The expression of an EMT regulator ZEB2 is controlled by a new destabilisation motif

Thesis submitted for the Degree of

Doctor of Philosophy

at the University of Leicester

By

Wan Ratmaazila Wan Makhtar BSc MSc

Department of Cancer Studies

University of Leicester

2015

The expression of an EMT regulator ZEB2 is controlled by a new destabilisation motif

Wan Ratmaazila Wan Makhtar

ABSTRACT

Epithelial-mesenchymal transitions (EMT) are genetic re-differentiation programs which generate motile and invasive cells during normal embryonic development and cancer metastasis. ZEB1 and ZEB2 are two-handed zinc finger transcription repressors which belong to the Zinc finger E-box (ZEB) protein family. These proteins inhibit transcription of E-cadherin gene and induce EMT *in vitro*. However, the regulation of *ZEB* genes is still incompletely understood. Given an important role for both proteins in EMT in normal development and cancer metastasis, better understanding of molecular mechanisms controlling their expression is an important challenge. Intriguingly, a comparative expression analysis of ZEB1 and ZEB2 proteins in several cancer cell lines indicated that ZEB1 protein was expressed at much higher level than ZEB2 mostly in mesenchymal carcinoma cells. On the other hand, these cell lines contained high levels of ZEB2 and ZEB1 mRNA. Therefore, we suggested that an additional specific control mechanism limiting ZEB2 protein synthesis exists; and we addressed this issue in greater detail. We showed that a protein motif adjacent to the smad-binding domain within ZEB2 protein induced ribosome stalling and compromised ZEB2 translation. The activity of this motif was dependent on triplets of rare codons, Leu(UUA)-Gly(GGU)-Val(GUA). However, introducing these rare codons in the ZEB1 region had no effect on ZEB1 protein expression. In conclusion, my study suggests that rare codon play regulatory role in the context of appropriate protein structures and may influence configuration of EMT programs in cancer cells.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank to my supervisor Dr Eugene Tulchinsky, for his valuable guidance and advice. His willingness to motivate me contributed tremendously to my PhD journey.

Also, I would like to take this opportunity to thank to the EMT group, past and present, Dr Marina Kriajevska, Dr Emre Sayan, Dr Gareth Browne and Mrs Karen Kulbicki for their helps, supports and unlimited advices throughout my PhD. Special thank you also go to Gina Tse, Andrew Irvine, Louise Hill and Youssef Saeed who always give their great friendship!

My warmest thanks go to my family, especially to Mama and Papa for their understandings and supports in completing this PhD. Special thanks also to all my friends; Dona Touma, Roslina Mat Yazid, Malinna Jusoh, Farra Aidah Jumuddin, Hasbullah Daud and Mohd Harizal Senik for their supports and jokes that took me through my bad times.

This PhD journey would not have been possible without my husband, Mohd Ridzuan Junoh, who was always there for me listening to my complaints and frustrations. You are my best friend and everything!

I am also very grateful to the Malaysian of Higher Education (MOHE) and the University Sains Malaysia for giving me this fellowship that helps me to archive my dreams.

Lastly, I offer my regards and blessings to all of those who were involved direct or indirectly during the completion of my PhD.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

%	- Percentage
µF	- Microfarad
µg	- Microgram
µg/µl	- Microgram/microliter
µl	- Microliter
µM	- Micromolar
∞	- Infinity
aa	- amino acid
APS	- Ammonium persulfate
ATCC	- American Type Culture Collection
ATP	- Adenosine-5'-triphosphate
BCA	- Bicinchoninic acid
bHLH	- Basic helix-loop-helix
BMP	- Bone morphogenetic protein
bp	- Base pair
BSA	- Bovine serum albumin
Ca^{2+}	- Calcium
CBD	- C-terminal catenin binding domain
CC	- Common codon
CCC	- Cadherin-catenin complex
cDNA	- Complementary deoxyribonucleic acid
ChiP	- Chromatin immunoprecipitation
CHX	- Cycloheximide
CID	- CtBP-interacting domain
cm	- Centimetre
Cm^2	- Centimetre squared
CMV	- CMV (cytomegalovirus)
CO_2	- Carbon dioxide
CRC	- Colorectal cancer
Cre	- Cyclisation recombination
CtBP	- C-terminal binding protein
DMEM	- Dulbecco's modified eagle media
DMSO	- Dimethyl sulphoxide
DNA	- Deoxyribonucleic acid
E6AP	- Homologous to E6-Associated Protein
EBV	- Epstein-Barr virus
ECL	- Enhanced chemiluminescence
ECM	- Extracellular matrix
EDTA	- Ethylenediaminetetraacetic acid
EGFR	- Enhanced green fluorescence protein
EMT	- Epithelial-mesenchymal transition
EMT-r	- Epithelial-mesenchymal transition regulator
ER	- Endoplasmic reticulum
EtBr	- Ethidium bromide
FBS	- Foetal bovine serum

FBXL14 - F-box and leucine-rich repeat protein 14
FGF - Fibroblast growth factor
GAPDH - Glyceraldehyde-3-phosphate dehydrogenase
GFP - Green fluorescent protein
GSK-3β - Glycogen synthase kinase-3β
HCl - Hydrochloric acid
HD - Homeodomain
HECT - homologous to E6-associated protein C-terminus
HEK-293 - Human embryonic kidney-293
HIF-1 - Hypoxia inducible factor-1
HRP - Horseradish peroxidase
HSF1 - Heat shock factor 1
hTERT - Human telomerase reverse transcriptase
Ig - Immunoglobulin
IHC - Immunohistochemical
JMD - Juxtamembrane domain
kb - Kilobase
kDa - Kilodalton
LB - Luria broth
Lgl2 - Lethal giant larvae
LOH - Loss of heterozygosity
LoxP - Locus of X-over of P1
MAPK - Mitogen-activated protein kinase
MDCK - Madin Darby Canine Kidney
MDM2 - murine double minute 2
MET - Mesenchymal-epithelial transition
mg/ml - Milligram/millilitre
miR/miRNA - MicroRNA
ml - Millilitre
mM - Milimolar
MMP - Matrix metalloproteinase
MOPS - 3-(N-morpholino) propanesulfonic acid
mRNA - Messenger ribonucleic acid
MW - Molecular weight
N/A - Not available
NaCl - Sodium chloride
NAT - Natural antisense transcription
NCBI - Nucleotide basic local alignment search tool
NCI - National Cancer Institute
NEAA - Non-essential amino acid
ng - Nanogram
NSCLC - Non-small lung carcinoma
^{o}C - Degree Celsius
OCT4 - Octamer-binding transcription factor 4
OGT - O-Glycosyltransferase
ORF - Open reading frame
P/CAF - P300/CBP-associated factor

p53 - Tumour protein 53
PAK1 - p21- activated kinase 1
PBS - Phosphate buffered saline
PC2 - Polycomb-2 protein
PCR - Polymerase chain reaction
PIAS - Protein inhibitor of activated STAT
PKB - Protein kinase B
PS - Penicillin-streptomycin
PTM - Post-translational modification
PVDF - Polyvinylidene fluoride
RanBP2 - Ran-binding protein 2
Rb - Retinoblastoma protein
RC - Rare codon
RCC - Renal cell carcinoma
RING - really interesting new gene
RNA - Ribonucleic acid
RNase - Ribonuclease
rpm - Revolutions per minute
RPMI - Rosewell Park Memorial Institute
R-Smads - Receptor-regulated SMADs
RT-PCR - Real-time polymerase chain reaction
SBD - Smad-binding domain
SCC - Squamous cell carcinoma
SDS - Sodium dodecyl sulphate
SDS-PAGE - Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SENP - Sentrin/SUMO-specific protease
SIP1 - Smad-interacting protein 1
SiRNA - Short interfering ribonucleic acid
SM - Smooth muscle
SNAIL1 - Snail homolog 1
SNAIL2 - Snail homolog 2
SOX2 - SRY (sex determining region Y)-box 2
SUMO - Small ubiquitin-like modifier
TAE - Tris acetic acid EDTA
TBS-T - Tris-buffered saline-Tween
TCA - Trichloroacetic acid
TE - Tris-EDTA
TEMED - Tetramethylethylenediamine
TGF-β - Transforming growth factor-beta
TM - Transmembrane domain
Tm - Annealing temperature
tRNA - Transfer RNA
TSS - Transcriptional start site
TWIST1 - Twist-related protein 1
Ub - Ubiquitin
UBA6 - Ubiquitin-like modifier activating enzyme 6
Ubc9 - Ubiquitin-conjugating 9-E2 conjugating enzyme

UBE1	- Ubiquitin-activating enzyme 1
UK	- United Kingdom
UPS	- Ubiquitin-proteasome system
UTR	- Untranslated region
UV	- Ultraviolet
V	- Volt
v/v	- Volume by volume
VHL	- von-Hippel-Lindau
w/v	- Weight by volume
WB	- Western blotting
ZEB1	- Zinc finger E-box-binding homeobox 1
ZEB2	- Zinc finger E-box-binding homeobox 2
ZHX1	- Zinc finger homeodomain 1
Ω	- Ohm

CHAPTER 1:

INTRODUCTION

1.1 CANCER

Cancer is a disease of uncontrolled cell growth in which normal cells acquire genetic and epigenetic alterations that transform them to malignant tumour cells. A review by D. Hanahan and R. Weinberg suggested that these alterations occur in nearly all cancers and include the capability of cancer cells to:

1. Sustain chronic proliferation
2. Evade growth suppressors
3. Resist apoptosis
4. Archive replicative immortalisation
5. Induce angiogenesis
6. Activate mechanisms for invasion and metastasis
7. Induce genome instability and mutation
8. Enhance tumour-promoting inflammation
9. Re-program glycolysis energy metabolism
10. Evade immune destruction (Hanahan and Weinberg, 2011).

From the clinical point of view, the most important acquired capability is the invasion and metastasis, which cause the death of patients in the majority of cancer cases. Metastasis is a multi-step process defined as the ability of tumour cells to migrate from its original site and to colonise distant organs (Spano et al., 2012). During metastasis, cancer cells detach from the solid primary tumour, invade neighbouring tissue and penetrate through the basement membrane and extracellular matrix (ECM). Following angiogenesis, cancer cells intravasate into a blood or lymphatic vessel, circulate around the body and halt within the capillary bed of the secondary organ. Finally, cancer cells extravasate through the vessel and generate a secondary tumour (Fidler, 2002, Weinberg, 2008, Spano et al., 2012) (Figure 1.1).

Many aspects of molecular interactions and mechanisms underlying the metastatic progression have been identified. A process referred as the epithelial-mesenchymal transition (EMT) has been proposed as a critical mechanism enabling epithelial cancer cells to invade and metastasise (Thiery, 2002, Thiery et al., 2009).

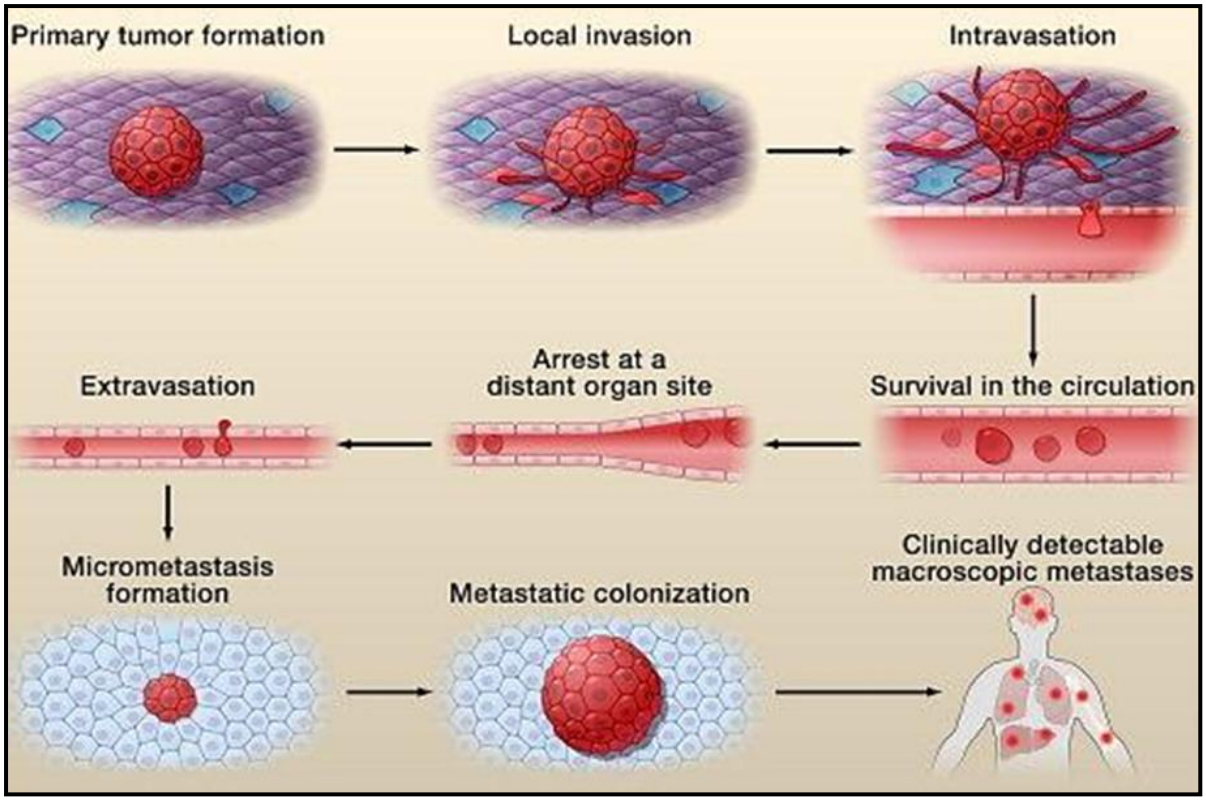


Figure 1.1: The metastasis cascade. During metastasis, tumour cells leave the primary tumour, invade into the local tissue, penetrate into blood vessel (intravasation), survive in the circulation and arrest in a new organ. At this point, tumour cells extravasate from the circulation into the nearby tissue, form micrometastasis and proliferate extensively to develop a metastatic colony and macroscopic metastases (Valastyan and Weinberg, 2011).

1.2 EMT

Epithelial cells are morphologically round-shaped and grow as clusters. These cells demonstrate an apical-basolateral polarisation and a typical basal lamina at the basal surface. The motility of epithelial cells is limited; they migrate collectively without damaging tissue architecture. In vivo and in vitro epithelial cells form sheets in which they are attached to each other by specialised junctions such as tight junctions, adherens junctions, gap junctions and desmosomes (Thiery and Sleeman, 2006, Christiansen and Rajasekaran, 2006). In contrast, mesenchymal cells have a spindle-shaped, fibroblast-like morphology; they contact neighbouring cells only focally, without forming organised cellular junctions. Mesenchymal cells do not exhibit the same apical-basolateral organisation as epithelial cells and have the potential to be highly migratory and invasive. Furthermore, the intermediate filaments in mesenchymal cells are composed of vimentin, whereas cytokeratins are

predominant components of intermediate filaments in epithelial cells (Thiery and Sleeman, 2006) (Figure 1.2a).

EMT is a highly conserved cellular program that allows polarised epithelial cells to convert into mesenchymal cells (Kalluri and Neilson, 2003). This conversion is accompanied by massive changes in gene expression pattern, which include down-regulation of epithelial markers such as E-cadherin (epithelial cadherin), occludins and desmoplakin and up-regulation of mesenchymal markers such as vimentin, smooth muscle actin, fibronectin and collagen (Christiansen and Rajasekaran, 2006). The reverse process, a mesenchymal-epithelial transition (MET) is believed to take part in the development of the metastatic growth by allowing cancerous cells to re-convert into epithelial cells and integrate into distant organ (Yang and Weinberg, 2008) (Figure 1.2b).

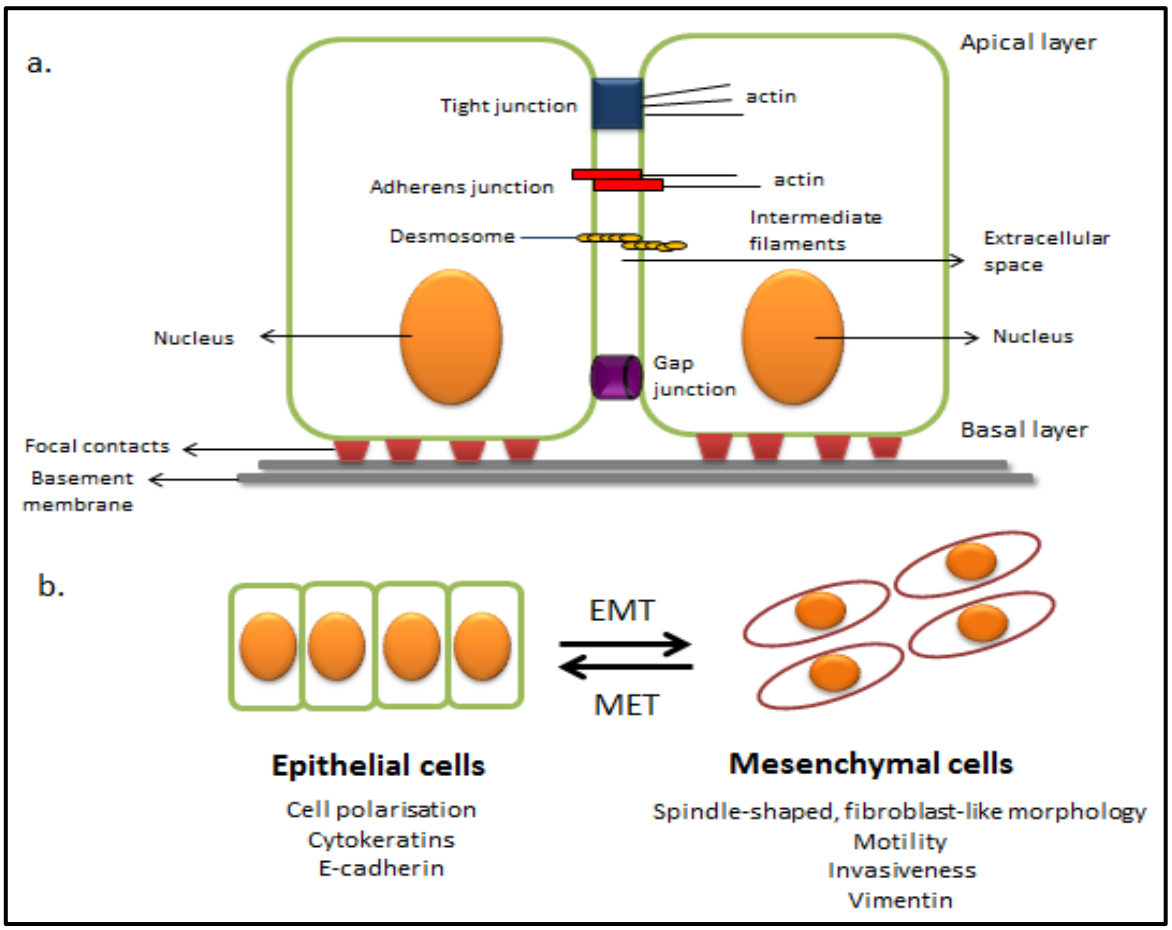


Figure 1.2: Epithelial junctional complexes and the processes of EMT and MET. (a) The junctional complexes within an epithelial cell are responsible for epithelial polarisation and tissue integrity. Tight junctions are located at the top of the apical layer, act as a barrier to passage variety of molecules between adjacent epithelial cells and are connected to the actin cytoskeleton. Adherens junctions consist of the trans-membrane protein E-cadherin, located immediately below the tight junctions and also are linked to the actin. Desmosomes

are made up of proteins such as desmosomal cadherins, the armadillo and the plakin family of cytolinkers. Desmosomes are connected to the intermediate filaments and involved in the maintenance of cellular shape and tissue integrity. Finally, gap junctions are intercellular membrane domains which connect neighbouring cells and allow the transmission of small molecules such as ions, amino acids and other metabolites between cells. (b) During EMT, epithelial cells lose their characteristic cell-cell adhesion structures, change their polarity and down-regulate expression of E-cadherin and other epithelial markers. This results in the formation of migratory and invasive mesenchymal cells. The reverse process, MET reverts the mesenchymal cells back into cells with an epithelial phenotype (Figure drawn by W.R. Wan Makhtar).

The process of EMT was first recognised as a feature of embryogenesis (Wu and Zhou, 2008). During implantation, the selected cells within the embryonic region undergo EMT to initiate a proper placental formation and anchorage. Following implantation, the embryo undergoes dramatic changes in size and shape leading to the formation of a trophoblast, hypoblast, epithelialised epiblast and amniotic cavity. During gastrulation, the epithelialised epiblast cells located in the early primitive streak experience EMT to generate three embryonic germ layers including endoderm (inner layer), mesoderm (middle layer) and ectoderm (outer layer), basic body plan and the primitive gut. After gastrulation, the neural crest forms at the boundary between the epidermal and neural territories with neural crest cells also undergoing EMT allowing cells to migrate. At the destination sites, neural creast cells differentiate into different cell types including craniofacial structures, melanocytes, endocrine cell and most of the peripheral nervous system (Thiery et al., 2009, Kalluri and Weinberg, 2009).

EMT is important not only during embryogenesis, but also during other normal and pathological events, such as wound healing, tissue repair and organ fibrosis. During these events, EMT generates fibroblasts that encourage tissue repair following trauma and injury. However, tissue fibrosis may occur if the EMT signals are persistent during this process. In addition, EMT is also associated with carcinoma invasion and metastasis. EMT alters adhesive properties and stimulates the motility mechanism of the cancer cells, allowing these cells to metastasise to the secondary organ. Apart of inducing carcinoma cell motility and invasiveness, EMT also affects other functions essential for cancer progression including premature senescence and apoptosis. Furthermore, EMT also allows cancer cells to acquire stem cell-like properties, accumulate resistance to chemotherapy and immunotherapy,

induce inflammation-initiated tumour and escape immune surveillance and immunosuppression (Thiery and Sleeman, 2006, Mani et al., 2008, Thiery et al., 2009, Kalluri and Weinberg, 2009, Sayan et al., 2009, Hill et al., 2013).

1.2.1 EMT and E-cadherin

Loss of E-cadherin function is a hallmark of EMT. A large cadherin superfamily genes comprising approximately 80 members has been identified in human genome (Tepass et al., 2000). E-cadherin, P-cadherin (placental cadherin) and R-cadherin (retinal cadherin) are prototypical examples of classical type l cadherins (Baranwal and Alahari, 2009) . E-cadherin (encoded by the *CDH1* gene) is 120 kilodaltons (kDa) calcium (Ca^{2+})-depending transmembrane glycoprotein predominantly found on the surface of epithelial cells, whereby it allows strong homophilic binding between neighbouring cells. E-cadherin interacts with the members of the catenin family such as α-catenin, β-catenin and p120-catenin to form the cadherin-catenin complex (CCC) (van Roy and Berx, 2008). The p120-catenin and β-catenin contain armadillo repeats that bind directly to E-cadherin in adherens junctions, whereas the association between α-catenin and E-cadherin is indirect and occurs via β-catenin. The homologue of β-catenin, γ-catenin (also known as plakoglobin) may substitute for β-catenin to bind to E-cadherin. As α-catenin interacts with actin-binding proteins, such as α-actinin or vinculin, these interactions thus link E-cadherin to the actin cytoskeleton (Niessen, 2007) (Figure 1.3).

The E-cadherin promoter has been studied in various mammals, and the analyses of human and murine found that both promoters contain two positive regulatory elements, a CCAAT-box and two AP-2 binding sites located in a GC-rich region. The murine E-cadherin also contains the palindromic sequence comprising of the two E-boxes (5'-CANNTG-3') known as E-pal element. However, the E-pal element is not conserved in the human promoter and contains only one E-box (Hennig et al., 1996, Peinado et al., 2004) (Figure 1.4). It has been reported that binding of zinc finger transcription factors to the E-boxes within the E-cadherin promoter causes suppression of E-cadherin expression both *in vitro* and *in vivo*, thereby leading to the disruption of cell-cell adhesions and contributing to EMT (Batlle et al., 2000, Cano et al., 2000, Comijn et al., 2001, Smit and Peeper, 2010).

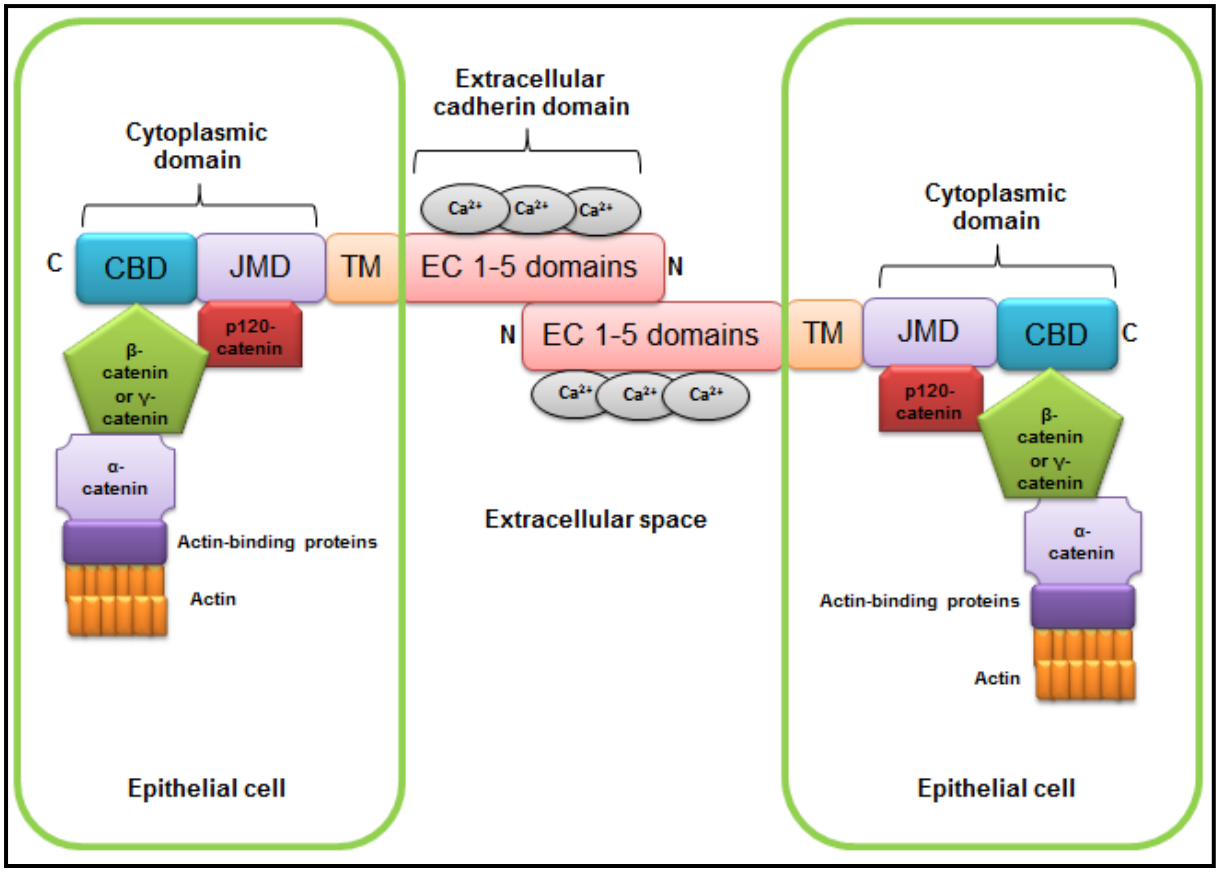


Figure 1.3: Structure of E-cadherin in adherens junction of epithelial cells. E-cadherin consists of five extracellular cadherin domains (EC 1-5) that bind to several Ca^{2+} ions, a transmembrane domain (TM) and a cytoplasmic domain containing juxtamembrane domain (JMD) and C-terminal catenin binding domain (CBD) which governs binding to the actin cytoskeleton through interactions with p120-catenin and β-catenin or γ-catenin. The interaction of α-catenin to β-catenin or γ-catenin links the E-cadherin complex to the actin cytoskeleton (Figure drawn by W.R. Wan Makhtar).

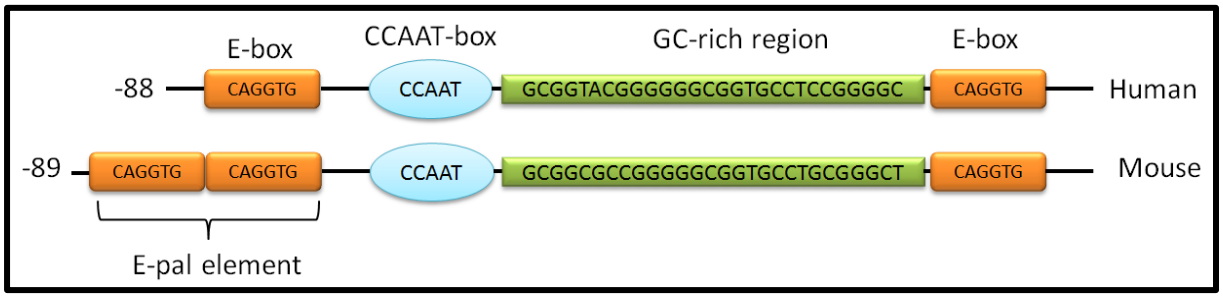


Figure 1.4: The human and mouse E-cadherin promoter region. Both E-cadherin promoters contain several functional cis-elements such as E-boxes, CCAAT-box and GC-rich region, whereas the E-pal element is present only in the mouse promoter (Peinado et al., 2004). Sequences were obtained from http://www.ncbi.nlm.nih.gov/nuccore/AY341818.1 for human and http://www.ncbi.nlm.nih.gov/nuccore/M81449.1 for mouse.

As a key component of adherens junction, E-cadherin is important during normal embryonic development, tissue morphogenesis and normal and pathological involving EMT (Baranwal and Alahari, 2009). The importance of E-cadherin during the embryonic development and tissue morphogenesis was proven *in vivo*, because *CDH1* knockout mouse embryos were unable to form normal trophectodermal epithelium during embryogenesis and died around the time of implantation (Larue et al., 1994). Furthermore, E-cadherin is crucial for the maintenance of the pluripotent state of mouse embryonic stem cells, with loss of E-cadherin expression resulted in the acquisition of EMT. It has also been reported that exogenous expression of E-cadherin is able to replace the Octamer-binding transcription factor 4 (OCT4) during cellular conversion of mouse embryonic fibroblasts into induced pluripotent stem cells, a process similar to MET (Chen et al., 2010, Soncin and Ward, 2011, Redmer et al., 2011, Lowry, 2011).

In order to initiate the process of cancer invasion, cancer cells must detach from the primary tumour and adopt migratory properties that allow them to move into the surrounding stroma. In tumours of an epithelial origin, this initial detachment from the primary tumour is thought to be largely mediated by down-regulation of E-cadherin containing adherens junctions. Down-regulation of E-cadherin itself is a common feature of a variety of cancers and can be caused by loss of heterozygosity (LOH), mutations, epigenetic silencing, transcriptional silencing or increased endocytosis and proteolysis (Berx and van Roy, 2009). The germline mutations in *CDH1* have been reported in familial gastric cancer and in many cases, these mutations are accompanied by the promoter hyper-methylation of the second allele, leading to a loss of *CDH1* expression (Guilford et al., 1998, Machado et al., 2001). Furthermore, somatic genetic aberrations of E-cadherin, including LOH and mutations have been identified in several human cancers such as diffuse gastric cancer, tumours of ovary, endometrium, lobular breast, bladder, stomach cancer, thyroid and prostate cancer (Becker et al., 1994, Berx et al., 1995, Latil et al., 1997, Elo et al., 1997, Strathdee, 2002).

Early Immunohistochemical (IHC) studies on the expression and distribution of E-cadherin in different cancer types have demonstrated its role in cancer progression and invasiveness. IHC analysis of the squamous-cell carcinomas (SCCs) of the head and neck revealed that expression of E-cadherin is significantly higher in differentiated regions of a tumour but less in de-differentiated tumour areas (Schipper et al., 1991). Such examples were also

recognised in colorectal carcinoma (Dorudi et al., 1993), thyroid (Brabant et al., 1993), renal cell (Katagiri et al., 1995), stomach (Shiozaki et al., 1991), prostate (Umbas et al., 1992) and breast carcinoma (Oka et al., 1993). Similarly, *in vitro* analysis of several human carcinoma cell lines showed that E-cadherin was well expressed in all differentiated carcinoma cells but not in carcinoma cell lines that had undergone de-differentiation. Expression of an exogenous E-cadherin in the differentiated carcinoma cell lines also inhibited invasive capacity of these cell lines (Frixen et al., 1991). In contrast, repression of E-cadherin in a non-invasive epithelial cell lines by an antisense method resulted in the formation of a fibroblast-like morphology, with enhanced invasion (Vleminckx et al., 1991). In addition, forced expression of E-cadherin in mouse intestine epithelia reduces migration and inhibited cell proliferation (Hermiston et al., 1996). Similarly, re-expression of E-cadherin in rat and human E-cadherin-negative prostate cancer cell lines also inhibited the invasive potential and proliferation of these cell lines (Luo et al., 1999, Sasaki et al., 2000). Furthermore, *in vivo* analysis of a transgenic mouse model of pancreatic β-cell cancer (Rip1Tag2) showed that E-cadherin is expressed in well-differentiated adenomas but lost in invasive carcinomas (Perl et al., 1998). When E-cadherin expression was maintained in double transgenic mice that expressed E-cadherin in pancreatic β-cells (Rip1Tag2; Rip1E-cad) the incidence of carcinoma was significantly reduced. However, when E-cadherin expression was lost in double transgenic mice expressing a dominant negative form of E-cadherin (Rip1Tag2; Rip1dnE-cad), the incidence of carcinoma increased dramatically and 25% of the double transgenic mice developed metastatic lesions. These results provided first clear evidence that E-cadherin has a key suppressive role in the transition from adenoma to invasive carcinoma (Perl et al., 1998). This view was supported in later studies. The knockdown of TP53 gene (coding for the tumour suppressor protein p53) alone resulted in large mammary carcinomas that were generally not invasive. However, a compound mutant mouse model of invasive lobular carcinoma with conditional deletion of E-cadherin in combination with epithelium-specific TP53 knock-down (K14Cre; CDH1F/F; Trp53F/F) developed invasive and metastatic mammary carcinomas.

1.2.2 EMT regulators

Various signalling pathways such as transforming growth factor-β (TGF-β), fibroblast growth factor (FGF), Notch and Wnt have been implicated in control of EMT. The majority of these

signals activate several transcriptional repressors also known as EMT regulators (EMT-r). Those include the ZEB (Zinc-finger-E-box-binding protein 1 (ZEB1) and Zinc-finger-E-box-binding protein 2 (ZEB2)), SNAIL (Snail homolog 1 (SNAIL1) and Snail homolog 2 (SNAIL2)), and basic helix-loop-helix (bHLH) (Twist-related protein 1 (TWIST1) and Twist-related protein 2 (TWST2)) families. These transcription factors have a potential to down-regulate epithelial genes, such as E-cadherin, thereby promoting EMT process (Hill et al., 2013). In recent years significant progress has been achieved in understanding of how the expression and function of different EMT-r is affected by different pathways implicated in cancer (see below for details).

1.2.3 EMT-r and cancer

As indicated above, EMT has a role in normal development and oncogenesis. An EMT regulator, SNAIL is a zinc finger transcription factor that binds to E-boxes in the promoters of the target genes and acts as a transcriptional repressor (Nieto, 2002). SNAIL was the first EMT-r that has been shown to directly repress *CDH1* gene (Cano et al., 2000). During embryonic development, mice that lack of *SNAI1* gene die at E8.5 and show defects during gastrulation and mesoderm formations. These mutant mice form an abnormal mesoderm layer that expresses mesodermal markers but maintains epithelial features such as adherens junctions due of the presence of E-cadherin (Carver et al., 2001).This suggest that, the loss of E-cadherin expression via SNAIL is important in early developmental processes.

In vitro analysis has shown that SNAIL2 is also able to inhibit endogenous E-cadherin and *CDH1* gene construct in a dose dependent manner in human breast cancer cell lines (Hajra et al., 2002). Furthermore, expression of SNAIL2 has also been implicated in development of melanoma with SNAIL2 detected in benign naevi and in malignant melanoma. Knockdown of SNAIL2 in transformed melanocytes slightly reduced the size of the primary tumour but significantly decreased the number of metastases. This suggests that, SNAIL2 has an important role in metastatic melanoma (Gupta et al., 2005).

