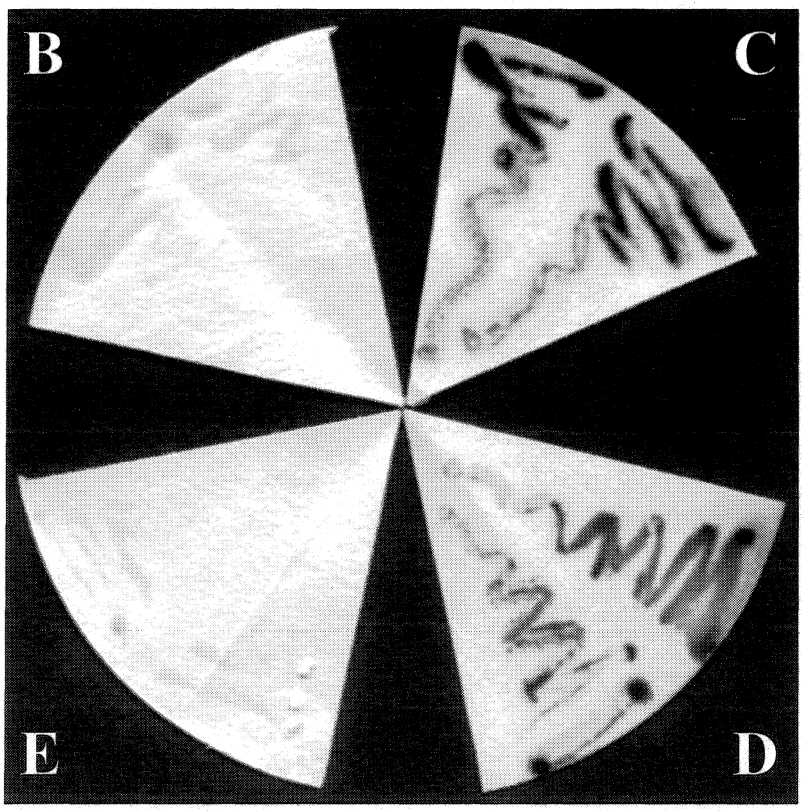
(A) The diagram illustrates the deletion mutants of hMI-ER1 fused to GAL4 DBD in pAS2-1 vector. The individual domains are identified in the legend below the diagram, and the hMI-ER1 amino acid residues encoded by each construct are listed. Yeast two-hybrid β-gal filter-lift assays of the transformant with pAS2-1-hMI-ER1α(325-433) and pACT2-HSP40(43-340) **(B)**; pAS2-1-hMI-ER1(287-410) and pACT2-HSP40 **(C)**; pAS2-1-hMI-ER1(287-357) and pACT2-HSP40(43-340) **(D)**; pAS2-1-hMI-ER1(287-330) and pACT2-HSP40(43-340) **(E)**.

A



pAS2-1-α(325-433)
and
pACT2-HSP40(43-340)

pAS2-1-(287-410)
and
pACT2-HSP40(43-340)

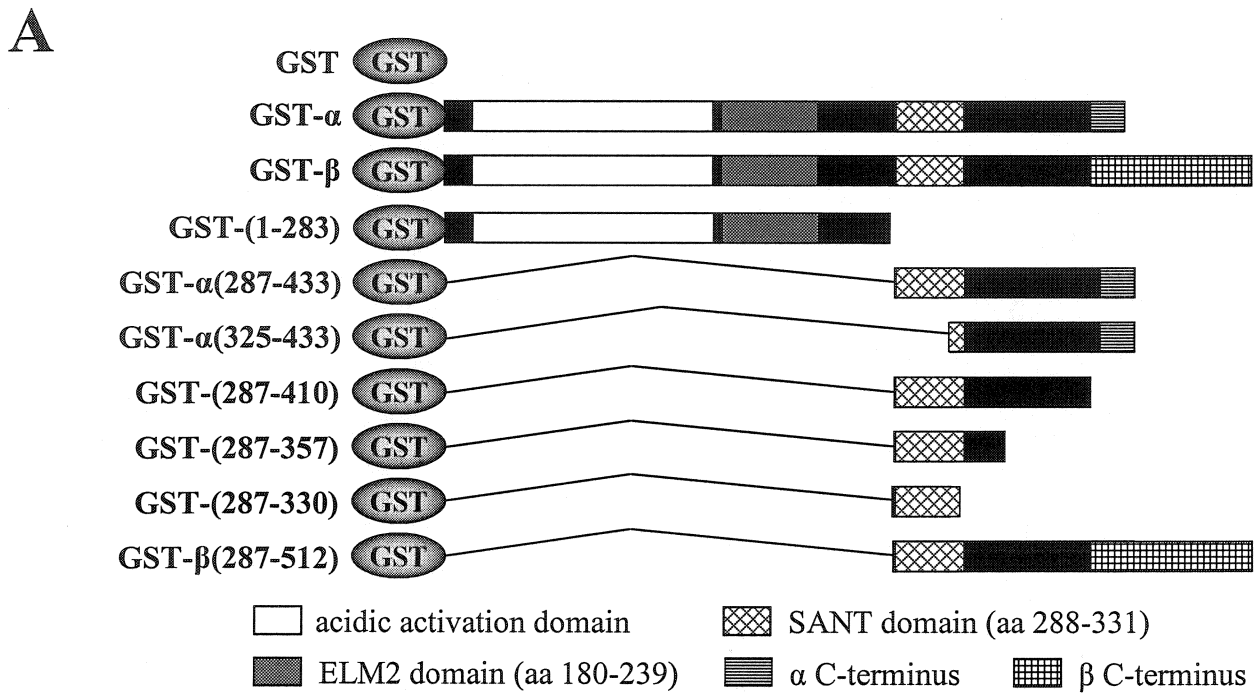


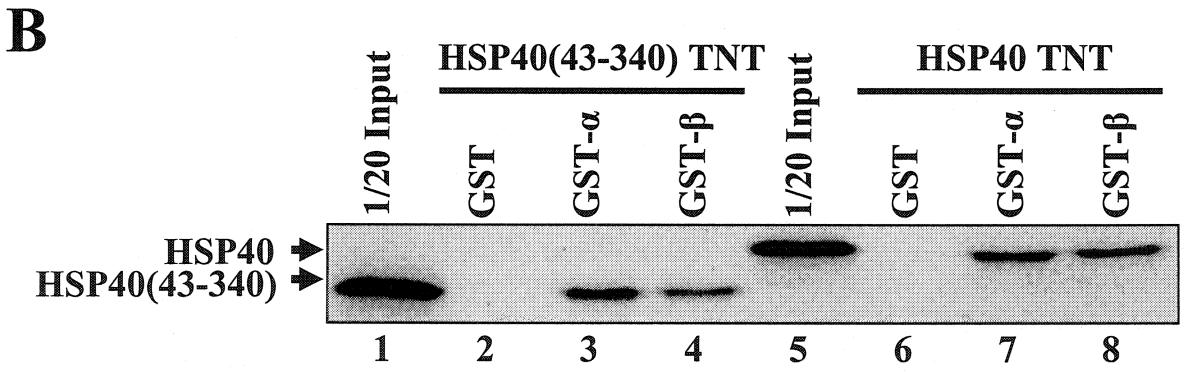
pAS2-1-(287-330)
and
pACT2-HSP40(43-340)

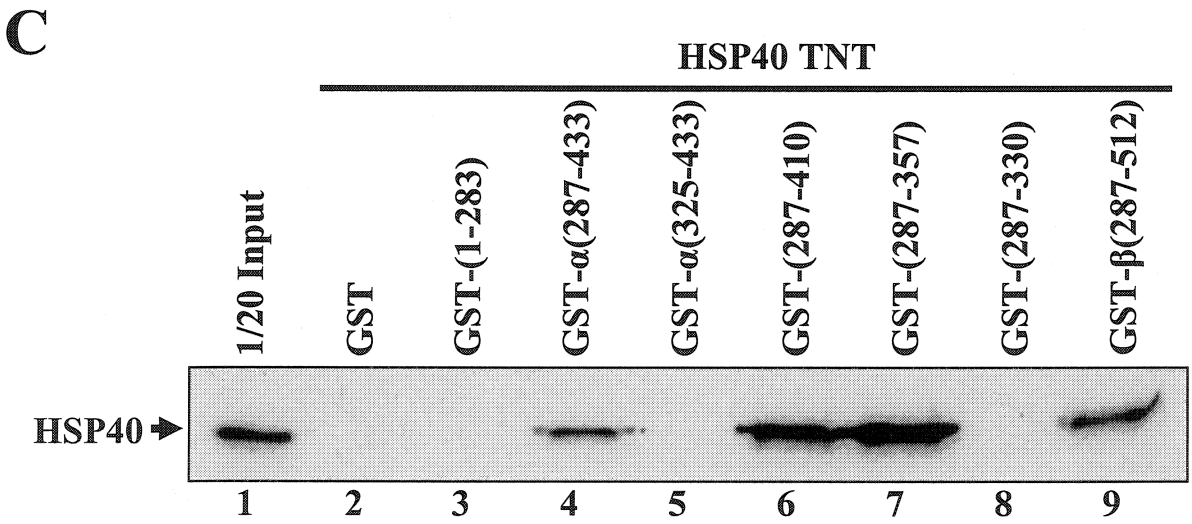
pAS2-1-(287-357)
and
pCT2-HSP40(43-340)

Figure 5.5 The region of hMI-ER1 containing SANT domain interacts with HSP40 *in vitro*.

(A) The diagram illustrates the deletion mutants of hMI-ER1 fused to GST in pGEX-4T-1 vector. The individual domains are identified in the legend below the diagram, and the hMI-ER1 amino acid residues encoded by each construct are listed. **(B)** ^{35}S-labelled TNTs programmed with cDNA encoding HSP40(43-340) was loaded directly on the gel (1/20 of input) (lane 1) or incubated with GST (lane 2), GST-α (lane 3), and GST-β (lance 4). ^{35}S-labelled TNTs programmed with cDNA encoding full-length HSP40 was loaded directly on the gel (lane 5) or incubated with GST (lane 6), GST-α (lane 7), and GST-β (lance 8). Proteins were resolved by SDS-PAGE and visualized by autoradiography. Shown is a representative autoradiograph, and the positions of the HSP40(43-340) and the full-length HSP40 are indicated by arrows. **(C)** ^{35}S-labelled TNTs programmed with cDNA encoding full-length *HSP40* was loaded directly on the gel (1/20 of input) (lane 1) or incubated with GST (lane 2), GST-(1-283) (lane 3), GST-(287-433) (lance 4), GST-(325-433) (lane 5), GST-(287-410) (lane 6), GST-(287-357) (lane 7), GST-(287-330) (lane 8), and GST-(287-512) (lane 9). Proteins were visualized by SDS-PAGE and autoradiography. Shown is a representative autoradiograph, and the positions of the full-length HSP40 are indicated by arrows.


A
GST
GST-α
GST-β
GST-(1-283)
GST-α(287-433)
GST-α(325-433)
GST-(287-410)
GST-(287-357)
GST-(287-330)
GST-β(287-512)
acidic activation domain
SANT domain (aa 288-331)
ELM2 domain (aa 180-239)
α C-terminus
β C-terminus


B
HSP40(43-340) TNT
HSP40 TNT
1/20 Input
GST
GST-α
GST-β
HSP40
HSP40(43-340)
1 2 3 4 5 6 7 8


C
HSP40 TNT
1/20 Input
GST
GST-(1-283)
GST-α(287-433)
GST-α(325-433)
GST-(287-410)
GST-(287-357)
GST-(287-330)
GST-β(287-512)
HSP40
1 2 3 4 5 6 7 8 9

hMI-ER1 (Figure 5.5C, lane 4, 6, 7 and 9). Further deletion of this region at the C-terminal end to remove residues from aa 331 to 357 (^{331}K to ^{357}T), which includes a highly conserved lysine (K) of SANT domain at position 331 and the C-terminal 26 aa, completely abolished interaction with HSP40 *in vitro* (Figure 5.5C, lane 8). HSP40 did not interact with the GST control (Figure 5.5C, lane 6), or the N-terminal aa 1-283 deletion mutant of hMI-ER1 (Figure 5.5C, lane 3), or the C-terminal aa 325-433 deletion mutant of hMI-ER1 (Figure 5.5C, lane 5). These results are consistent with the results from the yeast two-hybrid system analysis, and demonstrated the importance of the SANT domain-containing region from aa 287-357 for the interaction with HSP40 both in the yeast two-hybrid system and in *in vitro* GST-pull-down assays.

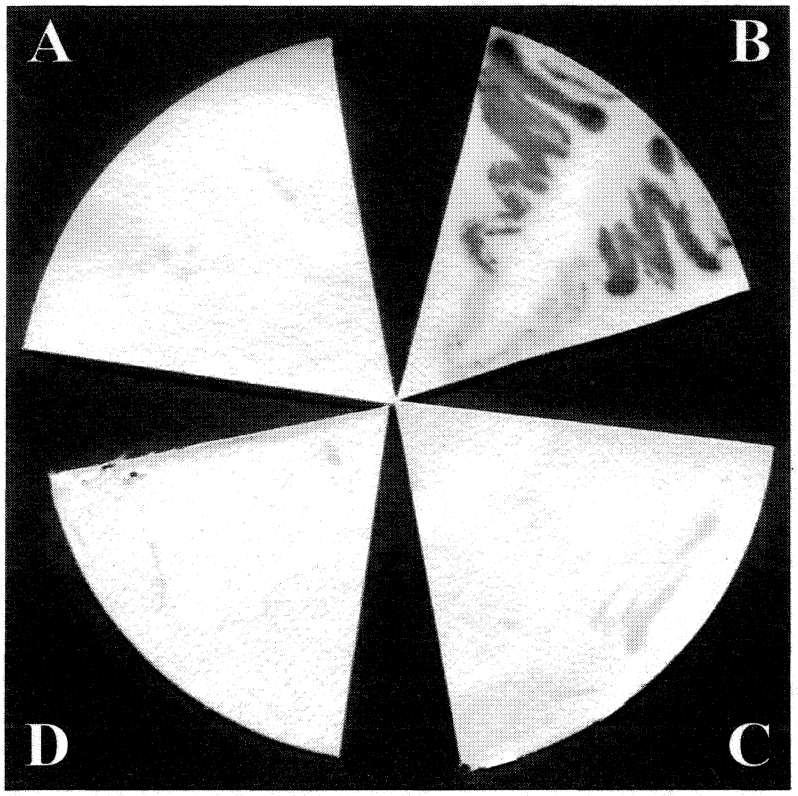
5.3.3 TRAF-binding domain (TRABID) was found to interact with hMI-ER1 in yeast and *in vitro*

Yeast two-hybrid β-gal filter lift assays revealed that another positive clone No. 3801 specifically interacts with pAS2-1-hMI-ERIα(287-433) (Figure 5.6B), but not with empty vector (Figure 5.6C), or a non-related protein pAS2-1-Lamin (Figure 5.6D). The cDNA insert of positive clone No. 3801 was found to contain 2173 bp sequence. Subsequent DNA sequencing analysis followed by BLAST at http://www.ncbi.nlm.nih.gov/BLAST/ revealed that the sequence of the positive clone No. 3801 was identical to *TRABID* mRNA (Figure 5.6E, lower blue bar). The function of TRABID is so far unclear. TRABID shares sequence and structural similarity with A20

Figure 5.6 Yeast two-hybrid β-gal filter-lift assays verified that pAS2-1-hMI-ER1α(287-433) interacts specifically with positive clone No. 3801 (human *TRABID*). pAS2-1-hMI-ER1α(287-433) was the bait and pAS2-1 empty vector and pAS2-1-Lamin were negative controls to verify the specificity of the interaction between pAS2-1-hMI-ER1α(287-433) and positive clone No. 3801. Yeast two-hybrid β-gal filter-lift assays of the transformant with pAS2-1 empty vector and pACT2 empty vector **(A)**; pAS2-1-hMI-ER1α(287-433) and positive clone No. 3801 **(B)**; pAS2-1 empty vector and positive clone No. 3801 **(C)**; pAS2-1-Lamin and positive clone No. 3801 **(D)**; **(E)** positive clone No. 3801 (lower blue bar) is identical to previously known sequences of *TRABID* (higher blue bar). Translation start codon ATG and stop codon TAG are indicated below each bar, and corresponding nt is given.

pAS2-1
and
pACT2

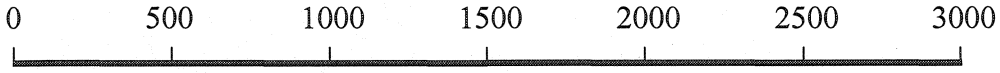
pAS2-1-α(287-433)
and
positive clone 3801



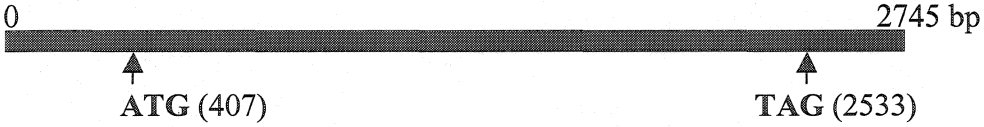
pAS2-1-Lamin
and
positive clone 3801

pAS2-1
and
positive clone 3801

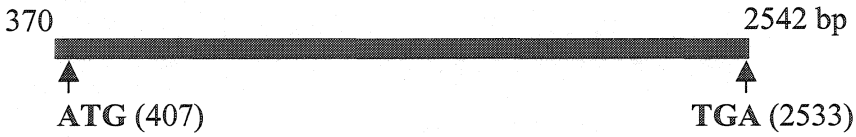
E



AJ252060 GI:6634458
Homo sapiens mRNA for TRAF-binding protein domain (TRABID)



Positive clone 3801



protein (Evans *et al.*, 2001), which was shown to be involved in TNF receptor associated factor (TRAF) signalling pathways (Heyninck and Beyaert, 1999; Lee *et al.*, 2000; Song *et al.*, 1996). C-immunoprecipitation studies indicated TRABID was able to interact with TRAF6 (Evans *et al.*, 2001).

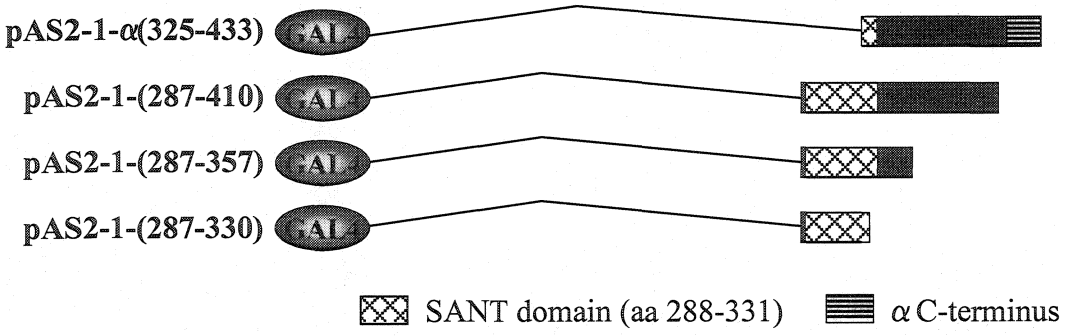
To determine which region of hMI-ER1α interacts with TRABID, four deletion mutants, made in frame with GAL4 DBD in the pAS2-1 vector, were utilized (Figure 5.7A). Yeast two-hybrid β-gal filter lift assays revealed that pACT2-TRABID did not interact with the C-terminal aa 325-433 deletion mutant of hMI-ER1α (Figure 5.7B). The sequence required for interaction with pACT2-TRABID maps to hMI-ER1 aa 287-357, which includes the SANT domain (Figure 5.7C and D). Further deletion of this region at the C-terminal end to remove residues from aa 331 to 357 (^{331}K to ^{357}T), which includes a highly conserved lysine (K) of SANT domain at position 331 and the C-terminal 26 aa, completely abolished interaction with pACT2-TRABID in yeast (Figure 5.7E).

To confirm the interaction between hMI-ER1 with TRABID *in vitro*, the *TRABID* cDNA was subcloned into pCMV-Tag2C vector (pCMV-Tag2C-TRABID) (see 5.2.2) which was used to synthesize protein *in vitro*. *In vitro* translated TRABID showed two bands (Figure 5.8B, lane 1), which may result from two alternative in-frame translation start codons in the construct. GST pull-down assays revealed that both *in vitro* translated TRABID products specifically associated with GST-α (Figure 5.8B, lane 3) or GST-β (Figure 5.8B, lane 4), but not with GST control (Figure 5.8B, lane 2), indicating that wild-type hMI- ER1α and β isoforms interact with TRABID.

Figure 5.7 Yeast two-hybrid β-gal filter-lift assays revealed that a region containing SANT domain of hMI-ER1 interacts specifically with TRABID.

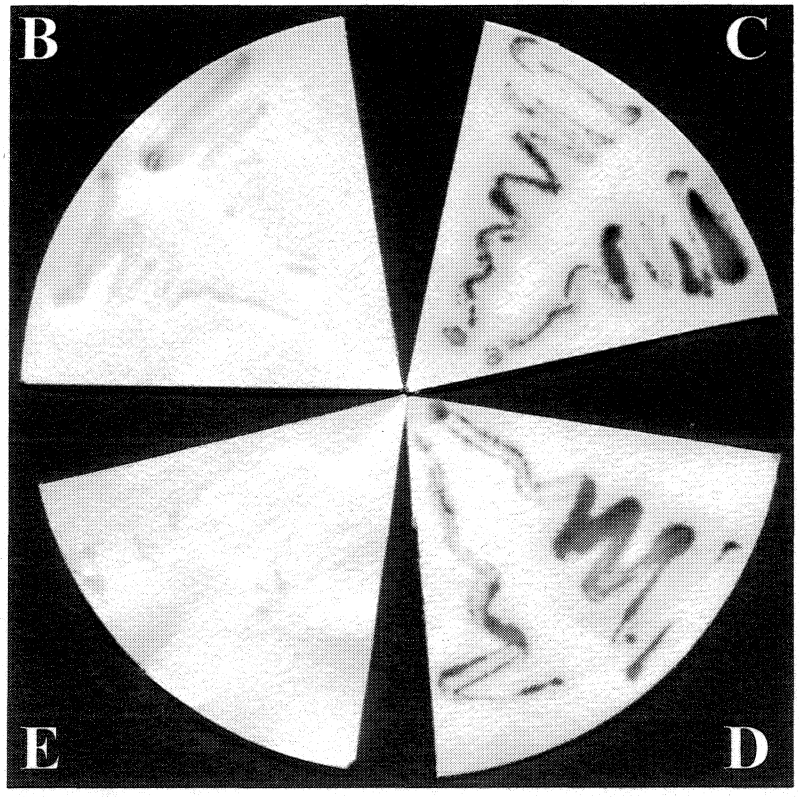
(A) The diagram illustrates the deletion mutants of hMI-ER1 fused to GAL4 DBD in pAS2-1 vector. The individual domains are identified in the legend below the diagram, and the hMI-ER1 aa residues encoded by each construct are listed. Yeast two-hybrid β-gal filter-lift assays of the transformant with pAS2-1-hMI-ER1α(325-433) and pACT2-TRABID **(B)**; pAS2-1-hMI-ER1(287-410) and pACT2-TRABID **(C)**; pAS2-1-hMI-ER1(287-357) and pACT2-TRABID **(D)**; pAS2-1-hMI-ER1(287-330) and pACT2-TRABID **(E)**.

A



pAS2-1-α(325-433)
and
pACT2-TRABID

pAS2-1-(287-410)
and
pACT2-TRABID

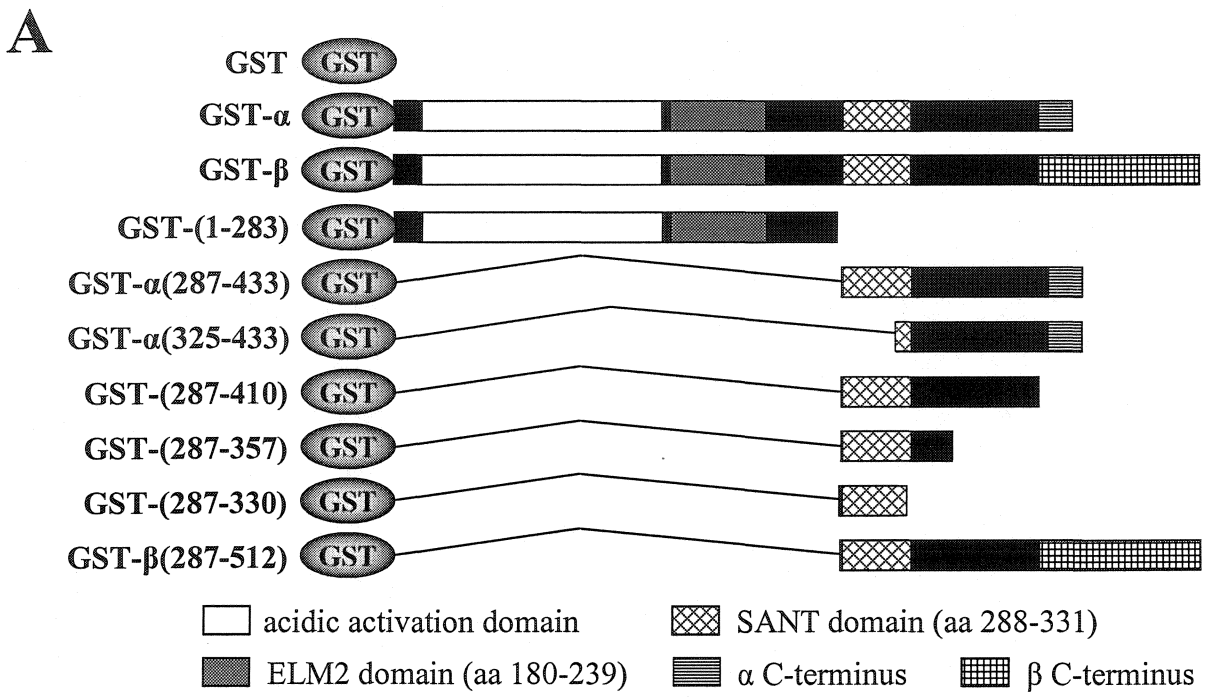


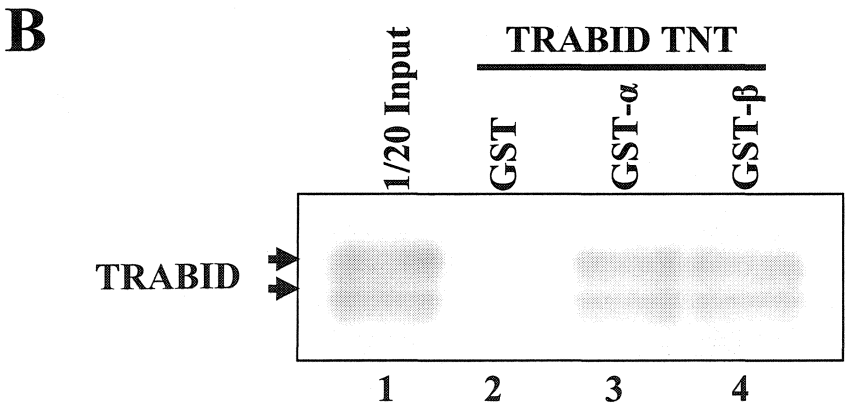
pAS2-1-(287-330)
and
pACT2-TRABID

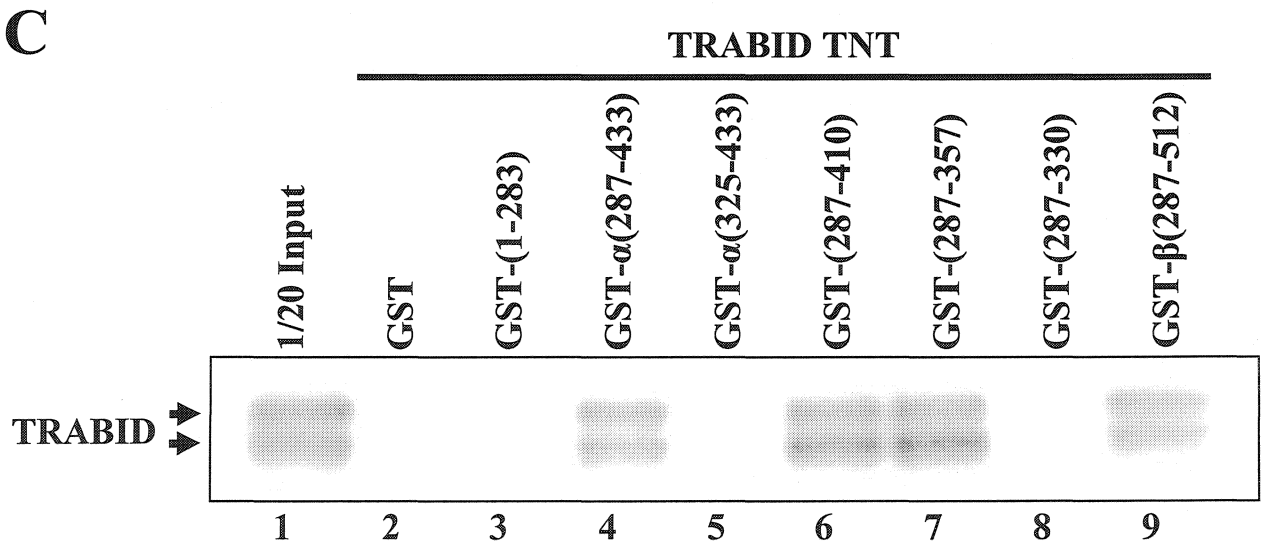
pAS2-1-(287-357)
and
pCT2-TRABID

Figure 5.8 The region of hMI-ER1 containing SANT domain interacts with TRABID *in vitro*.

(A) The diagram illustrates the deletion mutants of hMI-ER1 fused to GST in pGEX-4T-1 vector. The individual domains are identified in the legend below the diagram, and the hMI-ER1 amino acid residues encoded by each construct are listed. **(B)** ^{35}S-labelled TNTs programmed with cDNA encoding TRABID was loaded directly on the gel (1/20 of input) (lane 1) or incubated with GST (lane 2), GST-α(1-433) (lane 3), and GST-β(1-512) (lance 4). Proteins were resolved by SDS-PAGE and visualized by autoradiography. Shown is a representative autoradiograph, and the positions of TRABID is indicated by arrows. **(C)** ^{35}S-labelled TNTs programmed with cDNA encoding TRABID was loaded directly on the gel (1/20 of input) (lane 1) or incubated with GST (lane 2), GST-(1-283) (lane 3), GST-(287-433) (lance 4), GST-(325-433) (lane 5), GST-(287-410) (lane 6), GST-(287-357) (lane 7), GST-(287-330) (lane 8), and GST-(287-512) (lane 9). Proteins were resolved by SDS-PAGE and visualized by autoradiography. Shown is a representative autoradiograph, and the positions of TRABID is indicated by arrows.


A
GST
GST-α
GST-β
GST-(1-283)
GST-α(287-433)
GST-α(325-433)
GST-(287-410)
GST-(287-357)
GST-(287-330)
GST-β(287-512)
acidic activation domain
SANT domain (aa 288-331)
ELM2 domain (aa 180-239)
α C-terminus
β C-terminus


B
TRABID TNT
1/20 Input
GST
GST-α
GST-β
TRABID
1
2
3
4


C
TRABID TNT
1/20 Input
GST
GST-(1-283)
GST-α(287-433)
GST-α(325-433)
GST-(287-410)
GST-(287-357)
GST-(287-330)
GST-β(287-512)
TRABID
1
2
3
4
5
6
7
8
9

Further GST pull-down assays with deletion mutants of hMI-ER1 (Figure 5.8A) and TRABID (Figure 5.8C, lane 1) confirmed that the sequences required for interaction with TRABID were located in aa 287-357 (Figure 5.8C, lane 4, 6, 7 and 9). Further deletion of this region at the C-terminal end to remove residues from aa 331 to 357 (^{331}K to ^{357}T), which includes a highly conserved lysine (K) of the SANT domain at position 331 and the C-terminal 26 aa, completely abolished interaction with TRABID *in vitro* (Figure 5.8C, lane 8). TRABID did not interact with the GST control (Figure 5.8C, lane 6) or the N-terminal aa 1-283 deletion mutant of hMI-ER1 (Figure 5.8C, lane 3), or the C-terminal aa 325-433 deletion mutant of hMI-ER1 (Figure 5.8C, lane 5). These results are consistent with the results from the yeast two-hybrid system analysis, and demonstrated the importance of the SANT domain-containing region from aa 287-357 for interaction with TRABID both in the yeast two-hybrid system and *in vitro*.

5.4 Discussion

From the yeast two-hybrid library screening, hMI-ER1 was found to specifically interact with HSP40 (Figure 5.3) TRABID (Figure 5.10). Although the interactions between hMI-ER1 and HSP40 or TRABID were also confirmed by *in vitro* GST pull-down assays, it is still possible that hMI-ER1 interacts with HSP40 or TRABID indirectly through other proteins. Since there are multiple endogenous proteins in yeast cells and in the reticulocyte TNT lysates, we cannot at present exclude this possibility. Alternative protein-protein interaction assays using both purified proteins and then Western blot analysis may address this question.

In this chapter, I demonstrated that the SANT domain-containing region (aa 287-357) of hMI-ER1 interacted with HSP40 and TRABID, and the C-terminal deletion of the last ^{331}K of the SANT domain together with downstream 26 aa eliminated both hMI-ER1 interactions with HSP40 and TRABID (Figure 5.4, 5.9, 5.11, and 5.12). These results suggest that the SANT domain-containing region spanning residues aa 287-357 is an important interaction region for hMI-ER1 to interact with HSP40 and TRABID. However, whether an intact SANT domain or the region spanning the SANT domain and its C-terminal 26 aa is critical for the interaction with HSP40 and TRABID remains unclear. Site-directed mutagenesis and/or additional deletion mutagenesis to this SANT domain-containing region (aa 287-357) could be employed to address this question.

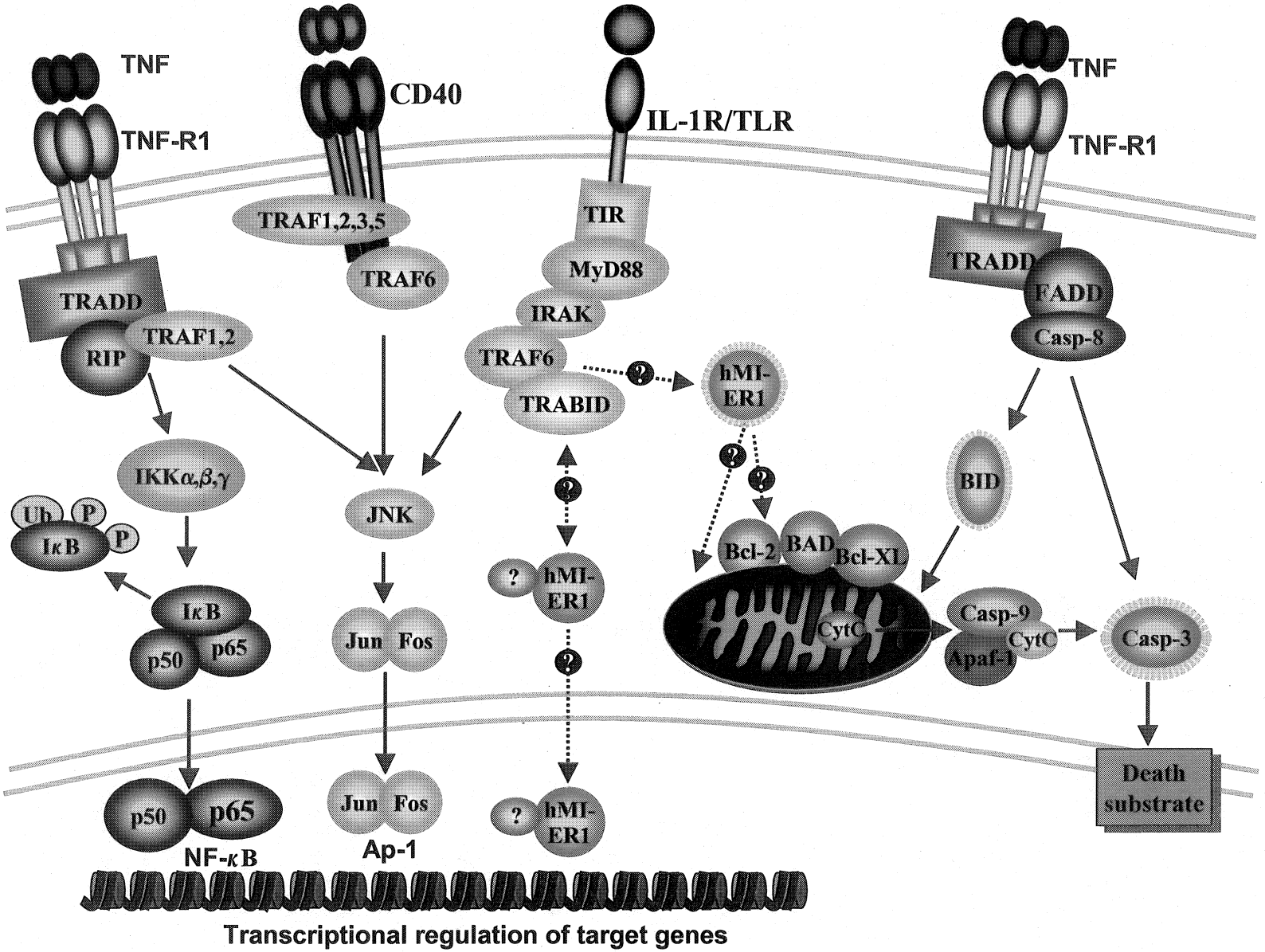
HSP40 plays a key role with other proteins such as HSP70 in the folding, translocation and degradation of proteins in eukaryotic cells (reviewed in Fink, 1999; Hartl, 1996). Another important cellular function of HSP40 is to direct the assembly of multiple protein complexes, such as nuclear hormone receptor (NR) complexes (Hernandez *et al.*, 2002). Current experiments in the laboratory reveal that hMI-ER1α and β interact with estrogen receptor α (ERα) (Savicky *et al.*, personal communication). Given the important role of HSP40 in NR complex assembly, it is possible that HSP40 is critical for the assembly of ERα complexes containing hMI-ER1. Once in the ERα complexes, hMI-ER1α and β may regulate ERα transcriptional function. There is mounting evidence that ERα plays a crucial role in normal breast development and is also linked to development and progression of mammary carcinoma (Hortobagyi, 1998; Osborne, 1998). However, the molecular mechanisms that regulate ERα transcriptional

function are not fully understood. The regulation of the transcriptional activity of ERα by hMI-ER1α and β in breast cancer may have potential clinical significance.

Although the precise function of TRABID remains unknown, co-immunoprecipitation studies indicated TRABID was able to interact with TRAF6 (Evans *et al.*, 2001). Mammalian TRAFs have emerged as the major signal transducers for the TNF receptor superfamily and the interleukin-1 receptor/Toll-like receptor (IL-1R/TLR) superfamily (reviewed in Chung *et al.*, 2002). TRAF molecules are recruited to signalling complexes in response to related extracellular signals such as cytokine treatment (Wallach *et al.*, 1999). As shown in Figure 5.9, there are at least three distinct ways that TRAF proteins can be recruited to and activated by ligand-engaged receptors. Those that contain an intracellular death domain, such as TNFR1, first recruit an adapter protein, TRADD, *via* a death-domain and death-domain interaction (Hsu *et al.*, 1995). TRADD then serves as a central platform of the TNFR1 signalling complex, which assembles Fas-associating protein with death domain (FADD) and caspase-8 for the induction of apoptosis (Hsu *et al.*, 1996), TRAFs2 and RIP (Hsu *et al.*, 1996; Stanger *et al.*, 1995) for survival signalling. Members of the TNF receptor superfamily that do not contain intracellular death domains, such as CD40, recruit TRAFs directly *via* short sequences in their intracellular tails (Cheng *et al.*, 1995; Pullen *et al.*, 1998; Rothe *et al.*, 1994). Members of the IL-1R/TLR superfamily contain a protein interaction module known as the TIR domain (Xu *et al.*, 2000), which sequentially recruits MyD88 and IRAKs (Cao *et al.*, 1996; Muzio *et al.*, 1997; Wesche *et al.*, 1997; Wesche *et al.*, 1999). IRAKs in turn associate with TRAF6 to elicit signalling (reviewed in O'Neill, 2002).

Figure 5.9 TRAF signalling pathways and the possible functional implication of TRABID and hMI-ER1 interaction (adapted from Chung *et al.*, 2002).

Membrane-proximal events and downstream signalling events for TRAFs in TRAF signalling. hMI-ER1α and β may be activated directly when they are recruited into TRABID-TRAF6 complexes. Alternatively, hMI-ER1 may regulate the TRAF6-signalling pathways through association with TRABID.



Transcriptional regulation of target genes

TRAF signalling appears to be mediated mainly through the activation of transcription factors of the NF-κB and AP-1 family. NF-κB is sequestered in the cytoplasm of non-stimulated cells by binding to inhibitory κB (IκB) molecules. Signalling cascades triggered by cytokine stimulation lead to phosphorylation of IκB by the IκB kinase (IKK) (reviewed in Karin and Ben-Neriah, 2000; Stancovski and Baltimore, 1997). Phosphorylation, followed by ubiquitination and proteosomally mediated degradation of IκB occurs and NF-κB is released and translocated to the nucleus to activate transcription (reviewed in Karin and Ben-Neriah, 2000; Stancovski and Baltimore, 1997). Signalling cascades triggered by cytokine stimulation also lead to the activation of AP-1 by mitogen-activated protein (MAP) kinases (Karin, 1996). The stimulation of AP-1 activity by MAP kinases may result in expression of AP-1 target genes involved in both cell survival and cell death (Shaulian and Karin, 2001).

The interaction of hMI-ER1 with TRABID suggested that hMI-ER1 may regulate the TRAF signalling pathways through association with the TRABID-TRAF6 complex (Figure 5.9). Alternatively, it is possible that TRABID-TRAF6 signalling complexes may be important for recruitment and activation of hMI-ER1. hMI-ER1 itself may be activated by post-translational modifications such as phosphorylation when it is recruited into TRABID-TRAF6 complexes. Although computer-assisted analysis predicted multiple phosphorylation sites in hMI-ER1, preliminary studies indicate that hMI-ER1 may not be phosphorylated in cells under normal growth conditions (Paterno *et al.*, personal communication). However, it remains possible that hMI-ER1 would be

phosphorylated under a certain cellular condition, such as TNFR activation. Once activated, hMI-ER1 may function as a transcription regulator on specific target gene promoters or it may directly participate in apoptotic pathways (Figure 5.9). The hMI-ER1 sequence was predicted to have a putative BH3 domain in its N-terminal region (Paterno *et al.*, personal communication), so it may be a novel BH3-only protein. Proteins of the Bcl-2 family are critical regulators of caspase activation and apoptosis, and they contain regions of homology (BH1-4) (reviewed in Bouillet and Strasser, 2002). The BH3 domain is essential for the binding of BH3-only pro-apoptotic proteins to the anti-apoptotic members of the Bcl-2 family and for their ability to induce cytochrome c release from mitochondria (reviewed in Bouillet and Strasser, 2002). For example, the BH3-only protein BAD induces cytochrome c release from mitochondria when it binds to BCL-X_L or BCL-2 (Kelekar *et al.*, 1997; Yang *et al.*, 1995; Zha *et al.*, 1997). Most interestingly, the BH3-only protein, BID protein, is activated by death receptor signalling pathway, and is translocated to the mitochondria where it induces cytochrome c release, thus providing cross talk from the cell death receptor pathways to the mitochondria death pathways (reviewed in Scorrano and Korsmeyer, 2003). In response to cytochrome c binding, APAF-1 in the cytoplasm can form a complex with and activate initiator caspase-9 to initiate apoptosis (Li *et al.*, 1997; Zou *et al.*, 1997). It may be possible that TRABID-TRAFs complexes recruit and activate hMI-ER1, which then translocate to the mitochondria and induces apoptosis, analogous to the function of BID transmitting the apoptotic signal from the cell death receptors to the mitochondria (Figure 5.9).

CHAPTER 6 hMI-ER1α AND β WERE FOUND TO INTERACT PHYSICALLY WITH THE TUMOUR SUPPRESSOR RB

6.1 Introduction

Previous studies revealed that overexpression of hMI-ER1 in NIH 3T3 fibroblast cells suppressed colony forming efficiency (Paterno *et al.*, unpublished data). This assay has been previously used to demonstrate, for example, the growth-suppressive activities of p53 and $p16^{INK4A}$ (Baker *et al.*, 1990; Lin *et al.*, 1996a). As shown previously in colony forming assays with p53 and $p16^{INK4A}$, the inhibitory effect of hMI-ER1 in colony formation efficiency could be due to an apoptotic-inducing activity or to an ability to inhibit cell proliferation (Baker *et al.*, 1990; Lin *et al.*, 1996a).

Proliferation of mammalian cells is controlled by the cell cycle regulatory machinery. The cell cycle regulatory machinery, composed of cyclins, cyclin-dependent kinases (CDKs) and their inhibitors (CKIs), controls cell cycle progression at particular checkpoints, directing the cell towards proliferation, growth arrest or apoptosis (reviewed in Alberts *et al.*, 2002). The most thoroughly characterized G1 checkpoint operates through CDK-mediated phosphorylation of Rb and other 'pocket' proteins, which release bound transcription factors of the E2F family and allow them to activate genes essential for S phase (reviewed in Qu *et al.*, 2003). Thus the CKI-CDK-RB pathway plays a central role in determining whether a cell will proceed through the G_1 phase of the cell cycle (reviewed in Kaelin, 1999). Another important regulator of the cell cycle regulatory machinery is p53. p53 functions as a transcription factor that can activate expression of a

range of cellular genes involved in arrest of cell cycle and induction of apoptosis. p53 can arrest cell proliferation or induce apoptosis in response to DNA damage or overexpression of oncogenes such as E1A (Sherr, 1998). Both RB and p53 were found to form complexes containing HDAC1 in cells (Luo *et al.*, 2000; Magnaghi-Jaulin *et al.*, 1998). HDAC1 was shown to form multi-subunit complexes (reviewed in Ahringer, 2000; Cress and Seto, 2000; Thiagalingam *et al.*, 2003). Recently, the ELM2 and SANT domain-containing protein, MTA2, was shown to form complexes containing HDAC1 and p53 (Luo *et al.*, 2000). Since hMI-ER1 associates with HDAC1 enzymatic activity in cells (Chapter 3), it is possible that hMI-ER1 is present in the multi-subunit complexes containing RB and/or p53. The experiments described in this section were designed to investigate if an interaction between hMI-ER1 and RB and p53 was possible.

6.2 Materials and Methods

6.2.1 Cell culture

Human MCF-7 breast carcinoma cell line was obtained from the American Type Culture Collection and cultured at 37°C in 5% CO_2 in DMEM containing 10% FCS. MCF-7 Tet-On and HeLa Tet-On cell lines were obtained from Clontech, and cultured at 37°C in 5% CO_2 in DMEM containing 500 μg/ml G418 (Sigma) and 10% Tet system approved fetal bovine serum (Clontech). pTRE2, pTREα and pTRE-β HeLa Tet-On cell lines were established in our laboratory (see section 6.2.6), and cultured at

37°C in 5% CO_2 in DMEM containing 500 μg/ml G418, 200 μg/ml hygromycin B (Clontech) and 10% Tet system approved fetal bovine serum.

6.2.2 Plasmids and constructs

Myc-tagged plasmids (CS3+MT) (pMyc) containing either full-length *hmi-er1α* (pMyc-α) or β (pMyc-β) have been previously described (section 3.2.2). GST-hmi-er1α (GST-α) or GST-hmi-er1β (GST-β) fusion constructs were generated as described (section 4.2.2). Human p53 cDNA was obtained from Dr. J.G. Church, Memorial University of Newfoundland. Full-length human RB cDNA was obtained from Dr. B.L. Gallie, University of Toronto. A series of RB deletion mutations were generated by amplifying fragments encoding the appropriate amino acid residues of RB, using the primer pairs listed in Table 6.1. PCR products were cloned into pCR3.1. All plasmids were sequenced to verify the junctions and the RB sequence.

pTRE2 and pTK-Hyg vector were obtained from Clontech. pTRE2-hMI-ER1α (pTRE2-α) and pTRE2-hMI-ER1β (pTRE2-β) were generated by first amplifying entire coding sequence of either hMI-ER1α or β, using the primer pairs listed in Table 4.1 (section 4.2.2). PCR products were cloned into pCR3.1 and *BamHI*/*Xho*I fragments were then inserted into the *BamHI*/*Sal*I sites of the pTRE2 vector.

6.2.3 GST-fusion protein production

GST fusion proteins were expressed as described in section 2.2.6.

Table 4 Table 6.1 PCR primer pairs used for constructing mutating RB plasmids in pCR3.1 vector

Construct[a]	Forward primer	Reverse primer
RB(1-772)	5 -CACCATGGCGCCCAA AACCCCCCGAAAAACG-3	5 -CACCATGGACTCCAGTTA GGACTGTTATGAAC-3
RB(772-928)	5 -CACCATGGGCAGAGA CTGAAAACAAATATTTTG-3	5 -GCGGGATCCTCATTTCTC TTCCTTGTTTGAGGT-3
RB(373-772)	5 -CACCATGGCTCCAGT TAGGACTGTTATGAAC-3	5 -CACCATGGACTCCAGTTA GGACTGTTATGAAC-3
RB(373-573)	5 -CACCATGGCTCCAGT TAGGACTGTTATGAAC-3	5 -TCAAATAAGATCAAATAA AGGTGAATCTGAGAG-3
RB(571-772)	5 -CACCATGGATCTTATTAA ACAATCAAAGGACCGAG-3	5 -CACCATGGACTCCAGTTA GGACTGTTATGAAC-3
RB(645-772)	5 -CACCATGGCTACCTCTCT TTCACTGTTTTATAAAA-3	5 -CACCATGGACTCCAGTTA GGACTGTTATGAAC-3

[a] Deletion constructs were named according to the encoded amino acid residues of the RB protein

6.2.4 GST pull-down

GST pull-down assays were performed as described in section 4.2.6.

6.2.5 Co-IP and Western blot analysis

In vitro co-IP assays *were* performed as described in section 3.2.4 with either 10 µl anti-RB monoclonal antibody (Catalog No. 14001A: Clone No. G3-245, Santa Cruz Biotechnology, Inc.) or 10 µl anti-Myc monoclonal antibody 9E10.

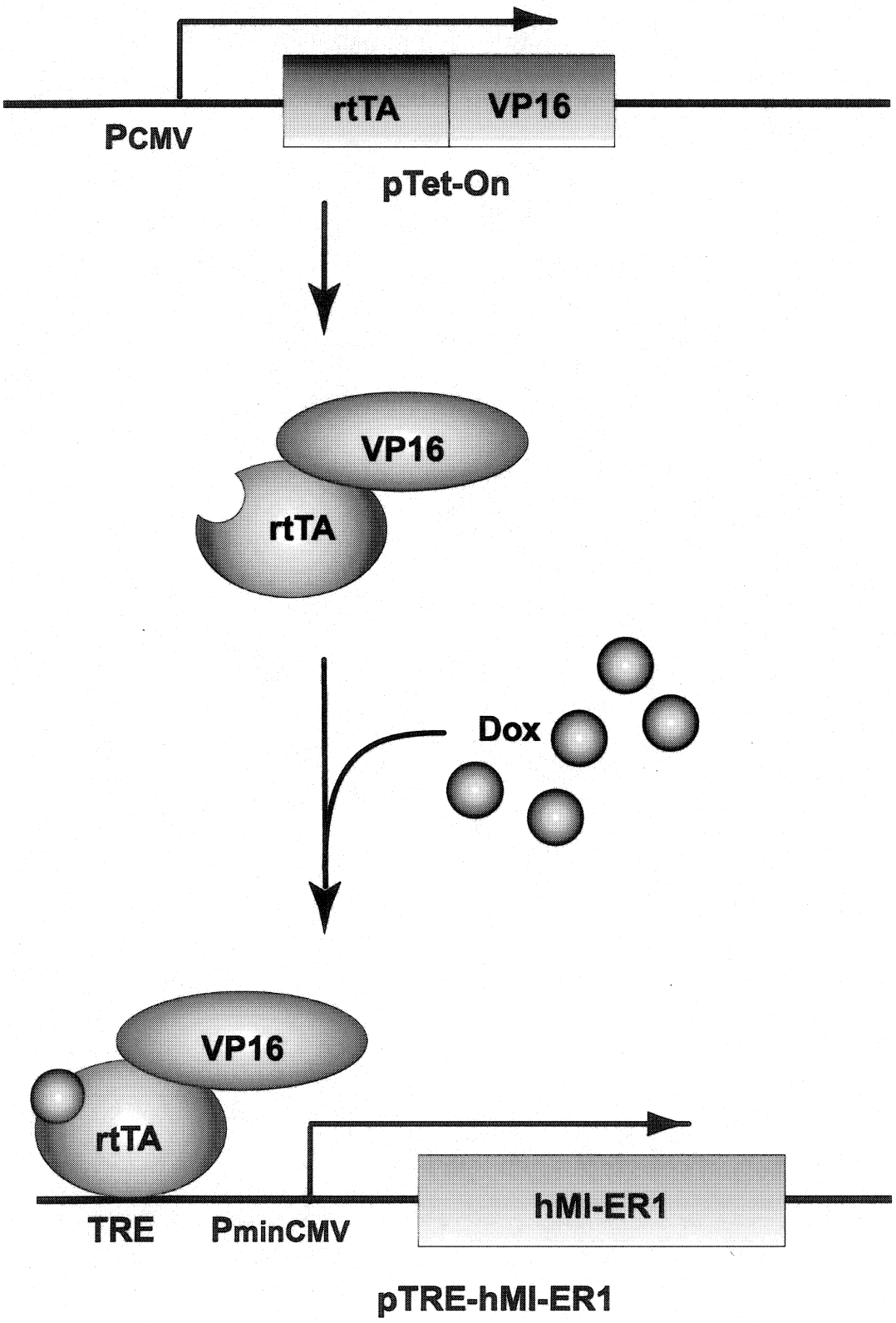
In vivo co-IP assays were performed in MCF-7 cells as described in section 3.2.4 with 10 µl anti-RB monoclonal antibody (Catalog No. 14001A: Clone No. G3-245, Santa Cruz Biotechnology, Inc.). Western blot analysis was performed as described in section 3.2.4 with anti-Myc monoclonal antibody 9E10 (1:100 dilution).

6.2.6 Establishment of hMI-ER1α and β Tet-On cell lines

The tetracycline induced (Tet-On) gene expression system (Clontech Inc.) takes advantage of the high specificity and affinity of the *E. coli* tetracycline repressor (TetR) for its operator sequence (*tet*O) induced by tetracycline and its derivative antibiotic doxycycline (Gossen *et al.*, 1995; Gossen and Bujard, 1992). MCF-7 Tet-On and HeLa Tet-On cells were stably transfected with the first regulatory pTet-On plasmid. This construct encodes a variant of the tTA protein, the reverse tetracycline transactivation (rtTA) protein (Clontech) (Figure 6.1). rtTA can bind to doxycycline and behaves as an activator of transcription (Figure 6.1).

Figure 6.1 Establishment of hMI-ER1 HeLa Tet-On cell lines.

HeLa Tet-On cells were stably transfected with the first regulatory pTet-On plasmid which encodes rtTA protein. The second regulatory plasmid is pTRE2-hMI-ER1, in which hMI-ER1 were under the control of the tetracycline response element (TRE). The TRE consists of *tet*O sequences immediate upstream of a minimal CMV promoter (PminCMV). In the absence of doxycycline, rtTA is unable to bind to the *tetO*, and there is no gene expression from the pTRE2-hMI-ER1. When doxycycline is present, however, doxycycline binds to rtTA and changes its conformation. rtTA then binds *tet*O to activate transcription of hMI-ER1.


rtTA
VP16
PCMV
pTet-On
VP16
rtTA
Dox
VP16
rtTA
TRE
PminCMV
hMI-ER1
pTRE-hMI-ER1

The second regulatory plasmid is pTRE2-hMI-ER1α (pTRE-α) or pTRE2-hMI-ER1β (pTRE-β), in which hMI-ER1α or hMI-ER1β were under the control of the tetracycline responsive element (TRE). The TRE consists of seven repeats of the *tet*O sequence immediately upstream of a minimal CMV promoter (Figure 6.1). The minimal CMV promoter does not contain any of the enhancer elements of the typical CMV promoter, and its activity is controlled by the TRE. In the absence of doxycycline, rtTA is unable to bind to the pTRE promoter, and therefore there is no gene expression from the pTRE2-hMI-ER1α or β. When doxycycline is present, however, doxycycline binds to rtTA and changes its conformation. rtTA then binds *tet*O to activate transcription of hMI-ER1α and β (Figure 6.1).

To establish hMI-ER1 MCF-7 Tet-On and HeLa Tet-On cell lines which can express hMI-ER1α or hMI-ER1β when doxycycline is present, MCF-7 Tet-On and HeLa Tet-On cells were co-transfected with the second regulatory plasmid pTRE2 empty vector, pTRE2-hMI-ER1α (pTRE2-α) or pTRE2-hMI-ER1β (pTRE2-β) and pTK-Hyg vector and selected by hygromycin, according to the protocol recommended by the manufacturer (Clontech). The stable clones were isolated and the induced overexpression of hMI-ER1α or β was screened using Western blot analysis using an anti-hMI-ER1 antibody produced in our laboratory (24).

6.2.7 Colony formation efficiency assays

For the colony formation efficiency of HeLa Tet-On cells, 150 cells/plate were seeded in 50 mm plates 18 h prior 5 μg/ml doxycycline (Clontech) induction. The medium was

changed every 2 d with fresh medium with or without the indicated concentration of doxycycline. Cell colonies were visualized by crystal violet staining 2 weeks later.

6.3 Results

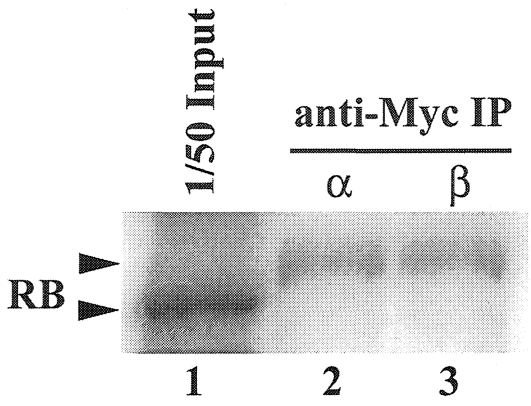
6.3.1 hMI-ER1α and β associated with RB *in vitro* and *in vivo*

Co-immunoprecipitation analysis was utilized to examine the ability of hMI-ER1α and β isoforms to physically associate with RB and p53 *in vitro*. RB and p53 were synthesized and ^{35}S-labelled *in vitro* and mixed with unlabelled *in vitro* translated Myc-hMI-ER1α (Myc-α) or Myc-hMI-ER1β (Myc-β), followed by immunoprecipitation with the anti-Myc antibody 9E10. RB was detected in both Myc-α and Myc-β immunoprecipitates by anti-Myc antibody 9E10 (Figure 6.2A, lane 2 and 3), indicating that RB associates with both isoforms *in vitro*. The RB immunoprecipitated with hMI-ER1α or β migrated to a higher position on the gel than the direct input of the full-length RB TNT (Figure 6.2A, arrow head), suggesting hMI-ER1α and β may be specifically associated with a modified RB form, such as phosphorylated RB. In the same assay conditions, p53 did not bind to hMI-ER1α and β, since p53 was not detected in Myc-hMI-ER1α or Myc-hMI-ER1β immunoprecipitates by anti-Myc antibody 9E10 (Figure 6.2B, lane 2 and 3).