Increased expression of SNAIL2 leads to the transcriptional repression of E-cadherin in Madin-Darby canine kidney (MDCK) cells (Bolos et al., 2003). Likewise, enforced expression of SNAIL2 in the rat bladder carcinoma NBT-II cell line resulted in the disappearance of desmoplakin and desmoglein. This was associated with an increase in cell spreading, cell-cell

separation and migration, suggesting that SNAIL2 may initiate a process resembling an EMT (Savagner et al., 1997). Further analysis has shown that the induction of SNAIL1 expression in a colon cancer cell line leads to a significant reduction in E-cadherin mRNA expression level, while reduced SNAIL expression in pancreatic cancer cell line results in restored E-cadherin expression (Batlle et al., 2000). Recent studies demonstrated that ZEB1, SNAIL1 and SNAIL2 proteins are able to down-regulate several epithelial markers including cell polarity proteins such as Crumbs3 and Lethal giant larvae (Lgl2) (Baranwal and Alahari, 2009). Taken together, these results suggest that SNAIL2 can induce loss of cell-cell adhesion in cancer cells by down-regulating the expression of either E-cadherin or desmosomal proteins or cell polarity genes depending on the cellular context.

Expression of SNAIL2 has also been found in association with poor survival in patients with several forms of cancer including lung cancer and in oesophageal squamous cell carcinoma (Shih et al., 2005, Uchikado et al., 2005).

ZEB1 expression has been studied in aggressive colorectal tumours and uterine cancers (Peinado et al., 2007). In addition, increased expression of ZEB2 in epithelial MDCK cells resulted in an invasive phenotype, showing the potential role of ZEB2 in the induction of EMT and the formation of migratory and invasive tumour cells (Comijn et al., 2001). Moreover, high level of ZEB2 has been detected in SCCs, gastric, bladder (Sayan et al., 2009) and pancreatic tumours (Peinado et al., 2007).

The expression of TWIST, another EMT regulator has been found to induce EMT in a breast cancer cell line. Inactivation of TWIST1 protein inhibits metastatic progression of mammary cancer cells (Yang et al., 2004). The expression of TWIST1 has also been identified in metastatic prostate cancer, oesophageal, SCCs and in hepato-cellular carcinomas (Peinado et al., 2007).

The transcription factor NF-κB is an activator of immune response; in which it functions through transcriptional control of genes encoding cytokines, cytokine receptors and cell adhesion molecules. The NF-κB has been found to play a role in controlling cell proliferation and oncogenesis (Orlowski and Baldwin, 2002). The activity of the NF-kB transcription factor family has been identified as an important mediator of EMT in mouse model of breast cancer progression. It has been shown that activation of NF-kB pathway promotes EMT by

increasing levels of SNAIL, TWIST and ZEB proteins and down-regulating E-cadherin expression (Baranwal and Alahari, 2009, Vandewalle et al., 2009).

The Wnt signalling pathway is an important developmental pathway that has recently been linked to the process of EMT. It was found that the links between Wnt pathway and ZEB1 induced nuclear accumulation of β-catenin, which leads to decrease E-cadherin expression and increase cell migration in CRC (Sanchez-Tillo et al., 2011). In addition, the canonical Wnt signalling pathway through an Axin2–GSK3β cascade regulates SNAIL1 activity in breast cancer cells (Yook et al., 2006).

The involved of the EMT-r and cancers written in the text are summarised in Table 1.1.

Table 1.1: Increased expression of EMT-r in cancers

EMT-r	Cancer type	Reference
SNAIL1	Colon	(Batlle et al., 2000)
SNAIL2	Breast	(Hajra et al., 2002)
	Melanoma	(Gupta et al., 2005)
	Bladder	(Savagner et al., 1997)
	Lung	(Shih et al., 2005)
	Oesophageal	(Uchikado et al., 2005)
ZEB1	Colon	(Peinado et al., 2007)
	Uterine	
ZEB2	Pancreatic	(Peinado et al., 2007)
	Squamous cell	(Sayan et al., 2009)
	Gastric	
	Bladder	
TWIST1	Breast	(Yang et al., 2004)
	Prostate	(Peinado et al., 2007)
	Oesophageal	
	Squamous cell	
	Hepato-cellular	
NF-κB	Breast	(Baranwal and Alahari, 2009, Vandewalle et al., 2009)
Wnt	Colon	(Sanchez-Tillo et al., 2011)
	Breast	(Yook et al., 2006)

1.3 THE ZEB FAMILY OF TRANSCRIPTION FACTORS

1.3.1 ZEB proteins structure

ZEB1 (δ-EF1, Nil-2-a, Tcf8, Bzp, Areb6, Meb1, Zfhep and Zfhx1a) and ZEB2 (Smad-interacting protein 1 (SIP1) and Zfhx1b) are complex transcription factors, characterised by the presence of several functional domains, including two zinc-finger clusters and separated by a homeodomain (figure 1.5). Both ZEB1 and ZEB2 exhibit high degree of sequence homology within N-terminal (four zinc-fingers) and C-terminal (three zinc-fingers) zinc finger domains (Vandewalle et al., 2009). Due to high level of homology in these domains, both proteins bind to the similar nucleotide sequences, E-boxes (CANNTG) located in the promoter region of E-cadherin and other target genes (Sekido et al., 1996, Remacle et al., 1999, Comijn et al., 2001, Vandewalle et al., 2005). Therefore, it is likely that ZEB1 and ZEB2 are able to control expression of the same or overlapping gene sets (Remacle et al., 1999). However, even though ZEB1 and ZEB2 show structural similarities and both act as potent EMT inducers; they may have dissimilar and even opposing roles. Indeed, whereas ZEB1 overrides replicative senescence, ZEB2 promotes this pathway by transcriptional repression of Human Telomerase Reverse Transcriptase (hTERT) (Browne et al., 2010).

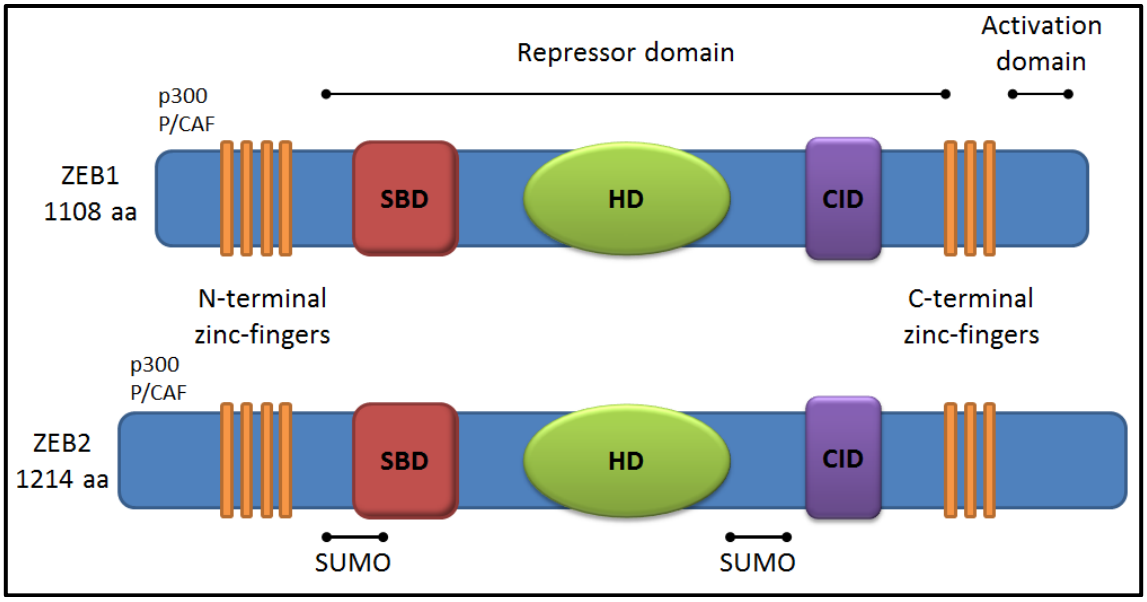


Figure 1.5: Structures of the human ZEB1 and ZEB2. The ZEB proteins contain N-terminal and C-terminal zinc-finger clusters, Smad-binding domain (SBD), centrally located homeodomain (HD) and C-terminal binding protein (CtBP)-interacting domain (CID). Protein region implicated in p300/PCAF binding and SUMOylation sites are indicated (Vandewalle et al., 2009). *Note: aa: amino acids.

Several important functional domains in central regions of ZEB proteins have been characterised. Those include C-terminal binding protein (CtBP)-interacting domain (CID) and Smad-binding domain (SBD). The central region also contains the POU-like homeodomain which does not bind to DNA but may be involved in interactions with other proteins important for transcriptional repression or activation (Smith and Darling, 2003, Vandewalle et al., 2009, Browne et al., 2010). Both ZEB1 and ZEB2 proteins can bind to activated R-Smads including R-Smads-1, R-Smads-2 and R-Smads-3 via the SBD. It was found that ZEB1-Smad signalling causes transcriptional activation, whereas ZEB2-Smad complex causes transcriptional repression (Postigo et al. 2003). Apparently, interaction of ZEB1 and ZEB2 with Smads show a potential role for ZEB proteins in bone morphogenetic protein (BMP) and TGF-β signalling (Verschueren et al., 1999). ZEB1 and ZEB2 proteins can also act as either transcriptional repressors or activators through different recruitment to transcriptional co-activators (p300 and P/CAF) and co-repressors (CtBP). CtBP binding occurs via CID and is required for transcriptional repression. Binding of ZEB1 to p300 and P/CAF prevents interaction of ZEB1 with CtBP thereby converting ZEB1 from a transcriptional repressor to an activator (Postigo et al., 2003). Subsequently, ZEB2 was also shown to interact with p300 and P/CAF (van Grunsven et al., 2006). Additionally, both ZEB1 and ZEB2 proteins may undergo post-translational SUMOylation, affecting the CtBP binding and repressive activity (Long et al., 2005).

1.3.2 ZEB1

ZEB1 was first identified as a nuclear factor which specifically binds and represses the lens-specific delta1-crystallin enhancer in chicken (Funahashi et al., 1991). The human ZEB1 protein is 1108 aa long (NCBI Reference Sequence: NP_001121600.1) and 1117 aa long in mouse (NCBI Reference Sequence: NP_035676.1). The human *ZEB1* gene (*ZFHX1a*) is located at chromosome 10p11.2 (Copeland et al., 1993) and consists of nine exons, with N-terminal and C-terminal zinc-finger clusters encoded by exons 5 to 7 and 8 and 9 respectively. The large exon 7 of ZEB1 corresponds to the central region of ZEB1 protein, which includes SBD, the HD and CID (Sekido et al., 1996).

1.3.2.1 Expression and function of ZEB1 during normal embryonic development

Interest in ZEB1 has increased as its role in the regulation of E-cadherin expression has been established in embryonic development and various epithelial malignancies. ZEB1 transcription factor is expressed during chicken embryogenesis, where it is detectable at post-gastrulation in the mesoderm, neuroectoderm and neural crest cells. Additionally, low expression of ZEB1 has been found in other ectodermal and endodermal tissues (Funahashi et al., 1993). Importance of ZEB1 in embryonic development has been demonstrated using ZEB1 knock-out mice. Homozygous ZEB1-/- mice develop in utero, however die shortly after birth due to ribs malformation and respiratory failure. They also show various defects including craniofacial, palate and limb defects, rib fusion, sternal defects and hypoplasia of intervertebral discs as well as severe T cell deficiency of the thymus (Takagi et al., 1998, Miyoshi et al., 2006, Vandewalle et al., 2009). The idea that ZEB1 has a role in skeletal tissue development and remodelling is in line with the finding, that ZEB1 regulates transcription of the osteocalcine gene (Sooy and Demay, 2002). Likewise, several studies suggested that ZEB1 function is related to collagen expression. ZEB1 repressed COL1a2 gene and decreased type ll collagen synthesis in mouse mesangial cells (Kato et al., 2007) and in chick mesenchymal and chondrocyte cells (Murray et al., 2000). Importantly, heterozygous mutation in *ZEB1* gene was found in patients with Posterior Polymorphous Corneal Dystrophy. The corneal abnormalities are thought to relate to the alterations in collagen synthesis, as demonstrated by increased expression of COL4a3 gene in the corneal epithelia of ZEB1 null mice (Krafchak et al., 2005, Liu et al., 2008c). In normal human tissues, ZEB1 mRNA was highly expressed in human bladder, uterus, skeletal muscle, aorta and thymus, with much less or no expression in liver and pancreas (Hurt et al., 2008). In addition, ZEB1 plays a role in smooth muscle (SM) cell differentiation. ZEB1 expression was found to bind and activate the SM α-actin promoter during differentiation of A404 SM cell line (Nishimura et al., 2006).

1.3.2.2 ZEB1 in cancer

Several studies addressed a role for ZEB1 in EMT in cancer cells. Over-expression of ZEB1 has been studied in several cell lines. The MCF10A cell line shows myoepithelial characteristics in nature, but upon over-expression of ZEB1, cell displayed several EMT traits

such as loss of E-cadherin and disorganised growth in three-dimensional matrigel (Chua et al., 2007). These findings suggest that, ZEB1 has a major role in controlling *CDH1* expression and is powerful EMT inducer in cancer cells.

In human breast cancer cell lines, ZEB1 has shown to be over-expressed in the highly metastatic MDA-MB-231 cell line, compared to the epithelial cells such as MCF-7 (Kirschmann et al., 1999). It has been found that ZEB1 knockdown in MDA-MB-231 cell line induced E-cadherin expression and re-establish epithelial features (Eger et al., 2005). In addition, in a panel of breast cancer cell line expression levels of ZEB1 and E-cadherin were inversely correlated (Aigner et al., 2007). The inverse relationship between ZEB1 and E-cadherin expression has also been observed in several NSCLC, bladder and prostate cancer cell lines (Ohira et al., 2003, Dohadwala et al., 2006, Sayan et al., 2009, Clarhaut et al., 2009). ZEB1 protein regulates not just *CDH1*, but also other adhesion molecules in NSCLC cells. For example, ZEB1 expression was found to associate with coxsackie-adenovirus receptor expression, which acts as an adhesion protein (Veena et al., 2009). In addition, ZEB1 has been shown to bind to the SEMA3F tumour suppressor gene promoter and down-regulate the semaphorin-3F protein. The semaphorin-3F is a tumour suppressor in lung cancer cell lines, and its loss in lung cancer tissues was correlated with poor pathological features (Clarhaut et al., 2009).

ZEB1 expression in relation to E-cadherin has also been studied in human colorectal carcinomas (CRC) cell lines, with high ZEB1 expression was detected in SW480, SW620 and HCT116 cell lines. Knockdown of ZEB1 expression in SW480 cell line resulted in changing in cellular phenotype, increased expression of E-cadherin, nuclear export of β-catenin and its re-localisation to the cell membrane (Spaderna et al., 2008). Similar to the data obtained in human breast cancer cell lines, ZEB1 is associated with EMT in colon cancer cells *in vitro*. *In vivo*, the expression of ZEB1 at tumour/stroma interface correlated with the loss of basement membrane. These data suggested that ZEB1 may contribute to metastases through basement membrane loss (Spaderna et al., 2006). Consistent with this hypothesis, ZEB1 was found to bind the Lama3 promoter and repress Laminin-332 protein expression. Additionally, ZEB1 activated expression of matrix metalloproteinase (MMP), which are critically involved in the degradation of the basement membrane at the invasive front of colorectal tumours (Spaderna et al., 2006, Drake et al., 2010).

In addition, ZEB1 was highly expressed in tumour-associated stromal cells, with little or no expression in differentiated epithelial tumour areas. This suggests that stromal cells may represent epithelial-derived tumour cells that have experienced a ZEB1-dependent EMT and have the capacity to invade the stroma (Aigner et al., 2007). In a mouse xenograft model, in which the HCT116 CRC cells with depleted ZEB1 were injected into immunocompromised mice, growth rate of primary tumours was not affected by the knockdown, but the number of metastases was strongly reduced compared with the control (Spaderna et al., 2008). This indicates that ZEB1 does not regulate primary tumour growth but is important for metastasis.

ZEB1 protein expression was found to be regulated by vitamin D3 metabolites. Treatment of SW480 cells with vitamin D3 increased cystatin D expression accompanied by a decrease in ZEB1 and increased expression of E-cadherin (Alvarez-Diaz et al., 2009). This suggests that cystatin D is induced by vitamin D3, may play a role in controlling ZEB1 expression in human CRC.

A role for ZEB1 protein in tumour development is not restricted to the activation of cell invasion programs and execution of canonical EMTs. Interplay between ZEB1 and cell cycle regulation has been addressed in several studies. The retinoblastoma protein (Rb) is a tumour suppressor protein which is dysfunctional in many cancer types. Knockdown of Rb in CRC cell line, DLD-1 was associated with increased expression of ZEB1 and loss of E-cadherin, showing a potential of ZEB1 interaction with Rb tumour suppressor pathway (Arima et al., 2008). Furthermore, ZEB1 was also shown to repress expression of p16 and p21. It was found that ZEB1 inhibition by cyclin-dependent kinase (cdk) inhibitors leads to permanent cell cycle arrest and premature senescence in mouse embryonic fibroblasts (Liu et al., 2008b, Liu et al., 2013).

It was found that expression of ZEB1 was induced by nicotine derivative (α7-nAChR) during cigarette smoking. Treatment of the CRC cell lines HT-29 and DLD-1 with α7-nAChR increased ZEB1 expression, down-regulated E-cadherin, enhanced migratory and increased invasive properties in both cell lines (Wei et al., 2009).

ZEB1 and miR-200 (an important component of ZEB1 regulatory network) have been studied in pancreatic carcinoma (Wellner et al., 2009). In pancreatic cancer, activation of Notch

signalling pathway results in ZEB1 up-regulation. The underlying mechanism involves activation of JAG1, MAML2 and MAML3, which occur via inhibition of miR-200 family members targeting these Notch pathway components. On the other hand, ZEB1 is directly activated by Notch and miR-200 genes are repressed by ZEB1. This complex interplay between Notch and ZEB1-regulated EMT pathway represents a classical self-enforcing feedback control loop. Indeed, initial activation of ZEB1 through RAS or TGF-β pathways results in the repression of miR-200 leading to Notch activation, further increase in ZEB1 and EMT (Brabletz et al., 2011). In other study, Pro-HB-EGF, a protein involved in EGF signalling negatively correlated with ZEB1 expression. It was found that the expression of the uncleaved form of Pro-HB-EGF on the membrane of pancreatic cancer cells stabilised E-cadherin and thus the epithelial phenotype by inhibiting ZEB1 expression (Wang et al., 2007).

Oncogenic RAS is known to cooperate with TGFβ in the induction of EMT (Janda et al., 2002). Mechanistically, hyperactivation of RAS/ERK2 signalling induced expression of AP-1 family member FRA1 (FOS related antigen 1). Activation of FRA1 was in turn necessary for the increased expression of ZEB1 and ZEB2 leading to a full EMT (Shin and Blenis, 2010). Interestingly, in malignant melanoma, activation of MAPK by oncogenic NRAS or BRAF also results in increased expression levels of FRA1. However, in this background, accumulation of FRA1 results in a switch in the expression of ZEB proteins, with ZEB1 and ZEB2 being up- and down-regulated respectively. This consequence of this switch is enhanced tumourigeneicity of melanoma cells (Caramel et al., 2013).

In addition, the relationship between ZEB1 and canonical Wnt pathway has been studied. A Wnt signalling pathway was found to regulate ZEB1 activity and promotes an invasive growth of colorectal tumours (Sanchez-Tillo et al., 2011).

1.3.3 ZEB2

ZEB2 was first discovered from a two-hybrid yeast screen through its ability to bind to Smad (Verschueren et al., 1999). The human *ZEB2* gene (*Zfhx1b*) is located at the chromosome 2q22 and the human ZEB2 protein is 1214 aa long (1215 aa long in mouse) (Acun et al., 2011, Vandewalle et al., 2009, Conidi et al., 2013). In contrast to *ZEB1* structure, *ZEB2* developed a divergent 5'-UTR (untranslated region), which possesses a highly complex

organisation, with the potential for multiple splicing products originating from several promoters. Analysis of the mouse ZEB2 5'-UTR identified nine un-translated exons (U1-U9) upstream of the first translated exon. These untranslated exons were spliced to the first translated exon; with no additional upstream in-frame start codons. In addition, three other promoters have also been identified, promoter 1 upstream of exon U1, promoter 2 upstream of exon U5 and promoter 3 upstream of exon 1. Furthermore, the gene coding for a natural antisense transcript (NAT) was also found in the region corresponding to the 5'UTR (Nelles et al., 2003). Heterozygote *ZEB2* mutations in humans have been implicated in the development of a rare congenital disease known as Mowat-Wilson syndrome, a hereditary condition characterised by cranio-facial abnormalities, mental retardation and several congenital malformation such as genital abnormalities, agenesis of the corpus callosum, poor hippocampal formation, congenital heart disease and Hirschsprung disease due to ZEB2 haploinsufficiency (Van de Putte et al., 2003, Vandewalle et al., 2009).

1.3.3.1 Expression and function of ZEB2 during normal embryonic development

Similar to ZEB1, ZEB2 also found to play an important role during the embryonic development. ZEB2 expression was detected in the neuroectoderm and neural crest derivatives in mouse embryos (Van de Putte et al., 2003). Homozygous *ZEB2*-/- mice are embryonically lethal with failed neural tube closure, abnormal enteric neural crest cell migration and formation of shortened somites. Furthermore, targeted knockout of *ZEB2* in neural crest cells resulted in defects in craniofacial structures, melanocytes, heart and neural system development. During differentiation of the neuroepithelium from the ectoderm, expression of E-cadherin was decreased, however E-cadherin expression persisted in the neuroepithelium of the homozygous *ZEB2*-/- mutant, suggesting that, ZEB2 is an important transcriptional repressor of E-cadherin (Van de Putte et al., 2003). In addition, compound ZEB1/ZEB2 null mice have more severe defects in neural tube development and loss of SOX2 (SRY (sex determining region Y)-box 2) expression. As the embryonic development of neural crest is a prototypical model for EMT, numerous defects in neural crest cell formation, migration and differentiation in the *ZEB2* mutant mice suggests a critical role for ZEB2 proteins during EMT (Miyoshi et al., 2006).

In addition to its critical role in the regulation of EMT processes, ZEB2 has been shown to promote replicative senescence, suppress DNA damage, induced apoptosis and cell cycle arrest at G1 phase of the cell cycle (Vandewalle et al., 2009). Mechanistically, cell cycle arrest is induced via transcriptional repression of cyclin D1 leading to the hypo-phosphorylation of Rb tumour suppressor protein (Mejlvang et al., 2007). Therefore, whereas ZEB1 appears to be controlled via Rb pathway (section 1.3.2.2), this pathway is downstream of ZEB2. These data indicate that ZEB proteins and Rb may constitute a composite regulatory mechanism controlling an equilibrium between cell proliferation and EMT. ZEB2 has also been identified as a cell survival protein, protecting cells from ultraviolet (UV) and cisplatin-induced apoptosis and DNA damage (Sayan et al., 2009). This role is, however, independent of cell cycle regulation.

1.3.3.2 ZEB2 in cancer

ZEB2 has been identified as another main contributor to malignant cancer progression. Increased expression of ZEB2 protein in the epithelial epidermoid carcinoma cell line A431, resulted in morphological changed of the cell phenotype from epithelial to fibroblast-like cells, down-regulated the expression of E-cadherin and α-catenin, and re-localised β-catenin from nucleus to the cytoplasm. Furthermore, expression of ZEB2 also induced the N-cadherin expression (Vandewalle et al., 2005). Using chromatin immunoprecipitation (ChIP) assays, ZEB2 was found to bind not only *CDH1*, but also plakophilin 2 and tight junction protein 3 (*ZO3*) promoters and to down-regulate expression of these genes in several human cancer cell lines (Vandewalle et al., 2005). These alterations enhanced loss of cells aggregation and the invasive capacity of A431 cells.

The relevance of the ZEB2 protein to cancer progression has been investigated in bladder carcinoma. It was found that, ZEB2 overexpression was an independent prognostic factor in bladder cancer and its expression positively correlated with a poor therapeutic outcome (Sayan et al., 2009). In addition, expression of ZEB2 was also found in renal cell (Fang et al., 2013), oesophageal, ovarian (Oztas et al., 2010) and breast carcinoma (Bindels et al., 2006, Elloul et al., 2005). Moreover, an inverse correlation between ZEB2 and E-cadherin expression has also been detected in several series of oral SCC samples (Maeda et al., 2005). Similarly, ZEB2 was detected in intestinal type gastric carcinomas, where its expression is

also inversely correlated with E- cadherin (Rosivatz et al., 2002). According to Elloul et al., ZEB2 may play a functional role also in ovarian carcinoma, where expression of ZEB2 mRNA was higher in effusion compared to primary tumours and solid metastases samples (Elloul et al., 2006).

Concomitant expression of ZEB1 and ZEB2 proteins has been addressed in several studies. In carcinoma cell lines, *ZEB1* and *ZEB2* mRNA are present in cells with mesenchymal phenotype lacking the expression of miR-200 family members. However, at the protein level, ZEB2 is rarely detectable in carcinoma cells (Sayan et al., 2009). In sarcoma and melanoma cell lines, ZEB1 and ZEB2 are often co-expressed, and in the latter background, they mutually repress each other (Caramel et al., 2013). Interestingly, in human head and neck cancer cells co-expression of ZEB1 and ZEB2 correlated with poor prognosis and survival in patients with this cancer (Chu et al., 2013).

A role for ZEB2 in tumours is not restricted to EMT in carcinoma cells, and ZEB2 was found to be involved in glioma progression. ZEB2 depletion by short interfering ribonucleic acid (siRNA) in U251 and U87 glioma cell lines resulted in the inhibition of cell proliferation, migration and invasion in these cell lines. Additionally, decreased expression of ZEB2 also promoted apoptosis in glioma cell lines (Qi et al., 2012). Epstein-Barr virus (EBV) is a human gammaherpesvirus associated with several types of malignancies in epithelial and B-lymphocytic cells, such as nasopharyngeal carcinoma post-transplant lymphoproliferative disease, Burkitt's lymphoma, Hodgkin's disease and some gastric cancers. It has been found that ZEB2 plays a potential role in maintaining the EBV latency by directly binding to EBV gene promoter, Zp. Additionally, reduced expression of ZEB2 resulted in re-activation of EBV lytic pathway (Ellis et al., 2010).

The von Hippel-Lindau (VHL) is a genetic cancer syndrome characterised by the development of clear cell type of renal cell carcinoma (RCC). This metastatic tumour develops upon somatic mutations in the VHL tumour suppressor gene. It has been found that, loss function of *VHL* gene reduced E-cadherin expression and induced hypoxia-inducible factor-1 (HIF-1) and ZEB2 protein. Furthermore, forced expression of *VHL* gene in the wild-type RCC VHL-/- cell line increased the expression of E-cadherin and down-regulated HIF-1 and ZEB2 levels. This finding suggests that, *VHL* gene may stimulates E-

cadherin transcription by HIF-1-mediated regulation of ZEB2 during the progression of RCC (Evans et al., 2007, Krishnamachary et al., 2006, Vandewalle et al., 2009, Banumathy and Cairns, 2010).

Increased expression of ZEB2 was detected in several hepatocellular carcinoma (HCC) cell lines, in which ZEB2 repressed *CDH1* and promoted cancer invasion by up-regulating matrix metalloproteinase (MMP) gene family such as *MMP-1, -2, -7* and *MT1-MMP* (Miyoshi et al., 2004, Acun et al., 2011). Furthermore, induction of ZEB2 expression in A431 cells resulted in G1 phase cell cycle arrest due to ZEB2-mediated transcriptional repression of cyclin D1, resulting in an increase in the level of hypo-phosphorylated and inactive Rb protein, thereby preventing progression through the cell cycle (Mejlvang et al., 2007). ZEB2 has also been shown to promote the expression of PTEN in malignant melanoma via a mechanism involving miRNA (Karreth et al., 2011). Furthermore, ZEB2 acts in the TGF-β pathway and mediates the TGF-β regulated repression of hTERT (Lin and Elledge, 2003). Likewise, ZEB2 was found to repress hTERT and induce senescence arrest in hepatocellular carcinoma-derived cells (Ozturk et al., 2006).

ZEB2 is activated by several signalling pathways. It was found that, TGF-β initiate intracellular signalling through activation of SMAD factors and triggering ZEB2 expression. A study by Shirakihara et al. found that, ZEB2 levels gradually increased in normal murine mammary gland (NMuMG) cells after treatment with TGF-β (Shirakihara et al., 2007). Furthermore, expression of ZEB2 was also found to be stimulated by RAS signalling pathway. An interplay between RAS and TGF-β signalling pathway is sufficient to cause EMT, and to induce cell scattering with loss of epithelial and gain of mesenchymal markers (Janda et al., 2002, Miyazono, 2009). Additionally, a link between ZEB2 and Wnt/JNK signalling has been studied. Inhibition of non-canonical Wnt signalling by *SFRP1* (secreted Frizzled related protein 1) induced degradation of the ZEB2 protein, in parallel with inhibition of JNK signalling activity and decreased proliferation and increased death of neuronal progenitors during the hippocampal development (Miquelajauregui et al., 2007). Furthermore, expression of ZEB2 can also be induced by exposure to collagen type I. It was found that, growing the pancreatic cell line, PANC-1 in the presence of type I collagen increased ZEB2 expression resulting in E-cadherin down-regulation (Peinado et al., 2007, Imamichi et al., 2007).

1.4 ZEB PROTEINS AND miRNAs

miRNAs are non-coding double stranded 20-25 nucleotides long RNA molecules that bind to the complementary 3'-UTR in mRNAs resulting in the translation block and/or mRNA degradation (Rushworth et al., 2012, Hill et al., 2013). Over the past decade of extensive research, it has become clear that this class of regulatory molecules controls virtually all aspects of cell biology including cell growth, survival, differentiation, motility and EMT. Several species of miRNAs including five members of miR-200 family (miR-200a, miR-200b, miR-200c, miR-141 and miR-429) and miR-205 were found to control EMT by directly repressing ZEB1 and ZEB2 proteins expression (Gregory et al., 2008a, Browne et al., 2010, Hurteau et al., 2007, Park et al., 2008). Additionally, miR-192 targets ZEB2, but not ZEB1 in mouse kidney cells (Kato et al., 2007).

Several studies have found an inverse correlation between miRNAs and ZEB proteins expression. Analysis of the panel of The National Cancer Institute (NCI) 60 cell lines, found that high expression of miR-200 family members correlated with the epithelial phenotype. Furthermore, forced expression of miR-200a or miR-200c resulted in the inhibition of ZEB1 and ZEB2 expression in parallel with an epithelial differentiation in the undifferentiated breast cancer cell line MDA-MB-231 (Park et al., 2008). Likewise, ectopic expression of miR-200c in lung cancer A549 cells reduced expression of ZEB1, increased E-cadherin levels and induced a transition from mesenchymal to epithelial morphology (Hurteau et al., 2007). Overall, a number of studies performed in several dedifferentiated carcinoma cell lines have shown that miR-200 family members have a potential to induce MET (Park et al., 2008, Korpal et al., 2008, Gregory et al., 2008a, Christoffersen et al., 2007, Li et al., 2010). Consistent with these observations, inhibition of miR-200 family increased ZEB1 and ZEB2 expression levels and promoted an EMT in MDCK and HCT116 cells (Park et al., 2008, Gregory et al., 2008a, Gregory et al., 2008b). Moreover, a study in meningioma cell lines has demonstrated that enforced expression of miR-200a down-regulated the expression of ZEB1 in these cells (Saydam et al., 2009). Likewise, in K-ras/p53-mutant lung adenocarcinoma model, in which miR-200 family are supressed in metastasis-prone tumour cells, enforced expression of miR-200 abrogated EMT and metastases in syngeneic mice (Roybal et al., 2011). The strong involvement of miR-200 in EMT regulation is explained by direct

repression of both *ZEB* genes as a result of the miRNA binding to eight and nine sites in the 3'UTRs of *ZEB1* and *ZEB2* mRNA respectively.

Several studies identified the presence of a double negative feedback loop, in which miR-200 family and ZEB proteins negatively control each other expression (figure 1.6). Repression of ZEB1 and ZEB2 by miR-200 family members represents the first arm of this regulatory circuit. The second arm was identified in studies showing that enforced expression of ZEB1 and ZEB2 suppressed expression of the miR-200 family thereby allowing an EMT to proceed (Bracken et al., 2008, Brabletz and Brabletz, 2010, Hill et al., 2013). The existence of this regulatory mechanism has been demonstrated in a study, in which protein tyrosine phosphatase Pez- or TGF-β-induced EMT was analysed in MDCK cells. In this model, miR-200 family members were suppressed in parallel with an increase in ZEB1 and ZEB2 expression. Mechanistically, this suppression occurs via direct binding of ZEB proteins to multiple E-boxes localised to the promoters of genes encoding miR-200 family members. Overexpression of miR-200 family members reversed this process, inducing MET and reducing ZEB1 and ZEB2 expression in MDCK-Pez cells (Gregory et al., 2008a, Gregory et al., 2008b).

Taken together, the ZEB/miR-200 feedback loop represents an important part of cellular signalling network that regulates EMT and MET equilibrium and subsequent cancer metastasis. Further studies on this regulatory system may contribute to the development of new therapeutic approaches for human cancer.

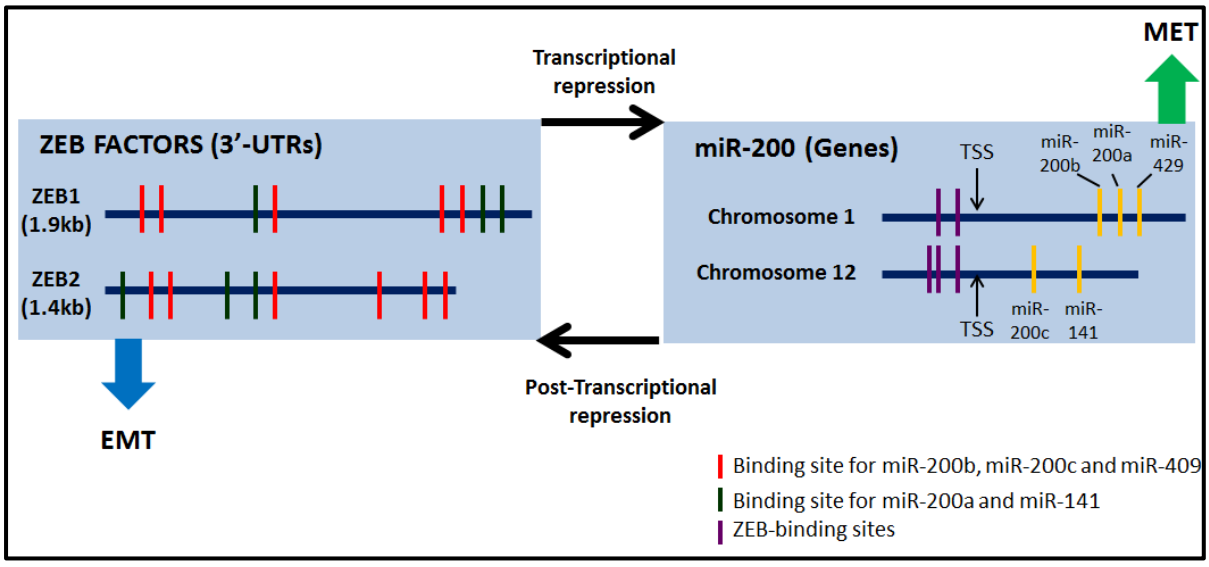


Figure 1.6: The ZEB/miR-200 double negative feedback loop. ZEB1 and ZEB2 3'-UTRs contain multiple sites that are targeted by miR-200 family members. Binding of mRNA-200 family members to these sites inhibit expression of ZEB1 and ZEB2 at the post-transcriptional level leading to a MET. Binding of ZEB factors to the E-boxes within transcriptional promotes of miR-200 genes represses their expression promoting EMT (Brabletz and Brabletz, 2010). *Note: TSS: transcriptional start site.