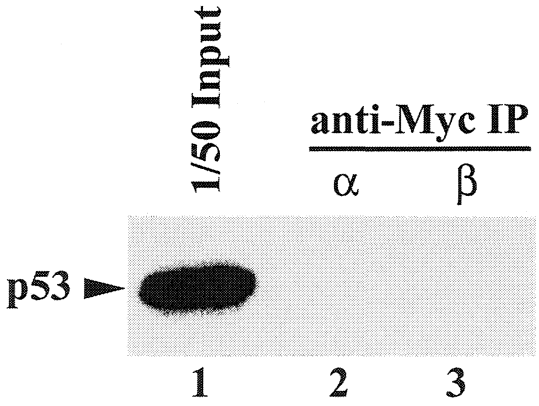
Figure 6.2 hMI-ER1α and β interact with RB *in vitro*.

(A) ^{35}S-labelled TNTs programmed with cDNA encoding RB were loaded directly on the gel (1/50 of input) (lanes 1) or incubated with unlabelled TNTs programmed with cDNA encoding Myc tagged hMI-ER1α (α), or Myc tagged hMI-ER1β (β) and subjected to IP with anti-Myc antibody 9E10 (lane 2-3). Proteins were resolved by SDS-PAGE and visualized by autoradiography. Shown is a representative autoradiograph, and the positions of RB are indicated by arrows. **(B)** ^{35}S-labelled TNTs programmed with cDNA encoding p53 were loaded directly on the gel (1/50 of input) (lanes 1)) or incubated with unlabelled TNTs programmed with cDNA encoding Myc tagged hMI-ER1α (α) or Myc tagged hMI-ER1β (β) and subjected to IP with anti-Myc antibody 9E10 (lane 2-3). Proteins were resolved by SDS-PAGE and visualized by autoradiography. Shown is a representative autoradiograph, and the positions of p53 are indicated by arrows. **(C)** ^{35}S-labelled TNTs programmed with cDNA encoding luciferase (Luc), Myc tagged hMI-ER1α [Myc-α (α)] or Myc tagged hMI-ER1β [Myc-β (β)] were loaded directly on the gel (1/50 of input) (lanes 1-3) or incubated with unlabelled TNTs programmed with *RB* cDNA and subjected to immunoprecipitation (IP) with anti-RB antibody (lanes 4-6). Proteins were resolved by SDS-PAGE and visualized by autoradiography. Shown is a representative autoradiograph, and the positions of the Myc-α or Myc-β and luciferase are indicated by arrows.

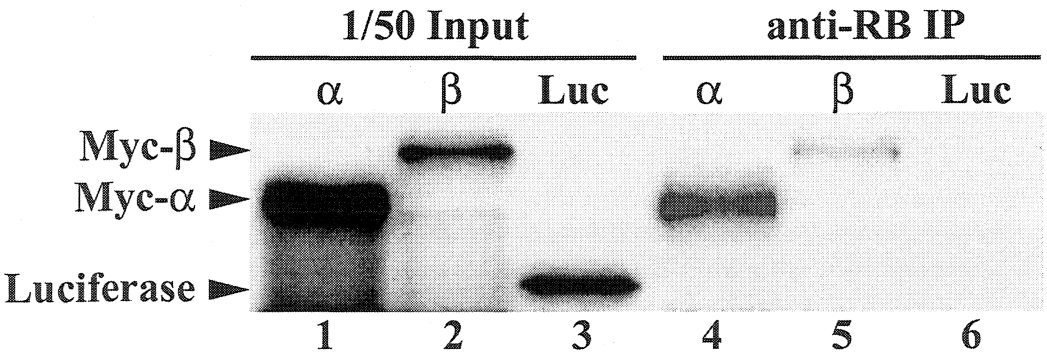
A


1/50 Input
anti-Myc IP
α
β
RB
1
2
3

B


1/50 Input
anti-Myc IP
α
β
p53
1
2
3

C


1/50 Input
anti-RB IP
α
β
Luc
α
β
Luc
Myc-β
Myc-α
Luciferase
1
2
3
4
5
6

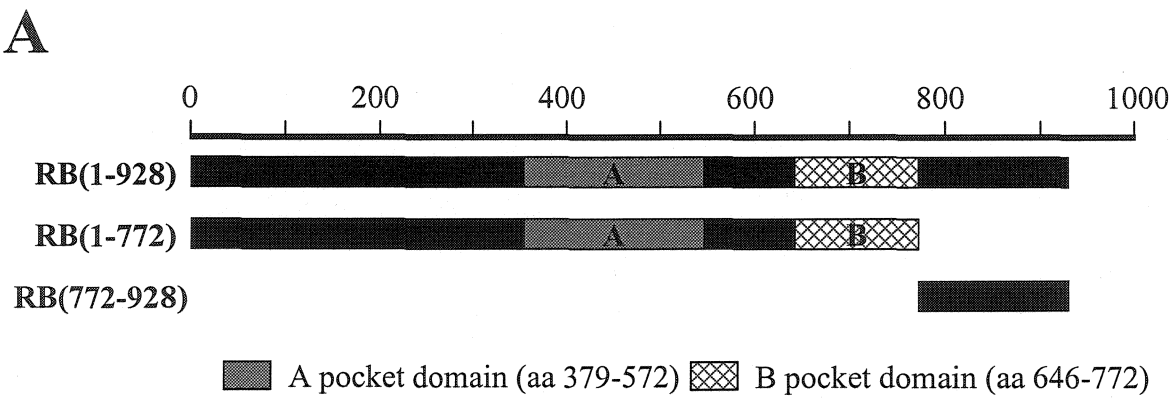
Further *in vitro* and *in vivo* studies were designed to focus on the interaction with RB. Reciprocal experiments, using anti-RB antibody for immunoprecipitation confirmed that both hMI-ER1α and β (Figure 6.2A) could associate with RB (Figure 6.2C, lane 4 and 5), while luciferase, a control protein, did not (Figure 6.2C, lane 6). The basic structure of RB consists of the N-terminal, the central region and the C-terminal region (Figure 6.3A) (reviewed in Morris and Dyson, 2001). The central region is termed the pocket , which folds into a pocket shape to offer a docking site for many interacting proteins such as E2F1 (reviewed in Morris and Dyson, 2001).

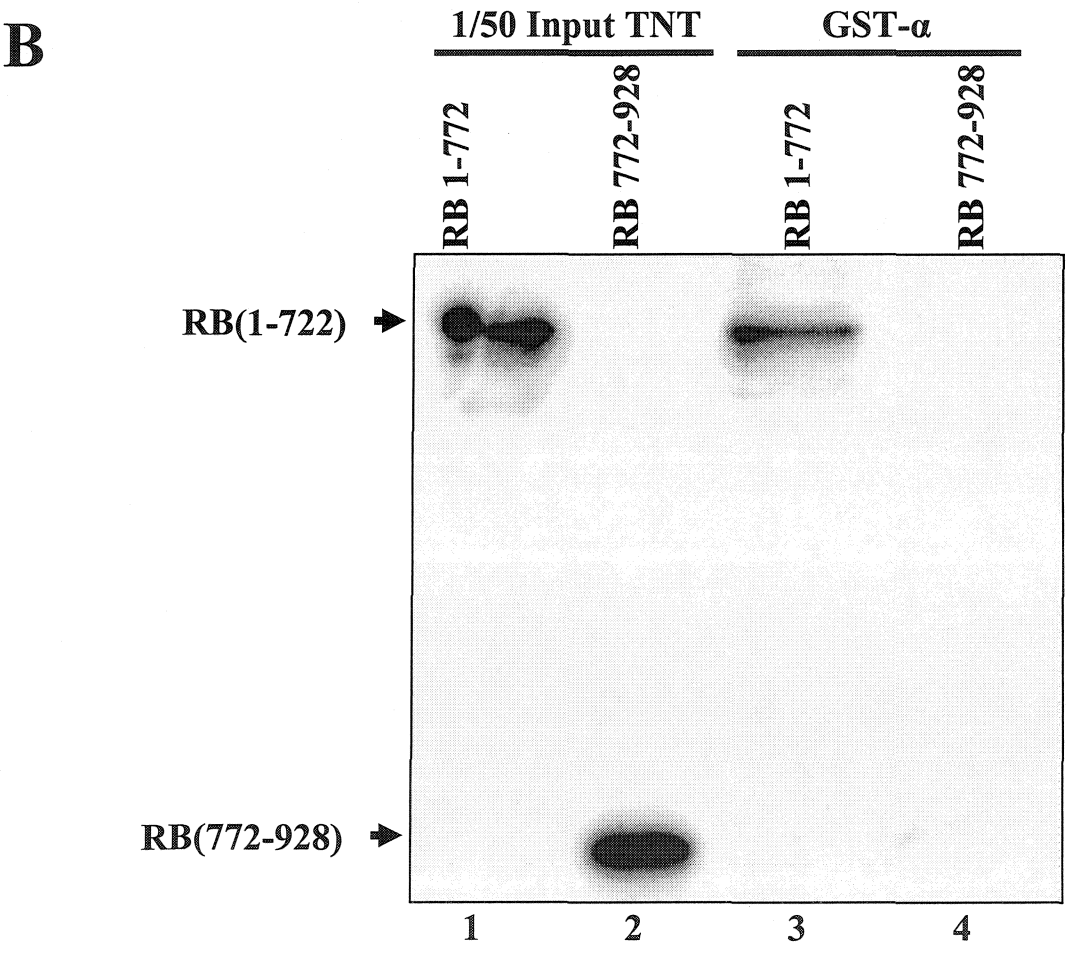
The pocket domain of RB contains two regions termed the A-and B-pocket domains, that are conserved both across species and among the related proteins (Chow and Dean, 1996). These domains are separated by a spacer region that varies in length and is not conserved (Chow and Dean, 1996). The human RB A-pocket domain spans residues 379 to 572 and the B-pocket domain spans residues 646 to 772 (reviewed in Morris and Dyson, 2001).

To determine which region of RB interacts with hMI-ER1, ^{35}S-labelled *in vitro* synthesized protein using templates encoding only the N-terminal region containing the A- and B-pocket domains or the C-terminal region alone were generated (Figure 6.3B, lane 1 and 2). GST pull-down assays revealed that the N-terminal region of RB, containing the A- and B-pocket domains, interacted with hMI-ER1 (Figure 6.3B, lane 3), while the C-terminal region, lacking both A- and B-pocket domains, showed no interaction with hMI-ER1 (Figure 6.3B, lane 4).

Figure 6.3 Mapping region of RB interacted with hMI-ER1 *in vitro*.

(A) The diagram illustrates the deletion mutants of RB. The protein pocket A and B domains are identified in the legend below the diagram, and the RB amino acid residues encoded by each construct are listed. **(B)** ^{35}S-labelled TNTs programmed with cDNA encoding RB(1-772) was loaded directly on the gel (1/50 of input) (lane 1) or incubated with GST-hMI-ER1α (GST-α) (lane 3); ^{35}S-labelled TNTs programmed with cDNA encoding RB(772-928) was loaded directly on the gel (lane 2) or incubated with GST-hMI-ER1α (GST-α) (lane 4). Proteins were resolved by SDS-PAGE and visualized by autoradiography. Shown is a representative autoradiograph, and the positions of RB(1-722) and RB(722-928) are indicated by arrows.


A
0
200
400
600
800
1000
RB(1-928)
A
B
RB(1-772)
A
B
RB(772-928)
A pocket domain (aa 379-572)
B pocket domain (aa 646-772)

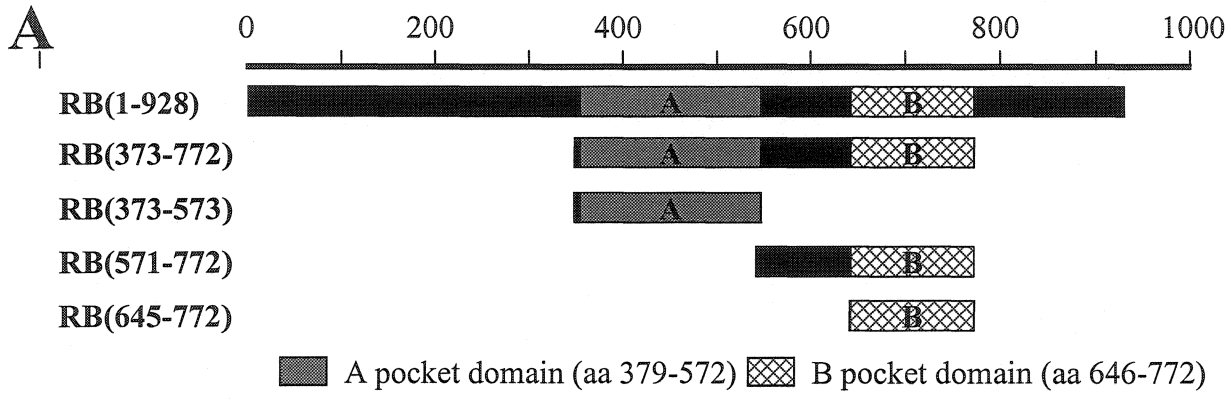

B
1/50 Input TNT
GST-α
RB 1-772
RB 772-928
RB 1-772
RB 772-928
RB(1-722)
RB(772-928)
1
2
3
4

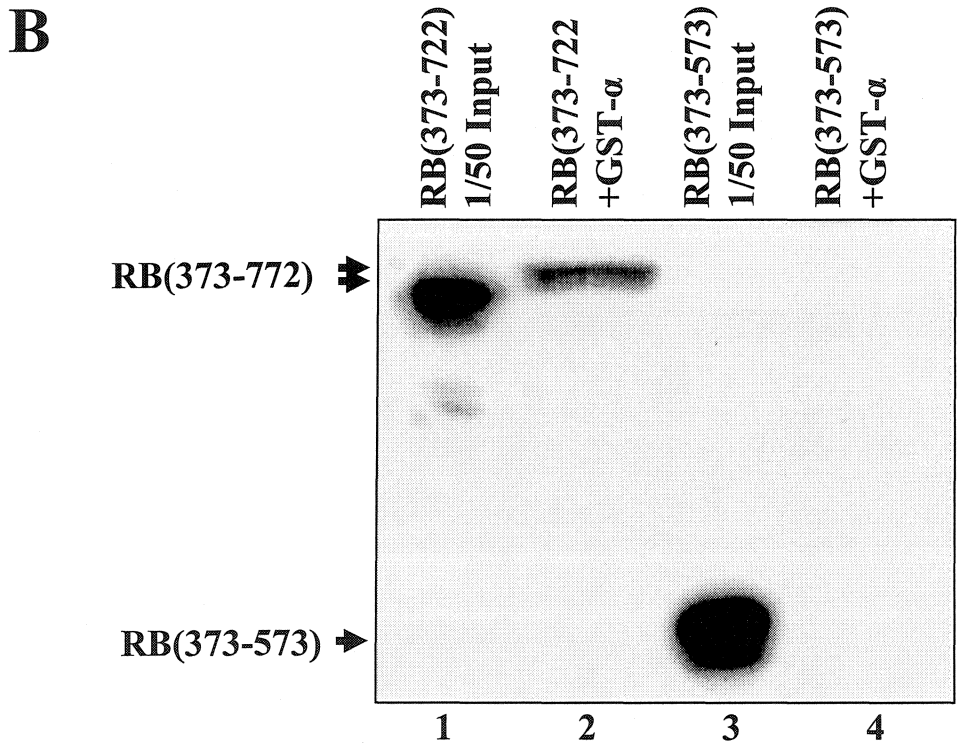
To determine whether the A- and B-pocket domains were required for the interaction with hMI-ER1, ^{35}S-labelled *in vitro* synthesized proteins using templates encoding deletion mutants were generated in the region containing the A- and B-pocket domains (Figure 6.4A; 6.4B, lane 1). GST pull-down assays revealed that this region containing both A- and B-pocket domains interacts with hMI-ER1 (Figure 6.4B, lane 4). Further experiments revealed that the A-pocket domain of RB was not required for interaction with hMI-ER1 (Figure 6.4B, lane 5 and 6), while the B-pocket domain (Figure 6.4C, lane 1 and 3) interacted with hMI-ER1 (Figure 6.4C, lane 2 and 3). Similar to the full-length RB, the RB deletion mutants in the GST-hMI-ER1 pull-down assays exhibited a higher molecular form on the gel than the direct input of the *in vitro* synthesized product (Figure 6.4, arrows), which suggested that hMI-ER1α and β may specifically associate with a post-translationally modified form of RB.

In vivo analysis involved transient expression of the Myc-tag empty vector, Myc-α or Myc-β in MCF-7 cells. Expression of Myc-α and Myc-β was verified by Western blot analysis with anti-Myc antibody 9E10 (Figure 6.5, lane 2 and 3). Cell extracts were subjected to immunoprecipitation with anti-RB antibody, followed by Western blot analysis with the anti-Myc antibody 9E10. Both hMI-ER1α and β isoforms were co-immunoprecipitated with endogenous RB by anti-RB antibody (Figure 6.5, lane 7 and 9), but not in the immunoprecipitates by the non-immune serum control (Figure 6.5, lane 5 and 6).

Figure 6.4 RB B-pocket domain (aa 645-772) interacted with hMI-ER1 *in vitro*.

(A) The diagram illustrates the deletion mutants of RB pocket domain. The protein pocket A and B domains are identified in the legend below the diagram, and the RB amino acid residues encoded by each construct are listed. **(B)** ^{35}S-labelled TNTs programmed with cDNA encoding RB(373-772) was loaded directly on the gel (lane 1) or incubated with GST-hMI-ER1α (GST-α) (lane 2); ^{35}S-labelled TNTs programmed with cDNA encoding RB(373-573) was loaded directly on the gel (1/50 of input) (lane 2) or incubated with GST-hMI-ER1α (GST-α) (lane 4); Shown is a representative autoradiograph, and the positions of RB(373-772) and RB(373-573) are indicated by arrows. **(C)** ^{35}S-labelled TNTs programmed with cDNA encoding RB(571-772) was loaded directly on the gel (1/50 of input) (lane 1) or incubated with GST-hMI-ER1α (GST-α) (lane 2); ^{35}S-labelled TNTs programmed with cDNA encoding RB(645-772) was loaded directly on the gel (lane 3) or incubated with GST-hMI-ER1α (GST-α) (lane 4). Proteins were resolved by SDS-PAGE and visualized by autoradiography. Shown is a representative autoradiograph, and the positions of RB(571-772) and RB(647-772) are indicated by arrows.


A
0
200
400
600
800
1000
RB(1-928)
RB(373-772)
RB(373-573)
RB(571-772)
RB(645-772)
A
B
A pocket domain (aa 379-572)
B pocket domain (aa 646-772)


B
RB(373-722) 1/50 Input
RB(373-722 +GST-α
RB(373-573) 1/50 Input
RB(373-573) +GST-α
RB(373-772)
RB(373-573)
1
2
3
4

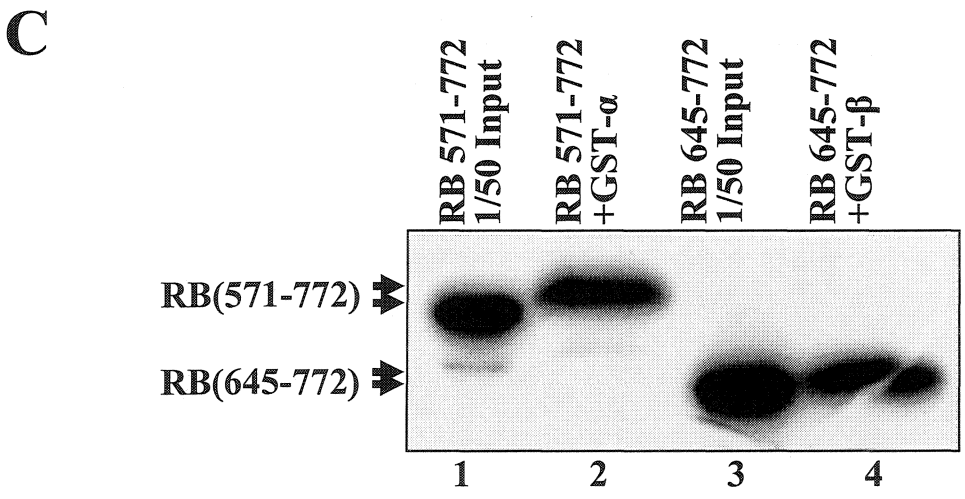
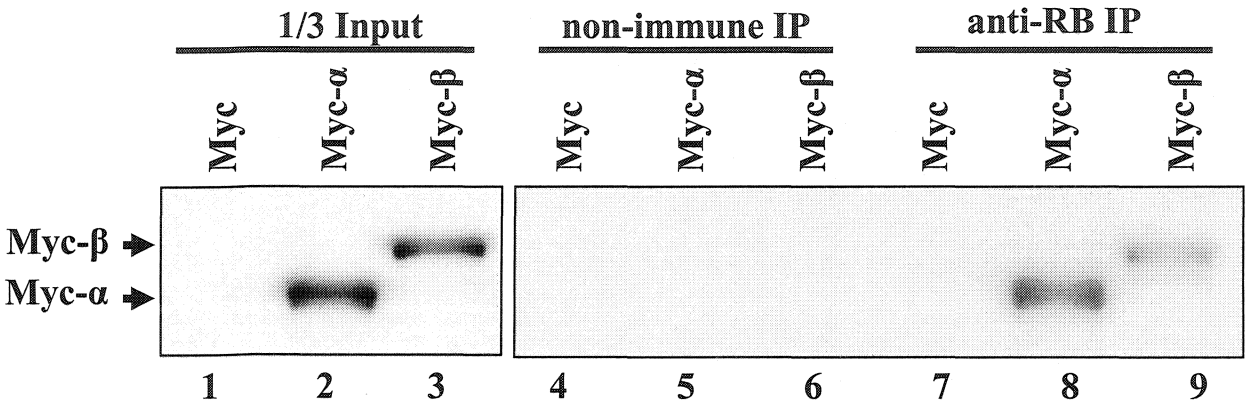

C
RB 571-772 1/50 Input
RB 571-772 +GST-α
RB 645-772 1/50 Input
RB 645-772 +GST-β
RB(571-772)
RB(645-772)
1
2
3
4

Figure 6.5 hMI-ER1α and β interact with RB *in vivo*.

Cell lysates from MCF-7 cells were transiently transfected with Myc-tag empty vector (pMyc), Myc-hMI-ER1α (pMyc-α) or Myc-hMI-ER1β (pMyc-β). Cell lysates were prepared and were either added directly to sample buffer (1/3 of input) (lanes 1-3) or subjected to IP with non-immune serum (lanes 4, 5, 6) or anti-RB antibody (lanes 7, 8, 9). Western blot (WB) analysis was performed using anti-Myc antibody 9E10. Shown is a representative Western blot, and the positions of the Myc-α and Myc-β proteins are indicated.


1/3 Input
non-immune IP
anti-RB IP
Myc
Myc-α
Myc-β
Myc
Myc-α
Myc-β
Myc
Myc-α
Myc-β
Myc-β
Myc-α
1
2
3
4
5
6
7
8
9

6.3.2 Overexpression of hMI-ER1α and β suppresses colony formation efficiency in hMI-ER1α and β HeLa Tet-On cell lines

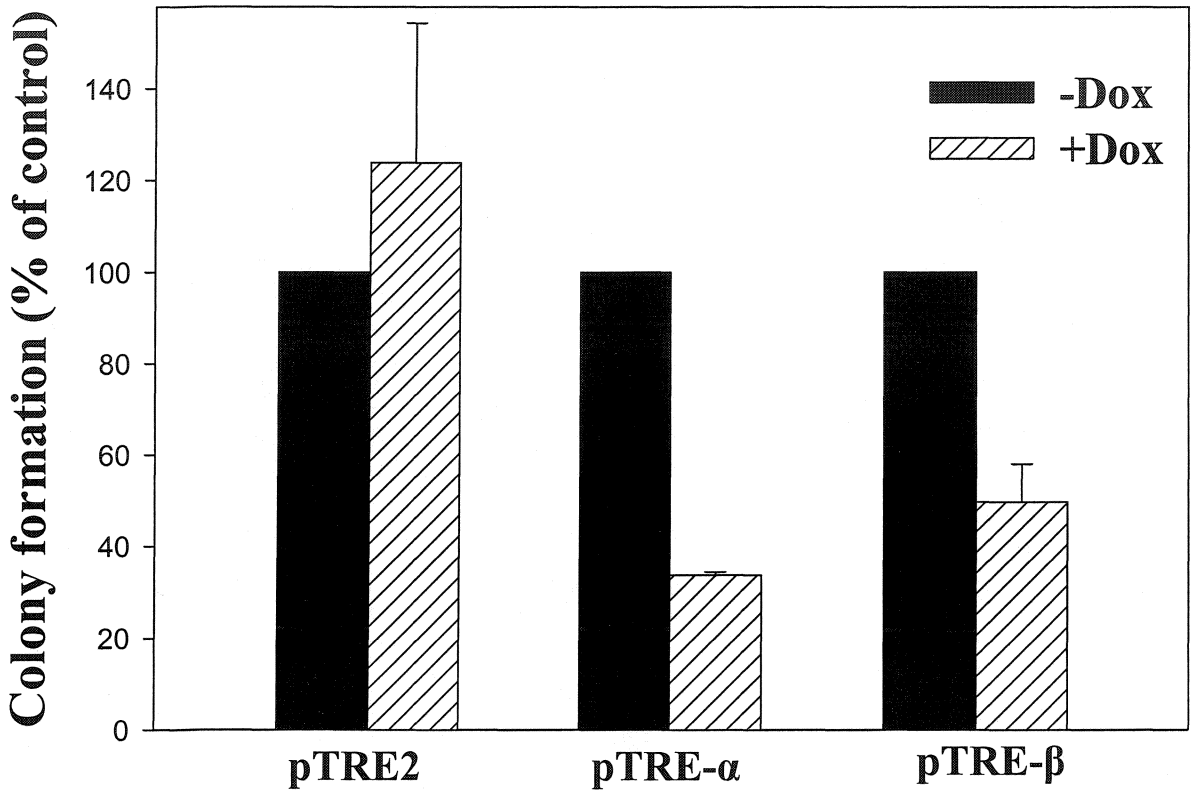
Previous studies revealed that overexpression of hMI-ER1α and β in NIH 3T3, cells suppressed cell colony formation efficiency (Paterno *et al.*, unpublished data). To further characterize the cell growth regulatory role of hMI-ER1α and β, experiments were designed to establish hMI-ER1α and β Tet-On cell lines. All hMI-ER1α and β Tet-On cells would express the same amount of hMI-ER1α and β when induced, thus excluding the possibility that the observed colony formation was an artifact of transient transfection. At the time this study was undertaken, MCF and HeLa Tet-On cell lines were commercially available (Clontech). hMI-ER1α and β HeLa Tet-On cell lines were successfully established in our laboratory (section 6.2.6). However, hMI-ER1α and β MCF-7 Tet-On cell lines have not yet been successfully established.

In those hMI-ER1α and β HeLa Tet-On cells, a preliminary study for colony formation efficiency assays was conducted after modulating the hMI-ER1α and β level while maintaining all other variables. Induced expression of hMI-ER1α or hMI-ER1α by doxycycline (Figure 6.6B) resulted in significant reduction of cell colony formation efficiency in hMI-ER1α or β HeLa Tet-On cells (Figure 6.6A). Maximum repression resulted in a reduction in colony formation efficiency to 38% and 14% of control levels by hMI-ER1α and hMI-ER1β, respectively. No significant effect on the colony formation efficiency of the pTRE2 control HeLa tet-on cells by addition of doxycycline in the culture medium (Figure 6.6A). These results further confirmed that overexpression of hMI-ER1α and β suppresses cell growth in HeLa cells.

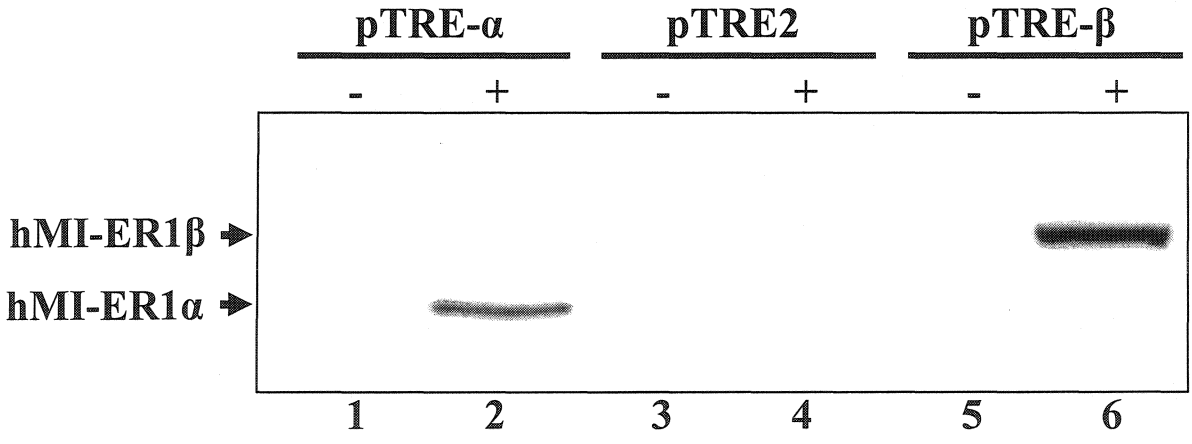
Figure 6.6 Induced overexpression of hMI-ER1α and β by doxycycline significantly suppressed the HeLa colony formation efficiency.