1.5 POST-TRANSLATIONAL MODIFICATION (PTM)

PTM is a process in protein biosynthesis that changes the characteristics of a protein by altering its chemical structure in a reversible manner (Deribe et al., 2010). The process of protein translation is carried out on the ribosome, where it translates the information from mRNA molecules into polypeptide chains. Following translation, the polypeptide chains must undergo further PTMs processes such as cleavage and other processes before it becomes fully functional (Cooper, 2000a). The most important PTMs that regulate all cellular functions are SUMOylation, ubiquitination, protein phosphorylation, glycosylation, nitrosylation and acetylation. Very often, two or more PTMs function in concert to regulate one pathway (Deribe et al., 2010).

1.5.1 SUMOylation

SUMOylation is a covalent modification that adds small ubiquitin (Ub)-like modifier (SUMO), a 100 amino acid polypeptide, to the acceptor lysine residues in the target proteins. In mammals, four SUMO proteins have been identified, SUMO-1, -2, -3 and -4. SUMO-1, SUMO-2 and SUMO-3 are characterized by different subcellular localisation patterns. SUMO-1 is usually conjugated to the substrate proteins whereas SUMO-2 and SUMO-3 (both with 50% similarity to SUMO-1) are found in a free form. In addition, SUMO-2 and SUMO-3 are capable of forming a long conjugated SUMOylation chains through their SUMO modification of the lysine residues, while SUMO-1 does not, and its attachment to SUMO-2 or SUMO-3 chains block their further elongation (Wilkinson and Henley, 2010, Wang and Dasso, 2009).

SUMOylation is a complex multistep process. The initial step in SUMO cycle is SUMO protein maturation, in which a member of the Sentrin/SUMO-specific protease (SENP) family is involved. The SENP enzyme cleaves the C-terminus of SUMO to expose a di-glycine motif which is essential for the conjugation of SUMO to the substrate protein. During the first step

of SUMO conjugation cycle, the E1 SUMO-activating enzyme (SAE) (SAE1/ SAE2 in mammals) forms a thioester bond between the active site cysteine residue of SAE2 and the C-terminal glycine residue of SUMO. This reaction is activated in an Adenosine-5'-triphosphate (ATP)-dependent manner. Activated SUMO is then transferred to Ubc9 (ubiquitin-conjugating 9-E2 enzyme) and to form a bond with the active site cysteine residue of the Ubc9 through a thioester linkage. Finally, the Ubc9-SUMO complex is bound by a SUMO-E3 ligase that catalyses the conjugation reaction of SUMO to the protein substrate. In addition, the Ubc9 catalyses the isopeptide bond between the ε-amino group of lysine in the protein substrate and the C-terminal carboxyl group of SUMO. The Ubc9 binds directly to the substrate proteins containing the consensus motif ψKxD/E (ψ is a large hydrophobic residue and x is any residue) of a protein substrate (figure1.7). However, the SUMOylation occurs also at lysine residues outside of this motif and not all the consensus ψKxD/E motifs are SUMOlyated. Several type of SUMO-E3 ligases have been identified, such as the protein inhibitor of activated STAT (PIAS), the nucleoporin Ran-binding protein 2 (RanBP2) or the Polycomb-2 protein (Pc2) (Kim and Baek, 2006, Garcia-Dominguez and Reyes, 2009, Wilkinson and Henley, 2010).

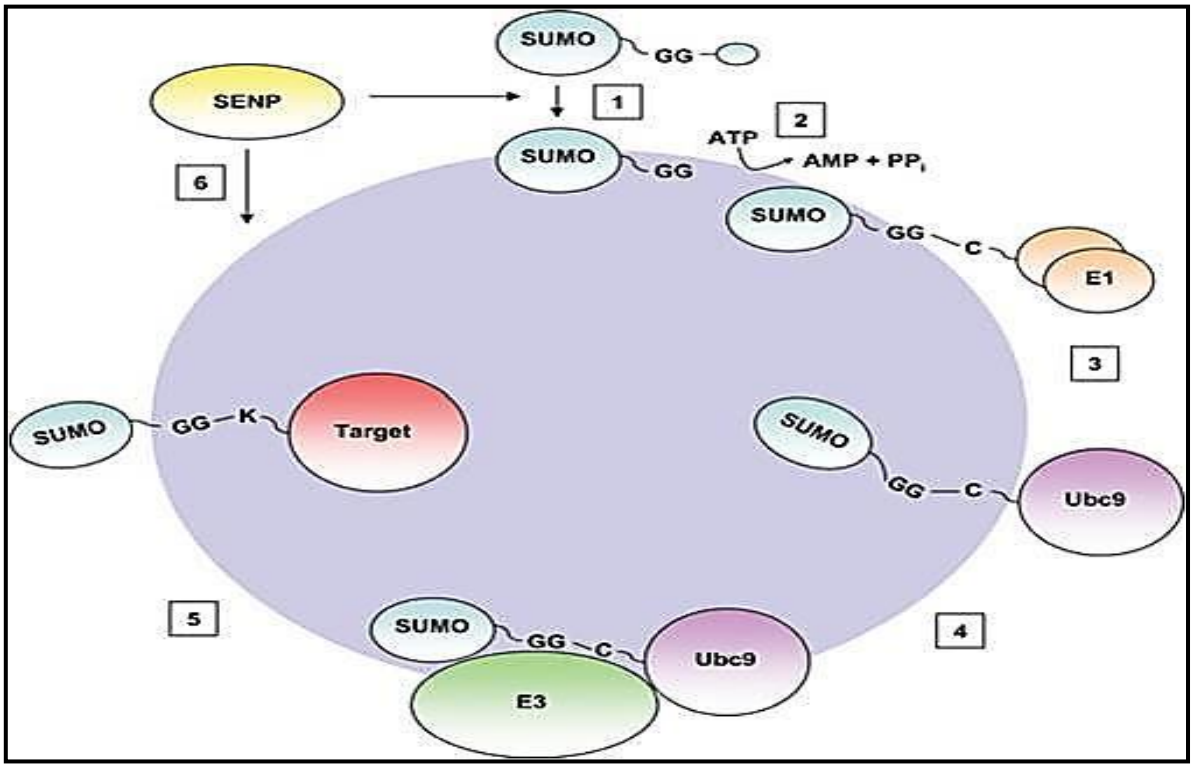


Figure 1.7: The SUMO cycle. (1). Maturation of SUMO by SENP protease. (2). ATP dependent activation of SUMO by E1-SUMO activating enzyme. (3). Activated SUMO binds to the Ubc9 conjugation enzyme, (4). SUMO-Ubc9 complex binds to the SUMO E3 ligase enzyme in order

the SUMOylation reaction to be catalysed (5). The SUMOylation of a protein substrate (6). De-SUMOylation is performed by SENP proteases that release SUMO and protein substrates (Wilkinson and Henley, 2010).

SUMOylation is a reversible process and the SUMOylation status of the target proteins are controlled via de-SUMOylation (also known as SUMO de-conjugation) (Gill, 2004). It has been found that, the ratio between SUMOylation and de-SUMOylation is critically important for protein function. Loss of this stability may be either deleterious for a cell or contribute to malignant transformation (Kim and Baek, 2006). In addition, similar to the SUMOlytaion, de-SUMOylation is also performed by SENPs and in an ATP-dependent manner (Wilkinson and Henley, 2010).

Currently, up to 100 proteins are identified as SUMOylation substrates and its number is continuously increasing (Garcia-Dominguez and Reyes, 2009). It was observed that SUMOylation can regulate protein stability, subcellular localisation, DNA repair, genome integrity and gene transcription (Garcia-Dominguez and Reyes, 2009, Wilkinson and Henley, 2010, Hay, 2005, Geoffroy and Hay, 2009). The potential role of SUMOylation in transcriptional regulation has been shown by studying transcription factors such as CtBP, p300, heat shock factor 1 (HSF 1) and p53 (Hay, 2005). Recently, SUMOylation of ZEB2 at two conserved lysine residues by Pc2 SUMO E3-ligase was identified (Long et al., 2005). This modification attenuates repression of E-cadherin promoter by interfering with CtBP/ZEB2 association (Long et al., 2005). In addition, SUMOylation of CtBP represses its nuclear import reducing the availability of CtBP for the interaction with ZEB2. Thus, SUMOylation may represent an important mechanism controlling ZEB2 repressor activity and, therefore, provides the EMT-MET stability in cancer progression and normal development. As SUMOylation sites as conserved in ZEB2 and ZEB1 proteins (Long et al., 2005), one can suggest that SUMOylation plays a similar role also in the control of ZEB1 function.

1.5.2 The ubiquitin-proteasome system (UPS)

Ubiquitination is a biochemical process that controls numerous aspects of protein function, such as protein degradation, protein-protein interactions and subcellular localisation (Pickart and Eddins, 2004). The UPS includes two main steps: (1) the attachment of a series of ubiquitin (Ub) molecules to the target protein by a process known as ubiquitination and

(2) degradation of the ubiquitinated proteins by a proteasome (Patterson et al., 2007, Mearini et al., 2008, Wang et al., 2008, Young et al., 2008, Glickman and Ciechanover, 2002). The Ub is an 8-kDa polypeptide comprising of 76 amino acids. It is attached to lysine residue in protein substrates via its C-terminal glycine (Schlesinger et al., 1975, Dikic and Robertson, 2012). Generally, ubiquitination involves three classes of enzymes, known as E1, E2 and E3 (Ciechanover, 1998) (Figure 1.8). The human genome encodes two E1 enzymes known as Ubiquitin-activating enzyme 1 (UBE1) and Ubiquitin-like modifier activating enzyme 6 (UBA6) (Groettrup et al., 2008). E2 enzymes are more abundant (about 30 to 40 orthologs exist in the human genome). They contain 150 amino acids central domains that comprises four β sheets and four α-helices and surrounded by active cysteine residues (Michelle et al., 2009). In addition, there are approximately 600 genes in human genome coding for E3 Ub ligases. These include a HECT (homologous to E6-associated protein C-terminus) and RING (Really Interesting New Gene) families. The RING type E3s act by bringing E2-bound Ub and the substrate protein close together to allow the ligation of Ub to the substrate. The RING type E3s also may mediate a conformational change of bound E2 that facilitates ubiquitin transfer (Passmore and Barford, 2004). Members of the RING type E3s include MDM2 (murine double minute 2, HDM2 in human) and MDMX (also known as MDM4, HDMX or HDM4 in human) (Metzger et al., 2012). Meanwhile, the HECT type E3s contains an active cysteine residue that forms a thioester bond with Ub before it is transferred to the substrate. The HECT type E3s include the Homologous to E6-Associated Protein (E6AP) (also known as UBE3A) (Metzger et al., 2012).

During the first step of ubiquitination, the C-terminus of Ub is activated by the E1 to form a thioester bond in an ATP-dependent manner. Following activation, the Ub is transferred from the E1 to the E2. Finally, the E3 binds to both Ub-charged E2 and a protein substrate and catalyses the transfer of the C-terminus of Ub to a lysine residue of the protein substrate to form an isopeptide bond. This process results in protein mono-ubiquitination (figure 1.9a). Furthermore, several lysine residues on the same protein substrate can be appended with Ub, resulting in multi-ubiquitination (Petroski and Deshaies, 2003, Petroski and Deshaies, 2005) (figure 1.9b). Additionally, E2 and E3 combinations can utilise lysines on the protein substrate-conjugated Ub, to catalyse further cycles of ubiquitination, resulting in poly-ubiquitination chains (figure 1.9c). The Ub itself contains seven lysines, in which can be

utilised during poly-ubiquitination to generate seven distinct poly-ubiquitin chains known as lysine-6 (K6), lysine-11 (K11), lysine-27 (K27), lysine-29 (K29), lysine-33 (K33), lysine-48 (K48) and lysine-63 (K63) chains. (Pickart and Eddins, 2004, Petroski and Deshaies, 2005, Dikic and Robertson, 2012). Importantly, different poly-ubiquitin chains have different functions, for example, K48-linked chains target proteins for proteasomal degradation and K63-linked chains function in kinase activation, DNA repair process, signal transduction and endocytosis (Passmore and Barford, 2004).

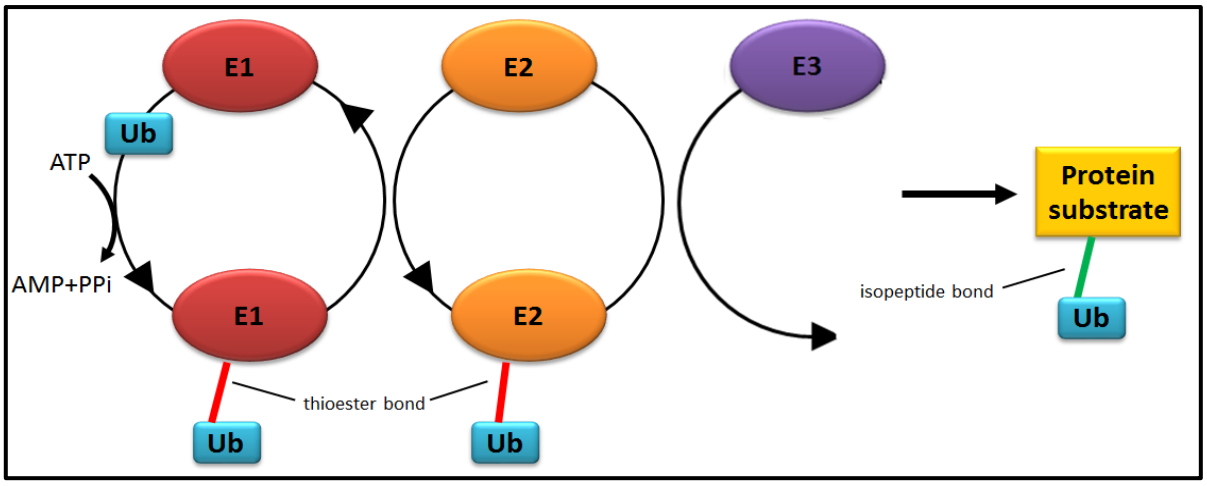


Figure 1.8: Enzymatic cascade of ubiquitination. The E1 enzyme is the Ub-activating enzyme, to which Ub is attached in an ATP-dependent manner to create a thioester bond. The E2 enzyme is the Ub-conjugating enzyme, to which the ubiquitin is transferred from the E1. Finally, the E3 is the Ub-ligase, in which facilitates the transfer process of the Ub to the protein substrate and forms an isopeptide bond (Dikic and Robertson, 2012).

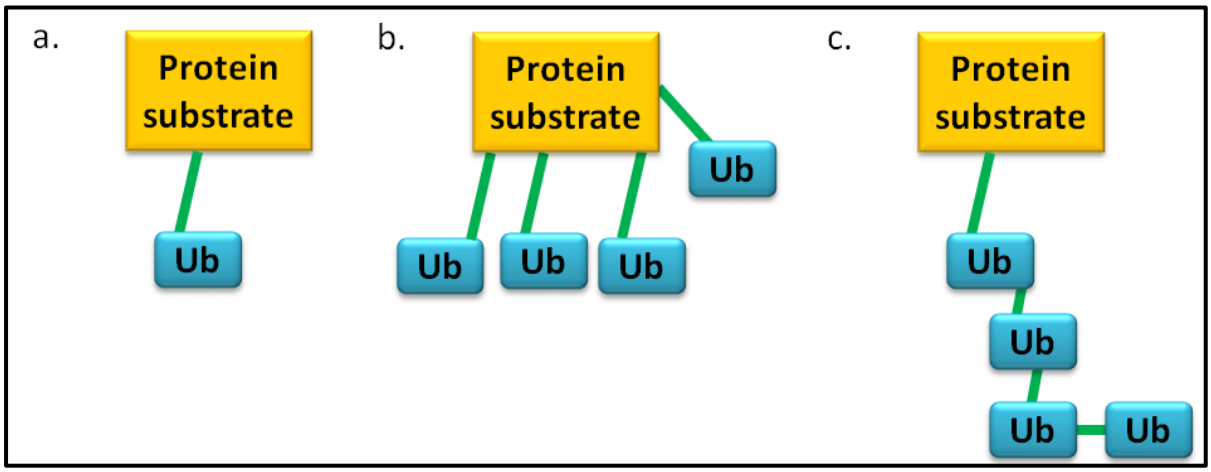


Figure 1.9: Different types of protein ubiquitination. Ubiquitination involves the attachment of Ub to lysine residues on protein substrates or itself, which can result in protein mono-ubiquitination, multi-ubiquitination or poly-ubiquitination. (a) Mono-ubiquitination is formed by the appending of one Ub molecule to a single lysine residue on the protein substrate (b) Multi-ubiquitination are formed by the binding of several Ub molecules at

several lysine residues on the same protein substrate and (c) Poly-ubiquitination is formed by the attachment of a chain of Ub molecules to the same lysine residues on the protein substrate (Figure drawn by W.R. Wan Makhtar).

The degradation of poly-ubiquitinated proteins is generally carried out by the 26S proteasome (Liu et al., 2008a). The 26S proteasome is a large multi-protein complex composed of two sub-complexes, a 20S core particle and a 19S regulatory particle (Glickman and Ciechanover, 2002). Following ubiquitination, the ubiquitinated proteins are recognised by the 26S proteasome leading to the degradation of these proteins to small peptides. Finally, the peptides are degraded to amino acids or used in antigen presentation (Lecker et al., 2006) (figure 1.10).

Many transcription factors involved in EMT such as SNAIL, TWIST, ZEB and E12/E47 are directly regulated by ubiquitination and the UPS. It has been found that the stability of SNAIL protein expression occurs in a UPS-dependent manner, when SNAIL is ubiquitinated by the RING type E3s such as MDM2 and βTrCP (Voutsadakis, 2012b). In addition, FBXL14 (F-box and leucine-rich repeat protein 14) is an E3 ligase that interacts with SNAIL and promotes its ubiquitination and subsequent proteasomal degradation (Vinas-Castells et al., 2010). TWIST is a substrate for caspases during apoptosis, and the products of the cleavage are rapidly degraded through the ubiquitin-proteasome pathway (Demontis et al., 2006).

The cullin family proteins (CUL1-3, CUL4A, CUL4B, CUL5 and CUL7) are scaffold proteins for the RING finger type E3 ligases that have crucial roles in the post-translational modification of cellular proteins involving ubiquitin (Sarikas et al., 2011, Fu et al., 2010). It was found that increased expression of ZEB1 and SNAIL2 by cullin7 down-regulated E-cadherin and increased invasion in human trophoblastic cell. Inhibition of cullin7 by proteasome inhibitor MG-132 treatment reduced expression of ZEB1 and SNAIL2, increased expression of E-cadherin and decreased invasiveness. These results suggest that the cullin7 E3 ligase is an important factor in regulating the EMT of trophoblastic cells (Fu et al., 2010). However, there is no reported evidence suggesting that ZEB2 is regulated via UPS pathways.

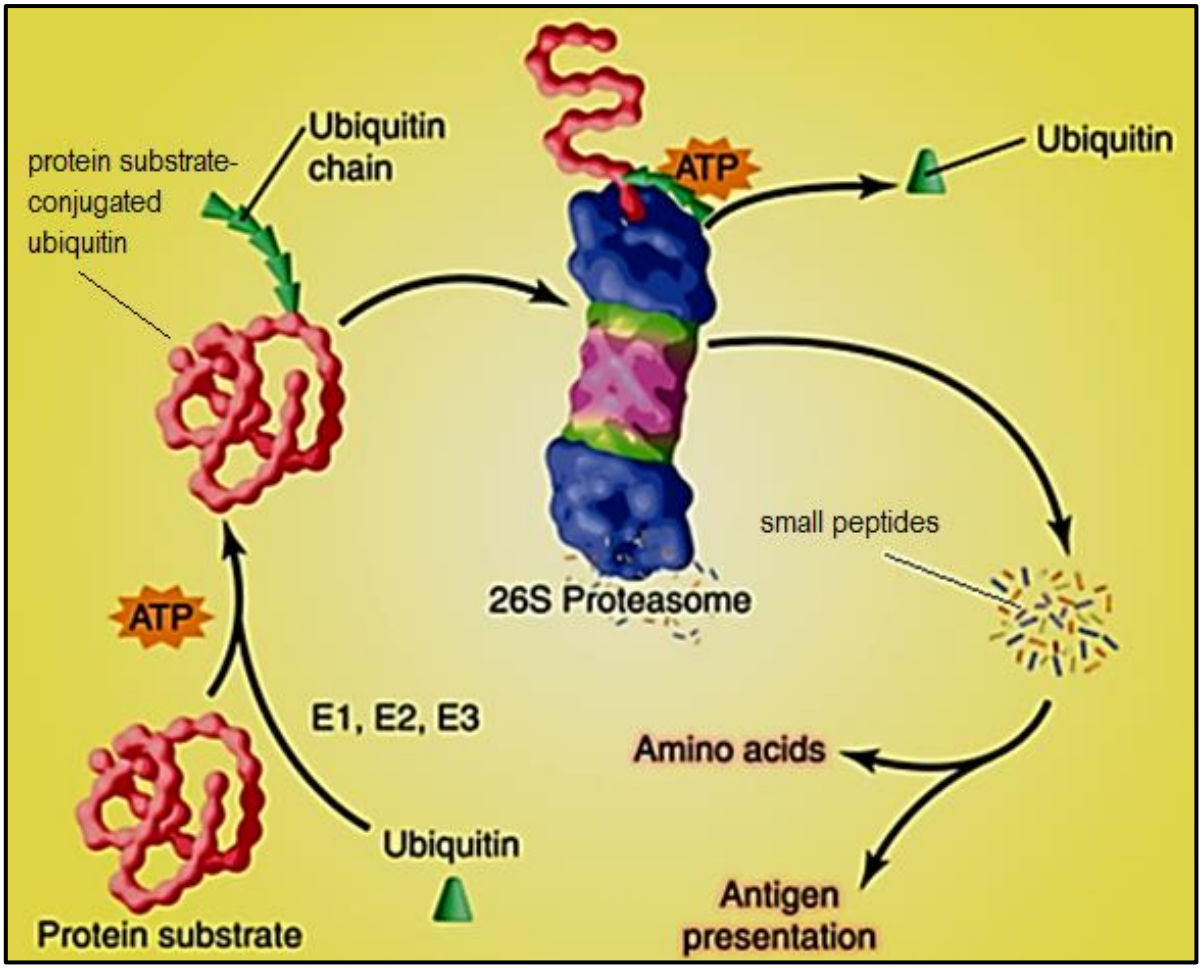


Figure 1.10: The Ub-proteasome system (UPS). Initially, the Ub is appended to the protein substrate by E1, E2 and E3 enzymes in an ATP-requiring reaction to generate the protein substrate-conjugated Ub. Next, the protein substrate-conjugated Ub is recognised by the 26S proteasome. Ub is removed and protein is digested into small peptides. These peptides are then degraded to the amino acids in the cytosol or nucleus or serve in antigen presentation (Lecker et al., 2006).

1.5.3 Protein phosphorylation

Protein phosphorylation is a reversible chemical process in which a phosphate group is added to a molecule (typically on serine, threonine or tyrosine residues). Protein phosphorylation is catalysed by protein kinases, which transfer phosphate groups from ATP to the hydroxyl groups on these residues. Protein phosphorylation is reversed by protein phosphatases, which catalyse the hydrolysis of phosphorylated amino acid residues (Cooper, 2000b). Phosphorylation alters protein structure and function and thereby modulates and controls intrinsic biological activity, subcellular location, stability, and interaction with other molecules (Cohen, 2000). It has been found that several EMT-inducing transcription factors can be regulated by phosphorylation. It was found that

phosphorylation of SNAIL by the p21-activated kinase 1 (PAK1) modulated its localisation and enhanced its transcriptional activity (Yang et al., 2005). Furthermore, SNAIL is sensitive to GSK-3β phosphorylation targeting it for ubiquitination and proteasomal degradation; therefore it is highly unstable with a half-life of approximately 25 minutes. Inhibition of the GSK-3β resulted in up-regulation of SNAIL expression, down-regulation of E-cadherin expression and EMT (Zhou et al., 2004). TWIST1 is another example of an EMT-controlling transcription factor, whose expression and function is regulated by phosphorylation. The protein kinase B (PKB/Akt) was found to phosphorylate TWIST1 at Serine-42 leading to the inhibition of p53 activity in response to DNA damage (Vichalkovski et al., 2010). Additionally, phosphorylation of TWIST1 at Serine-68 by p38, c-Jun N-terminal kinases and extracellular signal-regulated kinases1/2 increased expression of TWIST1 in human embryonic kidney 293 (HEK-293) and breast cancer cell lines. Consistent with these data, inhibition of Serine-68 phosphorylation induced degradation of the TWIST1 protein (Hong et al., 2011). Conversely, phosphorylation of ZEB1 and ZEB2 transcription factors still remains poorly studied. In a previous work, Costantiono and colleagues have demonstrated that ZEB1 is a phosphoprotein and its phosphorylation is cell type-specific (Costantino et al., 2002). However, neither the molecular pathways nor downstream protein kinases phosphorylating ZEB proteins are currently unknown. Likewise, the functional outcome of this phosphorylation remains to be studied. Apparently, better understanding of post-translational modifications of ZEB proteins will clarify the role of these EMT factors in cancer and is likely to open new avenues for therapy.

1.5.4 Protein glycosylation

The addition of a carbohydrate moiety to a protein molecule is referred to as protein glycosylation. It is a common post translational modification for protein molecules involved in cell membrane formation (Furukawa and Kobata, 1992, Varki, 1993). Protein glycosylation helps in proper folding of proteins, stability and secretion (Dwek et al., 2002). Protein glycosylation can be categorized in two main types: 1) N-linked glycosylation and 2) O-linked glycosylation (Roth et al., 2012). N-glycosylation is a well-known type of glycosylation in which N-glycans are directly attached to the amino group of asparagine in the endoplasmic reticulum (ER) (Ruddock and Molinari, 2006, Chesnokov et al., 2014). Meanwhile, O-linked glycosylation occurs via the addition of sugars to the hydroxyl group of serine or threonine

in the ER, Golgi, cytosol and nucleus (Martin, 2005, Stanley, 2011). Numerous proteins have been shown to be modified by intracellular glycosylation. Sp1 is a transcription factor that belongs to the family Sp/XKLF (specificity protein/Kruppel-like factor) family of zinc-finger motif containing transcription factors (Kaczynski et al., 2003). It was shown that O-glycosylated Sp1 is transcriptionally more active than unglycosylated Sp1 (Jackson and Tjian, 1988). In addition, glycosylation was suggested to stabilise Sp1 against proteasomal degradation (Han and Kudlow, 1997). Furthermore, O-glycosylation of Sp1 may change its interaction with other proteins of the transcription machinery in such manner that it favours the expression of diabetes related gene such as resistin gene (Chung et al., 2006). *C-myc* is a nuclear phosphoprotein that regulates gene transcription in cell proliferation, differentiation and cell death (Dang, 1999). It was found that c-myc is O-glycosylated at threonine 58. It was found that serum-starvation of the human promyelocytic leukemia cells (HL-60) led to the accumulation of O-glycosylated c-myc (Kamemura et al., 2002). Plakoglobin functions in both cell-cell adhesion, as a component of both adherens junctions and desmosomes and in cell signalling (Zhurinsky et al., 2000). It was found that plakoglobin has a single site for O-glycosylation modification in the N-terminal region within the glycogen synthase kinase-3β (GSK-3β) phosphorylation consensus sequence, suggesting that O-glycosylation might potentially irritate the GSK-3β mediated plakoglobin phosphorylation (Hatsell et al., 2003). A study also found that over-expression of O-Glycosyltransferase (OGT) in murine keratinocytes results in increased plakoglobin O-glycosylation levels and associated with increased numbers of plakoglobin based adherens junctions and desmosomes, which functionally resulted in increased in keratinocyte cell-cell adhesion (Hu et al., 2006). Finally, keratins are a major protein component of epithelial cytoskeleton. It has been demonstrated that O-glycosylation on K8 and K18 site-specific phosphorylation increase their solubility and degradation (Srikanth et al., 2010).

Few recent studies have established the participation of glycosylation with EMT-r .It has been found that EMT-r such as SNAIL induces differential glycosylation of Crumbs3a, which contribute to the loss of apical-basal polarity during EMT (Harder et al., 2012). The fucosyltransferase 4 (FUT4) is one of the enzymes that is involved in the transfer of GDP-fucose to the N-acetylglucosamines residue of glycoproteins. Knocked-down of FUT4 in MCF-7 and MDA-MB-231 cell lines increased expression of E-cadherin and reduced

expression of several mesenchymal markers including fibronectin, vimentin, N-cadherin, SNAIL, TWIST and ZEB1 (Yang et al., 2013). Several studies also link aberrant glycosylation to altered cadherin-dependent cell functions. It was found that alteration in E-cadherin structure by N-glycosylation disrupt epithelial architecture and increase tumour progression in breast tumour (Pinho et al., 2009a, Pinho et al., 2009b, Pinho et al., 2011, Vagin et al., 2008). In addition, N-cadherin contains eight N-glycosylation sequences in the extracellular segment that are post-transnationally modified. Ablating all eight sites in the N-cadherin structure increased intercellular junction stability and reduced cell migration in a scratch wound-healing assay (Guo et al., 2009). However, the links between glycosylation and ZEB proteins are yet unknown. Apparently, studies on this links may contribute to a better understanding of how cancer cells alter their differentiation status during tumour progression.

1.5.5 S-Nitrosylation

S-Nitrosylation is another form of PTM. It is a reversible reaction that involves the transfer of a nitric oxide (NO) moiety into thiol groups, to form S-nitrothiols (SNOs) (Anand and Stamler, 2012). It has been found that activation of nitric oxide synthase (NOS2) activity and NO signalling via S-nitrosylation modification stimulates EGFR/Src signalling, induce aggressiveness in breast cancer cells, activating several signal transduction pathways such as c-Myc, Akt, and β-catenin and inducing EMT and chemo resistance (Switzer et al., 2012). Conversely, high level of NO induced by DETANONOate inhibit the NFκB and SNAIL expression, up-regulates Raf-1 kinase inhibitor protein (RKIP) and reverses EMT in metastatic prostate cells *in vitro* and *in vivo* (Baritaki et al., 2010). However, there is no data showing whether S-Nitrosylation has any impact on the function of ZEB proteins.

1.5.6 Acetylation

Acetylation, another major PTM, is a chemical reaction that transfers an acetyl functional group to an active organic compound (Yang and Seto, 2008, Bannister and Miska, 2000). Several links between acetylation and EMT have been studied. Activation of p300/CBP, a histone acetyltransferase, by TGF-β1 mediated EMT through acetylation of Smad2 and Smad3 in human lung cancer cells. Conversely, inhibition of p300/CBP activity with EGCG treatment inhibited TGF-β1-induced EMT by suppressing the acetylation of Smad2 and

Smad3 leading to the expression of EMT-related genes in human lung cancer cells (Ko et al., 2013). Furthermore, the MAP3K4 was found to control the histone acetyltransferase CBP, in which acetylation of histones H2A and H2B by CBP is required to maintain the epithelial phenotype. It was found that combined loss of MAP3K4/CBP activity represses expression of epithelial genes and causes trophoblast stem cells to undergo EMT (Abell et al., 2011). Likewise, decreased in histone H3K27 acetylation on ZEB1 binding sites induced EMT in lung cancer cells (Roche et al., 2013). Additionally, inhibition of histone deacetylase activity by TGF-β1 induced suppression of EMT in renal epithelial cells (Yoshikawa et al., 2007).

1.6 THE CRE/LOX P RECOMBINATION SYSTEM

The Cre/loxP recombination system has enabled researchers to exploit the Cre enzyme in order to manipulate DNA both *in vivo* and *in vitro*. LoxP (locus of X-over of P1) was found in bacteriophage P1 and is 34 base pairs (bp) long. It consists of an 8 bp core element flanked by 13 bp of palindromic sequence (figure 1.11). The core element is a determinant of directionality. Meanwhile, the Cre (cyclisation recombination) recombinase is a 38 kDa protein from bacteriophage P1, which recognises and mediates site-specific recombination between LoxP sites. When two loxP sites are oriented in the same direction, the Cre recombinase induces the deletion of the DNA segment placed between them (figure 1.12a). However, if the two loxP sites are located at the opposite direction, the Cre-recombination event induces the inversion of the floxed sequences (figure 1.12b) (Sauer, 1998, Zhang and Lutz, 2002, Shashikant and Ruddle, 2003, Kuhn and Torres, 2002). Furthermore, if the loxP sites are located on different chromosomes, Cre recombinase mediates a chromosomal translocation (figure 1.12c) (Smith et al., 1995, Kuhn and Torres, 2002).

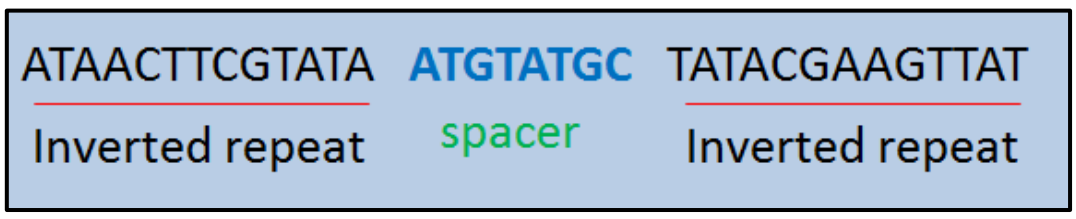


Figure 1.11: The structure of LoxP. LoxP contain an asymmetric 8 bp sequence in between with two sets of palindromic, 13 bp flanking sequences (Kuhn and Torres, 2002).

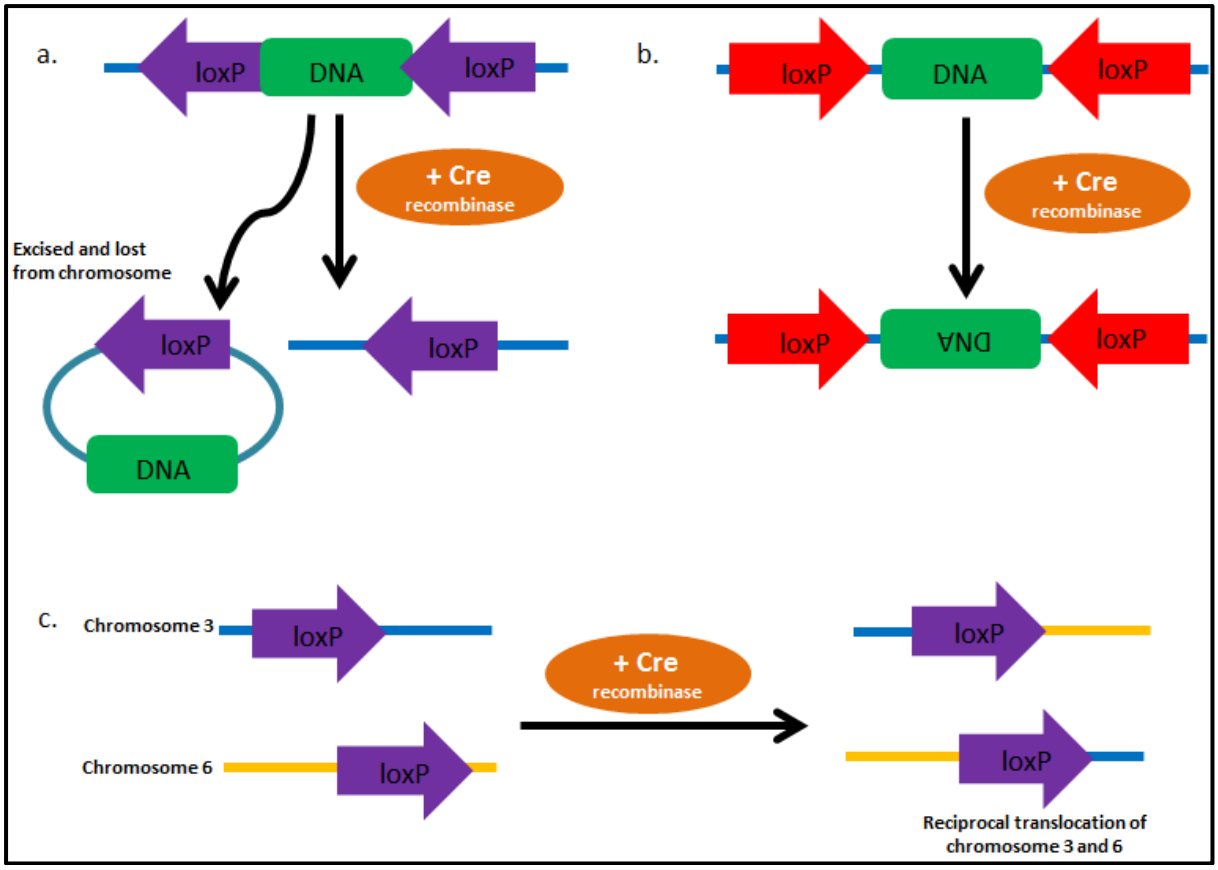


Figure 1.12: The Cre/LoxP mediated recombination. (a) If the loxP sites are located in the same direction on a same chromosome segment, Cre recombinase mediates a deletion of the DNA between the LoxP sites. Subsequently the deleted DNA fragment is degraded. (b) If the loxP sites are located in the opposite orientation on the same chromosome, Cre recombinase mediates the inversion of the flanked DNA segment. (c) If the loxP sites are positioned on different chromosomes, Cre recombinase mediates a reciprocal chromosomal translocation (*The Jackson Laboratory: Cre Expression Resource Web Site*).

To create a conditional gene knockout, loxP sites can be introduced to flank part of the coding region of the chosen gene, creating a floxed gene in a direct orientation. In mouse models, Cre-expressing and loxP-containing strains are developed independently. Multiple mouse strains with tissue-specific Cre expression can be generated and crossed with a loxP strain. Upon Cre expression, recombination occurs between the two LoxP sites, resulting in removal of the intervening DNA and consequently facilitating conditional ablation of gene function. This method is broadly used to generate tissue-specific knockouts (Kos, 2004). Furthermore, the conditional gene expression can also be generated by using Cre recombinase. In this case, a STOP sequence flanked by loxP sites is placed between the promoter and a gene to be expressed. In the absence of Cre, reporter gene expression is prevented by the STOP sequence. However, upon Cre recombinase expression, the floxed

STOP sequence is removed allowing gene of interest to be expressed (Jaisser, 2000) (figure 1.13). The pZ/EG vector which contains two loxP site is an example of a tool utilising this Cre/LoxP recombination system (Novak et al., 2000). By using the combination of the pZ/EG vector and Cre/loxP site-specific recombination strategy tissue-specific expression of genes of interests such as ZEB1 and ZEB2 can be archived.

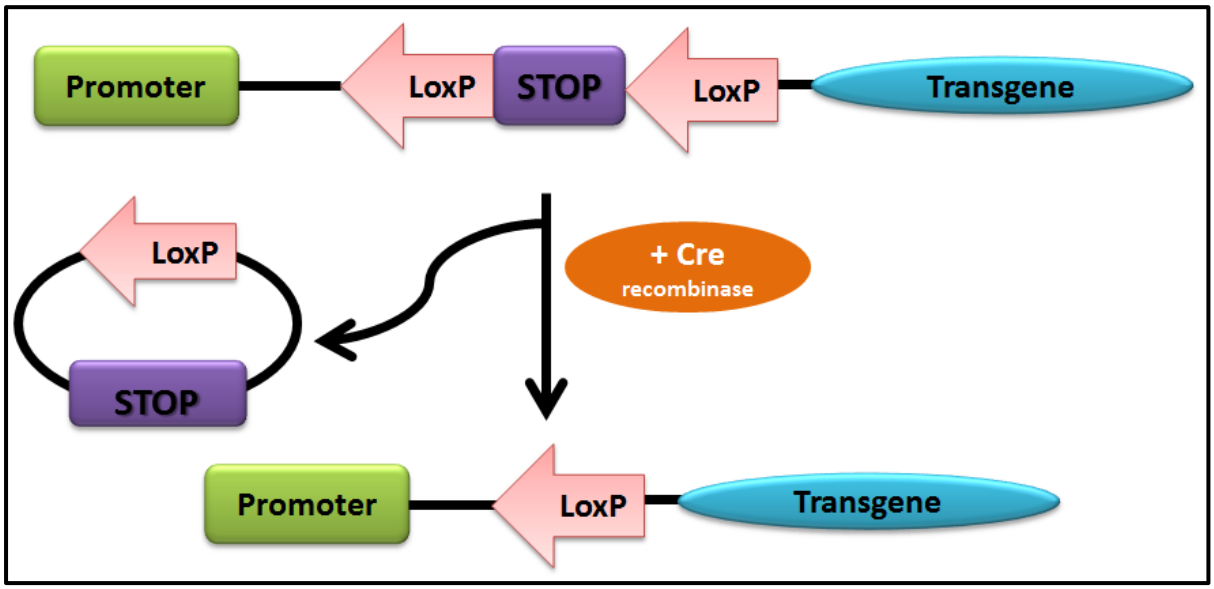


Figure 1.13: Inducible gene expression by using the Cre/LoxP recombination system. A STOP sequence is introduced in between LoxP elements to block the expression of the target transgene. When exposed to the Cre recombinase, the LoxP and STOP sequence are excised and the targeted transgene is expressed (Jaisser, 2000).

1.7 CONCLUSIONS

ZEB proteins play an important and widespread role in EMT in human cancer. However, regulation of *ZEB* genes is insufficiently studied. ZEB1 and ZEB2 transcription factors and several microRNA species (predominantly, miR-200 family members) form a double negative feedback loop that regulates EMT/MET equilibrium. These data, however, do not explain several observations pointing to a lack of co-expression of ZEB proteins in tumour tissues and cell lines (Sayan et al., 2009) and (EMT lab unpublished data). Understanding mechanisms ensuring differential expression of ZEB proteins is important, because accumulating evidence indicates dissimilar and possibly even opposing roles they play at certain phases of tumourigenesis. Indeed, ZEB2 transcriptionally represses cyclin D1 or hTERT resulting in cell cycle arrest or senescence or up-regulates PTEN via a posttranslational mechanism involving microRNA (Mejlvang et al., 2007, Ozturk et al., 2006,

Lin and Elledge, 2003, Karreth et al., 2011). These potentially oncosuppressive functions of ZEB2 have never been attributed to ZEB1, whose pro-proliferative, oncogenic and metastases-promoting features are well documented (Ansieau et al., 2008, Rhim et al., 2012).

1.8 OBJECTIVES

The objective of this work was to characterise regulatory mechanisms that may contribute to differential expression of ZEB proteins in human cancer. The focus of the research was on translational control of ZEB2 protein expression.

CHAPTER 2:

MATERIALS AND METHODS

2.1 MATERIALS

2.1.1 Cell lines

All cell lines were from the American Type Culture Collection (ATCC, Rockville, MD, USA). Details of cell lines are described in table 2.1.

Table 2.1: Description of cell lines used in this study

Cell line	Description	Reference
A431	Human epidermoid carcinoma of the vulva (epidermis) derived from 85-year old female.	(Giard et al., 1973)
MDA-MB-231	Human breast adenocarcinoma cell line derived in 1973 from 51-year old Caucasian woman with pleural effusion.	(Brinkley et al., 1980, Liu and Feng, 2010)
MDA-MB-468	Human breast adenocarcinoma cell line derived in 1977 from a 51-year old black woman with pleural effusion.	(Brinkley et al., 1980)
MCF-7	Human breast adenocarcinoma cell line derived in 1970 from a 69-year old Caucasian woman with a malignant pleural effusion.	(Soule et al., 1973)
T-47D	Human breast adenocarcinoma cell line isolated from a 54-year old woman with pleural effusion and infiltrating ductal carcinoma.	(Keydar et al., 1979, Ware et al., 2007)
ZR-75-1	Human breast adenocarcinoma cell line derived from a malignant ascetic effusion of a 63-year old Caucasian woman.	(Engel et al., 1978)
J82	Human bladder carcinoma cell line derived from a poorly differentiated and invasive transitional cell bladder carcinoma (stage 3) of 58-year old Caucasian male.	(Marshall et al., 1977, O'Toole et al., 1978)
T24	The highly invasive human bladder transitional cell carcinoma derived in 1970 from an 81-year old Caucasian woman.	(O'Toole et al., 1972, Ware et al., 2007, Kawanishi et al., 2008)
RT112	The poorly invasive human bladder carcinoma cell line derived in 1973 from Caucasian female.	(Benham et al., 1977, Kawanishi et al., 2008)
UM-UC-3	Bladder carcinoma cell line derived from a urinary bladder transitional cell carcinoma of human male.	(Grossman et al., 1986, Bellet et al., 1997)
A375P	Human malignant melanoma is a low-metastatic human melanoma cell line derived from a 54-years old female.	(Giard et al., 1973)
SaOS-2	Human osteosarcoma cell line derived in 1973 from an 11-years old Caucasian female.	(Fogh and Trempe, 1975)

U2OS	Human bone osteosarcoma cell line, originally known as the 2T line derived in 1964 from a moderately differentiated sarcoma on the tibia of a 15 year old Caucasian female.	(Ponten and Saksela, 1967, Raile et al., 1994)
H1299	Human non-small cell lung carcinoma cell line derived from metastatic site of lymph node of 43-years old Caucasian male	(Giaccone et al., 1992)
HEK293	Human embryonic kidney cell was generated in the early 70s by transformation of cultures of normal embryonic kidney cells from a health aborted foetus with sheared adenovirus type 5 DNA.	(Graham et al., 1977)

2.1.2 Chemicals, reagents and other items

All chemicals, reagents and items were analytical or molecular biology grade quality and purchased from Sigma-Aldrich, Dorset, UK or Fisher Scientific, Loughborough, UK unless otherwise stated. Details for other chemicals, reagents and items used in this study were listed in Appendix 1.

2.1.3 Antibodies

Table 2.2: Detail of primary antibodies used during this study

Antibody	Host	Dilution	Migration on SDS-PAGE (kDa)	Cat no.	Source	Reference
Green Fluorescent Protein (GFP)	Polyclonal rabbit IgG	1:1000	27	210-PS-IGFP	ImmunoKontact, UK	-
ZEB2[1]	Polyclonal rabbit IgG	1:3000	250	In house	In house	(Sayan et al., 2009, Oztas et al., 2010)
ZEB1 (H-102)	Polyclonal rabbit IgG	1:1000	250	sc-25388	Santa Cruz Biotechnology, USA	-
Cyclin D1(Clone EPR224)	Rabbit monoclonal	1:2000	36	04-1151	Milipore, UK	-
HA (Clone HA-7)	Mouse monoclonal IgG1	1:1000	Dependent upon the molecular weight of the HA tag being detected	H 3663-200UL	Sigma, USA	-
E-cadherin	Mouse monoclonal IgG2a, κ	1:2000	120	610181	BD Transduction, UK	-
Vimentin	Mouse monoclonal IgG1	1:1000	60	550513	BD Transduction, UK	-
α Tubilin (Clone B-5-1-2)	Mouse monoclonal IgG1	1:5000	55	T5168	Sigma-Aldrich, USA	-

*Note: [1] Available in the laboratory of Dr E. Tulchinsky, Department of Cancer Studies, University of Leicester, Leicester, UK.

Table 2.3: Detail of secondary antibodies used during this study

Antibody	Host	Dilution	Cat no.	Source
Anti-Mouse Immunoglobulins/HRP	Polyclonal goat	1:2000	P 0447	DAKO
Anti-Rabbit Immunoglobulins/HRP	Polyclonal goat	1:2000	P 0448	DAKO

2.1.4 Oligonucleotides

Table 2.4: Detail of oligonucleotide primers used during this study

Primer name	Sequence (5'-3')		Tm (°C)	Cycle no	Product size (bp)	Notes
	Forward	Reverse				
ZEB1	CACCAAGATCTGCGGATGGCCCCAGGTGTAAG	CCCGCGGTACCTTATGCATAGTCTGGGACGTCATATGGATAAGCTTCATTTGTCTTCTCTTCAGAC	68	25	3354	Primer set for cloning the full length mouse ZEB1 into pEGFP-C1. NCBI Reference Sequence: NM_011546.3
ZEB2	CACCAAGATCTAAGCAGCCGATCATGGCGGATG	CCCGCGGTACCTTATGCATAGTCTGGGACGTCATATGGATATTCCATGCCATCTTCCATATTGTCTTC	68	25	3648	Primer set for cloning the full length mouse ZEB2 into pEGFP-C1. NCBI Reference Sequence: NM_015753.3
ZEB1 (317-372aa)	CACCAAGATCTAAGATAGAGAATAAACCCCTTC	CCCGCGGTACCTCATGCATAGTCTGGGACGTCATATGGATATTGCACCACACCCTGAGGAG	68	25	219	Primer set for cloning the truncated mouse ZEB1 317-372 aa into pEGFP-C1. NCBI Reference Sequence: NM_011546.3
ZEB2 (372-437aa)	CACCAAGATCTAAGTTGGAAAATGGAAAACCAC	CCCGCGGTACCTCATGCATAGTCTGGGACGTCATATGGATACTGCATTGGACTC	68	25	242	Primer set for cloning the truncated mouse ZEB2 372-437 aa into pEGFP-C1. NCBI Reference Sequence:

		TGAGCAG				NM_015753.3
mRNA.ZEB1	ACAAGCGAGAGGATCATGGC	TCCTGCTTCATCTGCCTGAG	60	25	126	Primer set for the mRNA amplification of human ZEB1 (Full length mRNA for ZEB1; NCBI Reference Sequence: NM_001128128.2)
mRNA.ZEB2	CAACTGCCATCTGATCCGCTC	GTCTGGATCGTGGCTTCTGG	60	25	410	Primer set for the mRNA amplification of human ZEB2 (Full length mRNA for ZEB2; NCBI Reference Sequence: NM_014795.3
mRNA.E-cadherin	CAGCCACAGACGCGGACGAT	CAGCCTCCCACGCTGGGGTA	58	32	381	Primer set for the mRNA amplification of human E-cadherin (Full length mRNA for E-cadherin; GenBank Reference Sequence: AB025105.1)
mRNA.GADPH	TCTTTTGCGTCGCCAGCCGA	GCCAGCATCGCCCCACTTGA	60	25	312	Primer set for the mRNA amplification of human GAPDH (Full length mRNA for GAPDH; GenBank Reference Sequence: AB062273.1)
ZEB2 (STOP-383)	CACTTAGCATGTCT**T**AGCAGACAGGCTTAC	GTAAGCCTGTCTGCT**A**AGACATGCTAAGTG	55	18	N/A	Primer set for site-directed mutagenesis of STOP codon insertion at the position of 383 aa of mouse truncated ZEB2 372-437 aa. NCBI Reference Sequence: NM_015753.3
ZEB2 (STOP-400)	CCACTAGACTTCAATGACTA**G**AAAGTTCTTATGGCAACAC	GTGTTGCCATAAGAACTTT**C**TAGTCATTGAAGTCTAGTGG	55	18	N/A	Primer set for site-directed mutagenesis of STOP codon insertion at the position of 400 aa of mouse truncated ZEB2 372-437 aa. NCBI Reference Sequence: NM_015753.3
ZEB2 (STOP-415)	GGCAGCAGTCCCT**AA**ATGAACGGTGGG	CCCACCGTTCAT**TT**AGGGACTGCTGCC	55	18	N/A	Primer set for site-site-directed mutagenesis of STOP codon insertion at the position of 415 aa of mouse

						truncated ZEB2 372-437 aa. NCBI Reference Sequence: NM_015753.3
ZEB2 (STOP-426)	CCACCAGCCCTT**A**AGG TGTACACCC	GGGTGTACACCT**T**AAGGG CTGGTGG	55	18	N/A	Primer set for site-directed mutagenesis of STOP codon insertion at the position of 426 aa of mouse truncated ZEB2 372-437 aa. NCBI Reference Sequence: NM_015753.3
ZEB2 (ATG-400)	CACCAAGATCTGAGGA CC**ATG**TATCCATATGA CGTCCCAGACTATGCA TATAAAGTTCTTATGG CAACACATG	CCCGCGGTACCGCTGCAT TGGACTCTGAGCAGAT	68	25	158	Primer set for ATG codon insertion at the position of 400 aa of mouse truncated ZEB2 372-437 aa. NCBI Reference Sequence: NM_015753.3
ZEB2 (ATG-415)	CACCAAGATCTGAGGA CC**ATG**TATCCATATGA CGTCCCAGACTATGCA TTTATGAACGGTGGGC TTGGAGCC	CCCGCGGTACCGCTGCAT TGGACTCTGAGCAGAT	68	25	113	Primer set for ATG codon insertion at the position of 415 aa of mouse truncated ZEB2 372-437 aa. NCBI Reference Sequence: NM_015753.3
ZEB2 (ATG-426)	CACCAAGATCTGAGGA CC**ATG**TATCCATATGA CGTCCCAGACTATGCA CTTTAGGTGTACACCC ATCTGCTCAG	CCCGCGGTACCGCTGCAT TGGACTCTGAGCAGAT	68	25	89	Primer set for ATG codon insertion at the position of 426 aa of mouse truncated ZEB2 372-437 aa. NCBI Reference Sequence: NM_015753.3
ZEB2 (RC-CC: 426-428)	GCCACCAGCCCT**C**T**G**G G**C**GT**G**CACCCATCTGC TCAG	CTGAGCAGATGGGTGCA CGCCCAGAGGGCTGGTG GC	55	18 or 30	N/A	Primer set for site or multi-site-directed mutagenesis of mouse ZEB2 or ZEB2 (372-437aa), Rare (RC) codons to common codons (CC) at the location of 426-428 aa. NCBI Reference Sequence: NM_015753.3
ZEB2 (RC-CC: 439-441)	CCAATGCAGCAC**C**T**G**G G**C**GT**G**GGGATGGAAG	GGGGCTTCCATCCCCACG CCCAGGTGCTGCATTGG	55	18 or 30	N/A	Primer set for site or multi-site-directed mutagenesis of mouse ZEB2,

	CCCC					RC to CC at the location of 439-441 aa. NCBI Reference Sequence: NM_015753.3
ZEB1 (CC-RC: 367-369)	GTGGTGGCCCATT**AG GT**G**T**AACCAGTTCTCC TC	GAGGAGAACTGGTTACAC CTAATGGGCCACCAC	55	30	N/A	Primer set for site or multi-site-directed mutagenesis of mouse ZEB1 CC to RC at the location of 367-369aa. NCBI Reference Sequence: NM_011546.3
ZEB1 (CC-RC: 377-379)	CAGGGTGTG**T**T**AGGT** G**TA**GTTGTTCTGCC	GGCAGAACAACTACACCT AACACACCCTG	55	30	N/A	Primer set for site or multi-site-directed mutagenesis of mouse ZEB1 CC to RC at the location of 377-379aa. NCBI Reference Sequence: NM_011546.3
ZEB2-FP	GCAGGTAACCGCAAG TTCAAGTGC	N/A	N/A	N/A	N/A	Primer sequence used for sequencing analysis designed by GATC Biotech AG Company, Germany.
ZEB1-FP	CCATACGAATGCCCGA ACTG					
pEGFP-C2-FP	GATCACATGGTCCTGC TG					
pEGFP-FP	TTTAGTGAACCGTCAG ATC					
pEGFP-RP	AACAGCTCCTCGCCCT TG					

*Note: 1. all aa substitutions in the primer are shown with an **underline**; 2. N/A: not available

2.1.5 Plasmids

Table 2.5: Detail of plasmids used during this study

Plasmid	Vector	Insert	Notes	Source
pcDNA™3.1/TOPO®-HA-ZEB1	pcDNA™3.1/TOPO®; eukaryotic expression vector	Full length mouse ZEB1 (NCBI Reference Sequence: NM_011546.3)	ZEB1 cloned downstream of the CMV promoter and in between *BamHI* and *XhoI* sites (for plasmid map see Appendix 2A).	Dr E. Tulchinsky (EMT laboratory, University of Leicester, UK)
pcDNA™3.1/TOPO®-HA-ZEB2		Full length mouse ZEB2 (NCBI Reference Sequence: NM_015753.3)	ZEB2 cloned downstream of the CMV promoter and in between *BamHI* and *XhoI* sites (for plasmid map see Appendix 2B).	
pcDNA™3.1/TOPO®-HA-ZEB2 (STOP 305)		Full length mouse ZEB2 with STOP 305 codon insertion	pcDNA™3.1/TOPO®-HA-ZEB2-FL with STOP codon introduced at the position of 305 aa (for plasmid map see Appendix 2C).	Dr G. Browne (EMT laboratory, University of Leicester, UK)
pcDNA™3.1/TOPO®-HA-ZEB2 (STOP 372)		Full length mouse ZEB2 with STOP 372 codon insertion	pcDNA™3.1/TOPO®-HA-ZEB2-FL with STOP codon introduced at the position of 372 aa (for plasmid map see Appendix 2C).	
pcDNA™3.1/TOPO®-HA-ZEB2 (STOP 406)		Full length mouse ZEB2 with STOP 406 codon insertion	pcDNA™3.1/TOPO®-HA-ZEB2-FL with STOP codon introduced at the position of 406 aa (for plasmid map see Appendix 2C).	
pcDNA™3.1/TOPO®-HA-ZEB2 (STOP 437)		Full length mouse ZEB2 with STOP 437 codon insertion	pcDNA™3.1/TOPO®-HA-ZEB2-FL with STOP codon introduced at the position of 437 aa (for plasmid map see Appendix 2C).	
pcDNA™3.1/TOPO®-HA-		Full length mouse ZEB2 with STOP 505 codon insertion	pcDNA™3.1/TOPO®-HA-ZEB2-FL with STOP codon introduced at	

ZEB2 (STOP 505)			the position of 505 aa (for plasmid map see Appendix 2C).	
pcDNA™3.1/TOPO®-HA-ZEB2 (STOP 705)		Full length mouse ZEB2 with STOP 705 codon insertion	pcDNA™3.1/TOPO®-HA-ZEB2-FL with STOP codon introduced at the position of 705 aa (for plasmid map see Appendix 2C).	
pcDNA™3.1/TOPO®-HA-ZEB2 (STOP 904)		Full length mouse ZEB2 with STOP 904 codon insertion	pcDNA™3.1/TOPO®-HA-ZEB2-FL with STOP codon introduced at the position of 904 aa (for plasmid map see Appendix 2C).	
pcDNA™3.1/TOPO®-HA-ZEB2 (STOP 1023)		Full length mouse ZEB2 with STOP 1023 codon insertion	pcDNA™3.1/TOPO®-HA-ZEB2-FL with STOP codon introduced at the position of 1023 aa (for plasmid map see Appendix 2C).	
EGFP-ZEB1 (317-372aa)		Truncated mouse ZEB1 317-372 aa	Truncated ZEB1 cloned between *BglII* and *XhoI* of pEGFP-N2 (for plasmid map see Appendix 2D).	
EGFP-ZEB2 (372-437aa)		Truncated mouse ZEB2 372-437 aa	Truncated ZEB2 cloned between *BglII* and *XhoI* of pEGFP-N2 (for plasmid map see Appendix 2D).	
pZ/EG	pZ/EG; eukaryotic expression vector	N/A	The pZ/EG contains LacZ floxed by two loxP sites, driven by the chicken β-actin promoter and a cytomegalovirus (CMV) enhancer with enhanced green fluorescent protein (EGFP) downstream (for plasmid map see Appendix 2E).	Dr C. Lobe (University of Toronto, Canada)
pCre-pac	pCre-Pac; mammalian expression vector	N/A	This vector encodes Cre-recombinase and the puromycin resistance gene (pCre-pac) (for	Professor C. Pritchard (University of Leicester, UK)

			plasmid map see Appendix 2F).	
pEGFP-C1	pEGFP-C1; eukaryotic expression vector with N-terminal GFP	N/A	pEGFP-C1 encodes a red shifted variant of wild-type GFP which has been optimised for brighter fluorescence and higher expression in mammalian cells. (Excitation maximum= 488 nm; emission maximum= 507 nm). The multiple cloning sites (MCS) in pEGFP-C1 is between the EGFP coding sequences and the SV40 poly A. Inserts cloned in these cloning sites will result in N-terminally-tagged GFP fusion proteins (for plasmid Appendix see 2G).	Cat no: 6084-1, Clontech laboratories, UK.
pEGFP-N2	pEGFP-N2; eukaryotic expression vector with C-terminal GFP	N/A	pEGFP-N2 encodes a red shifted variant of wild-type GFP which has been optimised for brighter fluorescence and higher expression in mammalian cells. (Excitation maximum= 488 nm; emission maximum= 507 nm). The MSC in pEGFP-N2 is between the immediate early promoter of CMV and the EGFP coding sequence. Inserts cloned in these cloning sites will result in C-terminally-tagged GFP fusion proteins (for plasmid map see	Cat no: 6081-1, Clontech laboratories, UK.

			Appendix 2H).	
pmaxGFP®		N/A	Plasmid encoding maxGFP, a Green fluorescent protein from the copepod *Pontellina p*. This plasmid is under regulation of a CMV promoter element and is kanamycin resistant (for plasmid map see Appendix 2I).	Amaxa nucleofector kit, Cat no: VCA-1003V, Lonza, UK.

*Note: N/A: not available

All plasmid maps were shown in Appendix 2A to 2I

2.2 METHODS

2.2.1 Mammalian cell culture technique

2.2.1.1 Resuscitation of frozen cells from liquid nitrogen

Cell lines were taken from liquid nitrogen and rapidly thawed in a 37˚C water bath. Cells were then re-suspended in 4 ml complete culture medium and centrifuged at 1000 rpm for 5 minutes. After centrifugation, the supernatant was discarded and the cell pellet was re-suspended in 10 ml complete medium. Cells were then transferred to a 25 cm^2 cell culture flask. The human squamous epidermoid carcinoma cells A431, the HEK293 cell, the human bladder cancer cell lines T24 and UM-UC-3 and the human breast carcinoma cell lines MDA-MB-231, MDA-MB-468, MCF-7 and T-47D were cultured in DMEM high glucose with L-glutamine supplemented 10% FBS and 1% PS. The human osteosarcoma cell lines SaOS-2 and U2OS were cultured in McCoy 5A with L-glutamine supplemented 10% v/v FBS and 1% v/v PS. The human bladder carcinoma cells J82 and RT112 were cultured in DMEM high glucose with L-glutamine supplemented with 10% v/v FBS, 1% v/v PS and 1% v/v NEAA. The NSCLC cell line H1299, human breast carcinoma cell line ZR-75-1 and the human melanoma cells A375P was cultured in RPMI-1640 with L-glutamine supplemented 10% v/v FBS and 1% v/v PS. All cell lines were grown in a 100% humidified incubator at 37˚C in an atmosphere of 5% CO_2. Cells were split once they had reached approximately 70% confluence. All cells tested negative for various forms of *Mycoplasma*.

2.2.1.2 Cell passaging

Cells were rinsed twice with PBS (160 mM NaCl, 3mM KCl, 8 mM Na_2HPO_4, 1 mM KH_2PO_4, pH 7.3), and incubated in TE (0.5 mg/ml Trypsin, 0.22 mg/ml EDTA in PBS, pH 7.5) 5 minutes. Flasks were gently tapped to detach cells and pre-warmed media was added to neutralise the TE. Cells were pelleted at 1000 rpm for 5 minutes, re-suspended in complete media and re-plated (1:10 dilution) into a new culture flask. Volume of PBS, TE and media required for 25 cm^2, 75 cm^2 and 175 cm^2 flasks are detailed in table 2.6.

Table 2.6: Volume of PBS, TE and culture media require during cell passaging

Flask (cm^2)	PBS (ml)	TE (ml)	Complete culture media added to neutralise TE (ml)	Total volume of complete culture media (ml)
25	2	1	4	5
75	5	2	6	10
175	10	3	7	25

2.2.1.3 Cryopreservation and storage

Cells were rinsed twice with PBS, trypsinised with TE and centrifuged at 1000 rpm for 5 minutes. The cell pellet was re-suspended with 10 ml freezing media (10% v/v FBS, 10% v/v DMSO and 80% v/v complete media) and 1 ml of cells were aliquoted into individual cryo-tube. The tubes were placed in a cryo-container with 2-propanol and slowly frozen at -80˚C overnight. All tubes were then transferred to the liquid nitrogen tank the next day.

2.2.1.4 Determination of cell number

Cells were rinsed twice with PBS, trypsinised with TE and centrifuged at 1000 rpm for 5 minutes. The cell pellet was re-suspended with 1 ml media and 10 µl of cells were added to a Coulter Counter vial containing 10 ml of Isoton®II Diluent. The vial was placed into the Z2™Coulter analyser Z2™ (Beckman Coulter Inc, UK) and cell numbers were counted between a lower limit of 8 µm, upper limit of 20 µm and dilution factor of 1000.

2.2.1.5 Transfections

Transient transfection was used as a method of over-expressing proteins of interest in mammalian cells. Different transfection methods were used depending on the requirements of each cell line.

2.2.1.5.1 JetPRIME™

Transfection of H1299 and HEK 293 cells were performed using JetPRIME™ DNA transfection reagent. The day before transfection, cells (n = 2.5×10^6) were plated onto 6 cm dishes so cells were 60-80% confluent at the time of transfection. On the day of transfection, 4 µg of DNA was mixed with 200 µl of JetPRIME™ buffer and 8 µl of JetPRIME™ reagent. The

mixture was vortexed for 10 seconds, centrifuged at 1000 rpm for 1 minute and incubated at room temperature for 10 minutes. After incubation, the mixture was dropped evenly onto the cells. The media was changed after 4-6 hours. Cells were grown for 24 hours and used for further analysis. To check the efficiency of transfections, the expression vector pEGFP-C1 was transfected in parallel.

2.2.1.5.2 Electroporation

Transfection of A431, MDA-MB-231, MDA-MB-468 and SaOS-2 cells were carried out using the electroporation method. First, cells (n = 2.5×10^6) were pipetted into 1.5 ml eppendorf tube and centrifuged at 1000 rpm for 5 minutes. The supernatant was discarded and the pellet was re-suspended with 60 µl of Ingenio® electroporation solution. The cell suspension was mixed with 4 µg of DNA and the mixture was pipetted into 4 mm electroporation cuvette. The cuvette was placed in the GenePulser Xcell machine (BioRad, UK) and electroporated using 250/250 program (Voltage: 250 volts (V), Capacitance: 250 µF, Resistance: ∞ Ω). Following electroporation, cells were re-suspended in 1 ml culture medium and transferred into 6 cm dish containing 5 ml fresh media. Cells were grown in a humidified incubator at 37˚C in an atmosphere of 5% CO_2 and used for further analysis 24 hours post transfection.

2.2.1.6 Cell treatments

2.2.1.6.1 MG-132

The MG-132 stock (10 mM) was prepared in DMSO and kept at -20^{o}C. 24 hours post transfection, the old medium was aspirated and replaced with fresh medium containing MG-132 (10 µM) (Yuan et al., 2008). Cells were analysed at baseline (0), and then after 6, and 16 hours of incubation.

2.2.1.6.2 Chloroquine

The Chloroquine stock (100 mM) was prepared in sterile dH_2O and kept at -20^{o}C. 24 hours post transfection, the medium was aspirated and replaced with pre-warmed fresh medium containing Chloroquine (100 µM) (Myeku and Figueiredo-Pereira, 2011). Cells were incubated and analysed at baseline (0), after 6 and 16 hours of incubation.

2.2.2 Protein analysis technique

2.2.2.1 Preparation of protein lysates

Cells were washed twice with PBS and 1× Laemmli buffer (50 mM Tris-HCl (pH 6.8); 2% SDS; 10% (v/v) glycerol) was added into the dish. Cells were scraped into 1.5 ml eppendorf tube and lysates were sonicated using a MSE Soniprep 150 machine (MSE, Lower Sydenham, UK) for 15 seconds to disrupt chromosomal DNA. Lysates were then heated to 95^{o}C for 5 minutes and stored at -20^{o}C for later use.

2.2.2.2 Protein quantification

The protein concentrations of Laemmli buffer solubilised lysates were quantified using the Pierce BCA protein assay kit. First the working reagent was prepared by mixing 50 part of reagent A and 1 part of reagent B. About 200 µl of prepared reagent was pipetted to 25 µl of unknown protein samples and 25 µl of provided Bovine serum albumin (BSA) standards (concentration ranging 0-2 µg/µl) in 96 wells plate. The plate was incubated at 37^{o}C for 30 minutes and an absorbance reading at 562 nm was measured using the ELx808iu Ultra microplate reader (Bio-Tek Instruments Inc., USA). In order to determine the concentration of the unknown samples, the absorbance for the blank (0 µg/µl) was subtracted from each sample and a straight line was fitted to the graph of absorbance versus concentration. After quantification, lysates were equilibrated to 1 µg/µl and 1µl of 1× protein loading buffer (1% v/v β-mercaptoethanol and 0.006% w/v bromophenol blue) was added. For quantification of pulled-down lysates, a Bradford assay was used. 5 µl of unknown samples and 5 µl of provided BSA standards were added to 250 µl Bradford reagent, vortexed and incubated at room temperature for 10 minutes. Protein samples were measured at 595 nm and protein concentration was calculated as described above.

2.2.2.3 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

The percentage of the SDS-PAGE to use depends on the molecular weight (MW) of the protein of interest to be separated. Higher molecular weight proteins required a lower percentage of SDS-PAGE for better separation and identification. The resolving gel (table 2.7) was prepared and poured into the glass plate cassette. Water-saturated butanol (100

µl) was added on top of the resolving gel to produce a homogenous surface and prevent evaporation. After polymerisation, the gel was rinsed with dH_20 and filter paper was used to absorb the excess water. The stacking gel (table 2.7) was prepared and pipetted to the top of the resolving gel and immediately the 15-wells comb was inserted. After polymerisation, the comb was removed and the wells were rinsed with dH_20. The gel was placed in the tank and filled with 1× protein running buffer (25 mM Tris-HCl; 250 mM glycine; 0.1% (w/v) SDS, pH 8.3). A sample containing 20 µg of total protein was loaded onto the gel. About 3 µl of the PageRuler™ Plus Prestained Protein Ladder (250 kDa) was also loaded into a separate well. The gel was run at 120 V until both sample and the protein ladder reached the bottom part of the gel.

Table 2.7: The solutions for preparing resolving and stacking gels for SDS-PAGE

Resolving gel				
% of gels	**Volume**			
8%	**5 ml**	**10 ml**	**15 ml**	**20 ml**
dH_2O	2.3	4.6	7.0	9.3
30% acrylamide	1.3	2.7	4.0	5.3
1.5 M Tris (pH 8.8)	1.3	2.5	3.8	5.0
10% SDS	0.05	0.10	0.15	0.20
10% APS	0.05	0.10	0.15	0.20
TEMED	0.003	0.006	0.009	0.012
12%	**5 ml**	**10 ml**	**15 ml**	**20 ml**
dH_2O	1.7	3.3	5.0	6.6
30% acrylamide	2.0	4.0	6.0	8.0
1.5 M Tris (pH 8.8)	1.3	2.5	3.8	5.0
10% SDS	0.05	0.10	0.15	0.20
10% APS	0.05	0.10	0.15	0.20
TEMED	0.002	0.004	0.006	0.008
Stacking gel				
5%	**1 ml**	**2 ml**	**3 ml**	**4 ml**
H_2O	0.68	1.40	2.10	2.70
30% acrylamide	0.17	0.33	0.50	0.67
1.5 M Tris (pH 8.8)	0.13	0.25	0.38	0.50
10% SDS	0.01	0.02	0.03	0.04
10% APS	0.01	0.02	0.03	0.04
TEMED	0.001	0.002	0.003	0.004

2.2.2.3.1 Coomassie blue staining of SDS-PAGE

The SDS-PAGE was stained with 45% (v/v) methanol, 10% (v/v) acetic acid, 0.25% (w/v) Brilliant Blue G for 1 hour at room temperature. After incubation, the gel was de-stained for 2 hours with several changes of 45% (v/v) methanol and 10% (v/v) acetic acid and imaged using a Canon PowerShot digital camera (Canon UK Ltd.).