150 cells/plate HeLa tet-on cells were seeded into 50 mm plates. After 24 h of incubation cells were followed by changing the medium containing 5 μg/μl of doxycycline. The media with 5 μg/μl of doxycyline was changed every two days. After 10 days colonies were visualized by crystal violet staining. The average values and standard deviations from two independent experiments is shown **(A).** The amount of hMI-ER1α or β protein induced by in each sample was determined 24 h after addition of doxycycline by Western blot analysis using anti-hMI-ER1 antiserum. A representative Western blot is shown **(B).** The positions of hMI-ER1α and hMI-ER1β are indicated by arrows.

A


Colony formation (% of control)
140
120
100
80
60
40
20
0
-Dox
+Dox
pTRE2
pTRE-α
pTRE-β

B


pTRE-α
pTRE2
pTRE-β
-
+
-
+
-
+
hMI-ER1β
hMI-ER1α
1
2
3
4
5
6

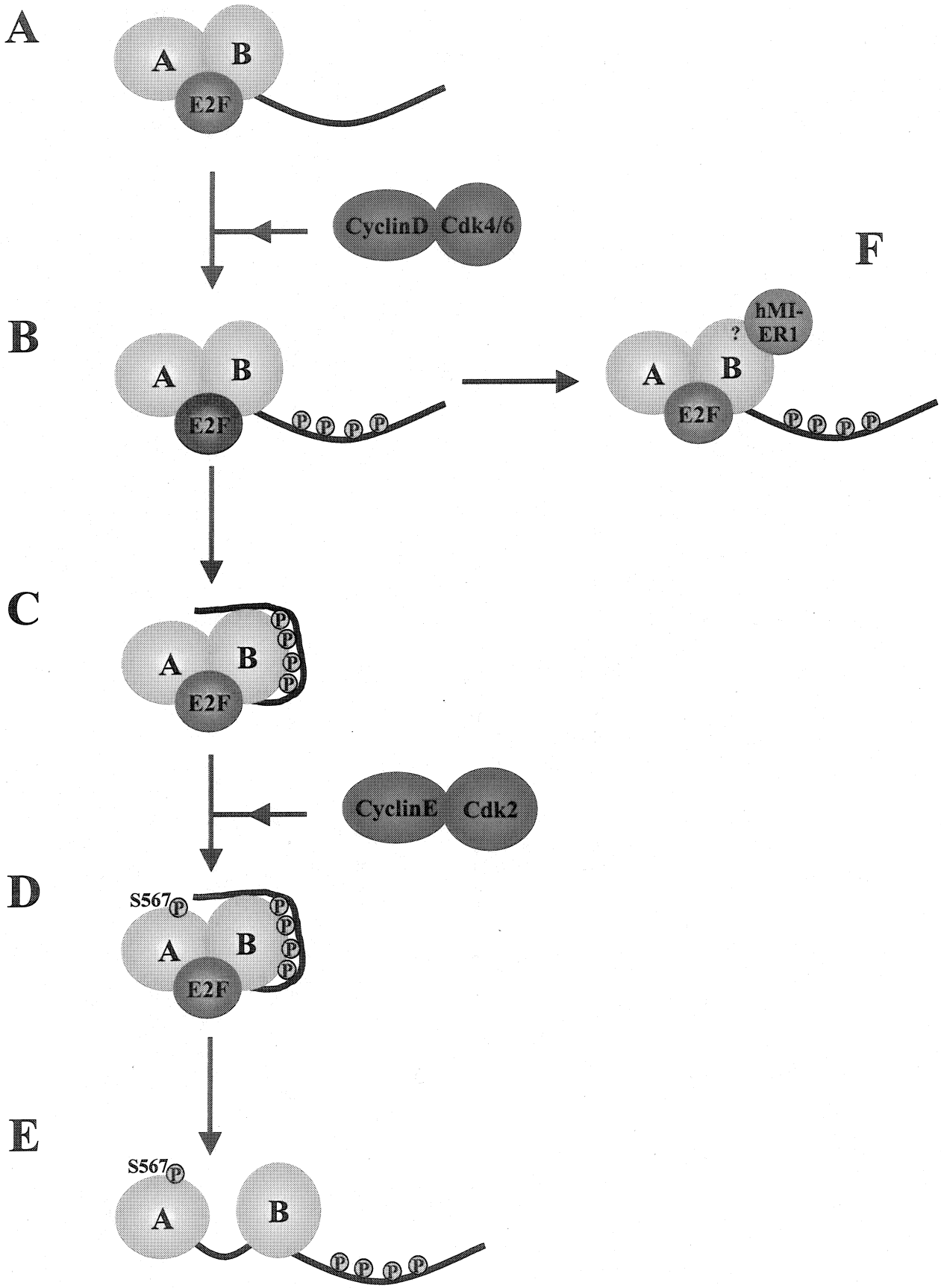
6.4 Discussion

The present study revealed that hMI-ER1 can interact with RB. However, it is possible that hMI-ER1 interacts with RB either directly or indirectly through other proteins in the multi-subunit complexes. Since there are multiple endogenous proteins in cells and in the reticulocyte TNT lysates, we cannot at present exclude these possibilities. Alternative protein-protein interaction assays using both purified proteins from bacterial recombinant products and then Western blot analysis may address this question.

The A-pocket domain and B-pocket domain of RB is critical in forming a transcription repression pocket for E2F (Chow and Dean, 1996; Harbour *et al.*, 1999). As shown in Figure 6.7, subsequent phosphorylation of RB, mediated by cyclin D-CDK4/6and cyclin E-CDK2 complexes, disrupts the transcription repression pocket for E2F and allows for the transcription of E2F target genes (reviewed in Harbour and Dean, 2000). In brief, cyclin D-CDK4/6 becomes active during early G1, and phosphorylates the C-terminal tail of RB (Figure 6.7B). This phosphorylated C-terminal tail then folds over RB pocket (Figure 6.7C). During cell cycle progression, cyclin E-CDK2 is activated and binds to the C-terminal tail and phosphorylates the pocket domain (S567 was the critical residue) (Figure 6.7D). This phosphorylation disrupts the pocket domain structure and releases E2F (Figure 6.7E) (Harbour *et al.*, 1999). In this chapter, I demonstrated that hMI-ER1 interacted specifically with the B-pocket domain (Figure 6.4). It is possible that the interaction of hMI-ER1 with the B-pocket domain stabilizes the interaction of A- and B-pocket domains to form the transcription repression pocket for

Figure 6.7 Intramolecular interactions of RB by sequential post-translational modification and the possible functional implication of RB and hMI-ER1 interaction (adapted from Harbour *et al.*, 1999).

(A) The A- and B-pocket domain interact with each other forming an pocket for the E2F binding, thus forming a transcriptional repressive status on the E2F target gene promoters. **(B)** cyclin D-CDK4/6 becomes active early during G1, and phosphorylates RB. **(C)** Phosphorylated C-terminal tail then folds over the RB pocket. **(D)** During the cell cycle progression, cyclin E-CDK2 is activated and binds to the C-terminal tail and phosphorylates the pocket domain (^{567}S was is the critical residue). **(E)** Phosphorylation of the pocket domain on ^{567}S was the critical residue disrupts the pocket domain structure and releases E2F. **(F)** hMI-ER1 binds to the B-pocket domain (modified form?) of RB. Binding to B-pocket domain by hMI-ER1 may inhibit the interaction of C-terminal domain of RB binding to the pocket domain, thus may prevent the phosphorylation of the pocket such as on S567 by cyclin E-cdc2 and thus stabilize the RB-E2F repressive complexes.


A
A
B
E2F
CyclinD
Cdk4/6
F
B
hMI-
ER1
?
A
B
E2F
C
A
B
E2F
CyclinE
Cdk2
D
S567
A
B
E2F
E
S567
A
B

E2F (Figure 6.7F), which requires further study. It would also be very interesting to know which region of hMI-ER1 interacts with RB. GST pull-down assays are ongoing in the Paterno and Gillespie laboratories in order to study the minimal region of hMI-ER1 required to interact with RB.

Although multiple cellular and viral proteins interact with RB, only a few proteins, such as Ski, were reported to interact specifically with the B-pocket domain (reviewed in Morris and Dyson, 2001). Association of Ski with RB represses RB activity, and facilitates cell-cycle progression (Medrano, 2003). It is possible that interaction of hMI-ER1 with RB inhibits the association of Ski with RB by competing for the binding sites in the RB B-pocket domain, thus removing the inhibitory effect on RB by Ski and enhancing RB transcriptional repression function. Thus, the interaction of hMI-ER1 with RB may represent a novel regulatory mechanism for RB transcriptional repression function. Since RB plays a pivotal role in the regulation of cell growth and tumour suppression, this preliminary finding may have potential clinical significance. Further experiments are required to explore this exciting hypothesis.

The full-length RB TNT immunoprecipitated with the hMI-ER1 TNT migrates at a higher relative molecular weight on the gel than the direct input of the RB TNT (Figure 6.2, arrows), suggesting that hMI-ER1 may specifically associate with a modified RB form. The same is true for the B-pocket domain-containing region mutants (Figure 6.4, arrows). Recent studies in our laboratory showed that hMI-ER1 recruits multiple enzymes, including CBP (Blackmore *et al.*, unpublished data) and HDAC1 (Chapter 3),

which may modify RB or other associated proteins. Alternatively, RB or RB deletion mutants may be modified directly by factors in the reticulocyte TNT lysates. Although RB can be phosphorylated (reviewed in Kaelin, 1999; Wright *et al.*, 1991) and acetylated (Chan *et al.*, 2001), those well characterized modification sites are not located on the B-pocket domain (Chan *et al.*, 2001; Kaelin, 1999). The mechanism of binding of hMI-ER1 with a relatively higher molecular weight form of RB B-pocket domain requires further investigation.

hMI-ER1 interacts both *in vitro* and *in vivo* with RB (Section 6.3.1), which plays an important role in cell proliferation regulation (reviewed in Qu *et al.*, 2003). Therefore, it is possible that hMI-ER1 inhibits cell proliferation through RB. Induced overexpression of hMI-ER1 represses colony formation efficiency in hMI-ER1 HeLa Tet-On cells (Figure 6.6). In these Tet-On cells, all cells express the same amount of hMI-ER1α and β when induced, thus excluding the possibility that the observed colony formation was an artifact of plasmid transfection. These preliminary results suggested hMI-ER1α and β may regulate cell proliferation. However, we cannot at present exclude that hMI-ER1 could suppress cell proliferation through other RB-independent mechanisms. Many different mechanisms could explain our observations. In HeLa cells, most wild-type RB is inactivated by HPV18 E7 oncoprotein (Goodwin and DiMaio, 2000). Repression of E7 in HeLa cells was shown to cause the reactivation of RB (Goodwin and DiMaio, 2000). It is, therefore, possible that overexpression of hMI-ER1 may enhance the activity of RB through suppression of the expression of E7 oncoprotein. Further studies to analyze RB protein levels as well as phosphorylation status before and

after the overexpression of hMI-ER1α and β induced by doxycycline would help to address this question.

Most recently, hMI-ER1α or β were found to bind to unacetylated histone protein H4 (Paterno *et al.*, unpublished data). Therefore, hMI-ER1 may directly bind to E2F-targeted gene promoters such as cyclin E and repress the transcription of those genes. Thus, overexpression of hMI-ER1α and β may result in misregulation of subsets of growth-related genes, which ultimately alter cell cycle arrest and/or apoptosis, independent of status RB.

hMI-ER1α and β were recently shown to form complexes containing HDAC1 and have a transcriptional repression function (Chapter 3). In this way, high levels of hMI-ER1α and β expression could compete for HDAC1 from other HDAC1 complexes and deregulate a number of other HDAC1 containing transcriptional regulatory complexes such as mSin3A and/or NURD complexes.

hMI-ER1α and β were also found to interact with the Sp1 transcription factor, and interfere with the DNA binding site recognition by Sp1 (Chapter 4). Therefore, overexpression of hMI-ER1α or β may affect a large set of Sp1 target genes, many of which play important roles in regulating cell growth and cell death (reviewed in Black *et al.*, 2001).

As shown previously in colony forming assays with p53 (Baker *et al.*, 1990), the inhibitory effect of hMI-ER1α and β in colony formation efficiency may be also due to an apoptotic-inducing activity. We cannot at present exclude the possibility that overexpression of hMI-ER1 may cause non-specific toxicity to the cells during colony

formation. However, overexpression of non-related proteins such as luciferase in HeLa cells showed no significant effect on cell colony formation efficiency (Yang *et al.*, 2001). Sequence analysis revealed a putative BCL-2 homology (BH)-3 (BH3) domain in the N-terminal region of hMI-ER1 (Paterno *et al.*, personal communication). Proteins of the Bcl-2 family are critical regulators of caspase activation and apoptosis, and they contain regions of homology (BH1-4) (reviewed in Bouillet and Strasser, 2002). The BH3 domain is essential for the binding of BH3-only pro-apoptotic proteins to the anti-apoptotic members of the Bcl-2 family and for their ability to kill cells (reviewed in Bouillet and Strasser, 2002). Therefore, hMI-ER1α and β may be novel BH3-only proteins and it is possible that hMI-ER1 may interact with Bcl-2 family proteins and modulate apoptotic pathways. Interestingly, hMI-ER1 was found to interact with TRABID, a component of death receptor signalling pathways (see Chapter 5). It is possible that hMI-ER1 transmits the signal from the cell death receptor signalling pathways through TRABID to the mitochondria so as to regulates apoptosis. Therefore, it is possible that overexpression of hMI-ER1 can induce apoptosis and kill the cell during colony formation.

Although multiple mechanisms arising from the overexpression of hMI-ER1α and β in HeLa Tet-On cells may suppress colony formation efficiency, the establishment of the inducible hMI-ER1α and β Tet-On cell lines would be very useful to further investigate the role of hMI-ER1 in cell growth regulation. For example, it would be interesting to analyze the effect of overexpression of hMI-ER1α and β on the cell proliferation rate of the hMI-ER1α and β HeLa Tet-On cells using approaches such as the

MTS assay (Hu *et al.*, 1999). Moreover, fluorescence activated cell sorting (FACS) would be very useful to analyze cell cycle arrest and/or apoptosis of the hMI-ER1α and β HeLa Tet-On cells, before and after doxycycline treatment.

Since hMI-ER1α and β may function in a cell type specific manner, it would be rewarding to establish more Tet-on cell lines, such as in MCF-7 and/or NIH 3T3 cells, to study the role of hMI-ER1α and β in cell growth regulation.

SECTION IV

CHAPTER 7 GENERAL DISCUSSION, FUTURE DIRECTIONS AND CONCLUSION

7.1 GENERAL DISCUSSION

7.1.1 The multiple layers of transcriptional regulation of promoter activity of *hmi-er1* P2 promoter

This thesis has described the molecular cloning of the *hmi-er1* P1 and P2 promoters (Chapter 2). Further characterization of the *hmi-er1* P2 promoter provided information on the transcription regulation of the *hmi-er1* P2 promoter (Chapter 2 and 4). Studies here bring new insights into molecular mechanisms of the regulation of *hmi-er1* P2 promoter activity. DNA *cis*-acting regulatory elements in the *hmi-er1* P2 promoter clearly represent an important mechanism for the specific expression of this gene. The transcriptional activation of the *hmi-er1* P2 promoter may depend on cell types as well as co-expression of transcription regulators such as members of the Sp1 family (Chapter 2). hMI-ER1 was found to repress its own promoter activity through interaction with and interference with GC box recognition by Sp1 (Chapter 4), thus forming a negative feedback loop to tightly control the activity of hMI-ER1α and β, which function as transcriptional regulators (Chapter 3 and 4).

The accessibility of promoter elements to transcriptional regulators is also regulated at the level of chromatin assembly. The *hmi-er1* P2 promoter has a CpG island, and may be subjected to DNA methylation. The state of gene methylation is associated with chromatin structure and transcriptional repression (Bird and Wolffe, 1999). Methyl-CpG binding proteins such as MeCP2 and MBD3 are shown to recruit chromatin remodelling and modification complexes, including the NURD complex (Ng *et al.*, 1999; Tyler and Kadonaga, 1999; Zhang *et al.*, 1999). It has been proposed that DNA methylation and histone deacetylation might work together to establish a repressive chromatin environment to silence gene expression (Cameron *et al.*, 1999). Aberrant methylation on the CpG island of cell growth regulatory gene promoters has been found in certain cancers (reviewed in Esteller and Herman, 2002; Nephew and Huang, 2003). For example, aberrant methylation of the promoter of the tumour suppressor genes $p16^{INK4A}$ is closely related to tumour growth characteristics (Bender *et al.*, 1998; Cameron *et al.*, 1999; Issa *et al.*, 1994). Deregulated expression of hMI-ER1 in some breast carcinoma (Paterno *et al.*, 1998; and data at http://genome-www5.stanford.edu/cgi-bin/SMD/source/expressionSearch?option=cluster&criteria=Hs.222746&dataset=2&organism=Hs) may be due to aberrant methylation of the promoter CpG island of *hmi-er1* in those cancers.

At the moment, knowledge of the transcriptional regulation of the *hmi-er1* is far from complete. The finding of two promoters (Chapter 2 and Paterno *et al.*, 2002) adds additional complexity to the expression regulation of *hmi-er1*. Moreover, there may also be other enhancers located beyond the DNA region we investigated in this study.

7.1.2 hMI-ER1-containing complexes

It is becoming increasingly clear that protein function is not isolated in cell, but rather proteins work together with other associated protein partners in multi-subunit complexes. For example, RB was shown to associate with members of the E2F family as well as other cellular proteins such as HDAC1 to form transcription repression complexes that function in cell growth regulation (reviewed in Morris and Dyson, 2001). Within a protein complex, each individual protein may have a particular specialized function that contributes to the overall activity of the complex. This specialized function may be dependent on the interaction with neighbouring proteins in the complex which may, for example, lead to modulation of protein activity through conformational changes and/or post-translational modifications. Multi-protein complexes are emerging as important entities of biological activity inside cells that serve to create functional diversity by contextual combination of gene products and, at the same time, organize the large number of different proteins into functional units.

In the present studies, hMI-ER1 was found to associate with HDAC1-containing complexes (Chapter 3). Prominent examples of HDAC1-containing complexes involving chromatin modification and transcriptional regulation are mSin3A, NURD and CoREST complexes. The mSin3A complex consists of several core protein components and some additional binding protein components. The core is built from mSin3, HDAC1, HDAC2, RBAP46, RBAP48, SAP18 and SAP30 proteins. mSin3A itself is a large multi-domain protein that most likely forms the scaffold upon which the remainder of the complex

assembles (Ayer, 1999). The core protein components of NURD complexes consist of Mi-2, HDAC1, HDAC2, RBAP46, RBAP48, and one of the MTA family proteins including MTA1, MTA2 and MTA3 (Fujita *et al.*, 2003; Guschin *et al.*, 2000; Knoepfler and Eisenman, 1999). The ATPase activity of Mi-2 can increase the efficiency of histone deacetylation by NURD complexes (Guschin *et al.*, 2000; Tong *et al.*, 1998; Xue *et al.*, 1998b), which suggests that chromatin remodelling facilitates the access of the HDACs in NURD to the histone tails. Interestingly, MTA1, MTA2 and MTA3, all containing ELM2 and SANT domains, have been recently shown to be components of each *bona fide* distinct NURD complexe (Fujita *et al.*, 2003; Tong *et al.*, 1998; Wade *et al.*, 1999; Xue *et al.*, 1998b; Yao and Yang, 2003; Zhang *et al.*, 1999). Thus, the NURD complexes consist of several distinct forms. Another distinct HDACs-containing complex is the CoREST complex. CoREST is an ELM2 and SANT domain-containing protein and forms unique complexes different from NURD and mSin3 complexes (You *et al.*, 2001).

The observations that each of MTA1, MTA2, MTA3, and CoREST form distinct HDAC1 complexes suggested that the ELM2 and SANT domain-containing proteins may be sufficient to perform an as yet unknown function (chromatin targeting?) in HDAC1-containing complexes. Indeed, most recently, the ELM2 and SANT domain-containing region of hMI-ER1 was found to bind specifically to unacetylated histone protein H4 (Paterno *et al.*, unpublished data). Thus, it would be possible that the ELM2 and/or SANT domain of MTA1, MTA2, MTA3, or CoREST have similar histone binding activity. It is possible that hMI-ER1 is also a key subunit of *bona fide* unique HDAC1-containing complex which is different from MTA1, MTA2, MTA3 and CoREST related

complexes. To confirm this hypothesis, immunoprecipitation and Western blot analysis could be useful to address the presence or the absence of molecular markers of those complexes mentioned above, such as MTA1, MTA2, MTA3 and CoREST, in the hMI-ER1-HDAC1 complex. Alternatively, this question could also be addressed by the combination of classical biochemical fractionation techniques for the enrichment of hMI-ER1 related complexes with the large-scale purification and identification of distinct hMI-ER1 complexes by mass spectrometry and bioinformatics (see section 6.2.4).

hMI-ER1 was also found to interact with multiple other proteins, including Sp1 (Chapter 4), RB (Chapter 5), HSP40 (Chapter 5), TRABID (Chapter 5), ERα (Savicky *et al.*, unpublished data), CBP (Blackmore *et al.*, unpublished data), and core histone proteins (Paterno *et al.*, unpublished data). It is possible that interaction of one of those proteins with hMI-ER1 may affect the binding of another interacting partner with hMI-ER1 through competition for the binding surface of hMI-ER1. Thus, hMI-ER1 may form multiple distinct complexes with each interacting partner. Alternatively, hMI-ER1 could assemble with all these interacting partners together to form a single transcription regulatory complex. Although the interactions of those proteins with hMI-ER1 were investigated by *in vitro* and *in vivo* protein-protein interacting assays, it is still possible that some of the interactions between them are indirect, or one may act as a bridge for binding of all the others. Since there are multiple endogenous proteins in the cells and in the reticulocyte TNT lysates, the interaction detected in these systems may be an indirect interaction through other protein(s) in cells or in reticulocyte TNT lysates. It would be interesting to know whether interactions of all these proteins with hMI-ER1 are direct or

indirect. Protein-protein interaction assays using both purified proteins and then Western blot analysis may address this question.

7.1.3 Does hMI-ER1 function as transcriptional repressor or activator?

hMI-ER1 was found to function as a transcriptional repressor through recruitment of HDAC1 and interacting with Sp1 and interference with GC box recognition by Sp1 (Chapter 3 and 4); however, it cannot be excluded that it may function as a transcriptional activator under other conditions. The N-terminal region of hMI-ER1 contains an acidic activation domain, which may be implicated in transcriptional activation. Indeed, hMI-ER1α fused to GAL4 DBD from pAS2-1-hMI-ER1 alone can activate reporter genes in yeast (Figure 5.2B), and this activity was located in the N-terminal region which contains the acidic activation domain (Figure 5.2C). These observations suggested that the N-terminal region of hMI-ER1 may recruit the transcriptional machinery in yeast, and thus is implicated in transcription activation. There is evidence that *Xenopus* MI-ER1 functions as a powerful transcriptional activator (Paterno *et al.*, 1997). Recently, the N-terminal region of hMI-ER1 was found to interact with the histone acetyltransferase CBP (Blackmore *et al.*, unpublished data). CBP acetylates histone tails and changes the conformation of chromatin structure, which is correlated with transcription activation (reviewed in Carrozza *et al.*, 2003; Roth *et al.*, 2001). Therefore, the interaction of hMI-ER1 with CBP further suggests that hMI-ER1 may function as transcriptional activator under certain conditions. Alternatively, it is possible that hMI-ER1 binds to CBP and

regulates its HAT activity. It is, of course, also possible that CBP interacts with hMI-ER1 and acetylates hMI-ER1 potentially regulating its functions.

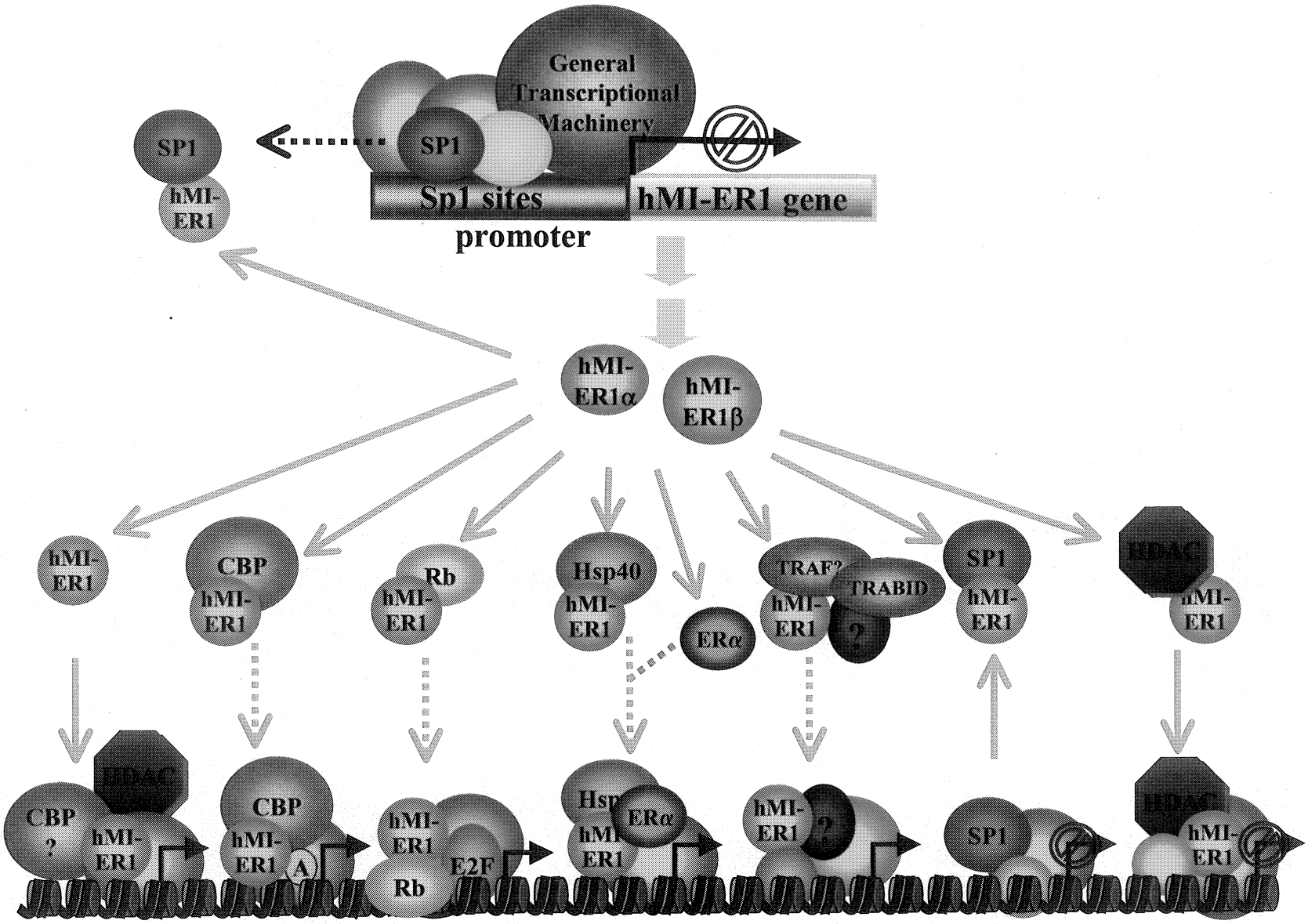
The ability to both repress and activate transcription has been reported for a number of other transcription factors, including the nuclear hormone receptors (reviewed in Aranda and Pascual, 2001). In most cases, differential recruitment of coactivator or corepressor complexes determines the different effect on transcription by nuclear hormone receptors (reviewed in Aranda and Pascual, 2001). It is possible that hMI-ER1 recruits distinct transcriptional regulatory complexes in response to different signals. The transcriptional activation or repression function of hMI-ER1 may be dependent on the subunits of the complexes recruited by hMI-ER1, for example, the presence of HDAC1 and/or CBP.

Recently, transcriptional activation and repression have been shown to be more closely integrated than previously thought. Examples of corepressors and coactivators coexisting in a single regulatory unit have been recently described. For example, the N-CoR and the coactivator CBP have been shown to bind the homeodomain heterodimer PBX-HOX simultaneously, while protein kinase A (PKA) stimulation of CBP has been found to facilitate the switch from transcriptional repression to activation in this system (Asahara *et al.*, 1999; Saleh *et al.*, 2000). Therefore, it is also possible that hMI-ER1 associates with corepressors, such as HDAC1, and coactivators, such as CBP, together in the same complex in cells, and transcription activation and/or repression may be controlled not only by the protein context in a particular cell type but also by signal transduction pathways, following growth factor or cytokine stimulation.

7.1.4 A model of the regulation of *hmi-er1* P2 promoter activity and putative transcriptional functions of hMI-ER1

Figure 7.1 illustrates a model is for the regulation of *hmi-er1* P2 promoter activity and transcriptional functions of hMI-ER1, based on the results obtained in the laboratory. Firstly, Sp1 binds to the *hmi-er1* P2 promoter of *hmi-er1* and activates transcription (Chapter 2). hMI-ER1 can recruit with HDAC1 and repress transcription (Chapter 3). hMI-ER1 also recruits transcription factor Sp1 and interferes with GC box recognition by Sp1, and could thereby regulate Sp1-target gene expression (Chapter 4). hMI-ER1 interacts with RB (Chapter 5), which may enhance RB s transcriptional repressive activity. Interaction of hMI-ER1 with HSP40 (Chapter 5) may be implicated in assembly with nuclear hormone receptors such as ERα. Most recently, hMI-ER1 was found to be specifically associated with ERα (Savicky *et al.*, unpublished data). Association of hMI-ER1 with ERα may modulate the transactivation functions of ERα and regulate essential ERα target gene expression. hMI-ER1 interacts with TRABID (Chapter 5), which may recruit hMI-ER1 into TRABID-TRAF signalling complexes and activate hMI-ER1 to function as a transcription regulator. hMI-ER1 recruits CBP (Blackmore *et al.*, unpublished data), which may oppose HDAC1 function to provide precise regulation of transcriptional activity. Most recently, hMI-ER1 was found to interact with unacetylated histone H4 (Paterno *et al.*, unpublished data), and thus may direct multiple subunit complexes to target promoters in response to distinct cellular signals. This discovery strongly supports the hypothesis that hMI-ER1 functions as an important transcription

Figure 7.1 Model of putative mechanisms of transcriptional regulation of hMI-ER1. Sp1 binds to the *hmi-er1* P2 promoter of *hmi-er1* and activate transcription. Interaction of hMI-ER1 with HDAC1 results in histone deacetylation and transcription repression. hMI-ER1 interacts with transcription factor Sp1 and interferes with GC box recognition by Sp1, thus may thereby regulate Sp1-target gene expression. Interaction of hMI-ER1 with RB may enhance or inhibit RB activity as transcription regulator. HSP40 may direct complexes assembling of hMI-ER1 with nuclear hormone receptors such as ERα, which may regulate essential ERα target gene expression. Interaction of hMI-ER1 with TRABID may bring hMI-ER1 into TRABID-TRAFs signalling complexes and activate hMI-ER1 as transcription regulators. Recruitment of CBP by hMI-ER1 may result in the acetylation of histone and activate target gene transcription. Binding of hMI-ER1 to the unacetylated histone protein H4unacetylated histone may direct multiple subunit complexes to the target promoter. hMI-ER1 was also found to interfere with DNA recognition by Sp1 to its own P2 promoter and repress its own P2 promoter activity, thus forming a negative feedback loop to tightly control the activity of hMI-ER1.



Transcriptional regulation of target genes

regulator and could target chromatin and regulate gene transcription. Finally, hMI-ER1 interferes with DNA recognition by Sp1 on its own promoter and represses its own gene promoter activity (Chapter 4), thus forming a negative feedback loop to tightly control the activity of hMI-ER1. This ensures that hMI-ER1 regulates the expression of other genes at appropriate times.

7.2 Future directions

This thesis has described the molecular cloning of *hmi-er1* promoters. Characterization of the *hmi-er1* P2 promoter provided information on its transcriptional regulation (Chapter 2). In the present study, hMI-ER1 was found to repress transcription on the TK minimal promoter of G5TKCAT through recruiting HDAC1 activity (Chapter 3). Moreover, hMI-ER1 was found to repress promoter activity on its own P2 promoter through interacting with Sp1 and interfering with GC box recognition by Sp1 (Chapter 4). Furthermore, multiple hMI-ER1 interacting protein partners were identified using different approaches (Chapter 5 and 6). Most interestingly, overexpression of hMI-ER1 was found to have a profound impact on cell growth (Chapter 6 and Paterno *et al.*, unpublished data). However, at the moment we are just able to see the tip of the iceberg. To fully understand the molecular mechanisms of hMI-ER1 expression and function, more experiments must be.

7.2.1 Further characterization of *hmi-er1* promoters

Cloning and characterization of the *hmi-er1* P2 promoter revealed that Sp1 binds to the promoter and plays a positive role in the regulation of promoter activity (Chapter 2). Besides Sp1 binding sites, other transcription regulator binding sites, including two TEF-2 sites, an HT4F site, and an ERE half-site, were also predicted in the P2 promoter (Chapter 3). Approaches such as EMSA (Chapter 2 and 4) and chromatin immunoprecipitation (ChIP) (Mazumdar *et al.*, 2001) can be used to investigate whether the related transcription factors such as TEF-2, HT4T and ERα bind to the promoter *in vitro* and *in vivo*. The ERE half-site in the *hmi-er1* P2 promoter, which suggests that *hmi-er1* may be an estrogen-responsive gene, is particularly interesting. Many estrogen-response genes, such as MTA3, were found to be downregulated in ER-negative breast tumours (Fujita *et al.*, 2003). Since hMI-ER1 was found to interact with ERα and regulate ERα-mediated transcription (Savicky *et al.*, unpublished data), hMI-ER1 may play an important role in estrogen-signalling pathways. It would be important to know whether the ERα-hMI-ER1 pathway is deregulated in hormone-resistant cancers, and whether the expression of hMI-ER1 is correlated to the status of ERα in breast carcinomas. It would therefore be very interesting to analyze the expression of hMI-ER1 and the correlation with ERα status in more breast carcinomas samples. This is presently under investigation in our laboratory.

To investigate whether other unpredicted transcription regulators bind to the *hmi-er1* P2 promoter region and regulate the promoter activity of *hmi-er1*, an expression cDNA library screening approach (Vinson *et al.*, 1988) may be conducted to identify and

isolate the *hmi-er1* promoter-binding protein(s) using the labelled P2 promoter as a probe. Moreover, it would be rewarding to characterize the *hmi-er1* P1 promoter using similar approaches.

7.2.2 Functional implication of the interactions with HSP40, TRABID and RB

Pioneering work in this study revealed that hMI-ER1 interacts with RB, HSP40 (Chapter 5), TRABID (Chapter 5) and RB (Chapter 6). These preliminary studies provided more clues to study the function of hMI-ER1. ERα plays a crucial role in normal breast development and is also linked to development and progression of mammary carcinoma (Hortobagyi, 1998; Osborne, 1998). The transcriptional activity of ERα is regulated by hMI-ER1 (Savicky *et al.*, personal communication). Since HSP40 may play a critical role in assembling an ERα-regulatory complex with hMI-ER1, the interaction of hMI-ER1 and HSP40 may be a critical event in regulation of ERα functions, and thus may have potential clinical significance. Further study of the interaction with HSP40 would involve investigating: (1) if hMI-ER1 interacts with HSP40 in mammalian cells and thus to exclude the possibility that the observed interaction was an artifact of overexpression; (2) if HSP40 directs ERα complex assembly with hMI-ER1; (3) the role of the hMI-ER1-HSP40-ERα pathway in ERα signalling; (4) if ER-resistant cancer results from the misregulation of hMI-ER1-HSP40-ERα pathway. Similarly,

hMI-ER1 interacts with TRABID, which may be involved in the regulation of cell death through the TRABID-TRAF signalling pathway. Further studies examining the

interaction with TRABID would involve determining: (1) whether hMI-ER1 interacts with TRABID in mammalian cells and thus to exclude the possibility that the observed interaction was an artifact of overexpression; (2) whether TRABID-TRAF signalling complexes recruit hMI-ER1 and activate hMI-ER1; and (3) whether overexpression of hMI-ER1 would sensitize cells to cytokine-induced apoptosis, through the TRABID-TRAF-dependent pathway.

Further study of the interaction with RB would involve investigating: (1) if endogenous complexes containing hMI-ER1 and RB exist in the cell, thus excluding the possibility that the observed interaction was an artifact of overexpression; (2) what post-translational modification of the RB B-pocket domain allow interactions with hMI-ER1; (3) which domain of hMI-ER1 interacts with RB; (4) whether hMI-ER1α and β affects the cell cycle-dependent interaction of RB with E2F on E2F binding sites; (5) whether hMI-ER1 affects the E2F target gene such as cyclin E expression; (6) whether hMI-ER1 binds to the E2F target promoters through RB-dependent and/or -independent mechanisms; and (7) whether the cell growth regulation by hMI-ER1 is dependent on the interaction with RB, i.e. a form of hMI-ER1 that fails to interact with RB may affect the cell growth regulation.

7.2.3 hMI-ER1 transcriptional target genes

The functional study of hMI-ER1 revealed that hMI-ER1 functions as a transcriptional regulator (Chapter 3 and 4; Savicky *et al.*, unpublished data; Blackmore *et al.*, unpublished data). Thus, a systematic approach may be used to identify the

transcriptional targets of hMI-ER1. The global view of gene expression profiles regulated by hMI-ER1α and β may provide further clues about their transcriptional target genes in cells. Microarray and/or serial analysis of gene expression (SAGE) approaches (Pollock, 2002) may be utilized to assess changes in RNA expression upon hMI-ER1α and β overexpression in cells. Recently, hMI-ER1α and β HeLa Tet-On cell lines were successfully established in the laboratory (Chapter 6). These cell lines would be very useful for the purpose of studying global gene expression profiles affected by induced hMI-ER1α and β overexpression. Recently, new technologies have been developed to identify novel target gene promoters for transcription factors in mammalian cells, such as modified ChIP (Johnson and Bresnick, 2002; Weinmann and Farnham, 2002).

7.2.4 Do hMI-ER1α and β complexes have distinct components?

Quite often proteins participate in more than one complex, or do so in different subcellular compartments. Different subcellular compartments contain compartment-specific subsets of gene products in order to provide suitable biochemical environments, in which they exert their particular function.

hMI-ER1β is a nuclear protein, while hMI-ER1α shuttles between the nucleus and cytoplasm, which may suggest distinct functions for the two isoforms (Paterno *et al.*, 2002 and unpublished data). Thus, the identification of subsets of protein complexes of hMI-ER1α and β at the subcellular level would also provide more clues towards the understanding of the precise function of hMI-ER1α and β. Interestingly, hMI-ER1α has the LXXLL NR binding motif in its C-terminal region, which may function to interact

with and sequester NR in the cytoplasm. Recently, a variant of MTA1, MTA1s, was found to contain the LXXLL motif and localize in the cytoplasm. MTA1s inhibits nuclear signalling by sequestering ER in the cytoplasm, which is distinct from the nuclear form of MTA1 (Kumar *et al.*, 2002). However, so far, proteins that have been identified as interacting partners, including ERα, have been found to interact with both hMI-ER1α and β. Therefore, isolation of the distinct subcellularly localized hMI-ER1α and β complexes may help to distinguish their different components and functions. Approaches to isolate the protein complexes include recombinant proteins, epitope-tagged proteins and antibody methods (reviewed in Bauer and Kuster, 2003). Since specific antibodies against the C-terminal portion of distinct hMI-ER1α or β have been generated in the laboratory, the antibody immunoprecipitation method for the isolation of hMI-ER1α or β related complexes would be very useful and feasible. Moreover, GST fusion proteins containing the distinct C-terminal portions of hMI-ER1α or β also have been generated in the laboratory. Therefore, GST pull-down isolation methods are also suitable for the isolation of distinct hMI-ER1α and β complexes in the laboratory. The combination of classic biochemical techniques for the isolation of hMI-ER1α or β related complexes from particular subcellular structures with the large-scale purification and identification of distinct hMI-ER1α and β complexes by mass spectrometry and bioinformatics would provide more information on their components.

7.2.5 Role in cell growth control and tumorigenesis

Overexpression of hMI-ER1α and β suppresses colony formation efficiency in multiple cell lines (Paterno *et al.*, unpublished data; Huang *et al.*, personal communication). Recently, hMI-ER1α and β HeLa Tet-On cell lines were established in the laboratory and preliminary studies showed that the induced expression of hMI-ER1α and β suppressed cell colony formation efficiency in these cells (Chapter 6). Further studies are required to determine whether upregulated hMI-ER1α and β induce inhibition of proliferation and/or cell death. It would be particularly interesting to know whether induced expression of hMI-ER1α and β by doxycycline would inhibit tumour growth in nude mice.

It would be rewarding to know which domain and functions of hMI-ER1 are critical for colony formation efficiency inhibition. The ELM2 and SANT domains of hMI-ER1 were shown to play a role in transcriptional repression through distinct mechanisms (Chapter 3 and 4). The mutant forms of aa ^{214}W and ^{227}FL in the ELM2 domain fail to recruit HDAC1 activity and repress transcription (Chapter 3). A deletion mutant retaining the SANT domain region (aa 287-357) was found to interact with Sp1, while further deletion of this region at the C-terminal end to remove residues from aa 331 to 357 (^{331}K to ^{357}T), which includes a highly conserved lysine (K) within the SANT domain at position 331 and the C-terminal 26 aa, completely abolished the interaction with Sp1 and the interference with the GC box recognition by Sp1 (Chapter 4). If the transcription repression of hMI-ER1 is important for its role in cell growth regulation, the mutant form of hMI-ER1 in the ELM2 domain and/or SANT domain may then affect the

cell growth regulation function by hMI-ER1. Most recently, mutagenesis analysis of the proline-rich region of *Xenopus* MI-ER1 (^{365}PSPPP) revealed that ^{365}P in the proline-rich region is required for the effects of xMI-ER1 on embryonic development and mesoderm induction (Teplitsky *et al.*, 2003). Therefore, it would be interesting to investigate the role of the ELM2 domain, the SANT domain, and the proline-rich PSPPP motif in cell growth regulation through deletion or point mutation of those conserved residues including aa ^{214}W, ^{227}FL, ^{331}K, and ^{376}PSPPP of hMI-ER1.

hMI-ER1 was found to have ubiquitous, but low level, expression in normal human tissues (Paterno *et al.*, 1998; Paterno *et al.*, 2002). Breast carcinoma cell lines and tumours, on the other hand, showed elevated levels (Paterno 1998), suggesting that *hmi-er1* expression is associated with the neoplastic state in human breast carcinoma. However, it remains to be determined whether the high levels of *hmi-er1* expression observed are causative or correlative in the tumorigenesis. That overexpression of hMI-ER1α and β in cancer cells suppresses colony formation efficiency is quite unexpected and may be a very complicated consequence of hMI-ER1α and β functions. However, it is clear now that there are at least 12 transcripts which encode 6 distinct hMI-ER1 protein isoforms. The expression of these transcripts is complex, as well as the subcellular localization of proteins in normal tissues and tumour tissues (Paterno *et al.*, unpublished data). Moreover, each protein isoform may have different, or even opposite functions. For example, p73, which is a member of the p53 family, was shown to have multiple isoforms which play different roles in cell growth regulation (reviewed in Irwin and Kaelin, 2001). Two of *hmi-er1* isoforms (N1α and N1β) contain a 25 aa cysteine-rich

region at the N-terminus (Paterno *et al.*, 2002), which may function in protein-protein interactions (Park *et al.*, 2001) or form zinc fingers (Laity *et al.*, 2001). It would be rewarding to investigate whether these cysteine-rich-region-containing isoforms have similar or distinct cell growth functions with hMI-ER1α (N3α) and hMI-ER1β (N3β) isoforms, which lack this cysteine-rich region and which were used in this study.

To obtain a better understanding of the role of hMI-ER1α and β in tumorigenesis, it would be rewarding to examine the protein expression and subcellular localization of each protein isoform of hMI-ER1, as well as correlate with other related molecules such as RB and ERα in more tumour and normal tissue samples. Tissue microarray approaches may be very useful for this purpose to examine a large number of samples in more efficient ways (Packeisen *et al.*, 2003).

7.2.6 Knockout specific hMI-ER1 isforms in mice

Although preliminary observation indicated that overexpression of hMI-ER1α and β suppresses colony formation efficiency (Paterno *et al.*, unpublished data), the role of physiological levels of these hMI-ER protein isoforms is unknown. Gene-targeting studies are which would address the role of mouse MI-ER1 in cell growth regulation. Initial studies would be to disrupt the mouse MI-ER1α or β isoforms by targeting the appropriate C-terminus-encoding exon, leaving potential expression of the other mouse MI-ER1 isoforms intact. Although exogenously expressed MI-ER1α or β protein isoforms both suppress colony formation efficiency, it is possible that MI-ER1α and β carry out specialized, non-overlapping functions *in vivo*.

Xenopus MI-ER1 was induced as an early response gene by FGF treatment and functions as negative regulator of the FGF pathway (Paterno *et al.*, 1997; Teplitsky *et al.*, 2003). An increased expression of hMI-ER1α induced by FGF was observed in mammalian cells (Paterno *et al.*, personal communication), suggesting that hMI-ER1 may play a role in the FGF signaling pathway in mammalian cells as well. Since FGF signalling pathways are involved in many biological processes including embryonic development and angiogenesis (reviewed in Wilkie *et al.*, 2002), it would be interesting to investigate whether mouse *mi-er1* knock-out would also affect those FGF signalling pathways.

7.3 Conclusion

The study of the *hmi-er1* P2 promoter has provided some insight into its structure and regulation. The functional study of hMI-ER1α and β isoforms revealed that hMI-ER1α and β function as transcriptional repressors on the G5TKCAT promoter by recruitment of HDAC1 *via* the ELM2 domain. hMI-ER1α and β also function as transcriptional repressors on their own promoter, through interaction with Sp1 and interference with the GC box recognition by Sp1 *via* the SANT-domain-containing region. In the present study, hMI-ER1α and β were also found to interact with HSP40, TRABID and RB, which have distinct functions. There is also preliminary evidence that hMI-ER1 plays a role in cell growth regulation. A likely scenario is that hMI-ER1 coordinates multiple interacting protein partners, thereby linking signal transduction, transcription and cell growth regulation. Further study of hMI-ER1α and β related complexes and function may help to decipher the composition of multi-protein signalling

complexes involving hMI-ER1α and β, which would eventually contribute to reconstruction of signal transduction pathways and benefit the specific therapeutic targeting of diseases such as cancer.

REFERENCES

Aasland,R., Stewart,A.F., and Gibson,T. (1996). The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB. Trends Biochem. Sci. *21*, 87-88.

Ahringer,J. (2000). NuRD and SIN3 histone deacetylase complexes in development. Trends Genet. *16*, 351-356.

Alberts,B., Johnson,A., Lewis,J., Raff,M., Roberts,K., and Walter,P. (2002). Molecular biology of the cell. New York : Garland Science).

Alexandropoulos,K., Cheng,G., and Baltimore,D. (1995). Proline-rich sequences that bind to Src homology 3 domains with individual specificities. Proc. Natl. Acad. Sci. U. S. A *92*, 3110-3114.

Allsopp,R.C., Vaziri,H., Patterson,C., Goldstein,S., Younglai,E.V., Futcher,A.B., Greider,C.W., and Harley,C.B. (1992). Telomere length predicts replicative capacity of human fibroblasts. Proc. Natl. Acad. Sci. U. S. A *89*, 10114-10118.

Antequera,F. and Bird,A. (1993). Number of CpG islands and genes in human and mouse. Proc. Natl. Acad. Sci. U. S. A *90*, 11995-11999.

Aranda,A. and Pascual,A. (2001). Nuclear hormone receptors and gene expression. Physiol Rev. *81*, 1269-1304.

Archakov,A.I., Govorun,V.M., Dubanov,A.V., Ivanov,Y.D., Veselovsky,A.V., Lewi,P., and Janssen,P. (2003). Protein-protein interactions as a target for drugs in proteomics. Proteomics *3*, 380-391.

Artandi,S.E. and DePinho,R.A. (2000). A critical role for telomeres in suppressing and facilitating carcinogenesis. Curr. Opin. Genet. Dev. *10*, 39-46.

Asahara,H., Dutta,S., Kao,H.Y., Evans,R.M., and Montminy,M. (1999). Pbx-Hox heterodimers recruit coactivator-corepressor complexes in an isoform-specific manner. Mol. Cell Biol. *19*, 8219-8225.

Ashraf,S.I. and Ip,Y.T. (1998). Transcriptional control: repression by local chromatin modification. Curr. Biol. *8*, R683-R686.

Auerbach,W. and Auerbach,R. (1994). Angiogenesis inhibition: a review. Pharmacol Ther. *63*, 265-311.

Ayer,D.E. (1999). Histone deacetylases: transcriptional repression with SINers and NuRDs. Trends Cell Biol. *9*, 193-198.

Baker,S.J., Markowitz,S., Fearon,E.R., Willson,J.K.V., and Vogelstein,B. (1990). Suppression of human colorectal carcinoma cell growth by wild-type p53. Science *249*, 912-915.

Baker,S.J. and Reddy,E.P. (1998). Modulation of life and death by the TNF receptor superfamily. Oncogene *17*, 3261-3270.

Ballestar,E. and Wolffe,A.P. (2001). Methyl-CpG-binding proteins. Targeting specific gene repression. Eur J Biochem. *268*, 1-6.

Barlev,N.A., Liu,L., Chehab,N.H., Mansfield,K., Harris,K.G., Halazonetis,T.D., and Berger,S.L. (2001). Acetylation of p53 activates transcription through recruitment of coactivators/histone acetyltransferases. Mol. Cell *8*, 1243-1254.

Bauer,A. and Kuster,B. (2003). Affinity purification-mass spectrometry. Powerful tools for the characterization of protein complexes. Eur. J. Biochem. *270*, 570-578.

Bender,C.M., Pao,M.M., and Jones,P.A. (1998). Inhibition of DNA methylation by 5-aza-2'-deoxycytidine suppresses the growth of human tumor cell lines. Cancer Res. *58*, 95-101.

Berger,S.L. (2002). Histone modifications in transcriptional regulation. Curr. Opin. Genet. Dev. *12*, 142-148.

Bestor,T.H. (1988). Cloning of a mammalian DNA methyltransferase. Gene *74*, 9-12.

Billon,N., Carlisi,D., Datto,M.B., van Grunsven,L.A., Watt,A., Wang,X.F., and Rudkin,B.B. (1999). Cooperation of Sp1 and p300 in the induction of the CDK inhibitor p21WAF1/CIP1 during NGF-mediated neuronal differentiation. Oncogene *18*, 2872-2882.

Bird,A.P. and Wolffe,A.P. (1999). Methylation-induced repression--belts, braces, and chromatin. Cell *99*, 451-454.

Birkenkamp,K.U. and Coffer,P.J. (2003). Regulation of cell survival and proliferation by the FOXO (Forkhead box, class O) subfamily of Forkhead transcription factors. Biochem. Soc. Trans. *31*, 292-297.

Black,A.R., Black,J.D., and Azizkhan-Clifford,J. (2001). Sp1 and kruppel-like factor family of transcription factors in cell growth regulation and cancer. J. Cell Physiol *188*, 143-160.

Blackburn,E.H. (2001). Switching and signaling at the telomere. Cell *106*, 661-673.

Blair,W.S., Bogerd,H.P., Madore,S.J., and Cullen,B.R. (1994). Mutational analysis of the transcription activation domain of RelA: identification of a highly synergistic minimal acidic activation module. Mol. Cell Biol. *14*, 7226-7234.

Bouillet,P. and Strasser,A. (2002). BH3-only proteins - evolutionarily conserved proapoptotic Bcl-2 family members essential for initiating programmed cell death. J Cell Sci. *115*, 1567-1574.

Bouwman,P. and Philipsen,S. (2002). Regulation of the activity of Sp1-related transcription factors. Mol Cell Endocrinol. *195*, 27-38.

Boyer,L.A., Langer,M.R., Crowley,K.A., Tan,S., Denu,J.M., and Peterson,C.L. (2002). Essential role for the SANT domain in the functioning of multiple chromatin remodeling enzymes. Mol Cell *10*, 935-942.

Brand,M., Yamamoto,K., Staub,A., and Tora,L. (1999). Identification of TATA-binding protein-free TAFII-containing complex subunits suggests a role in nucleosome acetylation and signal transduction. J. Biol. Chem. *274*, 18285-18289.

Braunstein,M., Sobel,R.E., Allis,C.D., Turner,B.M., and Broach,J.R. (1996). Efficient transcriptional silencing in Saccharomyces cerevisiae requires a heterochromatin histone acetylation pattern. Mol. Cell Biol. *16*, 4349-4356.

Brivanlou,A.H. and Darnell,J.E., Jr. (2002). Signal transduction and the control of gene expression. Science *295*, 813-818.

Brubaker,K., Cowley,S.M., Huang,K., Loo,L., Yochum,G.S., Ayer,D.E., Eisenman,R.N., and Radhakrishnan,I. (2000). Solution structure of the interacting domains of the Mad-Sin3 complex: implications for recruitment of a chromatin-modifying complex. Cell *103*, 655-665.

Buchwalter,G., Gross,C., and Wasylyk,B. (2004). Ets ternary complex transcription factors. Gene *324*, 1-14.

Burgess,W.H., Shaheen,A.M., Ravera,M., Jaye,M., Donohue,P.J., and Winkles,J.A. (1990). Possible dissociation of the heparin-binding and mitogenic activities of heparin-binding (acidic fibroblast) growth factor-1 from its receptor-binding activities by site-directed mutagenesis of a single lysine residue. J. Cell Biol. *111*, 2129-2138.

Bussolino,F., Mantovani,A., and Persico,G. (1997). Molecular mechanisms of blood vessel formation. Trends Biochem. Sci. *22*, 251-256.

Butler,J.E. and Kadonaga,J.T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. Genes Dev. *16*, 2583-2592.

Cameron,E.E., Bachman,K.E., Myohanen,S., Herman,J.G., and Baylin,S.B. (1999). Synergy of demethylation and histone deacetylase inhibition in the re- expression of genes silenced in cancer. Nat. Genet. *21*, 103-107.

Candau,R., Moore,P.A., Wang,L., Barlev,N., Ying,C.Y., Rosen,C.A., and Berger,S.L. (1996). Identification of human proteins functionally conserved with the yeast putative adaptors ADA2 and GCN5. Mol. Cell Biol. *16*, 593-602.

Cantrell,D.A. (2001). Phosphoinositide 3-kinase signalling pathways. J Cell Sci. *114*, 1439-1445.

Cao,Z., Henzel,W.J., and Gao,X. (1996). IRAK: a kinase associated with the interleukin-1 receptor. Science *271*, 1128-1131.

Carballada,R., Yasuo,H., and Lemaire,P. (2001). Phosphatidylinositol-3 kinase acts in parallel to the ERK MAP kinase in the FGF pathway during Xenopus mesoderm induction. Development *128*, 35-44.

Carrion,A.M., Link,W.A., Ledo,F., Mellstrom,B., and Naranjo,J.R. (1999). DREAM is a Ca2+-regulated transcriptional repressor. Nature *398*, 80-84.

Carrozza,M.J., Utley,R.T., Workman,J.L., and Cote,J. (2003). The diverse functions of histone acetyltransferase complexes. Trends Genet. *19*, 321-329.

Cesareni,G., Panni,S., Nardelli,G., and Castagnoli,L. (2002). Can we infer peptide recognition specificity mediated by SH3 domains? FEBS Lett. *523*, 38-44.

Chadjichristos,C., Ghayor,C., Kypriotou,M., Martin,G., Renard,E., Ala-Kokko,L., Suske,G., de Crombrugghe,B., Pujol,J.P., and Galera,P. (2003). Sp1 and Sp3 transcription factors mediate interleukin-1 beta down-regulation of human type II collagen gene expression in articular chondrocytes. J Biol Chem. *278*, 39762-39772.

Chan,H.M., Krstic-Demonacos,M., Smith,L., Demonacos,C., and La Thangue,N.B. (2001). Acetylation control of the retinoblastoma tumour-suppressor protein. Nat Cell Biol. *3*, 667-674.

Chang,C., Norris,J.D., Gron,H., Paige,L.A., Hamilton,P.T., Kenan,D.J., Fowlkes,D., and McDonnell,D.P. (1999). Dissection of the LXXLL nuclear receptor-coactivator interaction motif using combinatorial peptide libraries: discovery of peptide antagonists of estrogen receptors alpha and beta. Mol. Cell Biol. *19*, 8226-8239.

Cheng,G., Cleary,A.M., Ye,Z.S., Hong,D.I., Lederman,S., and Baltimore,D. (1995). Involvement of CRAF1, a relative of TRAF, in CD40 signaling. Science *267*, 1494-1498.

Cheng,S.W., Davies,K.P., Yung,E., Beltran,R.J., Yu,J., and Kalpana,G.V. (1999). c-MYC interacts with INI1/hSNF5 and requires the SWI/SNF complex for transactivation function. Nat. Genet. *22*, 102-105.

Chiang,C.M. and Roeder,R.G. (1995). Cloning of an intrinsic human TFIID subunit that interacts with multiple transcriptional activators. Science *267*, 531-536.

Chow,K.N. and Dean,D.C. (1996). Domains A and B in the Rb pocket interact to form a transcriptional repressor motif. Mol Cell Biol. *16*, 4862-4868.

Christofori,G. and Semb,H. (1999). The role of the cell-adhesion molecule E-cadherin as a tumour- suppressor gene. Trends Biochem. Sci. *24*, 73-76.

Chung,J.Y., Park,Y.C., Ye,H., and Wu,H. (2002). All TRAFs are not created equal: common and distinct molecular mechanisms of TRAF-mediated signal transduction. J. Cell Sci. *115*, 679-688.

Coultas,L. and Strasser,A. (2003). The role of the Bcl-2 protein family in cancer. Semin Cancer Biol. *13*, 115-123.

Coussens,L.M. and Werb,Z. (1996). Matrix metalloproteinases and the development of cancer. Chem. Biol. *3*, 895-904.

Craig,T.A., Benson,L.M., Venyaminov,S.Y., Klimtchuk,E.S., Bajzer,Z., Prendergast,F.G., Naylor,S., and Kumar,R. (2002). The methal binding properties of DREAM: evidence for calcium mediated changes in DREAM structure. J. Biol. Chem. *277*, 10955-10966.

Cress,W.D. and Seto,E. (2000). Histone deacetylases, transcriptional control, and cancer. J. Cell Physiol *184*, 1-16.

Cress,W.D. and Triezenberg,S.J. (1991). Critical structural elements of the VP16 transcriptional activation domain. Science *251*, 87-90.

Cunliffe,V. and Smith,J.C. (1994). Specification of mesodermal pattern in Xenopus laevis by interactions between Brachyury, noggin and Xwnt-8. EMBO J. *13*, 349-359.

Datta,P.K., Raychaudhuri,P., and Bagchi,S. (1995). Association of p107 with Sp1: genetically separable regions of p107 are involved in regulation of E2F- and Sp1-dependent transcription. Mol. Cell Biol. *15*, 5444-5452.

Daujat,S., Bauer,U.M., Shah,V., Turner,B., Berger,S., and Kouzarides,T. (2002). Crosstalk between CARM1 Methylation and CBP Acetylation on Histone H3. Curr. Biol. *12*, 2090-2097.

Davie,J.R. and Chadee,D.N. (1998). Regulation and regulatory parameters of histone modifications. J. Cell Biochem. Suppl *30-31*, 203-213.

Debbas,M. and White,E. (1993). Wild-type p53 mediates apoptosis by E1A, which is inhibited by E1B. Genes Dev. *7*, 546-554.

DePinho,R.A. and Wong,K.K. (2003). The age of cancer: telomeres, checkpoints, and longevity. J. Clin. Invest *111*, S9-14.