2.2.2.3.2 Western blotting

Following SDS-PAGE, resolved gel was rinsed with dH_2O and transfer to PVDF membrane using a wet transfer apparatus. First, the PVDF membrane was soaked in methanol for 1 minute and then into 1× transfer buffer (25 mM Tris, 250 mM glycine, 20% (v/v) methanol; pH 8.3) for 5 minutes . All the filter papers and the sponges were also soaked into 1× transfer buffer until used. The transfer apparatus was set up as in figure 2.1. Next, the prepared apparatus was placed into the tank and 1× transfer buffer was poured. Electrophoresis was carried out at 22 V overnight at room temperature. Following transfer, the apparatus was dissembled and the membrane was soaked into Ponceau S staining solution (0.1% (w/v) Ponceau S; 5% (v/v) acetic acid) for 5 minutes to confirm that the proteins had been transferred. The membrane was then washed in water for 10 minutes and once in TBS-T (20mM Tris-HCl (pH 8.0); 150 mM NaCl, 0.1% (v/v) Tween-20) to remove the stain and processed for immunoblotting. In order to block non-specific binding of the antibody, the PVDF membrane containing immobilised proteins was incubated in 5 % blotting solution (5% (w/v) dried milk powder in TBS-T) for 1 hour at room temperature. Next, primary antibody was diluted in 5 % blotting solution to an appropriate concentration and added to the membrane for 1 hour. After incubation, membrane was rinsed 3 times with 1× TBST. The secondary antibody conjugated with horseradish peroxidase (HRP) (also diluted in 5 % blotting solution) was added and the membrane was incubated at room temperature for 1 hour. After incubation, the membrane was washed 3 times for 10 minutes in 1× TBST. The Pierce ECL Western Blotting Substrate was used for protein detection. The reagent A and reagent B provided were mixed 1:1 and added to the membrane for 1 minute. The membrane was placed into the Hypercassette™ (Amersham Pharmacia Biotech, UK) which was covered with transparency film. Another transparency film was used to cover the top of the membrane and sealed with the sticky tape. The membrane was exposed to the CL xposure film and developed using AGFA Curix 60 film

developer (AGFA Healthcare, UK). The detected protein of interest will appear as dark regions on the developed film. All the developing works were done in the dark room. Films were scanned with a Canon LIDE 60 scanner using Arcsoft Photostudio 5 (Arcsoft, Inc., USA) software.

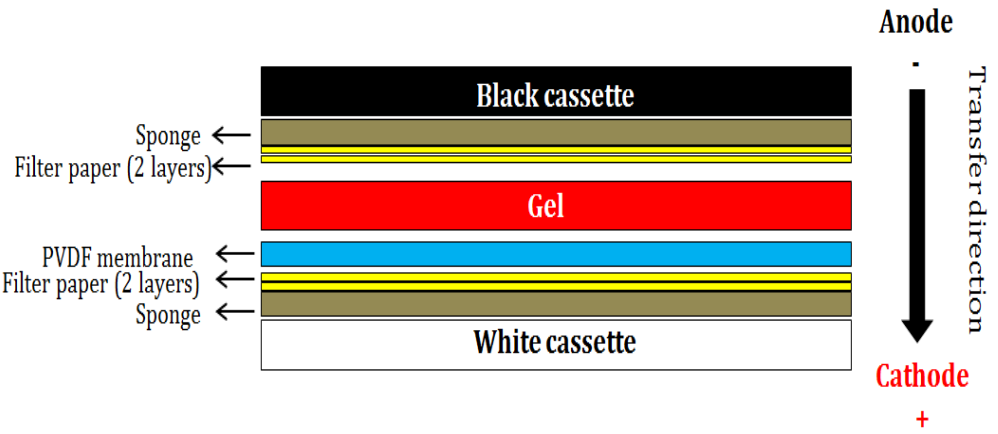


Figure 2.1: The gel and membrane setup for protein transfer.

2.2.2.4 Pull-down assay of GFP-fusion protein

The pull-down assay was performed using the GFP-Trap®_A kit by following the manufacturer's protocol. First, cells were rinsed twice with ice cold PBS and 500 µl of lysis buffer was added into the dish. The cells were scraped into 1.5 ml eppendorf tube and centrifuged for 10 minutes at 13000 rpm at 4°C. The supernatant was transferred to a new pre-cooled eppendorf tube and the pellet was discarded. The volume of the protein extract was adjusted to 1 ml by adding an appropriate volume of dilution buffer. An aliquot (200 µl) of the protein extract (referred as input) was diluted with 200 µl of 2× Laemmli buffer (100 mM Tris-HCl (pH 6.8); 4% SDS; 20% (v/v) glycerol). This input sample was kept for western blot analysis. Next, 30 µl of GFP-Trap®_A beads were re-suspended in 1 ml of ice-cold dilution buffer and centrifuged at 13000 rpm for 2 minutes at 4°C. The supernatant was discarded and the washing step was repeated 4 times. After washing, the remaining protein extract was mixed with the beads and incubated under constant mixing for 2 hours at 4°C. After incubation, the mixture was centrifuged at 13000 rpm for 2 minutes at 4°C. About 200 µl of the supernatant (referred to as non-bound) was taken and diluted with 200 µl of 2× Laemmli buffer. This non-bound sample was also kept for western blot analysis. The

remaining supernatant was discarded. Next, the beads were washed twice with 1 ml of ice cold wash buffer and re-suspended with 200 µl of 2× Laemmli buffer. The beads were then boiled for 10 minutes at 95°C and centrifuged for 2 minutes at 13000 rpm. The supernatant (referred to as bound) was transferred to a new eppendorf tube. The protein concentration of all samples (input, non-bound and bound) was measured using Bradford assay, diluted to a final concentration of 1 µg/µl, mixed with 1× protein loading buffer and finally 10 µg of the samples were resolved by SDS-PAGE.

2.2.2.5 Protein mass spectrometry

A sample containing 40 µg of GFP-fusion protein was loaded onto the gel. The gel was run at 120 V until the sample was separated about 5 cm from resolving gel. After separation, the gel area containing the protein sample was excised and transferred into a clean eppendorf tube. The sample was submitted to The Protein Nucleic Acid Chemistry Laboratory (PNACL), Hodgkin Building, University of Leicester, UK (Website: www.le.ac.uk/mrctox/pnacl) for protein mass spectrometry analysis.

2.2.2.6 Polysome profiling

2.2.2.6.1 Sucrose gradient preparation

The sucrose gradient was prepared according to the protocol borrowed from Dr E. Smith, The Protein Nucleic Acid Chemistry Laboratory (PNACL), Hodgkin Building, University of Leicester, UK. A sucrose density gradient was prepared by layering sucrose solutions of different density upon one another. First, 50 ml of 10%, 20%, 30%, 40% and 50% w/v sucrose solutions were prepared by dissolving an appropriate amount of sucrose in 5 ml of 10× polysome gradient buffer (3M NaCl; 150mM MgCl2; 150mM Tris-HCl (pH7.5); 1 mg/ml CHX; 10 mg/ml heparin and RNase-free H_2O as details below:

	Sucrose	**10× polysome gradient buffer**	
10%	5g	5ml	Make up to 50ml in RNase-free water
20%	10g	5ml	
30%	15g	5ml	
40%	20g	5ml	
50%	25g	5ml	

Following this, the gradient was then prepared by layering gradually less dense sucrose solution upon another. First, 2.15 ml of 50% sucrose solution was added into a clean Sorvall PA tube and immediately placed in -80^{o}C freezer for 30 minutes. Next, 2.15 ml of 40% sucrose solution was carefully added on top of 50% sucrose and freeze in -80^{o}C freezer for another 30 minutes. The freezing-layering process was repeated for 30%, 20% and 10%. Finally, the gradient was stored at -80^{o}C indefinitely covered with parafilm. (Note: CHX is a protein synthesis inhibitor that was used to arrest translating ribosomes during polysome profiling experiment (Huntzinger and Izaurralde, 2011). Heparin was used as unspecific RNAse inhibitors that allow us directly used total RNA (instead of polyA^{+} RNA) in polysome bound expression profiling experiments (del Prete et al., 2007))

2.2.2.6.2 Preparation of the cell extracts and polysome purification

Cells were incubated with CHX (100 µg/ml; freshly prepared in ethanol) for 3 minutes prior harvesting. The cells were then rinsed with ice cold PBS+CHX (10 mg/ml CHX in PBS) twice, scraped in 9 ml of PBS+CHX and transferred into 50 ml falcon tube. Next, the dish was rinsed with 5 ml PBS+CHX and the cells remaining were collected into the same 50 ml falcon tube. The cells were centrifuged for 5 minutes at 1400 rpm at 4°C and the supernatant was discarded. The pellet was re-suspended with 1 ml PBS+CHX, transferred into a new pre-cooled 1.5 ml eppendorf tube and centrifuged at 3000 rpm at 4°C for 4 minutes. After centrifugation, the supernatant was removed and the pellet was lysed with 1 ml 1× Polysome lysis buffer (10% v/v polysome gradient buffer (10×), 1% v/v Triton-X). Next, the nuclei were cleared by centrifugation at 13000 rpm for 1 minute at 4°C and the supernatant was removed and layered onto the top of the sucrose gradient. Next, the sample was separated by centrifugation in Sorvall WX Ultra Series Centrifuge (Thermoscientific, UK) at 38000 rpm for 2 hours at 4°C. After centrifugation, the lysosome sample was fractionated using ISCO Foxy Jr. Fractionator (Teledyne Isco, USA) and absorbance at 254 nm was monitored using UA-6 UV/VIS Detector machine (Teledyne Isco, USA). The fractions were collected into 11 fractions of approximately 1 ml each and immediately stored at -80°C.

2.2.2.6.3 Trichloroacetic acid (TCA) precipitation

In order to concentrate the protein fraction for analysis by SDS-PAGE, an equal volume of 20% w/v TCA (22 g TCA in dH_2O) solution was added into protein fraction sample. The

mixture was then incubated on ice for 30 minutes and centrifuged at 1300 rpm for 15 minutes at 4°C. Next, the supernatant was removed and the pellet was re-suspended with 300 µl of ice cold acetone. The mixture was again centrifuged at 1300 rpm for 5 minutes at 4°C. After centrifugation, the supernatant was discarded and the pellet was left to dry at room temperature for 5 minutes. Next, the pellet was re-suspended with 100 µl of 1× Laemmli buffer and heated at 65°C for 3 minutes. The protein lysate was then kept at -20°C for later use.

2.2.3 Molecular biology technique

2.2.3.1 Primer design

All primers used for polymerase chain reaction (PCR) and Reverse Transcription-PCR were designed using Primer 3 software (Rozen and Skaletsky, 2000) and the specificities for these primers were checked using BLAST (Nucleotide Basic Local Alignment Search Tool) programme from NCBI (Website: http://blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul et al., 1990). The lyophilised primers purchased from Sigma-Genosys, UK were dissolved in dH_20 to a master stock concentration of 100 µM as directed in the supplier's data sheet. An aliquot was taken and diluted (1:10) in dH_2O to make up a working concentration of 10 µM. All primers used for site and multi-site directed mutagenesis were designed using Primer X software (Website: http://www.bioinformatics.org/primerx/). The forward and reverse primers purchased from Sigma-Genosys, UK were centrifuged at 14000 rpm for 1 minute and re-suspended in dH_20 to a concentration of 1 µg/µl as directed in supplier's data sheet. The primers were than diluted (1:10) to make up a working concentration of 100 ng/µl.

2.2.3.2 PCR

The PCR reaction was prepared according to the KOD Hot Start Master Mix protocol. For PCR mixture, 25 ng template DNA, 0.3 µM forward and reverse primer and 25 µl 2× KOD Hot Start Master Mix were added to the PCR tube, made up to a total volume of 50 µl with DNase-free water and briefly centrifuged. The PCR reaction was carried out using the GeneAmp PCR system 2400 machine (Perkin Elmer, USA) under conditions as in table 2.8. The annealing temperatures (T_m) of the forward and reverse primers were calculated using the following formula and the lowest T_m of the pair of primers was usually used.

$$T_m = [4 \times \text{the number of C + G}] + [2 \times \text{the number of A + T}]$$

(C, G, A and T: number of C or G or A or T residues in the primer)

Table 2.8: PCR cycling conditions.

Step	Condition
1. Polymerase activation	95˚C for 2 minutes
2. Denature	95˚C for 20 seconds
3. Annealing	Lowest Primer T_m˚C for 10 seconds
4. Extension	70˚C for 10 seconds/kbp
Repeat steps 2-4	20-40 cycles

Following amplification, 1 µl of the PCR product was analysed by DNA gel electrophoresis to check for the presence of amplified DNA at the correct MW. The remaining product was then stored at -20°C until used.

2.2.3.3 Reverse transcription-PCR

2.2.3.3.1 Total RNA extraction

Total RNA extraction was performed by using a combination of RNA extraction from TRIzol® reagent homogenates (Chomczynski and Sacchi, 1987) combined with a column-based clean-up using the Qiagen RNeasy® plus mini kit. First, cells were rinsed twice with PBS and 1 ml of TRIzol® reagent was added into the dish. The cells were scraped using a cell lifter and the lysate was pipetted up and down several times for homogenisation. The lysate was then transferred to a sterile 1.5ml eppendorf tube and incubated at room temperature for 5 minutes. After incubation, 200 µl of chloroform was added to the tube, vortexed vigorously for 30 seconds and incubated for a further 3 minutes at room temperature. Next, the mixture was centrifuged at 13000 rpm at 4°C for 15 minutes and the upper aqueous phase containing the RNA was removed to a new tube. An equal volume of absolute ethanol was added to the RNA and mixed well by pipetting. The mixture was transferred to an RNeasy spin column and centrifuged at 10000 rpm for 1 minute. The flow-through in the collection tube was discarded and 700 µl of buffer RW1 was added to the column. The column was centrifuged at 10000 rpm for 1 minute and the flow-through was discarded. Then, 500 µl of buffer RPE was added and the column was again centrifuged at 10000 rpm for 1 minute. The flow-through was discarded and the washing step was repeated by the addition of 500 µl of

buffer RPE and centrifugation at 10000 rpm for 2 min. The collection tube was emptied and the column was re-centrifuged for 1 minute. Finally, the RNA was eluted in 50 µl of RNase-free water. The RNA was quantified using the NanoDrop ND-1000 Spectrophotometer (Thermoscientific, UK) and analysed by the formaldehyde RNA gel electrophoresis. The remaining RNA was stored at -80˚C until further used.

In this study, all RNA samples contained intact ribosomal RNA (rRNA), 28S (upper arrow) and 18S (lower arrow) as assessed by RNA gel electrophoresis (Figure 2.2). The intensities of the 28S bands were higher than that of 18 bands. Suggesting that, good quality total RNA samples have been successfully isolated during this experiment.

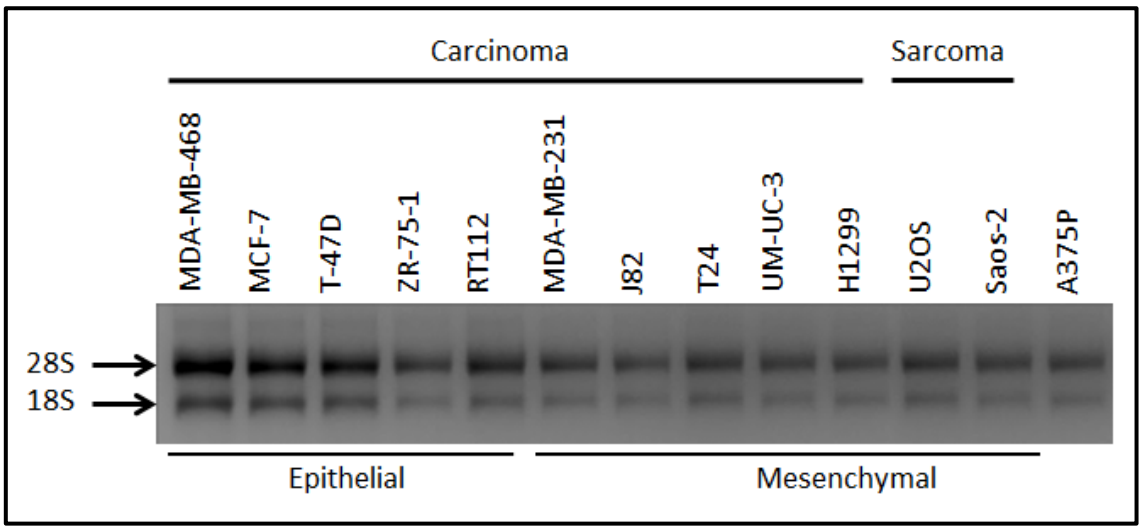


Figure 2.2: The determination of RNA quality. Approximately, 2 µg of the total RNA was analysed by 1% formaldehyde gel to visualise 28S and 18S rRNA; a good indication of RNA quality.

2.2.3.3.2 Generation of cDNA by reverse transcription-PCR

Total RNA was used for cDNA synthesis using the RevertAid^{TM} H Minus First Strand cDNA Synthesis Kit (Fermentas, UK). The standard 20 µl reaction setup was prepared according to the manufacturer's instructions as shown in table 2.9. The reaction mixture was mixed gently and centrifuged briefly for 5 seconds. The reaction was carried out by incubation at 42°C for 1 hour followed by heat inactivation at 70°C for 5 minutes. The cDNA was then used for PCR amplification.

Table 2.9: The first strand cDNA synthesis components.

Component	Amount (µl)
5× Reaction Buffer	4
Template RNA (1 µg)	1-5
Oligo $(dT)_{18}$ primer	1
RiboLock™ RNase Inhibitor (20 u/µl)	1
dNTP Mix (10 mM)	2
Revertaid™ H Minus M-MuLV Reverse Transcriptase (200 u/µl)	1
dH_2O	up to 20
Total reaction volume	**20**

2.2.4 Mutagenesis

2.2.4.1 Site-directed mutagenesis

The QuikChange II Site-Directed Mutagenesis Kit was used to generate mutations within the DNA sequence. For the synthesis reaction, 25 ng of template DNA, 125 ng forward and reverse primer, 5 µl 10× reaction buffer, 1 µl dNTP mix and 1 µl of *Pfu Ultra* HF DNA polymerase (2.5 U/µl) were added to a PCR tube, made up to total volume of 50 µl with DNase-free water and briefly centrifuged. The DNA was then amplified using the cycling parameters as in table 2.10.

Table 2.10: The site-directed mutagenesis cycling conditions.

Segment	Cycles	Temperature (˚C)	Time
1	1	95	30 seconds
2	18	95	30 second
		55	1 minute
		68	1 minute/kb of plasmid length

After amplification, 1 µl of *Dpn I* restriction enzyme (10U/µl) was gently mixed with the product and incubated at 37˚C for 1 hour to digest hemi/methylated-DNA (i.e. parental template). After digestion, the product was transformed into *E.Coli*, plated on selective-agar plate and incubated at 37°C overnight. The next day, 8 colonies were picked off and grown in 2 ml selective media overnight. The plasmid DNA was then isolated using a plasmid DNA mini-prep kit followed by sequencing to check for the desired mutations.

2.2.4.2 Multi-site directed mutagenesis

The QuikChange Lightning Multi-Site Directed Mutagenesis Kit was used to generate several mutations simultaneously within a DNA sequence. For this, 50 ng of template DNA, 100 ng forward and reverse primer, 2.5 µl 10× QuickChange Lightning Multi reaction buffer, 0.75 µl QuickSolution, 1 µl dNTP mix and 1 µl QuickChange Lightning Multi enzyme were added to a PCR tube, made up to total volume of 25 µl with DNase-free water and briefly centrifuged. The DNA was amplified using the cycling parameters as in table 2.11.

Table 2.11: The multi-site directed mutagenesis cycling conditions.

Segment	Cycles	Temperature (˚C)	Time
1	1	95	2 minutes
2	30	95	20 second
		55	30 seconds
		65	30 seconds/kb of plasmid length
3	1	65	5 minutes

Following amplification, 1 µl of *Dpn I* restriction enzyme was added to the amplification product and the reaction was carried out at 37˚C for 5 minutes. After reaction, the product was then used for bacterial transformation, plasmid DNA isolation and sequencing analysis.

2.2.5 Cloning

2.2.5.1 Restriction digestion of plasmid DNA

Restriction digestion was performed in order to use for insertion into expression vectors or to confirm the presence of the correct insert following isolation of plasmid DNA. For digestion reaction, 100 ng of template DNA, 2 µl 10× Digestion Buffer, 2 µl 10 × BSA, 1 µl restriction enzyme 1 (for double digestion, 1 µl of the second restriction enzyme was added) were mixed to eppendorf tube and top up to total volume of 20 µl with DNase-free water. The reaction was allowed to proceed at 37^{O}C for 1 hour. The product was then visualised on by agarose gel electrophoresis.

2.2.5.2 DNA purification

The digested DNA was purified using the Wizard® SV Gel and PCR Clean-Up System by following the manufacturer's protocol. An equal volume of membrane binding solution was added to the DNA product and transferred to the SV mini-column followed by incubation at

room temperature for 1 minute. After incubation, the SV mini-column was centrifuged at 14000 rpm for 1 minute and the flow-through was discarded. Next, 700 µl of membrane wash solution was added to the SV mini-column. The SV mini-column was centrifuged at 14000 rpm for 1 minute and the flow-through in the collection tube was emptied. The washing step was repeated by the addition of 500 µl of membrane wash solution and centrifugation at 14000 rpm for 5 minutes. The collection tube was again emptied and the SV mini-column was re-centrifuged for 1 minute. The SV mini-column was transferred to 1.5 ml centrifuge tube and 50 µl of Nuclease-Free Water was added to the middle of the column followed by incubation at room temperature for 1 minute. Finally, the SV mini-column was centrifuged at 14000 rpm for 1 minute to elute the DNA. To confirm successful elution and estimate the relative amounts of DNA recovered, 1 µl of the eluted DNA was analysed by on DNA gel electrophoresis. The eluted DNA was then used for ligation reaction.

2.2.5.3 Ligation of the DNA fragment and plasmid vector

The ligation reaction was setup as described in table 2.12 and incubated at room temperature for 1 hour. After ligation, 5 µl of the DNA product was analysed by DNA gel electrophoresis to assess the ligation efficiency. If successful, 50 ng of the DNA was transformed into *E.Coli.*

Table 2.12: The ligation components.

Component	Volume (µl)
5× ligation reaction buffer	4
Vector (50 ng)	1-5
Insert (25 ng)	1-5
T4 DNA ligase (5U/µl)	0.2 µl (0.1 unit)
dH_2O	up to 20
Total reaction volume	**20 µl**

2.2.5.4 Transformation of plasmid DNA into E.Coli.

The α-select chemically competent cells were thawed on ice. Approximately 50 ng of the ligation product was added to 50 µl of competent cells and incubated on ice for 30 minutes. After incubation, the reaction was heat-shocked for 30 seconds at 42˚C and chilled for 5 minutes on ice and diluted to 1 ml by addition of SOC media. The mixture was incubated at

37°C in a rotary shaking incubator at 450 rpm for 1 hour. After incubation, 100 µl of the transformation mixture was plated out on LB-agar (32 g of LB agar (Lennox L agar) per 1L of dH_2O) plates containing an appropriate antibiotic (either 50 µg/ml Ampicillin or 30 µg/ml Kanamycin) and incubated overnight at 37˚C.

2.2.5.5 Small scale isolation of plasmid DNA

Bacterial overnight cultures were prepared by picking a single bacterial colony from a LB-agar plate and inoculating 2 ml of LB broth (25 g of Luria Broth base (Miller's LB Broth Base) per 1L of dH_2O) containing an appropriate antibiotic (either 50 µg/ml Ampicillin or 30 µg/ml Kanamycin). The Macherey-Nagel mini-prep kit was used for plasmid isolation following the manufacturer's protocol. About 2 ml of the overnight cultures were centrifuged at 13000 rpm for 1 minute. The supernatant was discarded and the pellet was re-suspended in 250 µl of re-suspension buffer (buffer A1) with RNase A. Next, 250 µl of Lysis buffer (buffer A2) was added and the lysate was incubated at room temperature for 5 minutes followed by addition of 500 µl of Neutralization buffer (buffer A3). The lysate was then centrifuged at 13000 rpm for 10 minutes and the supernatant was loaded on the provided column. The column was centrifuged at 13000 rpm for 1 minute and the flow-through was discarded. Then, 600 µl of Wash buffer (buffer A4) was added and the column was again centrifuged at 13000 rpm for 1 minute. The flow-through was discarded and the column was re-centrifuged for 2 minutes. Next, the plasmid DNA was eluted in 50 µl of Elution buffer (buffer AE) (5 mM Tris pH 8.5). In order to confirm the presence of the correct plasmid during isolation, the plasmid was digested with an appropriate restriction enzyme. Next, if the desired DNA fragment is present, the DNA was further confirmed by sequencing analysis. For the mutagenesis, 30 ng of eluted DNA was directly sent for sequencing analysis to confirm the presence of the introduced mutation. The remaining DNA was stored at -20˚C until further needed.

2.2.5.6 Large scale isolation of plasmid DNA

The plasmid DNA that confirmed to contain the right DNA fragment or mutation was selected and re-transformed into *E.Coli*. The plasmid DNA was then isolated using the Macherey-Nagel maxi-prep kit by following the manufacturer's protocol. First, a starter culture was prepared by inoculating 2 ml of LB-broth containing either 50 µg/ml Ampicillin

or 30 µg/ml Kanamycin with a single colony picked from a LB-agar plate. The culture was grown for 8 hours in an incubator shaker at 37˚C and 300 rpm. Next, 300 ml of LB-broth containing an appropriate selection antibiotic as mentioned above was inoculated with 300 µl of prepared starter culture. The bacterial culture was grown overnight at 37˚C and 300 rpm in an incubator shaker. The next day, the overnight culture was pelleted by centrifugation for 20 minutes at 4˚C at 4000 rpm. The supernatant was discarded and the pellet was re-suspended in 12 ml of Re-suspension buffer (buffer RES) with RNase A. Then, 12 ml of Lysis buffer (buffer LYS) was added followed by incubation at room temperature for 5 minutes. In the meantime, 25 ml of Equilibration buffer (buffer EQ) was applied to the column filter. After incubation, 12 ml of Neutralization buffer (buffer NEU) was added to the lysate followed by inverting the tube 15 times until a homogeneous suspension was obtained. The lysate was then applied to the equilibrated column filter and allowed to empty. The column filter was washed with 15 ml of Equilibration buffer (buffer EQU). Next, the column filter was discarded and the column was washed with 25 ml of Wash buffer (buffer WASH). Plasmid DNA was eluted in 15 ml of Elution buffer (buffer ELU). To precipitate DNA, 10.5 ml of isopropanol was added to the eluted plasmid DNA. The mixture was vortexed and incubated for 2 minutes. The precipitation mixture was then passed through the NucleoBond® Finalizer membrane using the provided syringe. Finally, after a washing step with 5 ml of 70% ethanol and drying of the membrane, the plasmid DNA was eluted in 400 µl of Elution buffer (buffer TRIS). The plasmid DNA was then quantified and stored at -20˚C.

2.2.6 DNA or RNA quantification

DNA or RNA samples were quantified using the NanoDrop ND-1000 Spectrophotometer using the nucleic acid, DNA-50 or RNA-40 program respectively. First, the program was initialized with dH_2O and the blank reading was performed using the buffer used to elute the DNA or RNA during extraction. Next, 1.5 µl of DNA or RNA sample was pipetted onto pedestal and the purity of the DNA or RNA was analysed by monitoring the 260/280 reading.

2.2.7 Agarose gel electrophoresis

2.2.7.1 DNA gel electrophoresis

For analysis of plasmid DNA and DNA fragments, Tris-Acetate (TAE) gel electrophoresis was used. 100 ml of 1×TAE buffer (4.8 g/L Tris base, 1 mM EDTA, 1% v/v acetic acid) was mixed with 0.75-2.0 g agarose depending on the size of the DNA to be analysed, boiled in a microwave until dissolved and allowed to cool. Next, the EtBr (1 µg/ml) was added and gel was poured into gel electrophoresis apparatus. The gel was allowed to set. The DNA sample was mixed with 2 µl of 5× DNA loading buffer (0.1% (w/v) bromophenol blue; 50% glycerol (v/v), 100mM EDTA) and loaded onto the gel. The 1 kb or 100 bp DNA marker was also loaded onto the gel. The 1×TAE buffer was used as the running buffer. The gel was run for 30 minute at 100 V. Next, gel was visualised using UV-transilluminator (UVP BioDoc-H System, USA) and images captured using a Sony CCD Chip camera (Sony UK Ltd., UK).

2.2.7.2 RNA gel electrophoresis

To prepare the gel, 74 ml of dH_2O and 1 g of agarose was boiled and allowed to cool. Next, 16 ml of 37% formaldehyde and 10 ml of 10× MOPS (3-(N-morpholino) propanesulfonic acid) buffer was added and the gel was poured into gel electrophoresis apparatus. The gel was allowed to set. About 2 µg of RNA sample was mixed with 5 µl of 1× RNA loading buffer (50 % v/v formamide, 20 % v/v formaldehyde, 10 % v/v 10x MOPS buffer, 20% v/v glycerol dye) and loaded onto the gel. The 1×MOPS buffer was used as the running buffer. The gel was run for 40 minutes at 60 V and the band was visualised and imaged.

2.2.8 Sequencing reaction and alignment

The primers used for sequencing analysis were provided by GATC Biotech AG Company (Website: www.gatc-biotech.com). The reaction of 20 µl containing 30 ng/µl of DNA was prepared. The reaction samples were then sent to the GATC Biotech AG Company for sequencing analysis. The sequencing data obtained was then analysed on EMBOSS Pairwise Alignment Algorithm program from EMBL-EBI database (Website: http://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html).

2.2.9 Generation of ZEB1 and ZEB2 expression plasmids

2.2.9.1 Generation of the pZ/EG-HA-ZEB1 and pZ/EG-HA-ZEB2 expression plasmids

2.2.9.1.1 Restriction digestions of pcDNA™3.1/TOPO®-HA-ZEB1, pcDNA™3.1/TOPO®-HA-ZEB2 and pZ/EG

To excise the HA-tagged ZEB1 and HA-tagged ZEB2 gene fragments from pcDNA™3.1/TOPO®-HA-ZEB1 and pcDNA™3.1/TOPO®-HA-ZEB2, we used restriction enzymes which produced compatible sticky end within the pZ/EG expression vector as described in table 2.13.

Plasmid/vector name	Restriction enzymes	Expected bands (kb)
pcDNA™3.1/TOPO®-HA-ZEB2	*Bglll* and *Sall*	3.6, 2.3, 2.2 and 0.9
pcDNA™3.1/TOPO®-HA-ZEB1	*Bglll* and *Xhol*	4.6, 3.4 and 0.9
pZ/EG	*Bglll* and *Xhol*	11.0

Table 2.13: The restriction enzymes used for double digestion of pcDNA™3.1/TOPO®-HA-ZEB2, pcDNA™3.1/TOPO®-HA-ZEB1 and pZ/EG. The expected bands sizes were analysed and confirmed by Nebcutter version 2.0 online tools (New England Biolabs, UK). Website: (http://www.neb.uk.com/tools/index.aspx?req=nebcutter). kb:- kilo-base.

After double digestion (section 2.2.5.1), the restriction patterns were analysed by agarose gel electrophoresis. HA-tagged ZEB2 DNA was excised from the pcDNA™3.1/TOPO®-HA-ZEB2 (3.6 kb) when treated with *Bglll* and *Sall* and the digestion of pcDNA™3.1/TOPO®-HA-ZEB1 with *Bglll* and *Xhol* produced HA-tagged ZEB1 fragment of 3.4 kb. Additionally, complete digestion of pZ/EG vector with *Bglll* and *Xhol* produced a single band of 11.0 kb.

After digestion, all products were purified using the Wizard®SV Gel and PCR Clean-Up System as described in section 2.2.5.2. The prepared ZEB1 and ZEB2 insert and pZ/EG vector was then ligated (section 2.2.5.3). Chemically competent *E.coli* was then transformed with the ligation products, plated onto ampicillin-containing (50 µg/ml) agar plates and incubated overnight at 37^{o}C. Single colonies were then transferred into 2 ml LB broth with 50 µg/ml ampicillin and incubated overnight at 37^{o}C.

2.2.9.1.2 Confirmation of the correct insert for the pZ/EG-HA-ZEB1 and pZ/EG-HA-ZEB2

The plasmid DNA was isolated from the overnight cultures using the Macherey-Nagel mini-prep kit. The DNA samples were then digested with *BamHI* and analysed in 1% agarose gel. Concentration of the right pZ/EG-HA-ZEB2 and pZ/EG-HA-ZEB1 plasmids were calculated and adjusted to 30 ng/µl for sequencing. These plasmids were sequenced by GATC Biotech Company using the ZEB1-FP or ZEB2-FP primers as in table 2.4 (Section 2.1.4).

2.2.9.1.3 Sequencing analysis of pZ/EG-HA-ZEB1 and pZ/EG-HA-ZEB2

Raw sequencing data of pZ/EG-HA-ZEB1 and pZ/EG-HA-ZEB2 were compared with our predicted ZEB1 and ZEB2 sequence (Appendix 3A and 3B). All clones contained full length ZEB1 and ZEB2 ORF sequences without any mutations, along with the correct sequences for N-terminal HA tag, START and STOP codons. In addition, both of pZ/EG-HA-ZEB1 and pZ/EG-HA-ZEB2 plasmids contained identical Kozak elements.

2.2.9.2 Generation of pEGFP-C1-ZEB2 (372-437aa)-HA and pEGFP-C1-ZEB1 (317-372aa)-HA expression plasmids

2.2.9.2.1 Amplification of ZEB2 (372-437aa) and ZEB1 (317-372aa) genes from the pZ/EG-HA-ZEB2 and pZ/EG-HA-ZEB1 plasmids

ZEB2 (372-437aa) and ZEB1 (317-372aa) genes were amplified from the pZ/EG-HA-ZEB2 and pZ/EG-HA-ZEB1 respectively by using the forward and reverse primers as detailed in table 2.4 (Section 2.1.4).The reversed primers contained also a sequence for HA tag. In parallel, the pEGFP-C1 vector was digested with *Bglll* and *Kpnl* restriction enzymes, using the NEBuffer 2 + BSA as the sole reaction buffer. All products were then purified and 1 µl of the products were subjected to 1% agarose gel electrophoresis. After purification, the ligation reactions of the pEGFP-C1 vector with the ZEB2 (372-437aa) and ZEB1 (317-372aa) were carried out. α-select chemically competent *E.coli* cells were then transformed and plasmid DNA was isolated from individual colonies.

2.2.9.2.2 Confirmation of the correct insert for the pEGFP-C1-ZEB2 (372-437aa)-HA and pEGFP-C1-ZEB1 (317-372aa)-HA

The plasmid DNA samples were isolated from overnight bacterial cultures and analysed by digesting with *NdeI* and *AgeI* enzymes. After digestion, the products were subjected to 1% agarose gel. The correct pEGFP-C1-ZEB2 (372-437aa)-HA and pEGFP-C1-ZEB1 (317-372aa) plasmids were sequenced using the pEGFP-C2-FP primer as detailed in table 2.4 (Section 2.1.4).

2.2.9.2.3 Sequencing analysis of pEGFP-C1-ZEB2 (372-437aa)-HA and pEGFP-C1-ZEB1 (317-372aa)-HA

Raw sequencing data of pEGFP-C1-ZEB2 (372-437aa)-HA and pEGFP-C1-ZEB1 (317-372aa)-HA were compared with the predicted sequences. All selected clones contained C-terminally HA-tagged gene fragments corresponding to the ZEB2 (372-437aa) and ZEB1 (317-372aa) regions without any mutations in frame with EGFP (Appendix 4A and 4B). Both plasmids also contain identical Kozak. After sequencing, a large scale purification of plasmid DNA was then carried out. Plasmid DNA isolated were quantified and used for transfections. These expression plasmids are subsequently referred to as EGFP-ZEB2 (372-437aa) and EGFP-ZEB1 (317-372aa).

2.2.10 Generation of ZEB1 and ZEB2 mutant plasmids

2.2.10.1 Mapping of EGFP-ZEB2 (372-437aa) domain by STOP codons insertion

To map a domain within EGFP-ZEB2 (372-437aa), STOP codons (TAG or TAA) were introduced into the EGFP-ZEB2 (372-437aa) expression plasmid (Figure 2.3) by using the QuikChange II Site-Directed Mutagenesis Kit (section 2.2.4.1) and primer sets (table 2.4; section 2.1.4). After amplification, 10 µl of the products were subjected on 1 % agarose gel electrophoresis. In parallel, the products were also digested with *Dpn I* restriction enzyme to remove the non-mutated template DNA. After digestion, *E. coli* competent cells were transformed, and individual colonies were isolated and digested with *NdeI* and *AgeI* enzymes. All selected plasmids were then verified by sequencing using the pEGFP-C2-FP primer.