Doetzlhofer,A., Rotheneder,H., Lagger,G., Koranda,M., Kurtev,V., Brosch,G., Wintersberger,E., and Seiser,C. (1999). Histone deacetylase 1 can repress transcription by binding to Sp1. Mol. Cell Biol. *19*, 5504-5511.

Donehower,L.A., Godley,L.A., Aldaz,C.M., Pyle,R., Shi,Y.P., Pinkel,D., Gray,J., Bradley,A., Medina,D., and Varmus,H.E. (1995). Deficiency of p53 accelerates mammary tumorigenesis in Wnt-1 transgenic mice and promotes chromosomal instability. Genes Dev. *9*, 882-889.

Dynan,W.S. and Tjian,R. (1985). Control of eukaryotic messenger RNA synthesis by sequence-specific DNA- binding proteins. Nature *316*, 774-778.

Esteller,M. and Herman,J.G. (2002). Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. J Pathol. *196*, 1-7.

Evans,P.C., Taylor,E.R., Coadwell,J., Heyninck,K., Beyaert,R., and Kilshaw,P.J. (2001). Isolation and characterization of two novel A20-like proteins. Biochem. J. *357*, 617-623.

Fernandes,I., Bastien,Y., Wai,T., Nygard,K., Lin,R., Cormier,O., Lee,H.S., Eng,F., Bertos,N.R., Pelletier,N., Mader,S., Han,V.K., Yang,X.J., and White,J.H. (2003). Ligand-dependent nuclear receptor corepressor LCoR functions by histone deacetylase-dependent and -independent mechanisms. Mol Cell. *11*, 139-150.

Ferreira,R., Magnaghi-Jaulin,L., Robin,P., Harel-Bellan,A., and Trouche,D. (1998). The three members of the pocket proteins family share the ability to repress E2F activity through recruitment of a histone deacetylase. Proc Natl Acad Sci U S A. *95*, 10493-10498.

Ferreira,R., Naguibneva,I., Pritchard,L.L., Ait-Si-Ali,S., and Harel-Bellan,A. (2001). The Rb/chromatin connection and epigenetic control: opinion. Oncogene *20*, 3128-3133.

Ferrell,J.E., Jr. (1996). MAP kinases in mitogenesis and development. Curr. Top. Dev. Biol. *33*, 1-60.

Fidler,I.J., Singh,R.K., Yoneda,J., Kumar,R., Xu,L., Dong,Z., Bielenberg,D.R., McCarty,M., and Ellis,L.M. (2000). Critical determinants of neoplastic angiogenesis. Cancer J. *6 Suppl 3*, S225-S236.

Fields,S. and Song,O. (1989). A novel genetic system to detect protein-protein interactions. Nature *340*, 245-246.

Fields,S. and Sternglanz,R. (1994). The two-hybrid system: an assay for protein-protein interactions. Trends Genet. *10*, 286-292.

Figeys,D. (2002). Functional proteomics: mapping protein-protein interactions and pathways. Curr Opin Mol Ther. *4*, 210-215.

Fink,A.L. (1999). Chaperone-mediated protein folding. Physiol Rev. *79*, 425-449.

Freiman,R.N. and Tjian,R. (2003). Regulating the regulators: lysine modifications make their mark. Cell *112*, 11-17.

Fujita,N., Jaye,D.L., Kajita,M., Geigerman,C., Moreno,C.S., and Wade,P.A. (2003). MTA3, a Mi-2/NuRD complex subunit, regulates an invasive growth pathway in breast cancer. Cell *113*, 207-219.

Gill,G., Pascal,E., Tseng,Z.H., and Tjian,R. (1994). A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAFII110 component of the Drosophila TFIID complex and mediates transcriptional activation. Proc. Natl. Acad. Sci. U. S. A *91*, 192-196.

Gillespie,L.L., Paterno,G.D., Mahadevan,L.C., and Slack,J.M. (1992). Intracellular signalling pathways involved in mesoderm induction by FGF. Mech. Dev. *38*, 99-107.

Glass,C.K. and Rosenfeld,M.G. (2000). The coregulator exchange in transcriptional functions of nuclear receptors. Genes Dev. *14*, 121-141.

Goodman,R.H. and Smolik,S. (2000). CBP/p300 in cell growth, transformation, and development. Genes Dev. *14*, 1553-1577.

Goodwin,E.C. and DiMaio,D. (2000). Repression of human papillomavirus oncogenes in HeLa cervical carcinoma cells causes the orderly reactivation of dormant tumor suppressor pathways. Proc Natl Acad Sci U S A. *97*, 12513-12518.

Gorlich,D. (1998). Transport into and out of the cell nucleus. EMBO J. *17*, 2721-2727.

Gossen,M. and Bujard,H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc Natl Acad Sci U S A. *89*, 5547-5551.

Gossen,M., Freundlieb,S., Bender,G., Muller,G., Hillen,W., and Bujard,H. (1995). Transcriptional activation by tetracyclines in mammalian cells. Science *269*, 1766-1769.

Grant,P.A., Duggan,L., Cote,J., Roberts,S.M., Brownell,J.E., Candau,R., Ohba,R., Owen-Hughes,T., Allis,C.D., Winston,F., Berger,S.L., and Workman,J.L. (1997). Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. Genes Dev. *11*, 1640-1650.

Gregory,P.D. and Horz,W. (1998). Chromatin and transcription--how transcription factors battle with a repressive chromatin environment. Eur. J. Biochem. *251*, 9-18.

Gregory,R.C., Taxman,D.J., Seshasayee,D., Kensinger,M.H., Bieker,J.J., and Wojchowski,D.M. (1996). Functional interaction of GATA1 with erythroid Kruppel-like factor and Sp1 at defined erythroid promoters. Blood *87*, 1793-1801.

Gross,A., McDonnell,J.M., and Korsmeyer,S.J. (1999). BCL-2 family members and the mitochondria in apoptosis. Genes Dev. *13*, 1899-1911.

Grunstein,M. (1997). Histone acetylation in chromatin structure and transcription. Nature *389*, 349-352.

Gu,W. and Roeder,R.G. (1997). Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. Cell *90*, 595-606.

Guenther,M.G., Barak,O., and Lazar,M.A. (2001). The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3. Mol. Cell Biol. *21*, 6091-6101.

Guschin,D., Wade,P.A., Kikyo,N., and Wolffe,A.P. (2000). ATP-Dependent histone octamer mobilization and histone deacetylation mediated by the Mi-2 chromatin remodeling complex. Biochemistry *39*, 5238-5245.

Hanahan,D. and Folkman,J. (1996). Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. Cell *86*, 353-364.

Hanahan,D. and Weinberg,R.A. (2000). The hallmarks of cancer. Cell *100*, 57-70.

Harbour,J.W. and Dean,D.C. (2000). The Rb/E2F pathway: expanding roles and emerging paradigms. Genes Dev. *14*, 2393-2409.

Harbour,J.W., Luo,R.X., Dei Santi,A., Postigo,A.A., and Dean,D.C. (1999). Cdk phosphorylation triggers sequential intramolecular interactions that progressively block Rb functions as cells move through G1. Cell *98*, 859-869.

Hartl,F.U. (1996). Molecular chaperones in cellular protein folding. Nature *381*, 571-579.

Hawkins,P.T., Welch,H., McGregor,A., Eguinoa,A., Gobert,S., Krugmann,S., Anderson,K., Stokoe,D., and Stephens,L. (1997). Signalling via phosphoinositide 3OH kinases. Biochem. Soc. Trans. *25*, 1147-1151.

Hayes,J.J. and Hansen,J.C. (2001). Nucleosomes and the chromatin fiber. Curr. Opin. Genet. Dev. *11*, 124-129.

Heery,D.M., Hoare,S., Hussain,S., Parker,M.G., and Sheppard,H. (2001). Core LXXLL motif sequences in CREB-binding protein, SRC1, and RIP140 define affinity and selectivity for steroid and retinoid receptors. J. Biol. Chem. *276*, 6695-6702.

Heery,D.M., Kalkhoven,E., Hoare,S., and Parker,M.G. (1997). A signature motif in transcriptional co-activators mediates binding to nuclear receptors. Nature *387*, 733-736.

Herman,M.A., Ch'ng,Q., Hettenbach,S.M., Ratliff,T.M., Kenyon,C., and Herman,R.K. (1999). EGL-27 is similar to a metastasis-associated factor and controls cell polarity and cell migration in C. elegans. Development *126*, 1055-1064.

Hernandez,M.P., Chadli,A., and Toft,D.O. (2002). HSP40 binding is the first step in the HSP90 chaperoning pathway for the progesterone receptor. J. Biol. Chem. *277*, 11873-11881.

Heyninck,K. and Beyaert,R. (1999). The cytokine-inducible zinc finger protein A20 inhibits IL-1-induced NF- kappaB activation at the level of TRAF6. FEBS Lett. *442*, 147-150.

Hoeijmakers,J.H. (2001). Genome maintenance mechanisms for preventing cancer. Nature *411*, 366-374.

Hoey,T., Weinzierl,R.O., Gill,G., Chen,J.L., Dynlacht,B.D., and Tjian,R. (1993). Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. Cell *72*, 247-260.

Hood,J.K. and Silver,P.A. (2000). Diverse nuclear transport pathways regulate cell proliferation and oncogenesis. Biochim. Biophys. Acta *1471*, M31-M41.

Hortobagyi,G.N. (1998). Treatment of breast cancer. N Engl J Med. *339*, 974-984.

Hsieh,J.J., Zhou,S., Chen,L., Young,D.B., and Hayward,S.D. (1999). CIR, a corepressor linking the DNA binding factor CBF1 to the histone deacetylase complex. Proc. Natl. Acad. Sci. U. S. A *96*, 23-28.

Hsu,H., Shu,H.B., Pan,M.G., and Goeddel,D.V. (1996). TRADD-TRAF2 and TRADD-FADD interactions define two distinct TNF receptor 1 signal transduction pathways. Cell *84*, 299-308.

Hsu,H., Xiong,J., and Goeddel,D.V. (1995). The TNF receptor 1-associated protein TRADD signals cell death and NF- kappa B activation. Cell *81*, 495-504.

Hu,E., Chen,Z., Fredrickson,T., Zhu,Y., Kirkpatrick,R., Zhang,G.F., Johanson,K., Sung,C.M., Liu,R., and Winkler,J. (2000). Cloning and characterization of a novel human class I histone deacetylase that functions as a transcription repressor. J. Biol. Chem. *275*, 15254-15264.

Hu,M.C., Wang,Y.P., and Qiu,W.R. (1999). Human fibroblast growth factor-18 stimulates fibroblast cell proliferation and is mapped to chromosome 14p11. Oncogene *18*, 2635-2642.

Imbalzano,A.N., Kwon,H., Green,M.R., and Kingston,R.E. (1994). Facilitated binding of TATA-binding protein to nucleosomal DNA. Nature *370*, 481-485.

Imhof,A., Yang,X.J., Ogryzko,V.V., Nakatani,Y., Wolffe,A.P., and Ge,H. (1997). Acetylation of general transcription factors by histone acetyltransferases. Curr. Biol. *7*, 689-692.

Irwin,M.S. and Kaelin,W.G. (2001). Role of the newer p53 family proteins in malignancy. Apoptosis *6*, 17-29.

Issa,J.P., Ottaviano,Y.L., Celano,P., Hamilton,S.R., Davidson,N.E., and Baylin,S.B. (1994). Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon. Nat. Genet. *7*, 536-540.

Jans,D.A. and Hubner,S. (1996). Regulation of protein transport to the nucleus: central role of phosphorylation. Physiol Rev. *76*, 651-685.

Jenuwein,T. and Allis,C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

Johnson,D.E. and Williams,L.T. (1993). Structural and functional diversity in the FGF receptor multigene family. Adv. Cancer Res. *60*, 1-41.

Johnson,K.D. and Bresnick,E.H. (2002). Dissecting long-range transcriptional mechanisms by chromatin immunoprecipitation. Methods *26*, 27-36.

Johnson-Pais,T., Degnin,C., and Thayer,M.J. (2001). pRB induces Sp1 activity by relieving inhibition mediated by MDM2. Proc Natl Acad Sci U S A. *95*, 2211-2216.

Julenius,K., Robblee,J., Thulin,E., Finn,B.E., Fairman,R., and Linse,S. (2002). Coupling of ligand binding and dimerization of helix-loop-helix peptides: spectroscopic and sedimentation analyses of calbindin D9k EF-hands. Proteins *47*, 323-333.

Kaczynski,J., Zhang,J.S., Ellenrieder,V., Conley,A., Duenes,T., Kester,H., van Der,B.B., and Urrutia,R. (2001). The Sp1-like protein BTEB3 inhibits transcription via the basic transcription element box by interacting with mSin3A and HDAC-1 co- repressors and competing with Sp1. J. Biol. Chem. *276*, 36749-36756.

Kaelin,W.G. (1999). Functions of the retinoblastoma protein. Bioessays *21*, 950-958.

Kamat,A. and Carpenter,G. (1997). Phospholipase C-gamma1: regulation of enzyme function and role in growth factor-dependent signal transduction. Cytokine Growth Factor Rev. *8*, 109-117.

Kang,K.I., Devin,J., Cadepond,F., Jibard,N., Guiochon-Mantel,A., Baulieu,E.E., and Catelli,M.G. (1994). In vivo functional protein-protein interaction: nuclear targeted hsp90 shifts cytoplasmic steroid receptor mutants into the nucleus. Proc. Natl. Acad. Sci. U. S. A *91*, 340-344.

Karin,M. (1996). The regulation of AP-1 activity by mitogen-activated protein kinases. Philos. Trans. R. Soc. Lond B Biol. Sci. *351*, 127-134.

Karin,M. and Ben-Neriah,Y. (2000). Phosphorylation meets ubiquitination: the control of NF-[kappa]B activity. Annu Rev Immunol. *18*, 621-663.

Karlseder,J., Rotheneder,H., and Wintersberger,E. (1996). Interaction of Sp1 with the growth- and cell cycle-regulated transcription factor E2F. Mol. Cell Biol. *16*, 1659-1667.

Kastner,P., Perez,A., Lutz,Y., Rochette-Egly,C., Gaub,M.P., Durand,B., Lanotte,M., Berger,R., and Chambon,P. (1992). Structure, localization and transcriptional properties of two classes of retinoic acid receptor alpha fusion proteins in acute promyelocytic leukemia (APL): structural similarities with a new family of oncoproteins. EMBO J. *11*, 629-642.

Kay,B.K., Williamson,M.P., and Sudol,M. (2000). The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. FASEB J. *14*, 231-241.

Kelekar,A., Chang,B.S., Harlan,J.E., Fesik,S.W., and Thompson,C.B. (1997). Bad is a BH3 domain-containing protein that forms an inactivating dimer with Bcl-XL. Mol. Cell Biol. *17*, 7040-7046.

Kerr,J.F., Wyllie,A.H., and Currie,A.R. (1972). Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. Br J Cancer *26*, 239-257.

Kinzler,K.W. and Vogelstein,B. (1990). The GLI gene encodes a nuclear protein which binds specific sequences in the human genome. Mol. Cell Biol. *10*, 634-642.

Knoepfler,P.S. and Eisenman,R.N. (1999). Sin meets NuRD and other tails of repression. Cell *99*, 447-450.

Kornberg,R.D. (1999). Eukaryotic transcriptional control. Trends Cell Biol. *9*, M46-M49.

Kowenz-Leutz,E. and Leutz,A. (1999). A C/EBP beta isoform recruits the SWI/SNF complex to activate myeloid genes. Mol. Cell *4*, 735-743.

Kumar,R., Wang,R.A., Mazumdar,A., Talukder,A.H., Mandal,M., Yang,Z., Bagheri-Yarmand,R., Sahin,A., Hortobagyi,G., Adam,L., Barnes,C.J., and Vadlamudi,R.K. (2002). A naturally occurring MTA1 variant sequesters oestrogen receptor-alpha in the cytoplasm. Nature *418*, 654-657.

Kwon,H.S., Kim,M.S., Edenberg,H.J., and Hur,M.W. (1999). Sp3 and Sp4 can repress transcription by competing with Sp1 for the core cis-elements on the human ADH5/FDH minimal promoter. J. Biol. Chem. *274*, 20-28.

Lai,A., Lee,J.M., Yang,W.M., DeCaprio,J.A., Kaelin,W.G., Jr., Seto,E., and Branton,P.E. (1999). RBP1 recruits both histone deacetylase-dependent and -independent repression activities to retinoblastoma family proteins. Mol. Cell Biol. *19*, 6632-6641.

Laity,J.H., Lee,B.M., and Wright,P.E. (2001). Zinc finger proteins: new insights into structural and functional diversity. Curr Opin Struct Biol. *11*, 39-46.

Lawlor,M.A. and Alessi,D.R. (2001). PKB/Akt: a key mediator of cell proliferation, survival and insulin responses? J Cell Sci. *114*, 2903-2910.

Ledo,F., Kremer,L., Mellstrom,B., and Naranjo,J.R. (2002). Ca2+-dependent block of CREB-CBP transcription by repressor DREAM. EMBO J. *21*, 4583-4592.

Lee,D.K., Suh,D., Edenberg,H.J., and Hur,M.W. (2002). POZ domain transcription factor, FBI-1, represses transcription of ADH5/FDH by interacting with the zinc finger and interfering with DNA binding activity of Sp1. J. Biol. Chem. *277*, 26761-26768.

Lee,D.Y., Hayes,J.J., Pruss,D., and Wolffe,A.P. (1993a). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. Cell *72*, 73-84.

Lee,E.G., Boone,D.L., Chai,S., Libby,S.L., Chien,M., Lodolce,J.P., and Ma,A. (2000). Failure to regulate TNF-induced NF-kappaB and cell death responses in A20-deficient mice. Science *289*, 2350-2354.

Lee,J.S., Galvin,K.M., and Shi,Y. (1993b). Evidence for physical interaction between the zinc-finger transcription factors YY1 and Sp1. Proc. Natl. Acad. Sci. U. S A *90*, 6145-6149.

Lemmon,M.A. and Schlessinger,J. (1994). Regulation of signal transduction and signal diversity by receptor oligomerization. Trends Biochem. Sci. *19*, 459-463.

Li,E. (2002). Chromatin modification and epigenetic reprogramming in mammalian development. Nat. Rev. Genet. *3*, 662-673.

Li,E., Bestor,T.H., and Jaenisch,R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915-926.

Li,P., Nijhawan,D., Budihardjo,I., Srinivasula,S.M., Ahmad,M., Alnemri,E.S., and Wang,X. (1997). Cytochrome c and dATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade. Cell *91*, 479-489.

Lin,J., Chen,J., Elenbaas,B., and Levine,A.J. (1994). Several hydrophobic amino acids in the p53 amino-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus 5 E1B 55-kD protein. Genes Dev. *8*, 1235-1246.

Lin,J., Reichner,C., Wu,X., and Levine,A.J. (1996a). Analysis of wild-type and mutant p21WAF-1 gene activities. Mol. Cell. Biol. *16*, 1786-1793.

Lin,S.Y., Black,A.R., Kostic,D., Pajovic,S., Hoover,C.N., and Azizkhan,J.C. (1996b). Cell cycle-regulated association of E2F1 and Sp1 is related to their functional interaction. Mol. Cell Biol. *16*, 1668-1675.

Loeb,L.A. (2001). A mutator phenotype in cancer. Cancer Res. *61*, 3230-3239.

Luchman,H.A., Paterno,G.D., Kao,K.R., and Gillespie,L.L. (1999). Differential nuclear localization of ER1 protein during embryonic development in Xenopus laevis. Mech. Dev. *80*, 111-114.

Luo,J., Su,F., Chen,D., Shiloh,A., and Gu,W. (2000). Deacetylation of p53 modulates its effect on cell growth and apoptosis. Nature *408*, 377-381.

Lupas,A. (1996). Coiled coils: new structures and new functions. Trends Biochem Sci. *21*, 375-382.

Ma,J. and Ptashne,M. (1987). A new class of yeast transcriptional activators. Cell *51*, 113-119.

Macara,I.G. (2001). Transport into and out of the nucleus. Microbiol Mol Biol Rev. *65*, 570-594.

Macleod,D., Charlton,J., Mullins,J., and Bird,A.P. (1994). Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. Genes Dev. *8*, 2282-2292.

Magnaghi-Jaulin,L., Groisman,R., Naguibneva,I., Robin,P., Lorain,S., Le Villain,J.P., Troalen,F., Trouche,D., and Harel-Bellan,A. (1998). Retinoblastoma protein represses transcription by recruiting a histone deacetylase. Nature *391*, 601-605.

Martinez,E., Kundu,T.K., Fu,J., and Roeder,R.G. (1998). A human SPT3-TAFII31-GCN5-L acetylase complex distinct from transcription factor IID. J. Biol. Chem. *273*, 23781-23785.

Martinez-Balbas,M.A., Bauer,U.M., Nielsen,S.J., Brehm,A., and Kouzarides,T. (2000). Regulation of E2F1 activity by acetylation. EMBO J. *19*, 662-671.

Maser,R.S. and DePinho,R.A. (2002). Connecting chromosomes, crisis, and cancer. Science *297*, 565-569.

Masutomi,K. and Hahn,W.C. (2003). Telomerase and tumorigenesis. Cancer Lett. *194*, 163-172.

Mayer,B.J. (2001). SH3 domains: complexity in moderation. J Cell Sci. *114*, 1253-1263.

Mazumdar,A., Wang,R.A., Mishra,S.K., Adam,L., Bagheri-Yarmand,R., Mandal,M., Vadlamudi,R.K., and Kumar,R. (2001). Transcriptional repression of oestrogen receptor by metastasis-associated protein 1 corepressor. Nat Cell Biol. *3*, 30-37.

McCormick,F. (1991). GTP-binding proteins as oncogenes in human tumors. Environ. Health Perspect. *93*, 17-18.

McKenna,N.J., Lanz,R.B., and O'Malley,B.W. (1999). Nuclear receptor coregulators: cellular and molecular biology. Endocr. Rev. *20*, 321-344.

Medrano,E.E. (2003). Repression of TGF-beta signaling by the oncogenic protein SKI in human melanomas: consequences for proliferation, survival, and metastasis. Oncogene *22*, 3123-3129.

Mizzen,C.A., Yang,X.J., Kokubo,T., Brownell,J.E., Bannister,A.J., Owen-Hughes,T., Workman,J., Wang,L., Berger,S.L., Kouzarides,T., Nakatani,Y., and Allis,C.D. (1996). The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. Cell *87*, 1261-1270.

Mohammadi,M., Honegger,A.M., Rotin,D., Fischer,R., Bellot,F., Li,W., Dionne,C.A., Jaye,M., Rubinstein,M., and Schlessinger,J. (1991). A tyrosine-phosphorylated carboxy-terminal peptide of the fibroblast growth factor receptor (Flg) is a binding site for the SH2 domain of phospholipase C-gamma 1. Mol. Cell Biol. *11*, 5068-5078.

Momand,J. and Zambetti,G.P. (1997). Mdm-2: "big brother" of p53. J Cell Biochem. *64*, 343-352.

Moroianu,J. (1999). Nuclear import and export pathways. J Cell Biochem. *Suppl 32-33*, 76-83.

Morris,E.J. and Dyson,N.J. (2001). Retinoblastoma protein partners. Adv. Cancer Res. *82*, 1-54.

Muzio,M., Ni,J., Feng,P., and Dixit,V.M. (1997). IRAK (Pelle) family member IRAK-2 and MyD88 as proximal mediators of IL- 1 signaling. Science *278*, 1612-1615.

Naar,A.M., Beaurang,P.A., Zhou,S., Abraham,S., Solomon,W., and Tjian,R. (1999). Composite co-activator ARC mediates chromatin-directed transcriptional activation. Nature *398*, 828-832.

Naar,A.M., Lemon,B.D., and Tjian,R. (2001). Transcriptional coactivator complexes. Annu. Rev. Biochem. *70*, 475-501.

Narlikar,G.J., Fan,H.Y., and Kingston,R.E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. Cell *108*, 475-487.

Neely,K.E. and Workman,J.L. (2002). The complexity of chromatin remodeling and its links to cancer. Biochim. Biophys. Acta *1603*, 19-29.

Nephew,K.P. and Huang,T.H. (2003). Epigenetic gene silencing in cancer initiation and progression. Cancer Lett. *190*, 125-133.

Newell-Price,J., Clark,A.J., and King,P. (2000). DNA methylation and silencing of gene expression. Trends Endocrinol. Metab *11*, 142-148.

Ng,H.H., Zhang,Y., Hendrich,B., Johnson,C.A., Turner,B.M., Erdjument-Bromage,H., Tempst,P., Reinberg,D., and Bird,A. (1999). MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex [see comments]. Nat. Genet. *23*, 58-61.

Nicolson,G.L., Nawa,A., Toh,Y., Taniguchi,S., Nishimori,K., and Moustafa,A. (2003). Tumor metastasis-associated human MTA1 gene and its MTA1 protein product: role in epithelial cancer cell invasion, proliferation and nuclear regulation. Clin. Exp. Metastasis *20*, 19-24.

Nigg,E.A. (1997). Nucleocytoplasmic transport: signals, mechanisms and regulation. Nature *386*, 779-787.

Nomura,T., Khan,M.M., Kaul,S.C., Dong,H.D., Wadhwa,R., Colmenares,C., Kohno,I., and Ishii,S. (1999). Ski is a component of the histone deacetylase complex required for transcriptional repression by Mad and thyroid hormone receptor. Genes Dev. *13*, 412-423.

Nunez,G., Benedict,M.A., Hu,Y., and Inohara,N. (1998). Caspases: the proteases of the apoptotic pathway. Oncogene *17*, 3237-3245.

O'Neill,L.A. (2002). Signal transduction pathways activated by the IL-1 receptor/toll-like receptor superfamily. Curr. Top. Microbiol. Immunol. *270*, 47-61.

Ogryzko,V.V., Kotani,T., Zhang,X., Schiltz,R.L., Howard,T., Yang,X.J., Howard,B.H., Qin,J., and Nakatani,Y. (1998). Histone-like TAFs within the PCAF histone acetylase complex. Cell *94*, 35-44.

Ordentlich,P., Downes,M., Xie,W., Genin,A., Spinner,N.B., and Evans,R.M. (1999). Unique forms of human and mouse nuclear receptor corepressor SMRT. Proc. Natl. Acad. Sci. U. S. A *96*, 2639-2644.

Ornitz,D.M. and Itoh,N. (2001). Fibroblast growth factors. Genome Biol. *2*, REVIEWS3005.

Orphanides,G. and Reinberg,D. (2002). A unified theory of gene expression. Cell *108*, 439-451.

Osborne,C.K. (1998). Steroid hormone receptors in breast cancer management. Breast Cancer Res Treat. *51*, 227-238.

Owen,G.I., Richer,J.K., Tung,L., Takimoto,G., and Horwitz,K.B. (1998). Progesterone regulates transcription of the p21(WAF1) cyclin- dependent kinase inhibitor gene through Sp1 and CBP/p300. J. Biol. Chem. *273*, 10696-10701.

Packeisen,J., Korsching,E., Herbst,H., Boecker,W., and Buerger,H. (2003). Demystified...tissue microarray technology. Mol Pathol. *56*, 198-204.

Park,A.R., Cho,S.K., Yun,U.J., Jin,M.Y., Lee,S.H., Sachetto-Martins,G., and Park,O.K. (2001). Interaction of the Arabidopsis receptor protein kinase Wak1 with a glycine-rich protein, AtGRP-3. J Biol Chem. *276*, 26688-26693.

Paterno,G.D., Li,Y., Luchman,H.A., Ryan,P.J., and Gillespie,L.L. (1997). cDNA cloning of a novel, developmentally regulated immediate early gene activated by fibroblast growth factor and encoding a nuclear protein. J. Biol. Chem. *272*, 25591-25595.

Paterno,G.D., Mercer,F.C., Chayter,J.J., Yang,X., Robb,J.D., and Gillespie,L.L. (1998). Molecular cloning of human er1 cDNA and its differential expression in breast tumours and tumour-derived cell lines. Gene *222*, 77-82.

Paterno,GP., Ding,Z., Lew,Y., Nash,G., Mercer,F., and Gillespie,L. (2002). Genomic organization of the human mi-er1 gene and characterization of alternatively spliced

isoforms: regulated use of a facultative intron determines subcellular localization. Gene *295*, 79.

Pawson,T. (1995). Protein modules and signalling networks. Nature *373*, 573-580.

Pazin,M.J. and Kadonaga,J.T. (1997). What's up and down with histone deacetylation and transcription? Cell *89*, 325-328.

Peinado,H., Ballestar,E., Esteller,M., and Cano,A. (2004). Snail Mediates E-Cadherin Repression by the Recruitment of the Sin3A/Histone Deacetylase 1 (HDAC1)/HDAC2 Complex. Mol Cell Biol. *24*, 306-319.

Peterson,C.L. and Workman,J.L. (2000). Promoter targeting and chromatin remodeling by the SWI/SNF complex. Curr. Opin. Genet. Dev. *10*, 187-192.

Philipsen,S. and Suske,G. (1999). A tale of three fingers: the family of mammalian Sp/XKLF transcription factors. Nucleic Acids Res. *27*, 2991-3000.

Pollock,J.D. (2002). Gene expression profiling: methodological challenges, results, and prospects for addiction research. Chem Phys Lipids. *121*, 241-256.

Post,J.N., Gillespie,L.L., and Paterno,G.D. (2001). Nuclear localization signals in the Xenopus FGF embryonic early response 1 protein. FEBS Lett. *502*, 41-45.

Powers,C.J., McLeskey,S.W., and Wellstein,A. (2000). Fibroblast growth factors, their receptors and signaling. Endocr. Relat Cancer *7*, 165-197.

Prives,C. and Manley,J.L. (2001). Why is p53 acetylated? Cell *107*, 815-818.

Proudfoot,N.J., Furger,A., and Dye,M.J. (2002). Integrating mRNA processing with transcription. Cell *108*, 501-512.

Pullen,S.S., Miller,H.G., Everdeen,D.S., Dang,T.T., Crute,J.J., and Kehry,M.R. (1998). CD40-tumor necrosis factor receptor-associated factor (TRAF) interactions: regulation of CD40 signaling through multiple TRAF binding sites and TRAF hetero-oligomerization. Biochemistry *37*, 11836-11845.

Qu,Z., Weiss,J.N., and MacLellan,W.R. (2003). Regulation of the mammalian cell cycle: a model of the G1-to-S transition. Am J Physiol Cell Physiol. *284*, C349-C364.

Reed,C.J. (2000a). Apoptosis and cancer: strategies for integrating programmed cell death. Semin. Hematol. *37*, 9-16.

Reed,J.C. (2000b). Mechanisms of apoptosis. Am J Pathol. *157*, 1415-1430.

Richards,E.J. and Elgin,S.C. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. Cell *108*, 489-500.

Roth,S.Y., Denu,J.M., and Allis,C.D. (2001). Histone acetyltransferases. Annu. Rev. Biochem. *70*, 81-120.

Rothe,M., Wong,S.C., Henzel,W.J., and Goeddel,D.V. (1994). A novel family of putative signal transducers associated with the cytoplasmic domain of the 75 kDa tumor necrosis factor receptor. Cell *78*, 681-692.

Routledge,E.J., White,R., Parker,M.G., and Sumpter,J.P. (2000). Differential effects of xenoestrogens on coactivator recruitment by estrogen receptor (ER) alpha and ERbeta. J. Biol. Chem. *275*, 35986-35993.

Ryan,P.J. and Gillespie,L.L. (1994). Phosphorylation of phospholipase C gamma 1 and its association with the FGF receptor is developmentally regulated and occurs during mesoderm induction in Xenopus laevis. Dev. Biol. *166*, 101-111.

Ryan,P.J., Paterno,G.D., and Gillespie,L.L. (1998). Identification of phosphorylated proteins associated with the fibroblast growth factor receptor type I during early Xenopus development. Biochem. Biophys. Res. Commun. *244*, 763-767.

Ryu,S., Zhou,S., Ladurner,A.G., and Tjian,R. (1999). The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1. Nature *397*, 446-450.

Saleh,M., Rambaldi,I., Yang,X.J., and Featherstone,M.S. (2000). Cell signaling switches HOX-PBX complexes from repressors to activators of transcription mediated by histone deacetylases and histone acetyltransferases. Mol. Cell Biol. *20*, 8623-8633.

Sanchez,R., Nguyen,D., Rocha,W., White,J.H., and Mader,S. (2002). Diversity in the mechanisms of gene regulation by estrogen receptors. Bioessays *24*, 244-254.

Sartorelli,V., Puri,P.L., Hamamori,Y., Ogryzko,V., Chung,G., Nakatani,Y., Wang,J.Y., and Kedes,L. (1999). Acetylation of MyoD directed by PCAF is necessary for the execution of the muscle program. Mol. Cell *4*, 725-734.

Savill,J. and Fadok,V. (2000). Corpse clearance defines the meaning of cell death. Nature *407*, 784-788.

Scheid,M.P. and Woodgett,J.R. (2003). Unravelling the activation mechanisms of protein kinase B/Akt. FEBS Lett. *546*, 108-112.

Schmidt,W.M. and Mueller,M.W. (1999). CapSelect: a highly sensitive method for 5' CAP-dependent enrichment of full-length cDNA in PCR-mediated analysis of mRNAs. Nucleic Acids Res. *27*, e31.

Scorrano,L. and Korsmeyer,S.J. (2003). Mechanisms of cytochrome *c* release by proapoptotic BCL-2 family members. Biochem. Biophys. Res. Commun. *304*, 437-444.

Sellers,W.R. and Kaelin,W.G., Jr. (1997). Role of the retinoblastoma protein in the pathogenesis of human cancer. J. Clin. Oncol. *15*, 3301-3312.

Sharpless,N.E. and DePinho,R.A. (1999). The INK4A/ARF locus and its two gene products. Curr. Opin. Genet. Dev. *9*, 22-30.

Sharpless,N.E. and DePinho,R.A. (2002). p53: good cop/bad cop. Cell *110*, 9-12.

Sharpless,N.E. and DePinho,R.A. (2004). Telomeres, stem cells, senescence, and cancer. J. Clin. Invest *113*, 160-168.

Shaulian,E. and Karin,M. (2001). AP-1 in cell proliferation and survival. Oncogene *20*, 2390-2400.

Shay,J.W. and Bacchetti,S. (1997). A survey of telomerase activity in human cancer. Eur. J. Cancer *33*, 787-791.

Sherr,C.J. (1998). Tumor surveillance via the ARF-p53 pathway. Genes Dev. *12*, 2984-2991.

Sherr,C.J. (2001). The INK4a/ARF network in tumour suppression. Nat. Rev. Mol. Cell Biol. *2*, 731-737.

Shi,Y. (2002). Mechanisms of caspase activation and inhibition during apoptosis. Mol Cell. *9*, 459-470.

Shiloh,Y. (2003). ATM and related protein kinases: safeguarding genome integrity. Nat Rev Cancer. *3*, 155-168.

Sicinski,P. and Weinberg,R.A. (1997). A specific role for cyclin D1 in mammary gland development. J. Mammary. Gland. Biol. Neoplasia. *2*, 335-342.

Slootweg,M.C., de Groot,R.P., Herrmann-Erlee,M.P., Koornneef,I., Kruijer,W., and Kramer,Y.M. (1991). Growth hormone induces expression of c-jun and jun B oncogenes and employs a protein kinase C signal transduction pathway for the induction of c-fos oncogene expression. J Mol Endocrinol. *6*, 179-188.

Smith,J.C., Price,B.M., Green,J.B., Weigel,D., and Herrmann,B.G. (1991). Expression of a Xenopus homolog of Brachyury (T) is an immediate-early response to mesoderm induction. Cell *67*, 79-87.

Soderling,T.R. (1999). The Ca-calmodulin-dependent protein kinase cascade. Trends Biochem Sci. *24*, 232-236.

Solari,F., Bateman,A., and Ahringer,J. (1999). The Caenorhabditis elegans genes egl-27 and egr-1 are similar to MTA1, a member of a chromatin regulatory complex, and are redundantly required for embryonic patterning. Development *126*, 2483-2494.

Song,H.Y., Rothe,M., and Goeddel,D.V. (1996). The tumor necrosis factor-inducible zinc finger protein A20 interacts with TRAF1/TRAF2 and inhibits NF-kappaB activation. Proc. Natl. Acad. Sci. U. S. A *93*, 6721-6725.

Spencer,V.A. and Davie,J.R. (1999). Role of covalent modifications of histones in regulating gene expression. Gene *240*, 1-12.

Sporn,M.B. (1996). The war on cancer. Lancet *347*, 1377-1381.

Stancovski,I. and Baltimore,D. (1997). NF-kappaB activation: the I kappaB kinase revealed? Cell *91*, 299-302.

Stanger,B.Z., Leder,P., Lee,T.H., Kim,E., and Seed,B. (1995). RIP: a novel protein containing a death domain that interacts with Fas/APO-1 (CD95) in yeast and causes cell death. Cell *81*, 513-523.

Sterner,D.E., Wang,X., Bloom,M.H., Simon,G.M., and Berger,S.L. (2002). The SANT domain of Ada2 is required for normal acetylation of histones by the yeast SAGA complex. J. Biol. Chem. *277*, 8178-8186.

Stevaux,O. and Dyson,N.J. (2002). A revised picture of the E2F transcriptional network and RB function. Curr Opin Cell Biol. *14*, 684-691.

Stevens,C. and La Thangue,NB. (2003). E2F and cell cycle control: a double-edged sword. Arch Biochem Biophys. *412*, 157-169.

Strahl,B.D. and Allis,C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Sun,H. and Taneja,R. (2000). Stra13 expression is associated with growth arrest and represses transcription through histone deacetylase (HDAC)-dependent and HDAC-independent mechanisms. Proc. Natl. Acad. Sci. U. S. A *97*, 4058-4063.

Suzuki,T., Muto,S., Miyamoto,S., Aizawa,K., Horikoshi,M., and Nagai,R. (2003). Functional interaction of the DNA-binding transcription factor Sp1 through its DNA-binding domain with the histone chaperone TAF-I. J Biol Chem. *278*, 28758-28764.

Symonds,H., Krall,L., Remington,L., Saenz-Robles,M., Lowe,S., Jacks,T., and Van Dyke,T. (1994). p53-dependent apoptosis suppresses tumor growth and progression in vivo. Cell *78*, 703-711.

Szebenyi,G. and Fallon,J.F. (1999). Fibroblast growth factors as multifunctional signaling factors. Int. Rev. Cytol. *185*, 45-106.

Tanese,N., Saluja,D., Vassallo,M.F., Chen,J.L., and Admon,A. (1996). Molecular cloning and analysis of two subunits of the human TFIID complex: hTAFII130 and hTAFII100. Proc. Natl. Acad. Sci. U. S. A *93*, 13611-13616.

Teodoro,J.G. and Branton,P.E. (1997). Regulation of p53-dependent apoptosis, transcriptional repression, and cell transformation by phosphorylation of the 55-kilodalton E1B protein of human adenovirus type 5. J. Virol. *71*, 3620-3627.

Teplitsky,Y., Paterno,G.D., and Gillespie,L.L. (2003). Proline(365) is a critical residue for the activity of XMI-ER1 in Xenopus embryonic development. Biochem Biophys Res Commun. *308*, 679-683.

Thiagalingam,S., Cheng,K.H., Lee,H.J., Mineva,N., Thiagalingam,A., and Ponte,J.F. (2003). Histone deacetylases: unique players in shaping the epigenetic histone code. Ann. N. Y. Acad. Sci. *983*, 84-100.

Thompson,C.B. (1995). Apoptosis in the pathogenesis and treatment of disease. Science *267*, 1456-1462.

Thomson,S., Mahadevan,L.C., and Clayton,A.L. (1999). MAP kinase-mediated signalling to nucleosomes and immediate-early gene induction. Semin. Cell Dev. Biol. *10*, 205-214.

Thornberry,N.A. and Lazebnik,Y. (1998). Caspases: enemies within. Science *281*, 1312-1316.

Toh,Y., Pencil,S.D., and Nicolson,G.L. (1994). A novel candidate metastasis-associated gene, mta1, differentially expressed in highly metastatic mammary adenocarcinoma cell lines. cDNA cloning, expression, and protein analyses. J. Biol. Chem. *269*, 22958-22963.

Tong,J.J., Liu,J., Bertos,N.R., and Yang,X.J. (2002). Identification of HDAC10, a novel class II human histone deacetylase containing a leucine-rich domain. Nucleic Acids Res. *30*, 1114-1123.

Tong,J.K., Hassig,C.A., Schnitzler,G.R., Kingston,R.E., and Schreiber,S.L. (1998). Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. Nature *395*, 917-921.

Torchia,J., Glass,C., and Rosenfeld,M.G. (1998). Co-activators and co-repressors in the integration of transcriptional responses. Curr. Opin. Cell Biol. *10*, 373-383.

Tucker,C.L., Gera,J.F., and Uetz,P. (2001). Towards an understanding of complex protein networks. Trends Cell Biol. *11*, 102-106.

Tyler,J.K. and Kadonaga,J.T. (1999). The "dark side" of chromatin remodeling: repressive effects on transcription. Cell *99*, 443-446.

Vallian,S., Chin,K.V., and Chang,K.S. (1998). The promyelocytic leukemia protein interacts with Sp1 and inhibits its transactivation of the epidermal growth factor receptor promoter. Mol. Cell Biol. *18*, 7147-7156.

Vandromme,M., Gauthier-Rouviere,C., Lamb,N., and Fernandez,A. (1996). Regulation of transcription factor localization: fine-tuning of gene expression. Trends Biochem. Sci. *21*, 59-64.

Vinson,C.R., LaMarco,K.L., Johnson,P.F., Landschulz,W.H., and McKnight,S.L. (1988). In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. Genes Dev. *2*, 801-806.

Vogelstein,B. and Kinzler,K.W. (1993). The multistep nature of cancer. Trends Genet. *9*, 138-141.

Wade,P.A., Gegonne,A., Jones,P.L., Ballestar,E., Aubry,F., and Wolffe,A.P. (1999). Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation [see comments]. Nat. Genet. *23*, 62-66.

Wallach,D., Varfolomeev,E.E., Malinin,N.L., Goltsev,Y.V., Kovalenko,A.V., and Boldin,M.P. (1999). Tumor necrosis factor receptor and Fas signaling mechanisms. Annu. Rev. Immunol. *17*, 331-367.

Wallberg,A.E., Neely,K.E., Hassan,A.H., Gustafsson,J.A., Workman,J.L., and Wright,A.P. (2000). Recruitment of the SWI-SNF chromatin remodeling complex as a mechanism of gene activation by the glucocorticoid receptor tau1 activation domain. Mol. Cell Biol. *20*, 2004-2013.

Wang,L., Mizzen,C., Ying,C., Candau,R., Barlev,N., Brownell,J., Allis,C.D., and Berger,S.L. (1997). Histone acetyltransferase activity is conserved between yeast and human GCN5 and is required for complementation of growth and transcriptional activation. Mol. Cell Biol. *17*, 519-527.

Weinmann,A.S. and Farnham,P.J. (2002). Identification of unknown target genes of human transcription factors using chromatin immunoprecipitation. Methods *26*, 37-47.

Wesche,H., Gao,X., Li,X., Kirschning,C.J., Stark,G.R., and Cao,Z. (1999). IRAK-M is a novel member of the Pelle/interleukin-1 receptor-associated kinase (IRAK) family. J. Biol. Chem. *274*, 19403-19410.

Wesche,H., Henzel,W.J., Shillinglaw,W., Li,S., and Cao,Z. (1997). MyD88: an adapter that recruits IRAK to the IL-1 receptor complex. Immunity. *7*, 837-847.

White,E. (1996). Life, death, and the pursuit of apoptosis. Genes Dev. *10*, 1-5.

Wieczorek,E., Brand,M., Jacq,X., and Tora,L. (1998). Function of TAF(II)-containing complex without TBP in transcription by RNA polymerase II. Nature *393*, 187-191.

Wilkie,A.O., Patey,S.J., Kan,S.H., van den Ouweland,A.M., and Hamel,B.C. (2002). FGFs, their receptors, and human limb malformations: clinical and molecular correlations. Am J Med Genet. *112*, 266-278.

Williamson,M.P. (1994). The structure and function of proline-rich regions in proteins. Biochem. J. *297 (Pt 2)*, 249-260.

Wright,W.E., Binder,M., and Funk,W. (1991). Cyclic amplification and selection of targets (CASTing) for the myogenin consensus binding site. Mol Cell Biol. *11*, 4104-4110.

Wu,W.S., Xu,Z.X., Ran,R., Meng,F., and Chang,K.S. (2002). Promyelocytic leukemia protein PML inhibits Nur77-mediated transcription through specific functional interactions. Oncogene *21*, 3925-3933.

Xu,Y., Tao,X., Shen,B., Horng,T., Medzhitov,R., Manley,J.L., and Tong,L. (2000). Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. Nature *408*, 111-115.

Xue,Y., Wong,J., Moreno,G.T., Young,M.K., Cote,J., and Wang,W. (1998a). NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. Mol Cell. *2*, 851-861.

Xue,Y., Wong,J., Moreno,G.T., Young,M.K., Cote,J., and Wang,W. (1998b). NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. Mol. Cell *2*, 851-861.

Yamagoe,S., Kanno,T., Kanno,Y., Sasaki,S., Siegel,R.M., Lenardo,M.J., Humphrey,G., Wang,Y., Nakatani,Y., Howard,B.H., and Ozato,K. (2003). Interaction of histone acetylases and deacetylases in vivo. Mol. Cell Biol. *23*, 1025-1033.

Yan,C., Wang,H., Toh,Y., and Boyd,D.D. (2003). Repression of 92-kDa type IV collagenase expression by MTA1 is mediated through direct interactions with the promoter via a mechanism, which is both dependent on and independent of histone deacetylation. Biol Chem. *278*, 2309-2316.

Yang,E., Zha,J., Jockel,J., Boise,L.H., Thompson,C.B., and Korsmeyer,S.J. (1995). Bad, a heterodimeric partner for Bcl-XL and Bcl-2, displaces Bax and promotes cell death. Cell *80*, 285-291.

Yang,X.J., Ogryzko,V.V., Nishikawa,J., Howard,B.H., and Nakatani,Y. (1996). A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. Nature *382*, 319-324.

Yang,X.L., Li,D.M., Chen,W.L., and Xu,T. (2001). Human homologue of Drosophila lats, LATS1, negatively regulate growth by inducing G(2)/M arrest or apoptosis. Oncogene *20*, 6516-6523.

Yao,Y.L. and Yang,W.M. (2003). The metastasis-associated proteins 1 and 2 form distinct protein complexes with histone deacetylase activity. J Biol Chem. *278*, 42560-42568.

Yew,P.R., Liu,X., and Berk,A.J. (1994). Adenovirus E1B oncoprotein tethers a transcriptional repression domain to p53. Genes Dev. *8*, 190-202.

Yoshida,M., Kijima,M., Akita,M., and Beppu,T. (1990). Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A. J. Biol. Chem. *265*, 17174-17179.

Yoshinaga,S.K., Peterson,C.L., Herskowitz,I., and Yamamoto,K.R. (1992). Roles of SWI1, SWI2, and SWI3 proteins for transcriptional enhancement by steroid receptors. Science *258*, 1598-1604.

You,A., Tong,J.K., Grozinger,C.M., and Schreiber,S.L. (2001). CoREST is an integral component of the CoREST- human histone deacetylase complex. Proc. Natl. Acad. Sci. U. S. A *98*, 1454-1458.

Yu,J., Li,Y., Ishizuka,T., Guenther,M.G., and Lazar,M.A. (2003). A SANT motif in the SMRT corepressor interprets the histone code and promotes histone deacetylation. EMBO J. *22*, 3403-3410.

Yu,Y.B. (2002). Coiled-coils: stability, specificity, and drug delivery potential. Adv Drug Deliv Rev. *54*, 1113-1129.

Zacksenhaus,E., Jiang,Z., Hei,Y.J., Phillips,R.A., and Gallie,B.L. (1999). Nuclear localization conferred by the pocket domain of the retinoblastoma gene product. Biochim. Biophys. Acta *1451*, 288-296.

Zha,J., Harada,H., Osipov,K., Jockel,J., Waksman,G., and Korsmeyer,S.J. (1997). BH3 domain of BAD is required for heterodimerization with BCL-XL and pro-apoptotic activity. J. Biol. Chem. *272*, 24101-24104.

Zhang,Y., Ng,H.H., Erdjument-Bromage,H., Tempst,P., Bird,A., and Reinberg,D. (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. Genes Dev. *13*, 1924-1935.

Zhang,Y. and Reinberg,D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes Dev. *15*, 2343-2360.

Zhu,W.G., Srinivasan,K., Dai,Z., Duan,W., Druhan,L.J., Ding,H., Yee,L., Villalona-Calero,M.A., Plass,C., and Otterson,G.A. (2003). Methylation of adjacent CpG sites affects Sp1/Sp3 binding and activity in the p21(Cip1) promoter. Mol Cell Biol. *23*, 4056-4065.

Zou,H., Henzel,W.J., Liu,X., Lutschg,A., and Wang,X. (1997). Apaf-1, a human protein homologous to C. elegans CED-4, participates in cytochrome c-dependent activation of caspase-3. Cell *90*, 405-413.

APPENDICES

Appendice 1 Alignment of ELM2 domains by PFAM. Accession number PF01448

```
hMI-ER1/180-239     KEIMVGSM.FQAEIPVGICRYKENEKVYEND---DQLLWDPE.----- YLPEDKVIIFLKDASR---RTGDEK
MTA1_HUMAN/165-226  GEIRVGNR.YQADITDLLKEGEEDGRDQSRL---ETQVWEAH.-.--NPLTDKQIDQFLVVARS--VGTFARA
MTA1_MOUSE/165-226  GEIRVGNR.YQADITDLLKEGEEDGRDQSKL---ETKVWEAH.-.--NPLVDKQIDQFLVVARS--VGTFARA
MTA1_RAT/165-226    GEIRVGNR.YQADITDLLKDGEEDGRDQSKL---ETKVWEAH.-.--NPLVDKQIDQFLVVARS--VGTFARA
MTA2_HUMAN/145-206  GEIRVGCK.YQAEIPDRLVEGESDNRNQQKM---EMKVWDPD.-.--NPLTDRQIDQFLVVARA--VGTFARA
MTA2_MOUSE/145-206  GEIRVGCK.FQAEIPDRLAEGESDNRNQQKM---EMKVWDPD.-.--NPLTDRQIDQFLVVARA--VGTFARA
MTA3_HUMAN/148-209  GEIRVGPR.YQADIPEMLLEGESDEREQSKL---EVKVWDPN.-.--SPLTDRQIDQFLVVARA--VGTFARA
MTA3_MOUSE/148-208  GEIRVGPK.YQADIPDMLP-EDSDEREQSKL---EVKVWDPN.-.--SPLTDRQIDQFLVVARA--VGTFARA
EG27_CAEEL/224-281  YAIRVGTS.FQATLPPMAECSVGDDSDR------DELLYRPN.-.---SIESGEEEDYIKLARC--YRTYTLS
Q8BND7/174-234      REIMIGLE.YQAEIPPYLGEYNGDDEKAYEN-E-DQLLWHPG.-.---VLLESKVKEYLVETS---LRTGNEK
Q8BSS0/147-207      REIMIGLE.YQAEIPPYLGEYNGDDEKAYEN-E-DQLLWHPG.-.---VLLESKVKEYLVETS---LRTGNEK
Q86YG8/174-233      KEIMIGLQ.YQAEIPPYLGEYDGNEKVYENE---DQLLWCPD.-.---VVLESKVKEYLVETS---LRTGSEK
Q8NA13/111-170      KEIMIGLQ.YQAEIPPYLGEYDGNEKVYENE---DQLLWCPD.-.---VVLESKVKEYLVETS---LRTGSEK
Q9VWH3/90-153       EKIRVGRD.YQAVCPPLVPEAERRPEQMNER---ALLVWSPT.-.--KEIPDLKLEEYISVAKEKYGYNGEQA
Q8MSX7/1-38         --------.------------------MNER---ALLVWSPT.-.--KEIPDLKLEEYISVAKEKYGYNGEQA
Q9P2K3/24-86        VGMRVGAE.YQARIPEFDPGATKYTDKDN----GGMLVWSPY.-.--HSIPDAKLDEYIAIAKEKHGYNVEQA
Q8VDZ0/1-61         --MRVGAE.YQARIPEFDPGATKYTDKDN----GGMLVWSPY.-.--HSIPDAKLDEYIAIAKEKHGYNVEQA
Q8N5U3/57-119       VGMRVGAE.YQARIPEFDPGATKYTDKDN----GGMLVWSPY.-.--HSIPDARLDEYIAIAKEKHGYNVEQA
Q9JMK4/2-63         --IRVGTN.YQAVIPECKP--ESPARYSNKELK-GMLVWSPN.-.--HCVSDAKLDKYIAMAKEKHGYNIEQA
Q8C796/44-107       SMIRVGTN.YQAVIPECKP--ESPARYSNKELK-GMLVWSPN.-.--HCVSDAKLDKYIAMAKEKHGYNIEQA
Q8IZ40/44-107       SMIRVGTN.YQAVIPECKP--ESPARYSNKELK-GMLVWSPN.-.--HCVSDAKLDKYIAMAKEKHGYNIEQA
Q90WN5/50-113       GGMRVGLQ.YQAVVPEFDQ-EVAKNCQEREN--LGMLVWSPN.-.--QNISEAKLDEYISVAKEKHGYNMEQA
Q8CFE3/1-63         --MRVGPQ.YQAAVPDFDPAKLARRSQERDN--LGMLVWSPN.-.--QSLSEAKLDEYIAIAKEKHGYNMEQA
Q15044/14-78        GGMRVGPQ.YQAVVPDFDPAKLARRSQERDN--LGMLVWSPN.-.--QNLSEAKLDEYIAIAKEKHGYNMEQA
Q86VG5/100-164      GGMRVGPQ.YQAVVPDFDPAKLARRSQERDN--LGMLVWSPN.-.--QNLSEAKLDEYIAIAKEKHGYNMEQA
Q9UKL0/100-164      GGMRVGPQ.YQAVVPDFDPAKLARRSQERDN--LGMLVWSPN.-.--QNLSEAKLDEYIAIAKEKHGYNMEQA
Q9VHD1/1619-1679    SKVNLGST.YQAQIPSCRPLEEALRDSP-----GAEMMWNPE.-.--VQEDEKILMRYIDLSKS-SAVPMGSH
Q9U6X3/230-291      GEIRVGSR.YQCDIPAKLKDTATDDRKLEEL---ESLVWTPE.-.--HSLTDRKIDQFLVVSRS--IGTFARA
```

```
Q9VNF6/230-291      GEIRVGSR.YQCDIPAKLKDTATDDRKLEEL---ESLVWTPE.-.--HSLTDRKIDQFLVVSRS--IGTFARA
Q9VNF7/194-255      GEIRVGSR.YQCDIPAKLKDTATDDRKLEEL---ESLVWTPE.-.--HSLTDRKIDQFLVVSRS--IGTFARA
Q9PUN0/145-206      GEIRVGSK.YQAEIPDQLAEGESDNRNQQKM---EIKVWDPE.-.--NPLTDRQIDQFLVVARA--VGTFARA
Q7ZUE1/148-209      GEIRVGPR.FQADVPEMLQEGEADDRDQSKL---EMKMWDPE.-.--CPLTNKQIDQFLVVARA--VGTFARA
Q80UI1/161-222      GEIRVGNR.YQADITDLLKEGEEDGRDQSKL---ETKVWEAH.-.--NPLVDKQIDQFLVVARS--VGTFARA
Q86SW2/139-200      GEIRVGNR.YQADITDLLKEGEEDGRDQSRL---ETQVWEAH.-.--NPLTDKQIDQFLVVARS--VGTFARA
Q86TR6/79-140       GEIRVGNR.YQADITDLLKEGEEDGRDQSRL---ETQVWEAH.-.--NPLTDKQIDQFLVVARS--VGTFARA
O75046/16-75        GEIRVGPS.HQAKLPDLQP-FPSPDGDT-VTQH-EELVWMPG.-.----VNDCDLLMYLRAARS--MAAFAGM
Q9P2R6/284-343      GEIRVGPS.HQAKLPDLQP-FPSPDGDT-VTQH-EELVWMPG.-.----VNDCDLLMYLRAARS--MAAFAGM
Q96PN7/779-843      PRINIGLR.FQAEIPELQDISALAQDTHK-----ATLVWKPWp-eLENHDLQQRVENLLNLCCS-SALPGGGT
Q9NQ72/535-599      PRINIGLR.FQAEIPELQDISALAQDTHK-----ATLVWKPWp-eLENHDLQQRVENLLNLCCS-SALPGGGT
Q9NQ73/535-599      PRINIGLR.FQAEIPELQDISALAQDTHK-----ATLVWKPWp-eLENHDLQQRVENLLNLCCS-SALPGGGT
Q8BXJ2/785-849      PRINIGLR.FQAEIPELQDVSALAQDTHR-----ATLVWKPWp-eLENQALQQQVENLLNLCCS-SALPGGGT
Q810H8/784-848      PRINIGLR.FQAEIPELQDVSALAQDTHR-----ATLVWKPWp-eLENQALQQQVENLLNLCCS-SALPGGGT
Q8BUZ9/713-778      PRINVGTR.FQAEIPMMRD-RALAAFDPHK----ADLVWQPWeHlESSWEKQRQVDDLLTAACS-SIFPGAGT
Q7ZV74/167-229      KEIMVGSM.YQAETPVGLCKYKDNEKVYEND---DQLLWNPE.-.---LLPEEKVEEFLTEASKRSGDEAGVD
Q960H5/216-275      KIIMVGHD.YQAEIPEGLSQYGDILPYENE----DQLIWEPS.-.---QVSEREVEEYLAKIQE--TRSIVPP
Q9V4M4/272-331      KIIMVGHD.YQAEIPEGLSQYGDILPYENE---DQLIWEPS.-.---QVSEREVEEYLAKIQE.--TRSIVPP
Q8N344/159-220      KEIMVGPQ.FQADLSNLHLNRHCEKIYENE----DQLLWDPS.-.---VLPEREVEEFLYRAVKRRWHEMAGP
Q9ULM7/194-255      KEIMVGPQ.FQADLSNLHLNRHCEKIYENE----DQLLWDPS.-.---VLPEREVEEFLYRAVKRRWHEMAGP
SPR1_CAEEL/107-170  KEINVGTE.FQAKIADLNLNDKACNEDR-DDQ--DELIWNTP.-.--ETIDDEKLEAFIRESSDRYLIPIDRA
Q9LDD4/372-427      RCIKVGHQ.HQAQVDEWTESGVDSDSKW-----LGTRIWPPE.-.-----NSEALDQTLGNDL---VGKGRPD
O82364/372-428      PCALVGSK.FQAKVPEWTGITPESDSKW-----LGTRIWP--.-.---LTKEQTKANLLIERDR--IGKGRQD
Q84JT7/358-414      PCALVGSK.FQAKVPEWTGITPESDSKW-----LGTRIWP--.-.---LTKEQTKANLLIERDR--IGKGRQD
P91437/126-188      RSIRVDPVlFQADVPLFNE--ATVEESARED---DTILWT-Id-.QTNQPSDEVIDNYLKDVVG--LRKAHDQ
Q19306/20-82        RSIRVDPVlFQADVPLFNE--ATVEESARED---DTILWT-Id-.QTNQPSDEVIDNYLKDVVG--LRKAHDQ
O18691/129-193      LENAVGRE.HQAHVGPFKKLRGRDEEEDR-----DERMWCPPd-dDQNLDYDLLKNAYWRAIWR-QFENHIPM
```