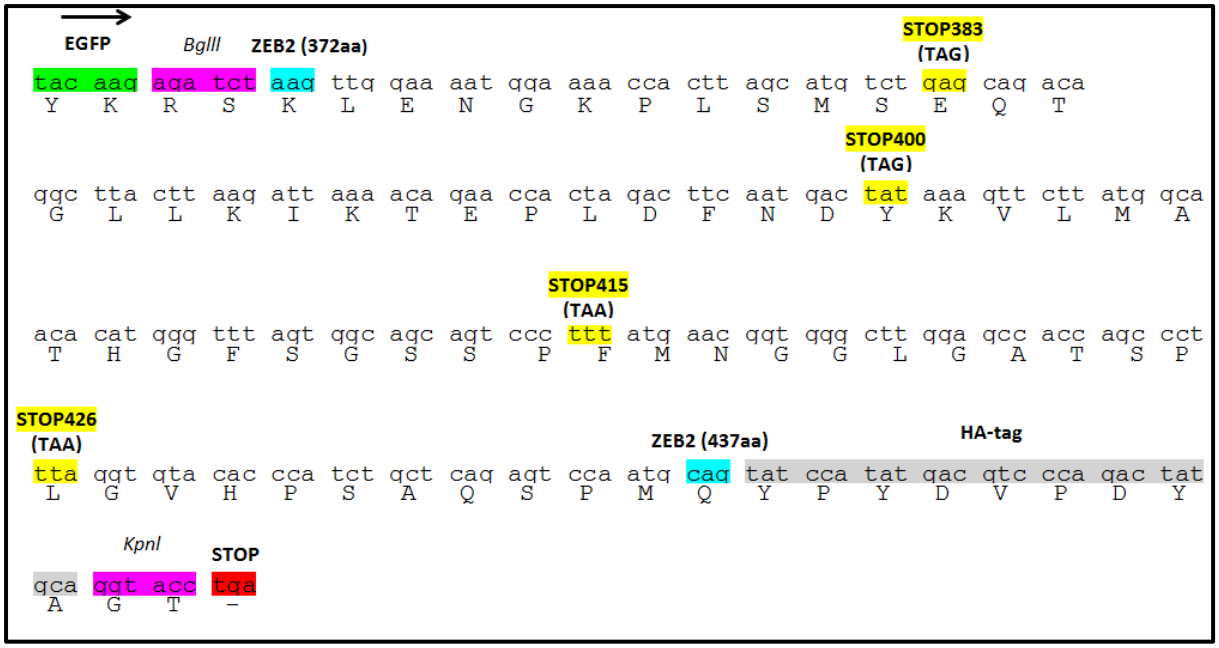


Figure 2.3: An illustration showing locations of STOP codons inserted at the position of 383aa, 400aa, 415aa and 426aa within EGFP-ZEB2 (372-437aa). These plasmids are subsequently named as EGFP-ZEB2 (372-437aa)-STOP383, EGFP-ZEB2 (372-437aa)-STOP400, EGFP-ZEB2 (372-437aa)-STOP415 and EGFP-ZEB2 (372-437aa)-STOP426.

Raw sequencing data of EGFP-ZEB2 (372-437aa)-STOP383, EGFP-ZEB2 (372-437aa)-STOP400, EGFP-ZEB2 (372-437aa)-STOP415 and EGFP-ZEB2 (372-437aa)-STOP426 were analysed and aligned with the predicted EGFP-ZEB2 (372-437aa) sequence. All plasmids contained STOP insertion at the right location (red circle) and no additional mutations (Appendix 5A, 5B, 5C and 5D). After sequencing, a large scale purification of plasmid DNA was carried out and used for transfection.

2.2.10.2 Mapping of the ZEB2 expression-limiting activity by START codon insertions

To map the region within ZEB2 (372-437aa)-EGFP, START (ATG) codons were introduced at the positions shown in figure 2.4. For this purpose, a DNA fragment corresponding to the ZEB2 (372-437aa region) was amplified using a set of forward primers containing START codon insertions and a reverse primer (for primer information see table 2.4; section 2.1.4). In parallel, the pEGFP-N2 vector was digested with *BglII* and *KpnI* enzymes. Next, all products were purified and analysed in 1% agarose gel electrophoresis. After purification, the ligation reactions of pEGFP-N2 vector with ZEB2 (372-437aa)-ATG400, ZEB2 (372-437aa)-ATG415 and ZEB2 (372-437aa)-ATG426 were carried out. After transformation of

competent cells, several individual colonies were selected, digested with *NgoMIV* enzyme and analysed in 1% agarose gel electrophoresis. All plasmids which give correct fragments when digested with *NgoMIV* enzyme were purified and verified by sequencing analysis using the pEGFP-FP primer as detailed in table 2.4; section 2.1.4.

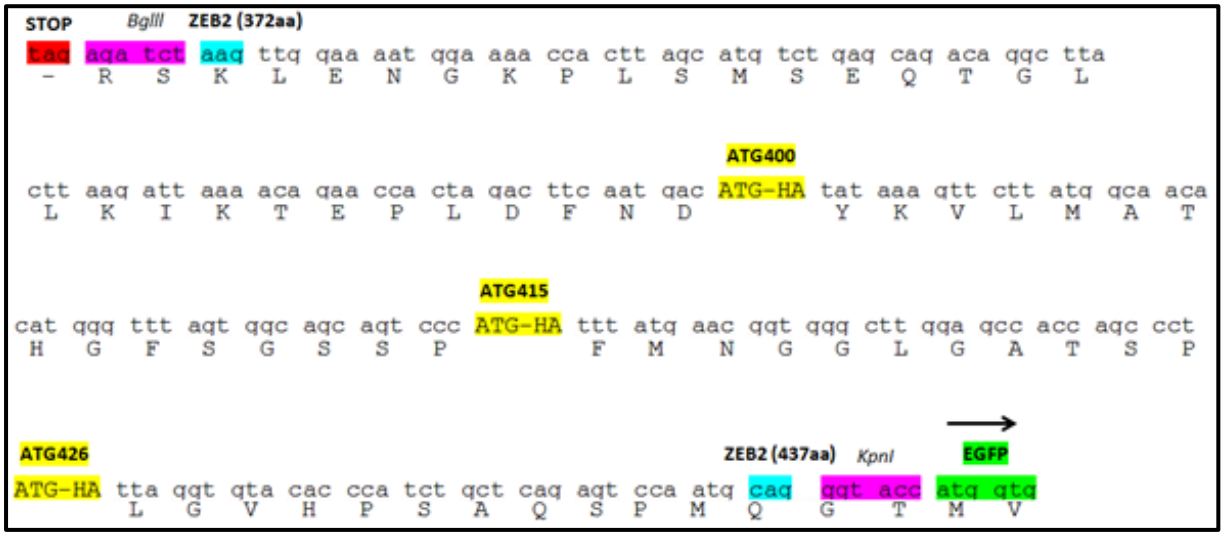


Figure 2.4: A scheme representing locations of ATG codons introduced at the position of 400aa, 415aa and 426aa within ZEB2 (372-437aa)-EGFP.

All raw sequencing data were analysed and compared with the DNA sequence corresponding to ZEB2 (372-437aa)-EGFP sequence. All plasmids contained ATG insertion at the right locations (red circle) (Appendix 6A, 6B and 6C). After sequencing, a large scale purification of plasmid DNA was carried out and used for transfection.

2.2.10.3 Changing of RC into common codon sequence (CC) within EGFP-ZEB2 (372-437aa) plasmid

We introduced the RC to CC mutation within EGFP-ZEB2 (372-437aa) plasmid by using the Quikchange II site-Directed mutagenesis kit and primer sets containing the codon substitutions as detailed in table 2.4 (Section 2.1.4). After amplification, 10 µl of the products were subjected to 1 % agarose gel electrophoresis. The products were then digested with *Dpn I* restriction enzyme to remove the non-mutated template DNA. After cloning procedures were completed, the plasmid DNA samples were isolated from the overnight bacterial cultures, quantified and sequenced by GATC Biotech Company using the pEGFP-C2-FP primer as in table 2.4 (Section 2.1.4). Raw sequencing data of EGFP-ZEB2 (372-

437aa)-(RC-CC) were compared with the wild type EGFP-ZEB2 (372-437aa) plasmid DNA sequence. As shown in appendix 7, all clones contained of RC (TTA GGT GTA) to CC (CTG GGC GTG) substitution (red circle). A large scale purification of plasmid DNA was carried out. Plasmid DNA isolated were then quantified and used for transfection.

2.2.10.4 Changing of RC into CC sequence within EGFP-ZEB2 plasmid

To introduce RC changes within ZEB2, our first step was to subclone the ZEB2-HA into pEGFP-C1 vector. To this end, the ZEB2-HA gene fragment was amplified from pZ/EG-HA-ZEB2. In parallel, the pEGFP-C1 vector was double digested with *BglII* and *KpnI* restriction enzymes. All products were purified and 1 µl of the DNA products were subjected to 1% agarose gel electrophoresis. After cloning procedures were completed, the plasmids were isolated from the overnight cultures and digested with *AgeI* and *HindIII* enzymes. After digestion, the products were analysed in 1% agarose gel. All selected plasmids were sequenced by GATC Biotech Company using the pEGFP-C2-FP primer as in table 2.4 (Section 2.1.4).

Raw sequencing data of pEGFP-C1-ZEB2-HA were aligned with the predicted pEGFP-C1-ZEB2-HA sequence. As shown in Appendix 8, all plasmids contained full length ZEB2 ORF sequence free of mutations, along with the correctly placed sequence for N-terminal HA tag and Kozak sequence. Plasmid DNA was then quantified and used for transfections. This plasmid is subsequently referred as EGFP-ZEB2.

Next, mutated RC triplets to CC within EGFP-ZEB2 construct by using the same method in section 2.2.10.3 using a primer sets detailed in table 2.4 (Section 2.1.4). After amplification, 10 µl of the products were subjected to 1 % agarose gel. The products were then digested with *Dpn I* enzyme and subjected to cloning. The selected colonies were then analysed by sequecing using ZEB2-FP primer as detailed in table 2.4 (Section 2.1.4).

Raw sequencing data of EGFP-ZEB2-(RC-CC) were aligned with the predicted EGFP-ZEB2 sequence. As shown in figure appendix 9, all clones analysed contained RC to CC mutations (red circles). After sequencing, a large scale purification of plasmid DNA was carried out and the concentration of the isolated DNA was determined.

2.2.10.5 Introducing CC to RC mutations into EGFP-ZEB1 plasmid

To introduce LGV codons clusters within ZEB1 ORF, we firstly generated the pEGFPC1-ZEB1-HA expression plasmid. To this end, the sequence of the full length ZEB1-ORF gene fragment was amplified using pZ/EG-HA-ZEB1 as template. The amplified fragment contained natural ZEB1-derived Kozak sequence followed by HA tag-encoding DNA. In parallel, the pEGFP-C1 vector was digested with *BglII* and *KpnI* enzymes. All products were then purified and 1 µl of the purified DNAs were subjected to 1% agarose gel electrophoresis. After the ligation reaction and subsequent transformation of competent bacteria, we selected several independent colonies for the analysis.

Next, the plasmids were isolated from the bacterial cultures and digested with *AgeI* and *HindIII* enzymes. After digestion, the products were analysed in 1% agarose gel electrophoresis. All samples produced fragments of expected sizes. Generated plasmids were analysed by sequencing using the pEGFP-C2-FP primer as in table 2.4 (Section 2.1.4).

Raw sequencing data of pEGFP-C1-ZEB1-HA were compared with the predicted pEGFP-C1-ZEB1-HA sequence. As shown in Appendix 10, only clone number 2 contained ZEB1 ORF sequence without any mutations, along with the correct Kozak element and sequence for N-terminal HA tag. A large scale purification of plasmid DNA was carried out. Plasmid DNA was then quantified and used for transfections. This plasmid is subsequently referred as EGFP-ZEB1.

After the generation of EGFP-ZEB1, we introduced the codons substitutions within EGFP-ZEB1 by using a similar method as in section 4.27 with primer sets as detailed in table 2.4 (Section 2.1.4). After amplification, 10 µl of the products were subjected to 1 % agarose gel electrophoresis. The products were digested with *Dpn I* enzyme, cloned and six colonies were selected for further analysed by DNA sequencing.

Raw sequencing data for EGFP-ZEB1-(CC-RC) were aligned with the EGFP-ZEB1 sequence. As shown in appendix 11, only one out of six samples contained both CC-RC mutations (red circles). This plasmid referred as EGFP-ZEB1-(CC-RC) was used in transfection experiments.

CHAPTER 3:

ROLE OF ZEB FAMILY MEMBERS IN CANCER CELL LINES

3.1 INTRODUCTION

Although both ZEB1 and ZEB2 proteins are known as master regulators of EMT, their roles in cancer are likely to be different (see Introduction). It is therefore important to understand molecular mechanisms controlling their expression in tumour cells. Previously, it has been shown that the majority of E-cadherin-negative carcinoma cell lines express high levels of ZEB1, but not ZEB2 (Sayan et al., 2009). On the other hand, all these cell lines were reported to be positive for both ZEB1 and ZEB2 mRNA (Park et al., 2008, Gregory et al., 2008a). However, the analyses on the correlation between ZEB protein and mRNA performed in one study remain limited. Therefore, the aim of the work in this chapter was to analyse the expression of ZEB1 and ZEB2 genes on RNA and protein levels in carcinoma cells.

3.2 RESULTS

3.2.1 Expression of ZEB1 and ZEB2 proteins in human cancer cell lines

To investigate correlation between ZEB protein and mRNA, we firstly analysed the endogenous expression of ZEB1 and ZEB2 proteins in several different human cancer cell lines via Western blotting. In this experiment, a human malignant melanoma cell line A375P was used as a control for expression of ZEB1 and ZEB2, as this cell line is known to express both proteins (data from EMT laboratory).

Morphologically, the selected cell lines represented two different categories, epithelial (MDA-MB-468, MCF-7, T-47D, ZR-75-1 and RT112) and mesenchymal (MDA-MB-231, J82, T24, UM-UC-3, H1299, U2OS and SaOS-2). To confirm the differentiation status of these cell lines, Western blot analysis was also carried out for the epithelial marker, E-cadherin. Figure 3.1 shows that, strong expression of E-cadherin staining was observed in all epithelial but not in the mesenchymal ZEB1 positive cell lines. Therefore, our data confirmed a perfect inverse correlation between ZEB1 and E-cadherin (Figure 3.1). In contrast, ZEB2 protein was detected at high level only in sarcoma cell line SaOS-2. Its expression was very low in mesenchymal carcinoma cell line H1299 and almost undetectable in other mesenchymal carcinoma cell lines, such as in MDA-MB-231, J82, T24 and UM-UC-3 (Figure 3.1).

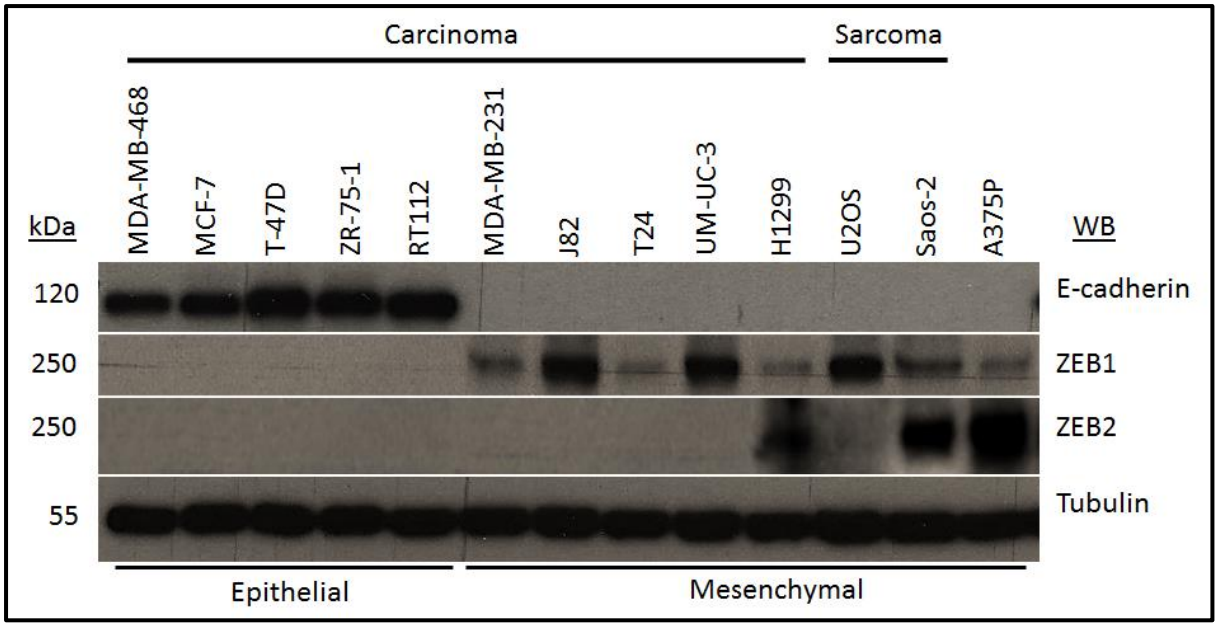


Figure 3.1: Protein expression of E-cadherin, ZEB1 and ZEB2 in a panel of human cancer cell lines. Western blotting was performed using cell lysates from untreated cells cultured to approximately 70 % confluency. The α-Tubulin was used as a protein loading control. WB: Western blot.

3.2.2 *ZEB1* and *ZEB2* gene transcription in cancer cell lines

We next examined *ZEB1* and *ZEB2* mRNA level in the same panel of cell lines. For this purpose, total RNA from MDA-MB-468, MCF-7, T-47D, ZR-75-1, RT112, MDA-MB-231, J82, T24, UM-UC-3, H1299, U2OS, SaOS-2 and A375P cell lines were extracted as described in section 2.2.3.3.1. As in previous experiment, A375P cell line was used as a control for the *ZEB1* and *ZEB2* expression.

The first-strand cDNA synthesis was performed using Revertaid™ H Minus M-MuLV Reverse Transcriptase and Oligo (dT)18 primer as described in section 2.2.3.3.2. After synthesis, approximately 25 ng of cDNA from each cell lines were used as templates in the PCR.

CDH1 was only detected in epithelial cell lines, such as MDA-MB-468, MCF-7, T-47D, ZR-75-1 and RT112. In contrast, *ZEB1* mRNA was only present in mesenchymal cell lines MDA-MB-231, J82, T24, UM-UC-3, H1299, U2OS and SaOS-2 (Figure 3.2). These findings are consistent with our expectations based on protein expression data (Figure 3.1), showing inverse correlation between E-cadherin and ZEB1 protein.

Likewise, *ZEB2* mRNA was present in all mesenchymal carcinoma cells such as MDA-MB-231, J82, T24, UM-UC-3 and H1299. However, at the protein level ZEB2 was barely detectable in

these cell lines (Figure 3.1). Consistent with the western blot data, *ZEB2* mRNA was not detected in all epithelial cell lines and in the mesenchymal sarcoma U2OS cell line.

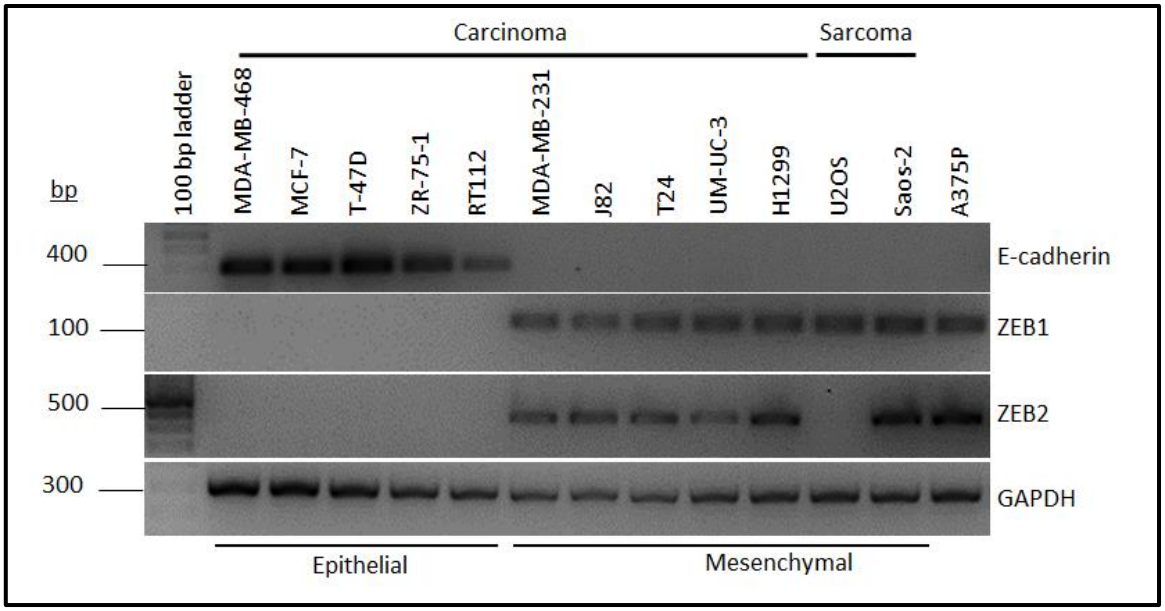
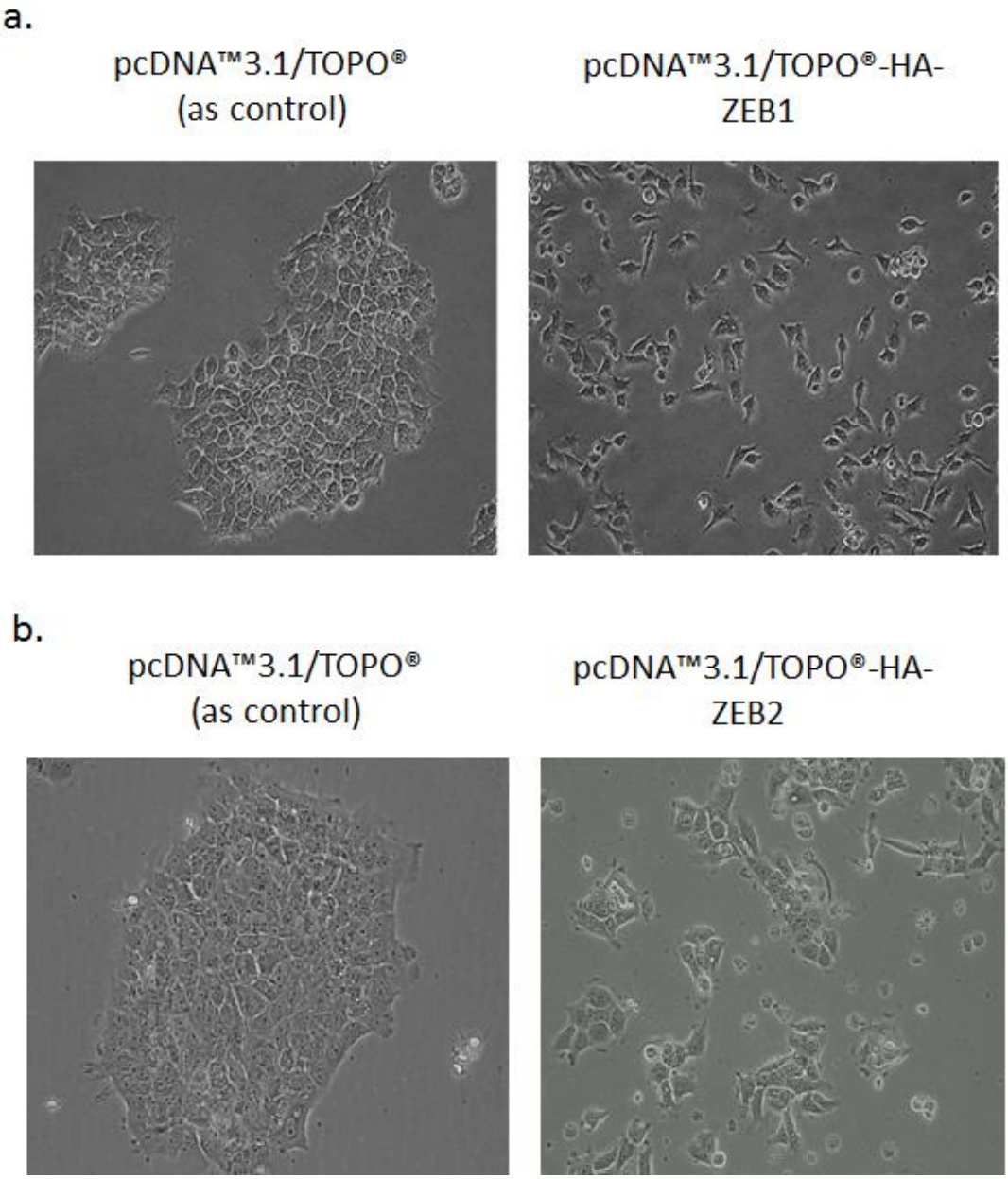


Figure 3.2: The mRNA levels of *CDH1*, *ZEB1* and *ZEB2* analysed by electrophoresis in 1× TAE agarose gel with GAPDH as a control gene. The 100 bp ladder was used as a molecular weight standard.

3.2.3 Expression of ZEB1 and ZEB2 proteins induce an EMT in A431 cells with different efficiencies

To assess the effectiveness of ZEB1 and ZEB2 proteins as EMT inducers, we transiently transfected A431 cells with 4 µg of pcDNA™3.1/TOPO®-HA-ZEB1, pcDNA™3.1/TOPO®-HA-ZEB2 or empty pcDNA™3.1/TOPO® (as a control) by using the electroporation technique, as described in section 2.2.1.5.2. 48 hours post-transfection, cells were harvested and lysates were analysed by Western blot. Figure 3.3a and figure 3.3b demonstrate that both ZEB1 and ZEB2 proteins were capable of inducing morphological alterations in A431 cells. The highly compact, clustered and rounded A431 cells were converted into a scattered, fibroblast-like phenotype, with loss of cell-cell contacts. However, the effect of ZEB1 on cell morphology was more profound (Figure 3.3a). Expectedly, both ZEB1 and ZEB2 factors activated a mesenchymal marker vimentin (Figure 3.3c and figure 3.3d). In contrast, transfection of pcDNA™3.1/TOPO®-HA-ZEB1, but not pcDNA™3.1/TOPO®-HA-ZEB2 in A431 cells decreased E-cadherin levels (Figure 3.3c). Although both ZEB1 and ZEB2 proteins are well-known transcriptional repressors of the E-cadherin gene, their relative ability to repress

transcription has not been compared in the same cell line. The observed differences might indicate that ZEB2 is a weaker transcriptional repressor than ZEB1. Alternatively, ectopic expression of ZEB1 might be higher than that of ZEB2, and concentration of a ZEB protein required for repression of E-cadherin is higher than that needed for the activation of vimentin gene.



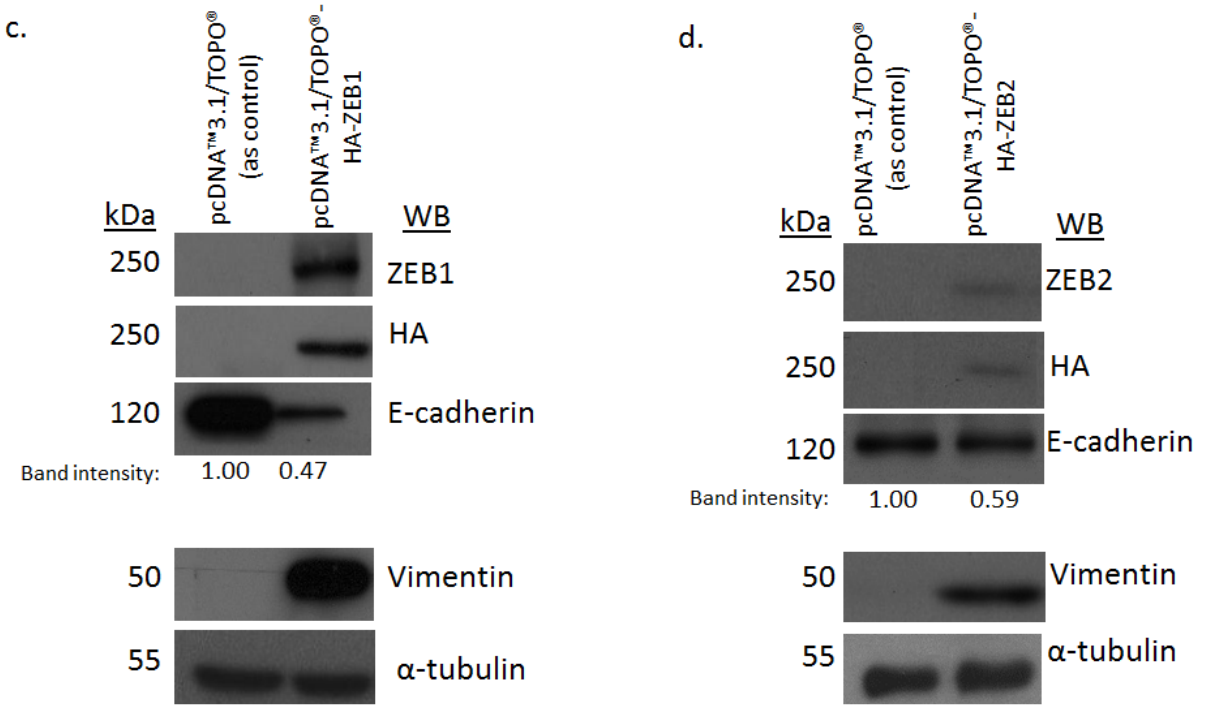


Figure 3.3: ZEB1 and ZEB2 repress E-cadherin and induce EMT in the A431 cells. (a and b) Phase contrast images of A431 cells undergoing an EMT due to expression of ZEB1 and ZEB2. The epithelial morphology of the cell was lost with cellular scattering and formation of a mesenchymal phenotype. (c and d) Western blotting analysis of A431 cells transfected with pcDNA™3.1/TOPO®-HA-ZEB1, pcDNA™3.1/TOPO®-HA-ZEB2 or empty pcDNA™3.1/TOPO® expression vectors. Membranes were probed with anti-ZEB1, anti-ZEB2, anti-HA, anti-E-cadherin or anti-vimentin antibodies. α-Tubulin staining was used as a protein loading control. WB: Western blot. (10× objective). The E-cadherin bands intensity was measured using Image J and normalised to α- tubulin loading control.

3.2.4 Expression of pZ/EG-HA-ZEB1 and pZ/EG-HA-ZEB2 in HEK 293 cell lines

Whereas *ZEB1* gene was expressed at both mRNA and protein levels in mesenchymal cell lines, ZEB2 protein was hardly detectable in all the cell lines of epithelial origin. On the other hand, *ZEB2* mRNA was present in E-cadherin-negative carcinoma cells. Ectopic expression of ZEB1 or ZEB2 in A431 cells induced EMT, but ZEB1 appeared to be a much more efficient EMT inducer. As ZEB2, but not ZEB1 promotes cell cycle arrest and senescence, the difference in their expression levels may result from the elimination of cells expressing endogenous or exogenous ZEB2. Alternatively, molecular mechanisms limiting ZEB2 expression at posttranscriptional level may be responsible.

To discriminate between these two possibilities, we applied the pZ/EG vector to generate constructs (see section 2.2.9.1), in which ZEB1 or ZEB2 are expressed upon removal of a stop cassette by cre recombinase. In these plasmids ZEB proteins are translated from bi-cistronic mRNAs, which encode also EGFP separated from ZEB open reading frames (ORFs) by internal ribosome entry sites (IRES). EGFP represented, therefore, an ideal internal control allowing comparing ZEB1 and ZEB2 expression levels upon transient expression in different cell lines. Both ZEB proteins encoded by pcDNA™3.1/TOPO-based vectors contained HA tags, and the same Kozak element derived from the ZEB1 gene.

To analyse whether ZEB1, ZEB2 and EGFP expression can be activated upon Cre delivery, transient transfection of HEK 293 cells was performed. HEK 293 cells were already used in our laboratory, they are easy to culture and to transfect using JetPRIME™ DNA transfection reagent (section 2.2.1.5.1).

HEK 293 cells (n= 2.6×10^6) were transfected with 2 µg of pZ/EG-HA-ZEB1 or pZ/EG-HA-ZEB2 with or without 2 µg of pCre-Pac. To monitor transfection efficiency, in parallel experiment HEK 293 cells were transfected with 1µg of pmaxGFP®. 24 hours post transfection, the expression of EGFP were assessed by fluorescence microscope and photographed. Appearance of green cells (Figure 3.4) in pCre-transfected cultures showed that the insertions of ZEB1 or ZEB2 sequences had no effect on the function of IRES.

Western blot analysis (Figure 3.5) also confirmed that Cre induced expression of ZEB1, ZEB2 and EGFP. Importantly, immunoblotting with an anti-HA antibody demonstrated higher expression of ZEB1 than ZEB2 in Cre-transfected HEK-293 cells.

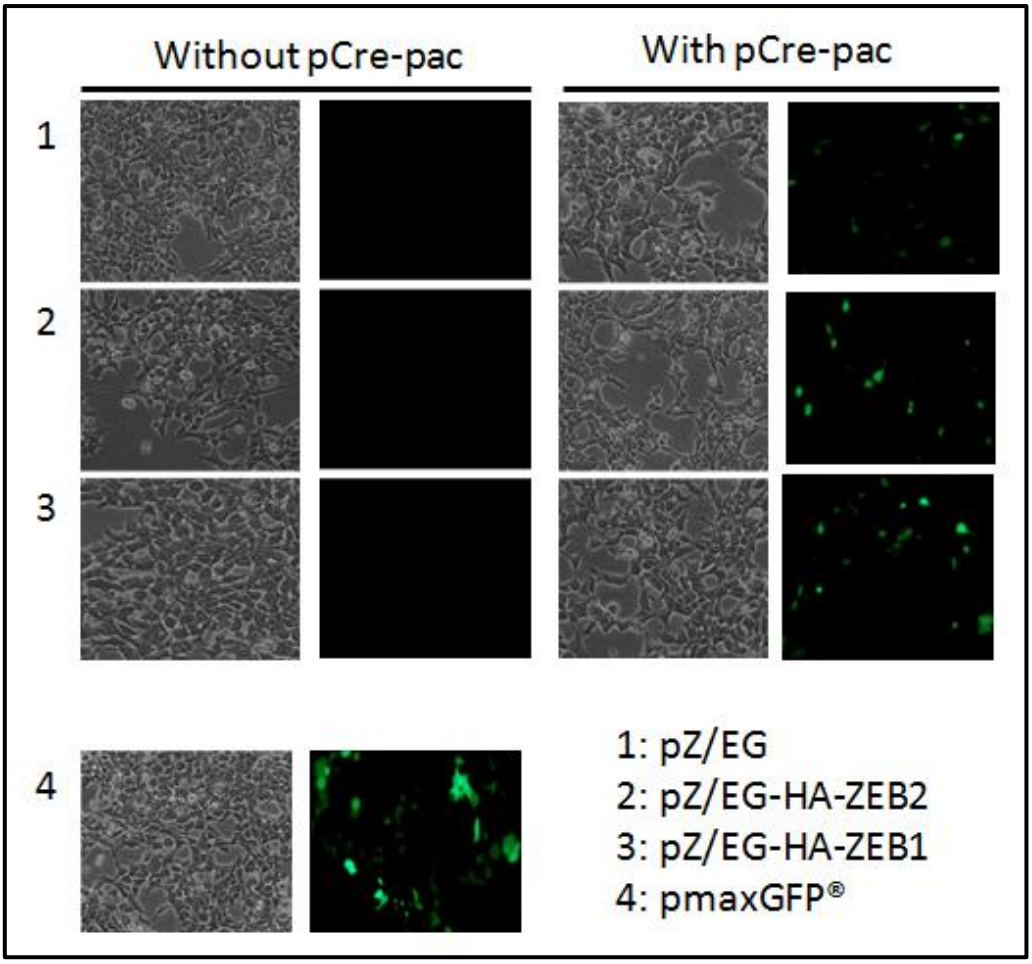


Figure 3.4: The appearance of green fluorescence colour in HEK 293 cells analysed by fluorescence microscope. Cre recombinase induced expression of EGFP in HEK 293 cells transfected with pZ/EG-HA-ZEB1, pZ/EG-HA-ZEB2 or empty pZ/EG vector. (20× objectives).