Appendice 2 Alignment of SANT domains by PFAM. Accession number PF00249

```
hMI-ER1/288-331      ---WTEEECRNFEQGLKAYGKD--------FHLIQANK-------.-VRTRSVGECVAFYYMWK
MTA1_HUMAN/285-331 MEEWSASEANLFEEALEKYGKDFTDIQ...QDFLPWKSL.............TSIIEYYYMWK
ADA2_HUMAN/72-118  DPSWTAQEEMALLEAVMDCGFG.......NWQDVANQMCTK.........TKEECEKHYMKHF
ADA2_YEAST/62-108  CPDWGADEELQLIKGAQTLGLG.......NWQDIADHIGSR.........GKEEVKEHYLKYY
EG27_CAEEL/334-380 VDNMTQDDAKKFAKGIKQLGKNFSRIH...RELLPHHSR.............EQLVSYYYLWK
YGN1_YEAST/557-602 EPSFTAVEIRKFEEAVEKFGSE........LRPVCEYVG.........TQPMSMIVRFYYNWK
O75046/125-171       EKCWTEDEVKRFVKGLRQYGKNFFRIR...KELLPNKET.............GELITFYYYWK
YDBJ_SCHPO/248-293 PDIWNEEQHSIFVQQFILHGKKFGKI....AEAVPGKNS.............KECVLHYYLTK
SNT1_YEAST/670-716 SDNFTDHEHSLFLEGYLIHPKKFGKI....SHYMGGLRSP.............EECVLHYYRTK
NCR1_HUMAN/437-482 MNVWTDHEKEIFKDKFIQHPKNFGLI....ASYLERKSV.............PDCVLYYYLTK
NCR2_HUMAN/429-474 MNMWSEQEKETFREKFMQHPKNFGLI....ASFLERKTV.............AECVLYYYLTK
YKZ6_CAEEL/215-260 LEEWSPEERSLFKSRQADHVKI........FHGLTEFFV.........DKTASDLVLFYYMNK
O61907/481-527       LEEWSTPEMNLFEDALDKVGKDFNEIR...AEYLPWKSI.............RDIVEYYYLMK
Q15044/103-148       PDEWTVEDKVLFEQAFSFHGKT........FHRIQQMLP.........DKSIASLVKFYYSWK
O42194/274-320       LSVWTEEECRNFEQGLKAYGKD........FHLIQANKVR........TRSVGECVAFYYMWK
Q20733/562-607       SVLWTPDEIYQFQDAIYQSEKD........FDKVAVELP.........GKSVKECVQFYYTWK
NCR1_HUMAN/625-670 TSRWTEEEMEVAKKGLVEHGRNWAAI....AKMVGTKSE.............AQCKNFYFNYK
P70413/224-268       VGKYTPEEIEKLKELRIKHGND........WATIGAALG..........RSASSVKDRCRLMK
TRF1_CRIGR/380-428 RQAWLWEEDKNLRSGVRKYGEG.......NWSKILLHYKFN.......NRTSVMLKDRWRTMK
TRF2_MOUSE/442-490 KQKWTIEESEWVKDGVRKYGEG.......NWAAISKSYPFV.......NRTAVMIKDRWRTMK
DOT6_YEAST/72-117  PSSWDPQDDLLLRHLKEVKKMG........WKDISQYFP.........NRTPNACQFRWRRLK
YBF4_YEAST/75-120  PSSWDPSDDIKLRHLKEIKNLG........WKEIAHHFP.........NRTPNACQFRWRRLK
O65366/59-104        KESLTQEEQILVINLQAKHGNK........WKKIAAEVP.........GRTAKRLGKWWEVFK
O64877/47-101        TQAWGTWEELLLACAVKRHGFG.......DWDSVATEVRSRS.SLSHLLASANDCRHKYRDLK
O48523/13-66         KQTWSTWEELLLACAVHRHGTE.......SWNSVSAEIQKLSPNLCSLTASA..CRHKYFDLK
Q9SJW0/17-68         RLRWTHELHERFVDAVAQLGGPD....RATPKGVLRVMG.......VQGLTIYHVKSHLQKYR
Q9S807/189-240       RLRWTPELHNRFVNAVNSLGGPD....KATPKGILKLMG.......VDGLTIYHIKSHLQKYR
Q9FH42/25-76         RLRWTADLHDRFVDAVAKLGGAD....KATPKSVLKLMG.......LKGLTLYHLKSHLQKYR
Q9FK47/47-98         RLKWTPDLHERFVEAVNQLGGGD....KATPKTIMKVMG.......IPGLTLYHLKSHLQKYR
```

```
Q9FG27/193-244      RIRWTQDLHEKFVECVNRLGGAD....KATPKAILKRMD.......SDGLTIFHVKSHLQKYR
Q9M0H0/227-278      RMRWTPELHEAFVEAVNSLGGSE....RATPKGVLKIMK.......VEGLTIYHVKSHLQKYR
Q9LSH8/243-294      RMRWTPELHESFVKAVIKLEGPE....KATPKAVKKLMN.......VEGLTIYHVKSHLQKYR
Q9MA23/25-76        RLRWTPELHRSFVHAVDLLGGQY....KATPKLVLKIMD.......VKGLTISHVKSHLQMYR
Q9SIZ5/84-135       RLRWTPELHICFLQAVERLGGPD....RATPKLVLQLMN.......VKGLSIAHVKSHLQMYR
O82801/56-104       RLRWTPDLHLRFVRAVERLGGQE....RATPKLVRQMMN.......IKGLSIAHVKSHLQ...
Q9FJV5/107-158      RMRWTSTLHAHFVHAVQLLGGHE....RATPKSVLELMN.......VKDLTLAHVKSHLQMYR
Q9ZPZ8/88-139       RLRWSSDLHDCFVNAVEKLGGPN....KATPKSVKEAME.......VEGIALHHVKSHLQKFR
Q9XEB2/16-67        RMHWTDDLDIRFIQVIEKLGGEES....ATPKRILSLMG.......VRDLTISHVKSHLQMYR
Q9FGT7/193-243      RVVWSQELHQKFVSAVQQLG.LDK....AVPKKILDLM.......SIEGLTRENVASHLQKYR
Q9FXD6/195-246      RVVWSFELHHKFVNAVNQIGCDHK....AGPKKILDLM.......NVPWLTRENVASHLQKYR
Q9FTT8/220-267      ...WTPELHRRFVQAVEQLG.IDK....AVPSRILELM.......GIECLTRHNIASHLQKYR
Q9LTH4/143-193      RLVWTPQLHKRFVDVVAHLG.IKN....AVPKTIMQLM.......NVEGLTRENVASHLQKYR
Q9LZJ8/213-268      RMQWTPELHHKFEVAVEKMGSLEK....AFPKTILKYMQEEL...NVQGLTRNNVASHLQKYR
Q9FX84/197-245      ...WNPELHRRFVDALQQLGGPG....VATPKQIREHMQ.......EEGLTNDEVKSHLQKYR
Q9FPE8/214-262      ...WSPELHRRFLHALQQLGGSH....VATPKQIRDLMK.......VDGLTNDEVKSHLQKYR
O81713/24-69        RERWTEDEHERFLEALRLYG........RAWQRIEEHIG.......TK..TAVQIRSHAQKFF
O04605/59-104       RENWTDQEHDKFLEALHLFD........RDWKKIEAFVG.......SK..TVVQIRSHAQKYF
O04322/118-165      GVPWTEEEHRLFLVGLQKLGKG.......DWRGISRNYVT.........SRTPTQVASHAQKYF
Q62187/564-608      KGRYNEEDTKKLKAYHSLHGND........WKKIGAMVA..........RSSLSVALKFSQIG
YD3A_SCHPO/55-105   RVKWTEKETNDLLRGCQIHGVG.......NWKKILLDERFH.....FTNRSPNDLKDRFRTIL
O24251/16-68        KQKWTAEEEEALKAGVKKHGMG.......KWKTILVDPDFATA...LTHRSNIDLKDKWRNLG
P78793/32-81        RKQFTPEEDERLLEGFFLHGPC........WTRISKDANLG.....LQNRRSTDLRDRFRNAF
REB1_KLULA/338-382  RGKWTPEEDAELARWCAEKE.G.......QWSN..IGKVLG........RMPEDCRDRWRNYV
Q12457/345-389      RGKWTPEEDQELARLCLEKE.G.......HWTE..VGKLLG........RMPEDCRDRWRNYM
REB1_SCHPO/313-357  RCVWSKEEDEELRKNVVEHG.K.......CWTK..IGRKMA........RMPNDCRDRWRDVV
O65366/4-53         RQRWRPEEDALLRAYVKEYGPR.......DWHLVTQRMN......KPLNRDAKSCLERWKNYL
O62208/356-402      HGRWTDQEDVLLVCAVSRYGAK.......DWAKVAQAVQ.........NRNDSQCRERWTNVL
Q00658/7-54         RGPWVPEEDQLLLQLVREQGPNNN......WVRISQHMH.........YRSPKQCRERYHQNL
MYB1_NEUCR/9-55     RGPWSAGEDQRLIKLVKDLGPGN.......WVNVARILG.........TRTPKQCRERWHQNL
Q9ZTD7/6-52         RGHWRPAEDEKLKDLVEQYGPHN.......WNAIALKLP.........GRSGKSCRLRWFNQL
```

```
Q42575/55-101        KGPWSKEEDDVLSELVKRLGARN.......WSFIARSIP.........GRSGKSCRLRWCNQL
O22179/13-59         KGPWSPEEDDLLQSLVQKHGPRN.......WSLISKSIP.........GRSGKSCRLRWCNQL
O04192/50-96         KGPWLPEQDEALTRLVKMCGPRN.......WNLISRGIP.........GRSGKSCRLRWCNQL
O65263/69-118        KGAWKKEEDELLSELVKDYMENDR....PPWSKISKELP.........GRIGKQCRERWHNHL
O01719/93-139        KGPWTTEEDERVVELVREHGPKR.......WSLISKFLV.........GRTGKQCRERWHNHL
MYB_DROME/136-182    KGPWTRDEDDMVIKLVRNFGPKK.......WTLIARYLN.........GRIGKQCRERWHNHL
O65528/26-72         KGPWSKEEDNTIIDLVEKYGPKK.......WSTISQHLP.........GRIGKQCRERWHNHL
MYBH_DICDI/201-247   KGAWTKDEDDKVIELVKTYGPKK.......WSDIALHLK.........GRMGKQCRERWHNHL
MYBA_CHICK/87-133    KGPWTKEEDQRVIELVQKYGPKR.......WSLIAKHLK.........GRIGKQCRERWHNHL
Q96463/12-59         GVPVDAAECKILVAHIHSHGHG.......NWRALPKQA.........GLLRCGKSCRLRWINYL
Q9ZTC3/10-57         KGAWTAEEDSLLRLCIDKYGEG.......KWHQVPLRA.........GLNRCRKSCRLRWLNYL
O23891/14-61         RGAWTAMEDDILVSYIAKHGEG.......KWGALPKRA.........GLKRCGKSCRLRWLNYL
MYBC_MAIZE/14-61     RGAWTSKEDDALAAYVKAHGEG.......KWREVPQKA.........GLRRCGKSCRLRWLNYL
O49538/14-62         KGPWSPEEDAKLKDYIENSGTGG......NWIALPQKI.........GLRRCGKSCRLRWLNYL
Q38850/25-72         RGPWTVEEDEILVSFIKKEGEG.......RWRSLPKRA.........GLLRCGKSCRLRWMNYL
O23892/22-69         RGPWTPEEDEVLARFVAREGCD.......RWRTLPRRA.........GLLRCGKSCRLRWMNYL
P81394/14-61         RGPWTEEEDQKLTSYVLKNGIQ.......GWRVIPKLA.........GLSRCGKSCRLRWMNYL
O49744/14-61         RGPWTPEEDQILVSFILNHGHS.......NWRALPKQA.........GLLRCGKSCRLRWMNYL
MYB2_PHYPA/14-61     RGPWTSEEDQKLVSHITNNGLS.......CWRAIPKLA.........GLLRCGKSCRLRWTNYL
O49649/14-61         KGPWTVEEDKKLINFILTNGHC.......CWRALPKLA.........GLRRCGKSCRLRWTNYL
Q43597/14-61         KGAWTPEEDKLLVDYIQVNGHG.......SWRLLPKLA.........GLNRCGKSCRLRWTNYL
Q39153/14-61         KGPWSAEEDRILINYISLHGHP.......NWRALPKLA.........GLLRCGKSCRLRWINYL
MYB1_MAIZE/16-63     RGSWTPQEDMRLIAYIQKHGHT.......NWRALPKQA.........GLLRCGKSCRLRWINYL
O04109/14-61         KGPWTPEEDIILVSYIQEHGPG.......NWRSVPINT.........GLMRCSKSCRLRWTNYL
O49782/15-62         KGAWTPEEDQKLLSYLNRHGEG.......GWRTLPEKA.........GLKRCGKSCRLRWANYL
O49765/14-61         KGAWTAEEDKKLISYIHEHGGG.......GWRDIPQKA.........GLKRCGKSCRLRWANYL
Q38739/14-61         KGPWTVDEDQKLLAYIEEHGHG.......SWRSLPLKA.........GLQRCGKSCRLRWANYL
O23618/15-62         KGPWLPEEDDKLTAYINENGYG.......NWRSLPKLA.........GLNRCGKSCRLRWMNYL
O04141/14-61         KGPWTPEEDEKLIAYIKEHGQG.......NWRTLPKNA.........GLSRCGKSCRLRWTNYL
Q40920/14-61         KGAWSAEEDSLLGKYIQTHGEG.......NWRSLPKKA.........GLRRCGKSCRLRWLNYL
MYBP_MAIZE/14-61     RGRWTAEEDQLLANYIAEHGEG.......SWRSLPKNA.........GLLRCGKSCRLRWINYL
```

```
MYB1_HORVU/14-61    KGAWTKEEDDRLTAYIKAHGEG.......CWRSLPKAA........GLLRCGKSCRLRWINYL
Q43524/14-61        RGLWSPEEDEKLIRYITSHGYG.......CWSEVPEKA........GLQRCGKSCRLRWINYL
MYB3_HORVU/16-63    KGLWSPEEDEKLYNHIIRHGVG.......CWSSVPRLA........ALNRCGKSCRLRWINYL
O49018/14-61        KGAWTAEEDRKLAEVITVHGAK.......RWKTIPSIA........GLNRCGKSCRLRWMNYL
O49017/15-62        KGSWTAEEDRRLAKYIEIHGAK.......RWKTIAIKS........GLNRCGKSCRLRWLNYL
O80883/20-67        KGPWTTTEDAILTEYVRKHGEG.......NWNAVQKNS........GLLRCGKSCRLRWANHL
O23893/43-90        KGPWTSWEDSILEKYIKKHGER.......NWKLVQKNT........GLLRCGKSCRLRWMNHL
Q39028/22-69        KGPWTEEEDAILVNFVSIHGDA.......RWNHIARSS........GLKRTGKSCRLRWLNYL
O50069/10-57        KGPWTEQEDILLVNFVHLFGDR.......RWDFVAKVS........GLNRTGKSCRLRWVNYL
P92986/20-67        KGPWTAEEDRLLIDYVQLHGEG.......RWNSVARLR........GLKRNGKSCRLRWVNYL
O49020/22-69        KGPWTEEEDSMLRAYVNIHGEG.......RWNAVARLS........GLRRTGKSCRLRWLNYM
Q39550/29-76        RGPWTVDEDFTLINYIAHHGEG.......RWNSLARFA........GLKRTGKSCRLRWLNYL
O22684/22-69        KGPWTMEEDLILINYIANHGDG.......VWNSLAKSA........GLKRTGKSCRLRWLNYL
GL1_ARATH/16-63     KGLWTVEEDNILMDYVLNHGTG.......QWNRIVRKT........GLKRCGKSCRLRWMNYL
BAS1_YEAST/169-214  LREWTLEEDLNLISKVKAYGTK........WRKISSEME.........FRPSLTCRNRWRKII
O65263/12-63        KPFFHLMQDQILTNVVKKYQGR.......NWKRIAECLPGSE....ENRRNDVQCQHRWLKVL
MYB_DROME/85-130    GKRWSKSEDVLLKQLVETHG.E.......NWEIIGPHFK.........DRLEQQVQQRWAKVL
O01719/41-87        RARWTKDEDDMLRQAIEVHGTL.......DWKLIGSFFP.........NRSELQCFHRWQKVL
MYBB_XENLA/31-77    KVKWTPEEDETLKALVKKHGQG.......EWKTIASNLN.........NRTEQQCQHRWLRVL
MYBB_CHICK/31-77    KVKWTQEEDEQLKMLVRHYGQN.......DWKFLASHFP.........NRSDQQCQYRWLRVL
MYBA_XENLA/35-80    KLRWTKDEDDKVKKLVEKHG.E.......DWGVVARHFI.........NRSEVQCQHRWHKVL
MYB_BOVIN/40-86     KTRWTREEDEKLKKLVEQNGTD.......DWKVIANYLP.........NRTDVQCQHRWQKVL
MYBA_CHICK/35-81    RVKWTRDEDEKLKKLVEQNGTD.......DWAFIASHLQ.........NRSDFQCQHRWQKVL
Q9ZPU2/40-86        KGGWTPEEDETLRRAVEKYKGK.......RWKKIAEFFP.........ERTQVQCLHRWQKVL
MYBH_DICDI/149-195  KGKWTSEEDQILIKAVNLHNQK.......NWKKIAEHFP.........DRTDVQCHHRYQKVL
BAS1_YEAST/116-161  KGKWTQEEDEQLLKAYEEHGPH........WLSISMDIP.........GRTEDQCAKRYIEVL
O15816/427-474      PNKWTKEESQNLIKLVTENGDK.......QWKKIATKLG........GGKTGAQCAQHWKRVL
O64502/30-76        IVTWSPEEDDILRKQISLQGTEN.......WAIIASKFN.........DKSTRQCRRRWYTYL
CEF1_YEAST/10-56    GGVWTNVEDQILKAAVQKYGTH.......QWSKVASLLQ.........KKTARQSELRWNEYL
Q21119/8-54         GGVWKNTEDEILKAAIMKYGKN.......QWSRIASLLH.........RKSAKQCKARWFEWL
CC5_SCHPO/6-52      GGAWKNTEDEILKAAVSKYGKN.......QWARISSLLV.........RKTPKQCKARWYEWI
```

```
O62208/301-350      SKEWSQDEDTKLIALTKITSING....HIQWDKVAQCMP.........GRTRQQVRTRFSHTL
CEF1_YEAST/62-106   FTEFSKEEDAQLLDLARELPNQ........WRTIADMM..........ARPAQVCVERYNRLL
CC5_SCHPO/58-102    KTEWSREEDEKLLHLAKLLPTQ........WRTIAPIV..........GRTATQCLERYQKLL
Q01073/213-264      KSSFTKEEDEFILDVVRKNPTKR..TTHTLYDEISHYVP.........NHTGNSIRHRFRVYL
O14108/327-371      RNAWEESELKKLYTLVEQEGTR........WNSIANKLG..........TSPAACMSQWRFVV
O22059/35-80        AVKMSEEEEDLISRMYKLVGDR........WELIAGRIP.........GRTPEEIERYWLMKH
Q00658/60-105       RDPISAEEGLAIERMVNEMGRC........WAEIARRLG.........NRSDNAVKNWWNGNM
MYB1_NEUCR/61-106   HGPMTQEEAAIIVREVDLKGPR........WADIARKLQ.........GRSDNAVKNYWNGLN
O64502/82-127       RGGWSPEEDTLLCEAQRLFGNR........WTEIAKVVS.........GRTDNAVKNRFTTLC
Q9ZTD7/58-103       RNPFTEEEEERLLAAHRIHGNR........WSIIARLFP.........GRTDNAVKNHWHVIM
Q9ZTD6/3-48         KRAFTEEEEFRLLAAHRAYGNK........WALISRLFP.........GRTDNAVKNHRHVIM
O65528/78-123       KNAWTQEEELTLIRAHQIYGNK........WAELMKFLP.........GRSDNSIKNHWNSSV
O65263/124-169      KSPWTREEELILVQAQRGNGNK........WAEIAKLLP.........GRTENNIKNHWNCSV
Q42575/107-152      RNSFTEVEDQAIIAAHAIHGNK........WAVIAKLLP.........GRTDNAIKNHWNSAL
O22179/65-110       HRGFTAEEDDTIILAHARFGNK........WATIARLLN.........GRTDNAIKNHWNSTL
O04192/102-147      RKPFSDEEEHMIMSAQAVLGNK........WSVIAKLLP.........GRTDNAIKNHWNSNL
MYBH_DICDI/253-298  KEAWSDEEDQIIRDQHAIHGNK........WAEIAKFLP.........GRTDNAIKNHWNSSM
MYB_DROME/188-233   KTAWTEKEDEIIYQAHLELGNQ........WAKIAKRLP.........GRTDNAIKNHWNSTM
MYBA_CHICK/139-184  KSSWTEAEDRVIYEAHKRLGNR........WAEIAKLLP.........GRTDNSIKNHWNSTM
O80883/73-118       KGSFTPDEEKIIIDLHAKLGNK........WARMASQLP.........GRTDNEIKNYWNTRM
O65898/13-58        KGAFSQEEEQLIVEMHAKMGNK........WAQMAEHLP.........GRTDNEIKNYWNTRI
O23893/96-141       KGAFSKEEENKIINLHRKMGNK........WSRMAADLP.........GRTDNEIKNYWNTRI
O65720/14-59        RGGFTDAEEDRIMELHSQLGNR........WSKIASHFS.........GRTDNEIKNHWNTKI
P81394/67-112       KGPLTEMEENQIIELHAHLGNR........WSKIALHIP.........GRTDNEIKNYWNTHI
O04110/67-112       RGLLSETEEKTVIDLHEQLGNR........WSKIASHLP.........GRTDNEIKNHWNTHI
Q9ZR24/33-78        RGPFSTEEEKLVIQLHGILGNR........WAAIASQLP.........GRTDNEIKNLWNTHL
Q43524/67-112       RGRFTPEEEKLIISLHGAVGNR........WAHIASHLP.........GRTDNEIKNYWNSWI
O49019/67-112       RGNFTSQEEEVIINLHATLGNR........WSLIASYLP.........GRTDNEIKNYWNSHL
O22264/67-112       RGNITPEEEELVVKLHSTLGNR........WSLIAGHLP.........GRTDNEIKNYWNSHL
Q9ZQS3/10-55        RGKFSVDEEELMIKLHALLGNR........WSLIAGRLP.........ERTDNEVKNYWNSHM
MYBC_MAIZE/67-112   RGNISYDEEDLIIRLHRLLGNR........WSLIAGRLP.........GRTDNEIKNYWNSTL
```

```
O22685/67-112        RGAFSQDEESLIIELHAALGNR........WSQIATRLP.........GRTDNEIKNFWNSCL
MYB3_HORVU/69-114    RGCFSQQEEDHIVALHQILGNR........WSQIASHLP.........GRTDNEIKNFWNSCI
Q9ZTE9/14-59         RGPFSEEEEETILTLHSSLGNK........WSRIAKYLP.........GRTDNEIKNYWHSYL
MYB2_PHYPA/67-112    RGIFSEAEENLILDLHATLGNR........WSRIAAQLP.........GRTDNEIKNYWNTRL
O04108/67-112        RGNFTADEEDLIVRLHNSLGNR........WSAIAAQMP.........GRTDNEIKNVWHTHL
O04140/67-112        RGNFSKEEEDTIIHLHELLGNR........WSAIAARLP.........GRTDNEIKNVWHTHL
MYB1_MAIZE/69-114    RGNFTDEEEEAIIRLHGLLGNK........WSKIAACLP.........GRTDNEIKNVWNTHL
Q9ZTC8/3-48          RGRFSFEEEETIIQLHGIMGNK........WSAIAARLP.........GRTDNEIKNYWNTHI
MYB1_HORVU/67-112    RGNFSHEEDELIIKLHSLLGNK........WSLIAGRLP.........GRTDNEIKNYWNTHI
O49782/68-113        RGEFTEDEERSIISLHALHGNK........WSAIARGLP.........GRTDNEIKNYWNTHI
O49765/67-112        RGEFSYEEEQIIIMLHASRGNK........WSVIARHLP.........KRTDNEIKNYWNTHL
Q02992/67-112        RGKFTLQEEQTIIQLHALLGNR........WSAIATHLP.........KRTDNEIKNYWNTHL
O23618/68-113        RGKFSDGEESTIVRLHALLGNK........WSKIAGHLP.........GRTDNEIKNYWNTHM
O64399/67-112        RGEFSPEEDDTIIKLHALKGNK........WAAIATSLA.........GRTDNEIKNYWNTNL
O04109/67-112        RGNFTAHEEGIIVHLQSLLGNR........WAAIASYLP.........QRTDNDIKNYWNTHL
Q39156/35-80         QRNFSKDEDDLILKLHALLGNR........WSLIAGRLP.........GRTDNEVRIHWETYL
Q9ZTC3/63-108        RGRLSNDEVDLLLRLHKLLGNR........WSLIAGRLP.........GRTANDVKNYWNTHL
GL1_ARATH/69-114     KGNFTEQEEDLIIRLHKLLGNR........WSLIAKRVP.........GRTDNQVKNYWNTHL
O50069/63-108        RGKMTPQEERLVLELHAKWGNR........WSKIARKLP.........GRTDNEIKNYWRTHM
O49020/75-120        RGNISLEEQLLILELHSRWGNR........WSKIAQHLP.........GRTDNEIKNYWRTRV
O22684/75-120        RGNITPEEQLIIMELHAKWGNR........WSKIAKHLP.........GRTDNEIKNFCRTRI
O49780/14-59         RGKFTPQEEEEIIQLHAVLGNR........WAAMAKKMQ.........NRTDNDIKNHWNSCL
O49770/14-59         HDSFSTQEEELIIECHRAIGSR........WSSIARKLP.........GRTDNDVKNHWNTKL
Q9ZTE4/14-59         HGDFTEEEDNIIYSLFASIGSR........WSVIAAHLQ.........GRTDNDIKNYWNTKL
O49538/68-113        HGGFSEEEDNIICNLYVTIGSR........WSIIAAQLP.........GRTDNDIKNYWNTRL
Q61866/451-507       SKNWSEDDLQLLIKAVNLFPAGTN....SRWEVIANYMNIHSS..SGVKRTAKDVISKAKSLQ
YMA9_CAEEL/215-264   EDDWSQAEQKAFETALQKYPKGTD....ERWERISEEIG.........SKTKKQVMVRFKQLA
Q94216/533-582       EDVWSATEQKTLEDAIKKHKSSDP....ERWEKISTEVG.........TKSKKACIRRFKYLV
DJC1_MOUSE/492-541   EEAWTQSQQKLLELALQQYPKGAS....DRWDKIAKCVP.........SKSKEDCIARYKLLV
O15816/480-525       KGSWDEAEEELLFQLVDKHGQS........WKNVAIEIK.........TRTDIQCRYQYFKAI
REB1_SCHPO/365-418   RNAWSLEEETQLLQIVAELRNREDLSSDINWTLVAQMLG.........TRTRLQCRYKFQQLT
```

```
SWI3_YEAST/524-569 DENWSKEDLQKLLKGIQEFGAD........WYKVAKNVG.........NKSPEQCILRFLQLP
O22811/225-270     KPEWSDKEILLLLEAVMHYG........DDWKKVASHVI.........GRTEKDCVSQFVKLP
YFK7_YEAST/312-358 KKNWSDQEMLLLLEGIEMYE........DQWEKIADHVGG........HKRVEDCIEKFLSLP
O13788/290-335     DDTWTAQELVLLSEGVEMYS........DDWAKVASHVN.........TKSVEECILKFLNLP
O14470/247-292     EKPWSNQETLLLLEAIETYG........DDWNQIALHVG.........SRTKEQCLIHFLQIP
Q9VF03/671-716     AREWTDQETLLLLEGLEMHK........DDWNKVCEHVG.........SRTQDECILHFLRLP
O23486/88-134      CPDWSADDEMLLLEGLEIYGLG.......NWAEVAEHVGTK.........SKEQCLEHYRNIY
O62208/409-455     NERFTLVEDEQLLYAVKVFGKG.......NWAKCQMLLP.........KKTSRQLRRRYLQLI
O22895/127-173     DSVWTKEETDQLFEFCQNFDLR........FVVIADRFP.........VSRTVEELKDRYYSVN
TBF1_YEAST/409-456 KRTWSKEEEEALVEGLKEV.GP.......SWSKILDLYGPG.......GKITENLKNRTQVQL
O42961/405-451     RRSWTKEEEEALLDGLDLVKGP.......RWSQILELYGPG.......GKKSEVLKYRNQV..
SNT1_YEAST/889-934 TSYWSVRESQLFPELLKEFGSQ........WSLISEKLG.........TKSTTMVRNYYQRNA
Q19953/77-125      AMSWTHEDEFELLKAAHKFKMG.......NWGEIAESIG.......RGRKDGHNCKDYFEKHF
O80435/19-68       GTKWTAEENKKFENALAFYDKDTP....DRWSRVAAMLP.........GKTVGDVIKQYRELE
TFC5_YEAST/417-462 TDPWTVEEMIKFYKALSMWG........TDFNLISQLY........PYR.SRKQVKAKFVNEE
O94481/365-410     PEKWNAMDTEKFYKALSQWG........TDFALIANMF........PTR.NRRQIKLKFKQEE
YDPA_SCHPO/71-116  GTYWSAEEKELFFQAVARNGKR.......DLDLIAYSIP.........SKSAVQIERYINAL.
Q15044/294-339     NARWTTEEQLLAVQAIRKYGRDFQAI....SDVIGNKSV.............VQVKNFFVNYR
SPR1_CAEEL/483-528 SQDWTQLERSQVIRCFNMYGAH........FEHIADVIG.........TKTPDQVYQFYLENQ
```