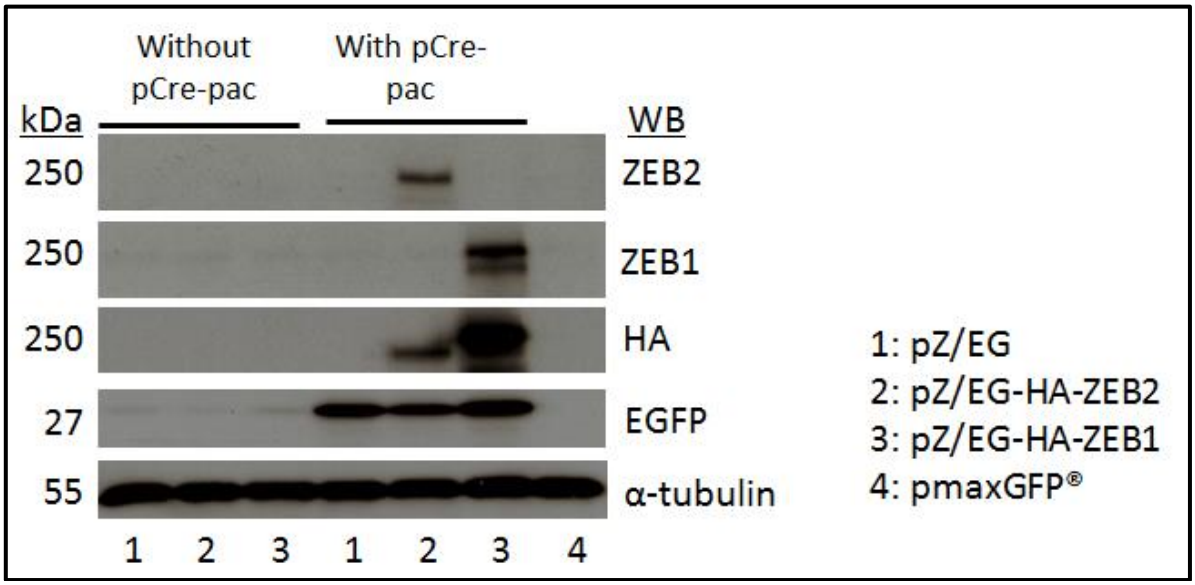


Figure 3.5: Western blot analysis demonstrates the levels of EGFP, ZEB1 and ZEB2 protein expression in HEK 293 cells transfected with empty pZ/EG vector, pZ/EG-HA-ZEB1 or pZ/EG-HA-ZEB2 (with or without pCre-pac). α-Tubulin staining was used as a protein loading control.

3.2.5 Expression of pZ/EG-HA-ZEB1 and pZ/EG-HA-ZEB2 in human cancer cell lines

In order to compare the expression levels of exogenous ZEB1 and ZEB2 in human cancer cell lines, MDA-MB-468, A431, H1299 and SaOS-2 cells were transiently transfected with the pZ/EG-HA-ZEB1, pZ/EG-HA-ZEB2 or an empty pZ/EG vector and protein level was analysed by immunoblotting.

ZEB1 expression was much higher than that of ZEB2 in all cell lines analysed. This was evident from the comparison of the signals obtained with the anti-HA antibody (Figure 3.6). Formally, we cannot exclude that HA tag was specifically cleaved off in transfected cells. However, C-terminally GFP-tagged ZEB1 and ZEB2 proteins demonstrated also unequal expression in transiently transfected carcinoma cell lines (not shown) suggesting that differences presented in Figure 3.6 are intrinsic characteristics of ZEB ORFs. This difference was particularly strong in epithelial carcinoma cell lines MDA-MB-468 and A431.

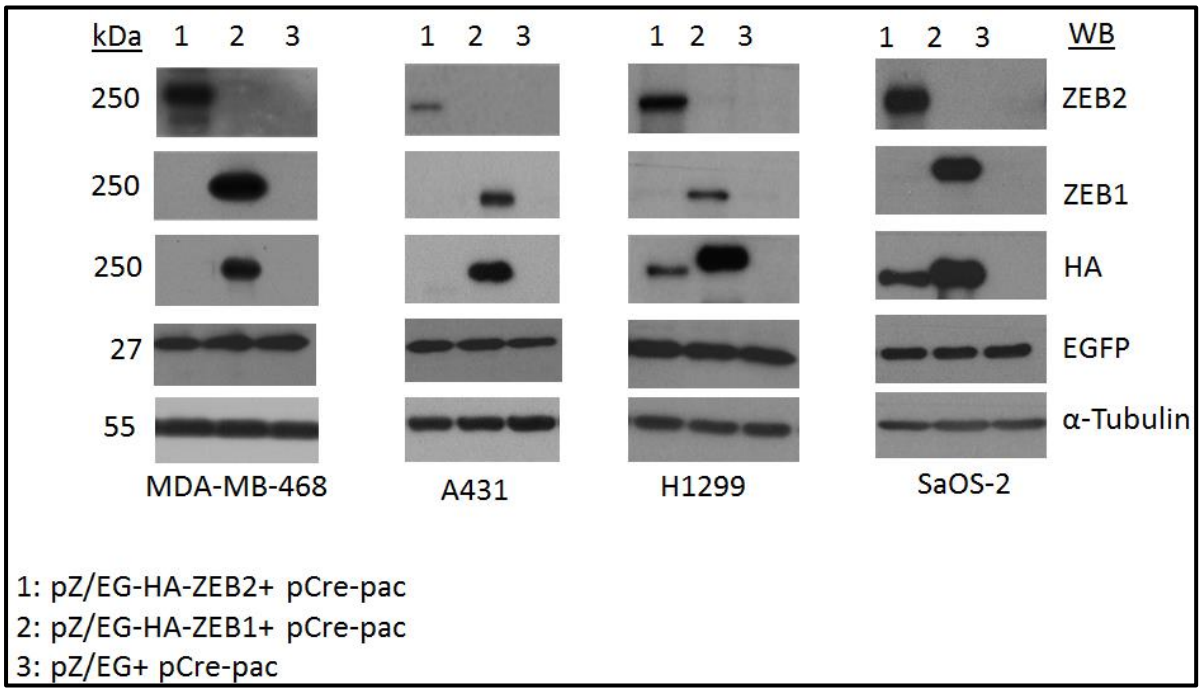


Figure 3.6: Western blot analysis showing expression of ZEB1, ZEB2, HA and EGFP in MDA-MB-468, A431, H1299 and SaOs-2 cell lines after induction of ZEB1 and ZEB2 expression using the Cre/LoxP recombination system. α-Tubulin staining was used as a protein loading control.

3.2.6 Conclusion

The pZ/EG vector contains a CMV-1E promoter that directs the expression of the loxP-flanked lacZ/neomycin resistance fusion gene. Restriction sites for *BglII* and *XhoI* endonucleases are immediately downstream of the second loxP site. Next to *XhoI* site is the IRES and EGFP gene (Figure 3.7). pZ/EG is a convenient vector for the analysis of relative protein abundance in transient transfections, since EGFP can be used to control for the transfection efficiency, as well as to monitor cell survival. Therefore, we cloned *HA-ZEB1* and *HA-ZEB2* into *BglII* and *XhoI* restriction sites within this vector. The restriction digestion using *BamHI* enzyme and the sequencing analysis have shown that, the cloning of HA-ZEB1 and HA-ZEB2 into the pZ/EG vector was successful. The pZ/EG-HA-ZEB1 and pZ/EG-HA-ZEB2 constructs expressed N-terminally HA-tagged open reading frames of ZEB1 and ZEB2 with identical Kozak sequences derived from *ZEB1* gene.

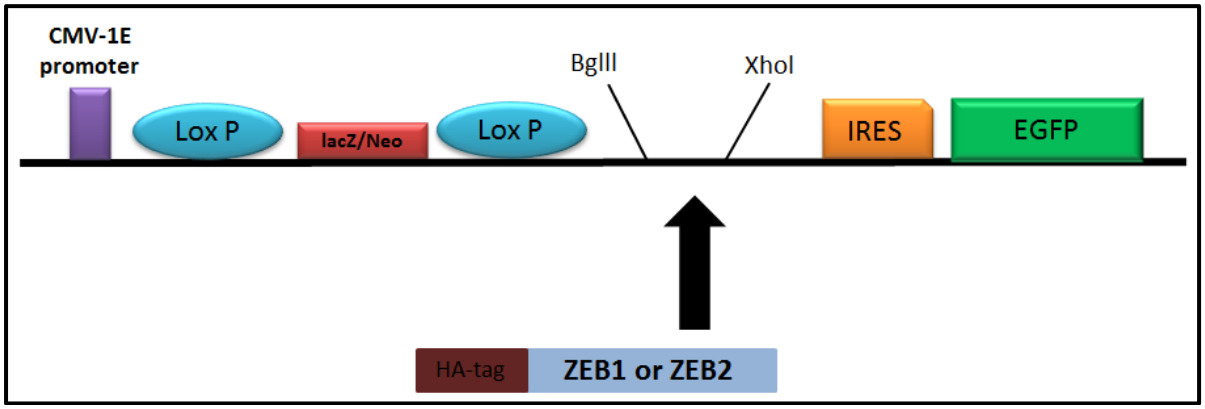


Figure 3.7: The pZ/EG expression vector. The pZ/EG vector contains of the CMV-1E promoter, loxP-flanked lacZ/neo, *BglII* and *XhoI* cloning sites, IRES and the EGFP gene. The HA-ZEB1 and HA-ZEB2 were cloned at the *BglII* and *XhoI* cloning sites.

Transfections experiments have shown that the expression of ZEB1, ZEB2 or EGFP occurs only in the presence of Cre demonstrating that Cre/loxP recombination system is functional in pZ/EG vector (Figure 3.8). Expression levels of EGFP were similar in cells transfected with the recombinant constructs or an empty vector. On the other hand, our data have shown that the expression of ZEB1 was much higher than that of ZEB2 particularly in epithelial carcinoma cell lines. Therefore, we concluded that: i) difference in ZEB protein expression levels in transfection experiments cannot be explained by selective inhibition of cell cycle progression or cell survival by ZEB2; ii) decreased level of ZEB2 is its intrinsic feature likely independent of miR (the inserts in pZ/EG vectors did not contain any 3'UTR sequences).

The existence of a posttranscriptional mechanism specifically regulating ZEB2 expression is in line with ZEB2 expression pattern we observed in a number tumour cell lines. We detected high ZEB2 mRNA levels in mesenchymal cells, but mild expression of ZEB2 gene on the protein level in the majority of the cell lines.

Taken together our data allowed us to propose that the expression of ZEB2 is repressed post-transcriptionally. Further experiments were undertaken to address underlying mechanisms.

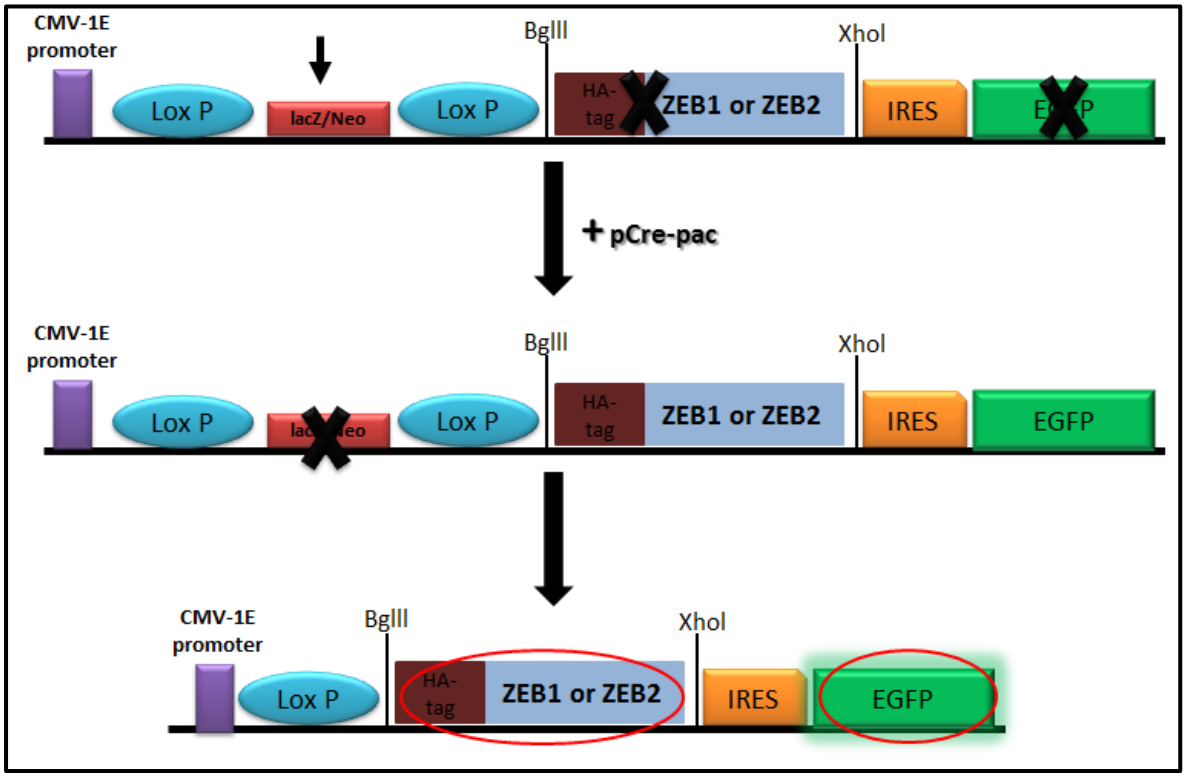


Figure 3.8: The Cre-mediated recombination of the pZ/EG expression vector. In the absence of Cre recombinase, the lacZ/Neo cassette blocks the transcription of the transgene from CMV promoter resulting in silencing of ZEB1, ZEB2 or EGFP expression. However, upon Cre delivery, the lacZ/Neo cassette is excised allowing their transcription (red circle).

CHAPTER 4: POST-TRANSCRIPTIONAL REGULATION OF ZEB2 IN CANCER CELL LINES

4.1 Mapping of ZEB2 ORF by STOP codon insertion

To further investigate mechanisms involved in the regulation of ZEB2 levels, we decided to map a domain within ZEB2 ORF that is involved in negative regulation of ZEB2 expression in cancer cell lines. For this purpose, we mapped STOP codons at nine different locations within the ZEB2 ORF (Figure 4.1a) in the context of pcDNA™3.1/TOPO®-HA-ZEB2 expression vector. Next, we transiently transfected these plasmids (for plasmids detail, see table 2.5; section 2.1.5) into A431 and H1299 cell lines.

After transfection, cells were incubated for 24 hours. The protein lysates were then prepared and analysed by Western blotting with anti-HA or anti-ZEB2 antibodies. In H1299 cells, the expression was increasing gradually with shortening of fragments. In transfected A431 cells, there was a sharp drop in expression following inclusion of the region between amino acids 372 and 437 of the ZEB2 ORF (marked with asterisk (*)), indicating that this region may have a particularly strong effect on ZEB2 protein expression (Figure 4.1b).

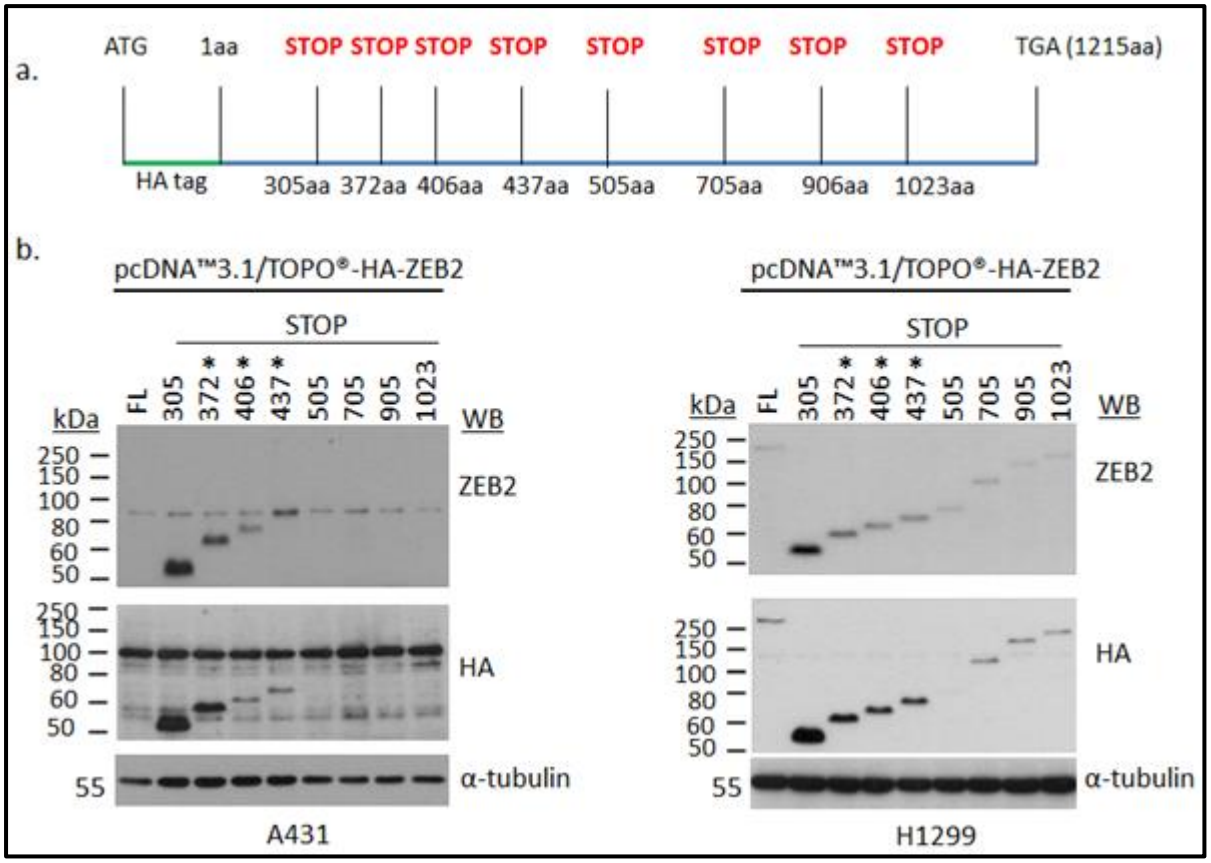


Figure 4.1: Mapping of ZEB2 ORF by STOP codon insertions. (a) An illustration showing the locations of STOP codon inserted in full length ZEB2. (b) Western blot analysis showing HA-ZEB2 expression levels in A431 and H1299 cell lines after transfected with pcDNA™3.1/TOPO®-HA-ZEB2 and a series of pcDNA™3.1/TOPO®-HA-ZEB2 (STOP) expression

plasmids. α-Tubulin staining was used as a protein loading control. Asterisk (*) indicates the locations of STOP codon that up-regulated ZEB2 expression in A431 and H1299 cells. *Note: FL: Full length

After demonstrating that the ZEB2 (372-437aa) region is important for ZEB2 expressions, we analysed if this region contains sequence similarity with ZEB1.

4.2 Analysis of sequence similarity between the ZEB2 (372-437aa) and ZEB1

We compared ZEB2 (372-437aa) and ZEB1 sequences by using the Lalign program, (Huang and Miller, 1991) (Website: http://www.ch.embnet.org/software/LALIGNform.html). We found that apart from conserved SUMOylation site IKTEP (Long et al., 2005), a region within ZEB2 (381-437aa) shows low sequence similarity (29.8 %) with the homologous area in ZEB1 (317-372aa) (Figure 4.2). Low degree of homology of between ZEB2 (372-437aa) and ZEB1 (317-372aa) regions suggest that this sequence may have ZEB2-specific regulatory role.

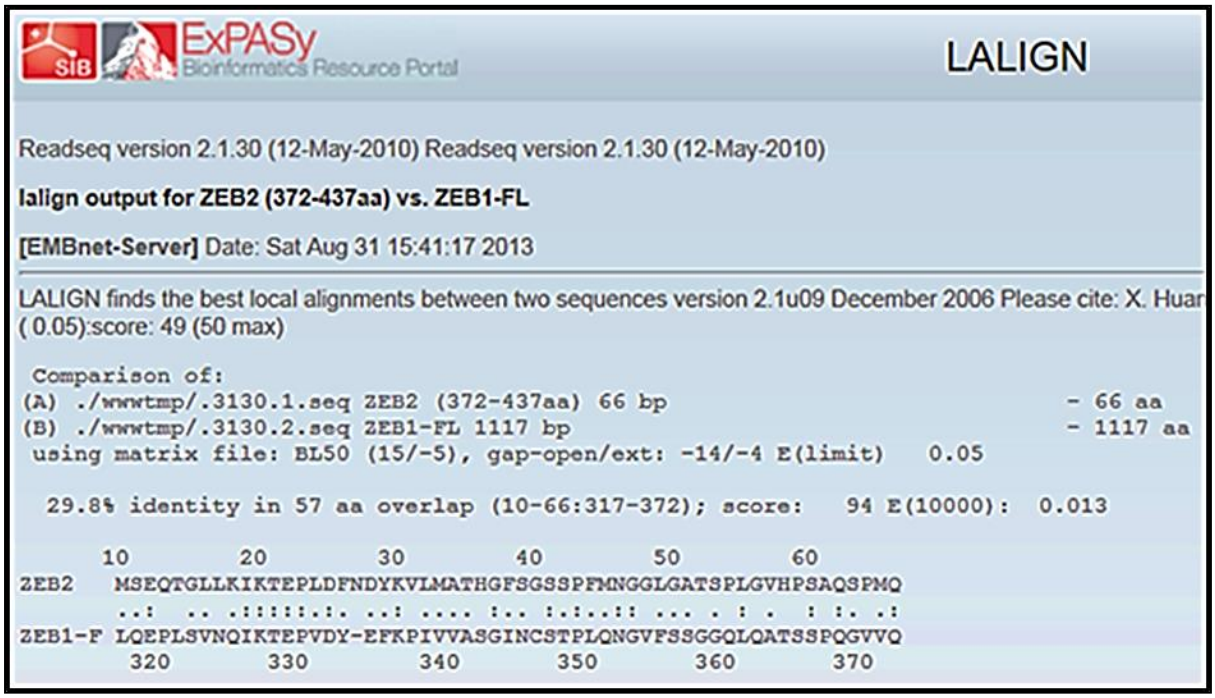


Figure 4.2: Sequence alignment of ZEB2 (372-437aa) and ZEB1 by Lalign program. ZEB2 (372-437aa) region was aligned with ZEB1 (317-372aa) sequence but has low sequence homology.

4.3 Ectopic expression of EGFP-ZEB2 (372-437aa) and EGFP-ZEB1 (317-372aa) in human cancer cell lines

To further investigate whether ZEB2 (372-437aa) and ZEB1 (317-372aa) regions regulate ZEB2 and ZEB1 expression in cancer cell lines, we aimed to generate plasmids expressing EGFP fused with ZEB2 (372-437aa) or ZEB1 (317-372aa) (see section 2.2.9.2). For this purpose, the pEGFP-C1 vector was used and the ZEB2 (372-437aa) and ZEB1 (317-372aa) gene fragments were cloned at the C-terminal EGFP within the pEGFP-C1 vector as illustrated in figure 4.3.

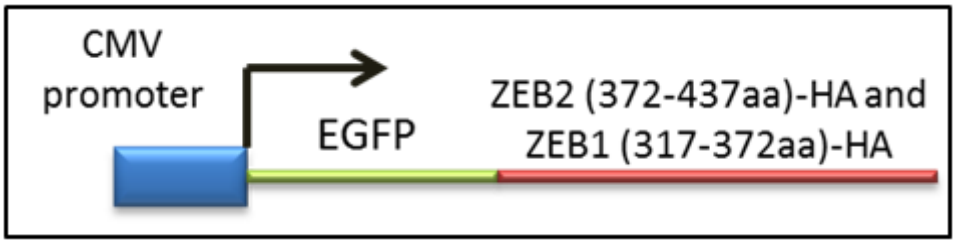


Figure 4.3: Illustration of the pEGFP-C1 vector. The ZEB2 (372-437aa) and ZEB1 (317-327aa) genes were expressed as C-terminal fusion with EGFP using the pEGFP-C1 vector.

To analyse the effect of ZEB2 and ZEB1 fragments on the EGFP protein expression in cancer cell lines, the generated EGFP-ZEB2 (372-437aa) and EGFP-ZEB1 (317-372aa) expression vectors were transiently transfected into MDA-MB-468, A431, H1299 and SaOS-2 cell lines. In two other vectors expressing ZEB-EGFP fusions were also transfected into these cell lines. These plasmids contained ZEB2 (372-437aa) and ZEB1 (317-372aa) located at N terminal of EGFP within pEGFP-N2 vector as illustrated in figure 4.4 (also detailed in table 2.5; section 2.1.5). Both ZEB2 (372-437aa)-EGFP and ZEB1 (317-372aa)-EGFP plasmids were generated made available by Dr G. Browne (EMT laboratory, University of Leicester, UK). 24 hours post transfection, cells were lysed and proteins were analysed via Western blotting using an anti-EGFP antibody. It was found that, ZEB2 (372-437aa) had a profound effect on EGFP expression and ZEB1 (317-327aa) effects was much in all cell lines (Figure 4.5). Furthermore, the C-terminal EGFP fusion has a stronger effect compared to N-terminal fusion. It remains currently unclear why the position of ZEB2-derived sequence is important for the expression of EGFP-ZEB2 fusion. We propose that the kinetics of the translation affects folding and degradation of a nascent peptide (see next sections). It is well established that protein

secondary structure is an important determinant of the efficacy of its proteasomal degradation. We therefore believe that location of ZEB2-derived fragment at the C-terminus of EGFP affects protein structure in a way that makes the fusion a better substrate for the proteasome.

As the effect of ZEB2 (372-437aa) was stronger in A431 and MDA-MB-468 cell lines as compared to H1299 and SaOS-2 cell lines, we suggested that ZEB2 expression maybe regulated in cell-type specific manner.

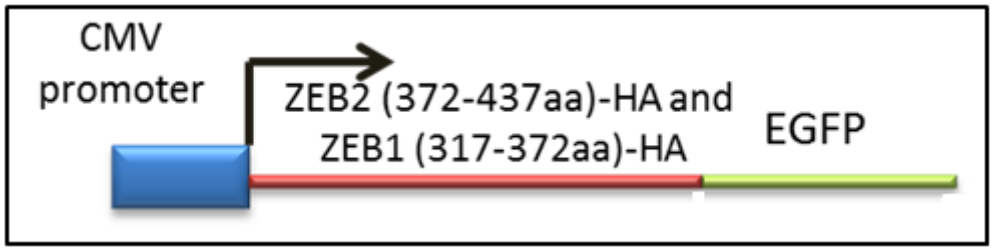


Figure 4.4: Illustration of the pEGFP-N2 vector. The ZEB2 (372-437aa) and ZEB1 (317-327aa) genes were cloned at the N-terminal EGFP within the pEGFP-N2 vector.

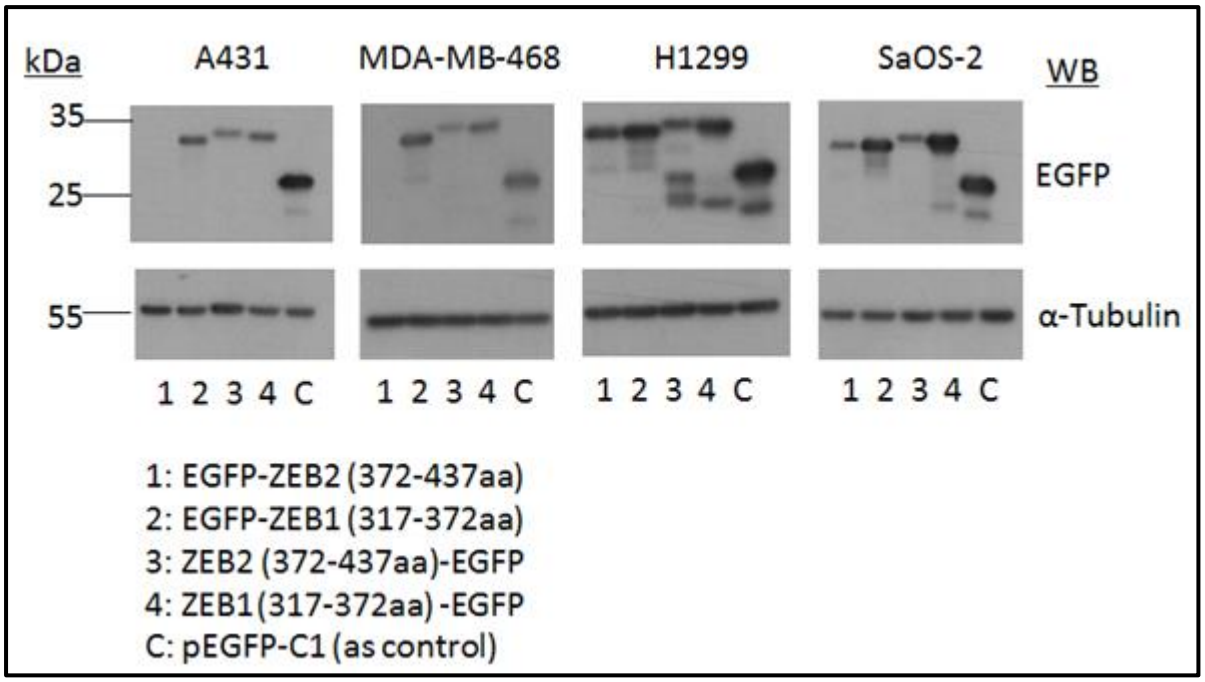


Figure 4.5: Western blot showing expression of EGFP fusions in MDA-MB-468, A431, H1299 and SaOs-2 cell lines after transfected with 4 µg of EGFP-ZEB2 (372-437aa), EGFP-ZEB1 (317-372aa), ZEB2 (372-437aa)-EGFP, ZEB1 (317-372aa)-EGFP vectors or empty and pEGFP-C1 (control). α-Tubulin staining was used as a protein loading control.

After demonstrating that ZEB2-derived fragment strongly reduced EGFP expression in cell line-specific manner, we aimed to map the ZEB2 (372-437aa) domain more precisely (see section 2.2.10.1) to investigate if there is any specific region in ZEB2 (372-437aa) that may control ZEB2 expression.

4.4 Expression of EGFP-ZEB2 (372-437aa)-STOP383, EGFP-ZEB2 (372-437aa)-STOP400, EGFP-ZEB2 (372-437aa)-STOP415 and EGFP-ZEB2 (372-437aa)-STOP426 in cancer cells

To analyse the effect of truncations within EGFP-ZEB2 (372-437aa) region on ZEB2 stability the expression level of the fusions, 4 µg of EGFP-ZEB2 (372-437aa)-STOP383, EGFP-ZEB2 (372-437aa)-STOP400, EGFP-ZEB2 (372-437aa)-STOP415, EGFP-ZEB2 (372-437aa)-STOP426, ZEB2 (372-237aa) and pEGFP-C1 (as control) expression plasmids were transfected into MDA-MB-468 cell line. 24 hours post transfection, cells were lysed and the expression of fusion proteins was analysed by Western blotting using an anti-EGFP antibody. When STOP codons were introduced at the positions 383aa, 400aa, 415aa or 426aa expression of the chimaeras was detectable (Figure 4.6). Expression level of EGFP-ZEB2 (372-437aa) was much lower than that of any of STOP codon insertion constructs indicating that the sequence between 426-437aa is necessary for the full effect of the ZEB2-derived sequence ZEB2 in MDA-MB-468 cell line.

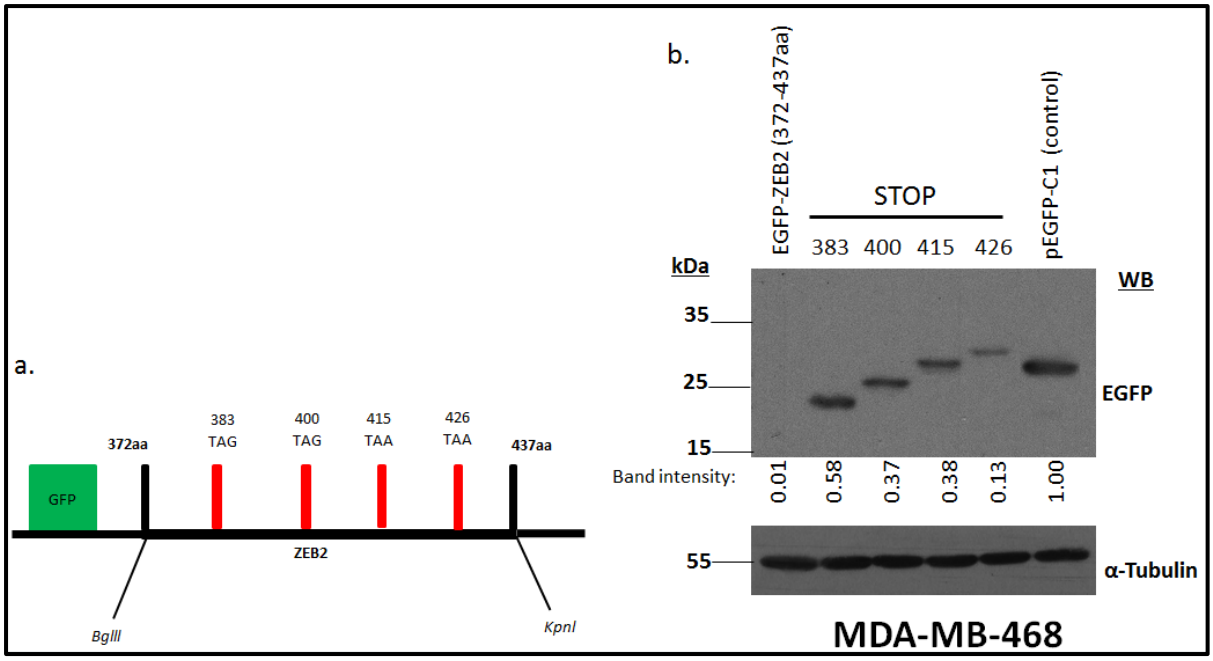


Figure 4.6: Mapping of a regulatory activity within the 372-437aa region by STOP codon insertions. (a) Illustration shows the locations of STOP codon inserted in EGFP-ZEB2 (372-437aa). (b) Western blot analysis shows EGFP expression levels in MDA-MB-468 cell line after transfection with EGFP-ZEB2 (372-237aa), EGFP-ZEB2 (372-437aa)-STOP383, EGFP-ZEB2 (372-437aa)-STOP400, EGFP-ZEB2 (372-437aa)-STOP415, EGFP-ZEB2 (372-437aa)-STOP426 expression constructs or pEGFP-C1 empty vector control. Membrane was probed with an anti-EGFP antibody. α-Tubulin staining was used as a protein loading control. The EGFP bands intensity was measured using Image J and normalised to α- tubulin loading control.

Next, we aimed to map the functional activity of the 372-437aa region from N-terminus. For these purposes, we introduced START codons followed by HA tag at different positions in the 372-437aa sequence within the ZEB2 (372-437aa)-EGFP expression vector (see section 2.2.10.2).

4.5 Expression of ZEB2 (372-437aa)-ATG400-EGFP, ZEB2 (372-437aa)-ATG415-EGFP and ZEB2 (372-437aa)-ATG426-EGFP in carcinoma cell line

To analyse the effect of the ATG codon insertion within ZEB2 (372-437aa)-EGFP region on ZEB2 expression, 4 µg of ZEB2 (372-437aa)-EGFP, ZEB2 (372-437aa)-ATG400-EGFP, ZEB2 (372-437aa)-ATG415-EGFP, ZEB2 (372-437aa)-ATG426-EGFP and pEGFP-N2 (as control) plasmids were transfected into MDA-MB-468 cell line. In 24 hours, cells were lysed and analysed by Western blotting. ZEB2 (372-437aa)-ATG400-EGFP was expressed at the level

higher than ZEB2 (372-437aa), but lower than ZEB2 (372-437aa)-ATG415-EGFP (figure 4.7). Insertion of the ATG codon at the position 400 decreased, but not diminished the expression-limiting activity of the fragment.

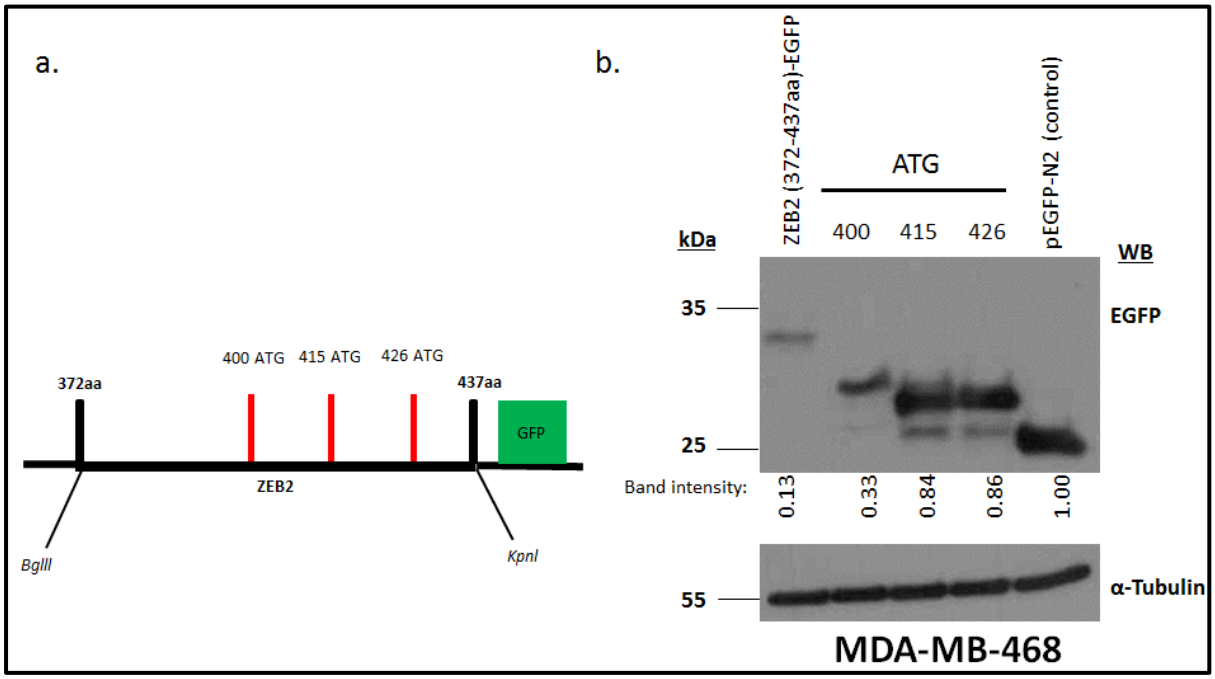


Figure 4.7: Mapping of the expression limiting activity within the 372-437 ZEB2 fragment by ATG codon insertions. (a) An illustration showing the location of ATG codons that have been inserted in ZEB2 (372-437aa)-EGFP (b) Western blot analysis showing the expression of ZEB2-EGFP fusions in MDA-MB-468 cells. Cells were transfected with a series of constructs containing artificial ATG codons as indicated, and protein lysates were analysed with an anti-EGFP antibody. α-Tubulin staining was used as a protein loading control. The EGFP bands intensity was measured using Image J and normalised to α- tubulin loading control.

4.6 Conclusion

372-437aa ZEB2 fragment strongly decreased EGFP expression when fused either N- or C-terminally in three different cell lines. We analysed the expression of two series of ZEB2-EGFP fusions containing artificial START or STOP codons in MDA-MB-468 cells. This approach allowed us to address how the N-terminal or C-terminal truncations affect the expression of fusion proteins and therefore to map expression-limiting activity of the 372-437aa fragment. The results presented in this chapter demonstrate that this activity is associated with the 383-437aa fragment (Figure 4.8).

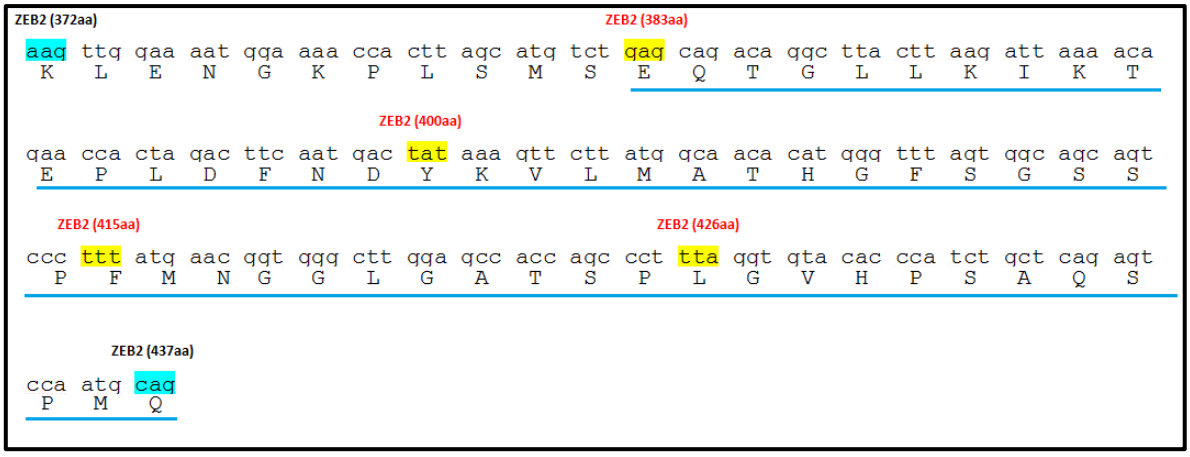


Figure 4.8: An illustration showing the ZEB2 (383-437aa) (blue line) domain that regulates the expression level of ZEB2-EGFP fusion proteins.

4.7 Fusion with ZEB2 (372-437aa) induces proteasomal rather than lysosomal degradation of EGFP

A decreased level of EGFP-ZEB2 chimera expression might be a result of degradation of the protein. We therefore investigated this possibility by treating the EGFP-ZEB2 (372-437aa)-transfected cells with lysosomal and proteasomal inhibitors. To this purpose, we transfected MDA-MB-468 cell lines with the EGFP-ZEB2 (372-437aa) expression vector or pEGFP-C1 as control. 24 hours after transfection, culture medium was replaced by media with or without lysosomal inhibitor, Chloroquine or proteasomal inhibitor, MG132. Cells were lysed after 6 or 16 hours of incubation and EGFP expression was then analysed in western blotting. EGFP-ZEB2 fusion was accumulated in cells treated with both inhibitors, although the effect of Chloroquine was not significant (Figure 4.9). On the other hand, EGFP level was not affected by any treatment. These data suggest the involvement of proteasome pathways in destabilisation of EGFP-ZEB2 fusion by the 372-437aa element.

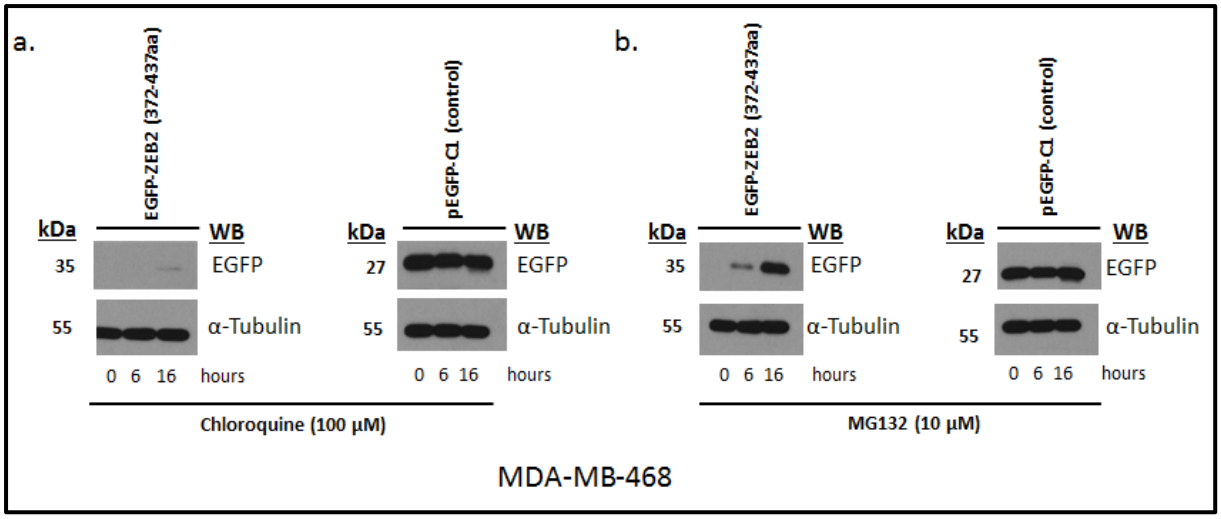


Figure 4.9: Western blot analysis of the expression of EGFP or EGFP-ZEB2 fusion proteins in transfected MDA-MB-468 cell line after treatment with lysosomal (Chloroquine (100µM)) or proteasomal (MG132 (10µM)) inhibitors. α-Tubulin staining was used as a protein loading control.

4.8 Pull-down assay and protein mass spectrometry of EGFP-ZEB2 (372-437aa)

Next, we proposed that proteins interacting within ZEB2 (372-437aa) domain may play a role in ZEB2 regulation. To this end, we decided to carry out a pull-down assay followed by protein mass spectrometric analysis to search for these potentially important interactions. Accordingly, MDA-MB-468 cell line was transiently transfected with EGFP-ZEB2 (372-437aa) or pEGFP-C1 (as control) and maintained in the presence of MG132 for 16 hours. After treatment, cells were lysed and pull-down assay was performed by using the GFP-Trap®_A kit (section 2.2.2.4). Following this, 10 µg of protein samples were separated by SDS-PAGE and analysed by Coomassie Blue staining (section 2.2.2.3.1) and Western blot analysis. Figure 4.10a shows cellular proteins were not detected in bound samples using Coomassie staining suggesting that immunoprecipitated samples were separated from the majority of cellular proteins. High levels of EGFP protein detected in samples bound by the anti-EFP antibody confirmed that the recombinant proteins were purified from transfected MDA-MB-468 cells (Figure 4.10b).

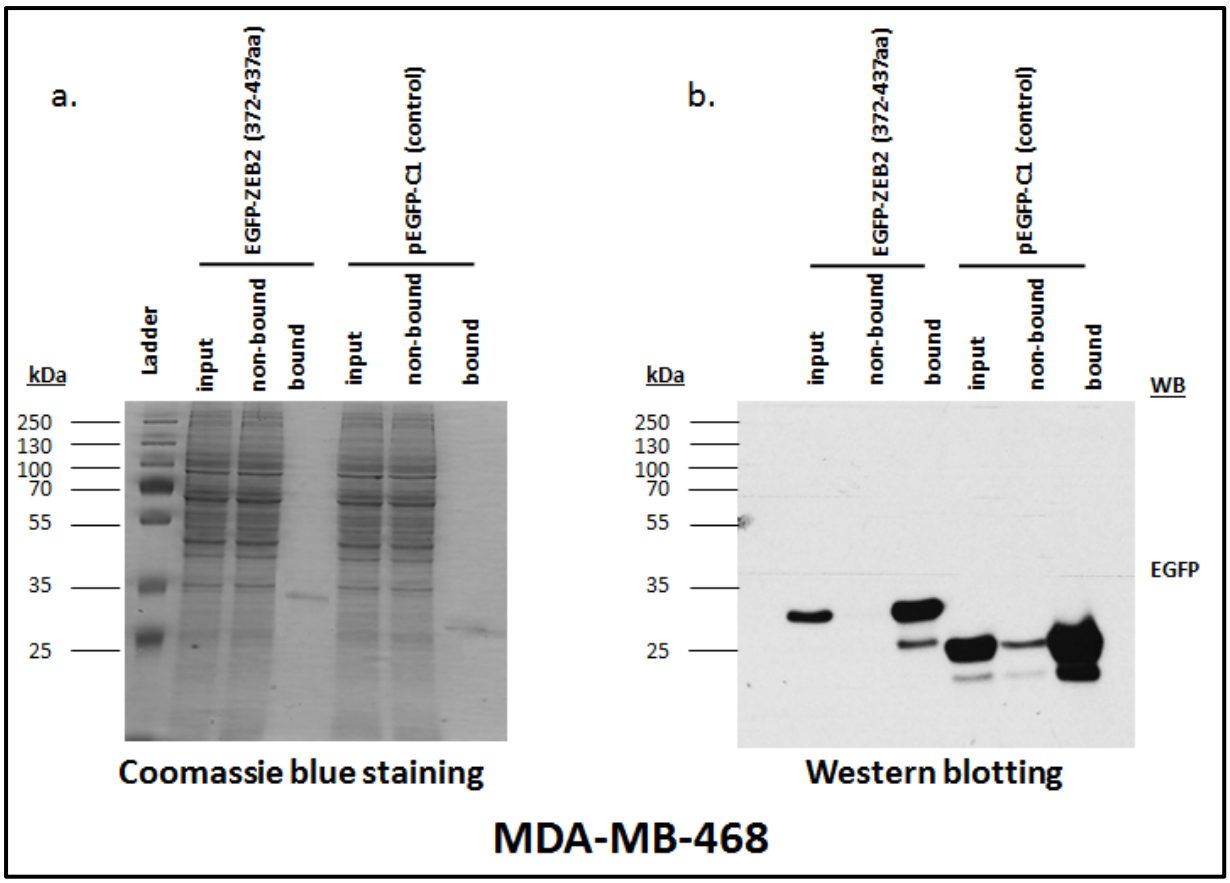


Figure 4.10: Pull-down assay of EGFP-ZEB2 (372-437aa) and pEGFP-C1 (control). Input, non-bound and bound samples were separated on 12 % SDS-PAGE and analysed by (a) Coomassie blue staining and (b) Western blotting. Results are representative of 2 independent experiments.

Following pull-down assay, 40 µg of bound proteins were separated on 12 % SDS-PAGE. After separation, the gel area containing the protein samples were excised and analysed by mass spectrometry (section 2.2.2.5). Following analysis, the lists of components interacted within EGFP-ZEB2 (372-437aa) and EGFP alone were compared. The data demonstrated that, several components of the aminoacyl tRNA synthetase complex (MSC) interacted with the EGFP-ZEB2 fusion, but not with EGFP alone. MSC is co-localised with elongating ribosomes and functions in protein translation by controlling synthetase activities during tRNA-ribosome attachments (David et al., 2011). The pull-down experiments have also shown that ZEB2 (372-437aa)-transfected cells interacted with several translation elongation factors 2 and four subunits of translation elongation factor 1 (Table 4.1). Recent studies in yeast proposed that elongating ribosomes, translation initiation and elongation factors, MCS and also proteasomes represent a supercomplex “a translasome”. This

supercomplex is proposed to link protein synthesis and degradation of improperly folded nascent peptides ensuring efficient and error-free translation (Sha et al., 2009).

Protein	Identified peptides	Expect	Score
Elongation factor 1-alpha 1	R.YEEIVK.E K.EVSTYIK.K K.QLIVGVNK.M R.TIEKFEK.E K.EVSTYIKK.I R.LPLQDVYK.I K.IGGIGTVPVGR.V K.STTTGHLIYK.C	0.0065 0.012 0.019 0.0058 0.079 0.00051 1.1e-005 0.0035	30 28 26 31 20 44 57 36
Elongation factor 1-gamma	K.AKDPFAHLPK.S K.STFVLDEFKR.K R.AVLGEVKLCEK.M M.AAGTLYTYPENWR.A K.AAAPAPEEEMDECEQALAAEPK.A	0.0001 1.1e-005 1.5e-005 0.046 1.5	49 59 59 22 3
Elongation factor 1-delta	K.LVPVGYGIR.K R.IASLEVENQSLR.G	0.0031 1.7e-006	36 69
Elongation factor 1-beta	R.SIQADGLVWGSSK.L	2.2e-005	58
Elongation factor 2	M.VNFTVDQIR.A	5.5e-006	64
Bifunctional aminoacyl-tRNA synthetase	R.LLSVNIR.V K.YYTLFGR.S	0.0032 0.0035	37 37
Glutaminyl-tRNA synthetase	R.LAWGQPVGLR.H K.AINFNFGYAK.A	0.00082 0.42	39 15
Lysyl-tRNA synthetase	K.ILDDICVAK.A M.AAVQAAEVKVDGSEPK.L	0.12 4.3e-007	21 75
Aminoacyl tRNA synthase complex-interacting multifunctional protein 2	R.VLSTVHTHSSVK.S R.SCENLAPFNTALK.L	0.0053 0.00013	32 48
Aspartyl-tRNA synthetase	R.GEEILSGAQR.I	0.018	29

Table 4.1: List of components identified by mass spectrometric analysis of ZEB2 (372-437aa). All protein listed were specific for ZEB2 (372-437aa) and not found in the EGFP control (The Protein Nucleic Acid Chemistry Laboratory (PNACL), Hodgkin Building, University of Leicester). Results are representative of 2 independent experiments. For full list of ZEB2 (372-437aa) and EGFP (control) see Attachment 17 in attached CD.

4.9 Polysome profiling of EGFP-ZEB2 (372-437aa)

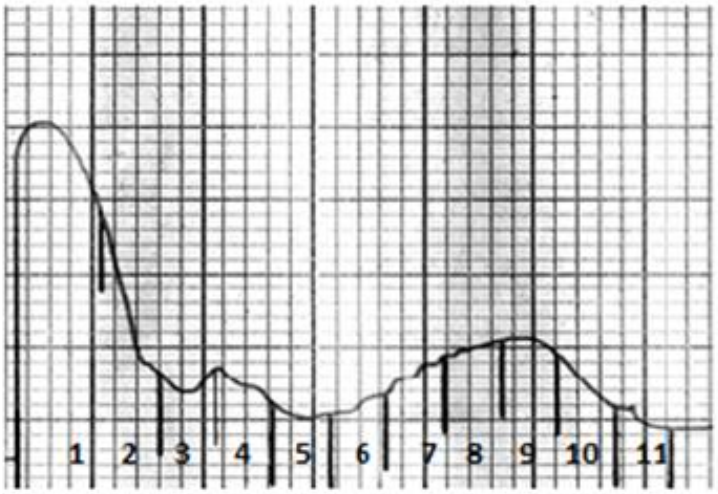
Our previous proteomics data did not reveal any E3 ligases in association with 372-437aa ZEB2 fragment. Instead, several components of the "translasome" such as MSC and the elongation factors were co-immunoprecipitated with the EGFP-ZEB2 fusion. We, therefore, proposed that ZEB2-derived fragment may induce ribosome stalling and slow down translation. To this end, we transfected MDA-MB-468 cells with either control EGFP-expression vector or a construct expressing EGFP-ZEB2 chimaera and performed ribosome fractionation of the lysates in sucrose gradient (section 2.2.2.6) followed by TCA

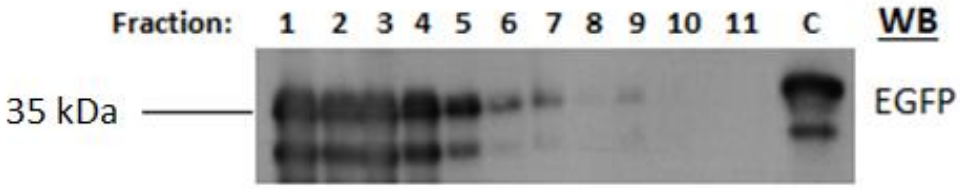
precipitation (section 2.2.2.6.3) and Western blotting. The major part of control EGFP was found free in cytosol or in fractions containing mono-ribosomes (fractions 1-4). In contrast, a significant enrichment of the EGFP-ZEB2 372-437aa fusion protein was found in the elongating polysomal fractions (fractions 7-11). Of note, the EGFP-ZEB2 372-437aa fusion protein detected in the polysomal fraction was present as a single band of a lower molecular weight than EGFP-ZEB2 372-437aa fusion protein found in fractions 1-4 (Figure 4.10a).

Pull-down and mass spectrometry experiments were carried out with MG132-treated cells. Therefore, we aimed to analyse the association of EGFP-ZEB2 with mono- and poly-ribosomes under conditions used for the proteomics experiments. Cells transfected with EGFP or EGFP-ZEB2 fusion were pre-treated with MG132 for 16 hours before harvesting, lysed and fractionated in sucrose gradient. Strikingly, treatment of MDA-MB-468 cells with MG132 for 16 hours completely abolished polysomal fraction, indicating that protein synthesis was inhibited under these conditions (Figure 4.11b). Nevertheless, immunoblotting analysis of the fractionated material has shown a difference in the distribution of EGFP and EGFP-ZEB2 fusion. Whereas a very large part of EGFP was represented as free cytosolic protein, nearly 50 % of EGFP-ZEB2 fusion was detected in fraction 4 corresponding to 80S mono-ribosomes. These data explain co-immunoprecipitation of different components of translational machinery with the EGFP-ZB2 fusion, and are therefore in accordance with the results of pull-down and mass spec experiments (Figure 4.11b).

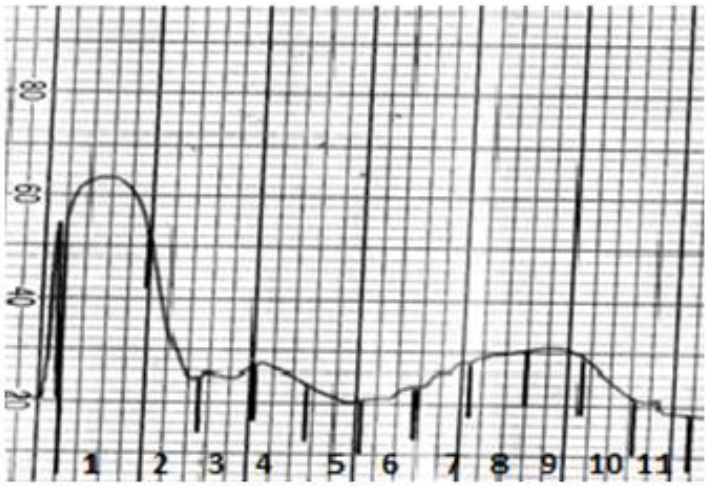
Overall, these data are in line with the hypothesis that 372-437aa ZEB2 fragment affects translational elongation through induction of ribosome stalling.

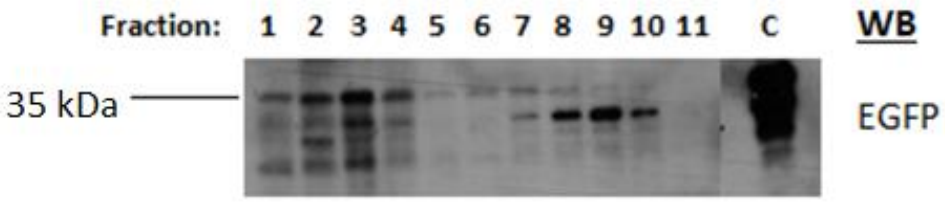
a. Untreated





pEGFP-C1





EGFP-ZEB2 (372-437aa)

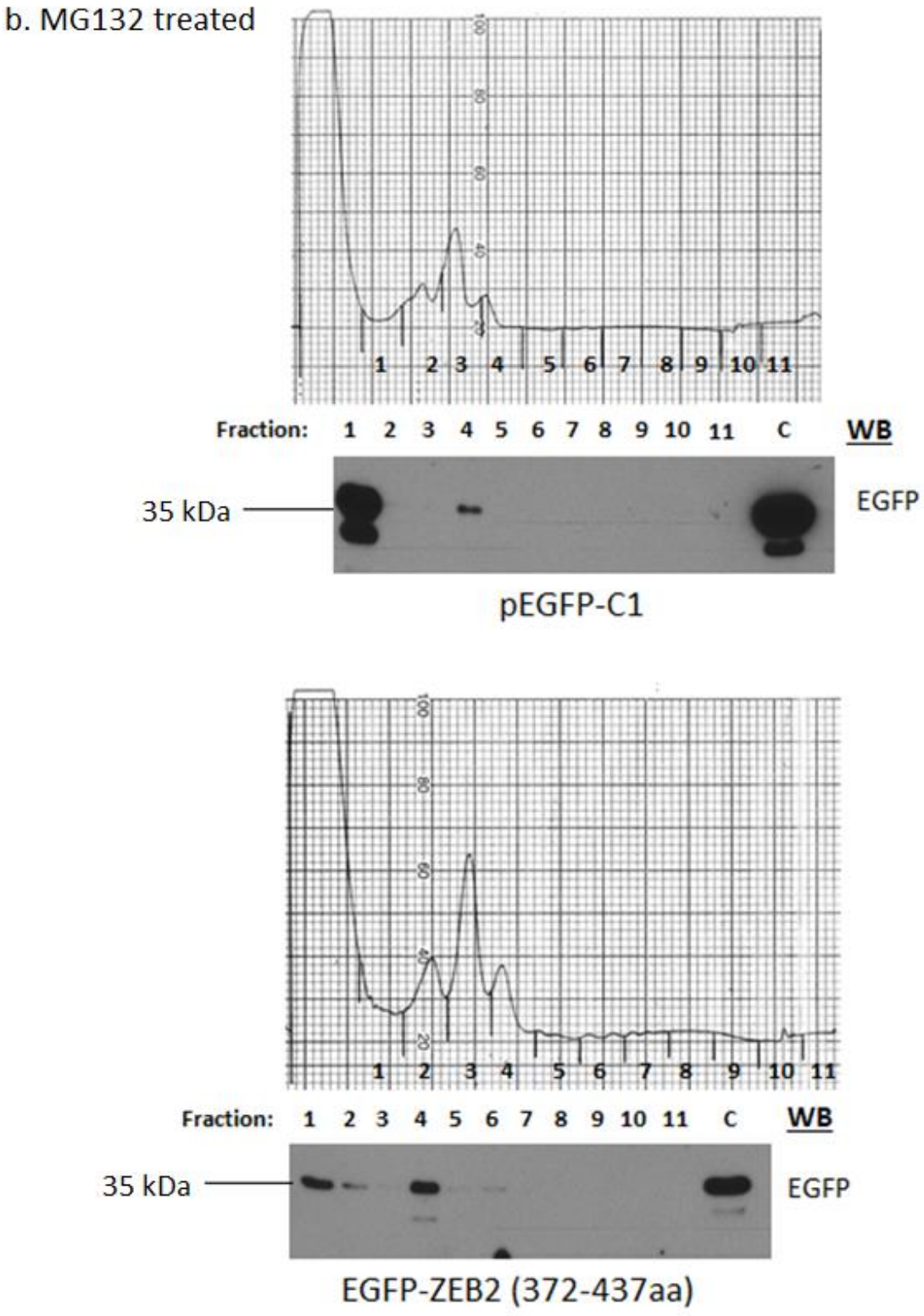


Figure 4.11: Ribosome profiling of ZEB2 (372-437aa) and pEGFP-C1 (a) untreated (b) MG132 treated MDA-MB-468 cell line.

4.10 ZEB2 (372-437aa) contains one rare codon (RC) cluster

Presence of rare codons (RC) is one of the factors which may cause ribosome stalling (Tsai et al., 2008, Robinson et al., 1984, Misra and Reeves, 1985, Li et al., 2006). Therefore, we investigated if low expression of EGFP-ZEB2 fusion in cancer cells is associated with the presence of RC. To this end, we analysed ZEB2 (372-437aa) ORF sequence using the Rare Codons' Search program (Website: http://www.molbiol.ru/eng/scripts/01_11.html). We identified that ZEB2 (372-437aa) contains a single RC cluster comprising three adjacent codons corresponding to the amino acids leucine (L_{426}), glycine (G_{427}) and valine (V_{428}) (codon frequencies 7.6, 10.8 and 7.1) (Figure 4.11). Interestingly, this RC cluster was a part of 426-437aa domain that was important for ZEB2 destabilisation activity of the ZEB2 (372-437aa) protein fragment (Figure 4.6; section 4.5). Knowing that the substitution of rare codons in a target sequence with the high frequency codons can improve protein expression levels (Angov, 2011), we aimed to replace the TTA GGT GTA triplet by synonymous codons CTG GGC GTG (codon frequencies 39.9, 22.4 and 28.3) (RC to CC mutation) (Figure 4.12).

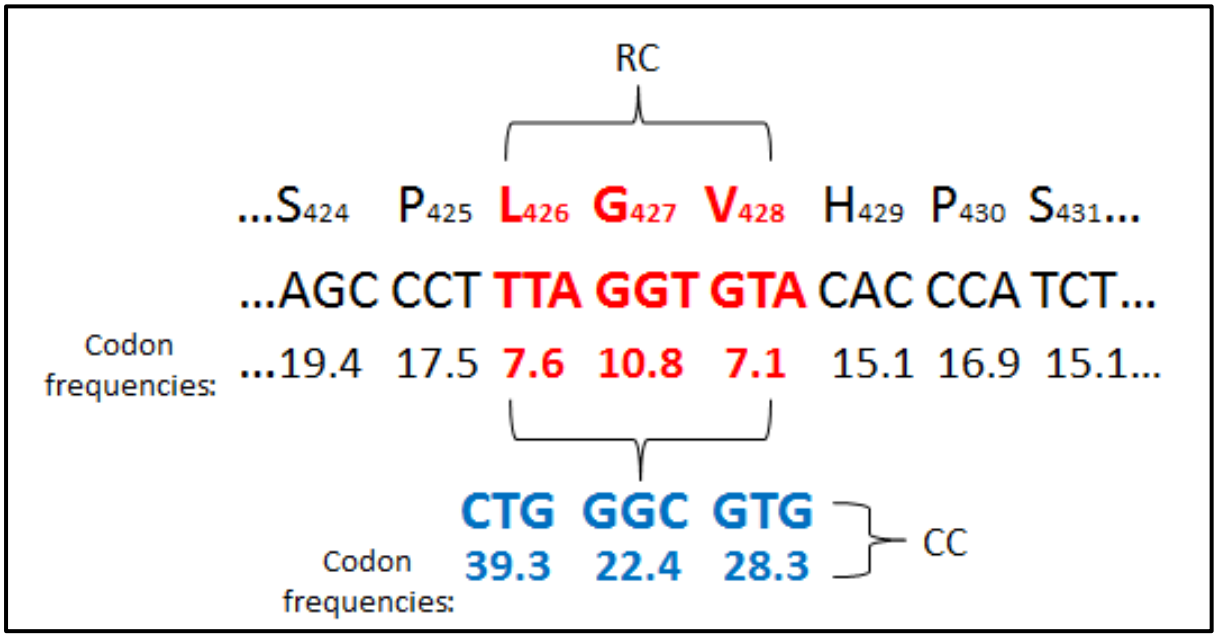


Figure 4.12: An illustration showing the location of RC cluster within ZEB2 (372-437aa) and it codon frequencies. Mutations of rare to common codons (CC sequence) are indicated in blue.

4.11 Expression of EGFP-ZEB2 (372-437aa)-(RC-CC) in cancer cell lines

To analyse whether the RC to CC substitutions within EGFP-ZEB2 (372-437aa) affects EGFP expression, we introduced the RC to CC mutation within EGFP-ZEB2 (372-437aa) plasmid (see section 2.2.10.3) and transfected the EGFP-ZEB2 (372-437aa)-(RC-CC), EGFP-ZEB2 (372-437aa), EGFP-ZEB2 (372-437aa)-STOP426 plasmids or pEGFP-C1 control into the MDA-MB-468, H1299 and SaOS-2 cell lines. 24 hours post-transfection cells were lysed and subjected to Western blot analysis. Western blot analysis was performed 24 hours post-transfection and demonstrated that, the synonymous substitutions of TTA GGT GTA codons by CTG GGC GTG sequence increased expression of EGFP-ZEB2 (372-437aa) in MDA-MB-468, H1299 and SaOS-2 cell lines (Figure 4.13). In all cell lines analysed, introducing of a stop codon at the position 426 influenced expression level much stronger than the RC to CC substitution (Figure 4.13). These data indicate that the 426-437 sequence impact on the expression level of the fusion protein likely through affecting its conformation and making the protein more vulnerable for proteasomal degradation.

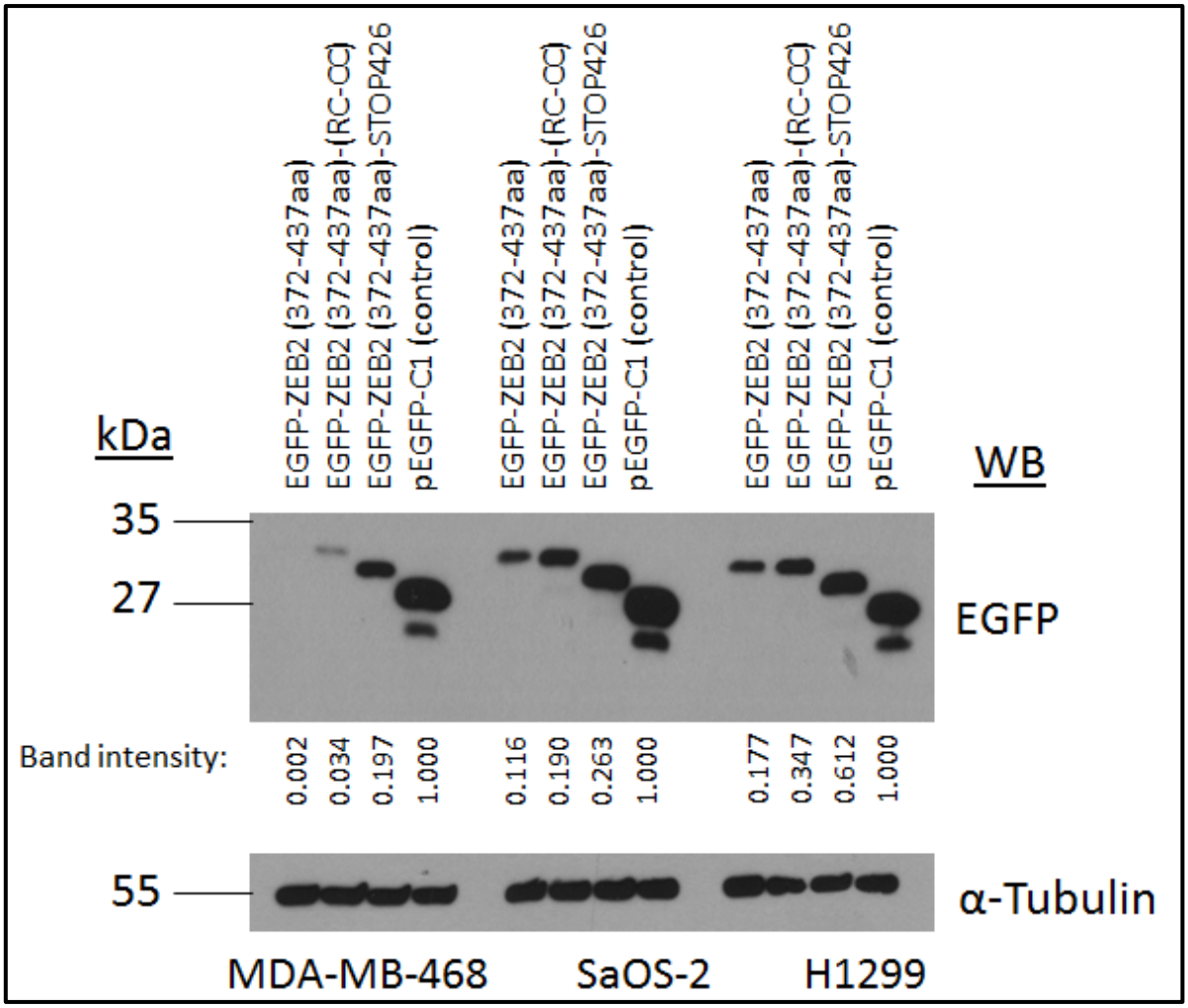


Figure 4.13: Western blot analysis demonstrated the level of EGFP protein expression in MDA-MB-468, SaOS-2 and H1299 cell lines transfected with EGFP-ZEB2 (372-437aa)-(RC-CC), EGFP-ZEB2 (372-437aa), EGFP-ZEB2 (372-437aa)-STOP426 and empty pEGFP-C1 expression plasmids. α-Tubulin staining was used as a protein loading control. The EGFP bands intensity was measured using Image J and normalised to α- tubulin loading control.

4.12 Full length ZEB2 contains two identical RC clusters (LGV)

We decided to analyse a role for LGV RC cluster in full length ZEB2. The analysis of ZEB2 ORF demonstrated that, ZEB2 contains the second RC cluster, $L_{439}G_{440}V_{441}$ positioned at the distance of just 10 amino acid residues from the first LGV triplet (Figure 4.14). The important role of this motif was shown by the analysis of the STOP codon series in MDA-468 cells, where we detected an increase in the expression of fusion protein after STOP codon insertion at the position 437aa (marked with asterisk (*)) (Figure 4.15). Therefore, we decided to substitute both TTA GGT GTA clusters with the CTG GGC GTG CC sequence

(Figure 4.14) within the full length ZEB2 expressing construct (see section 2.2.10.4) and analyse if these changes affect ZEB2 expression levels in cancer cell lines.

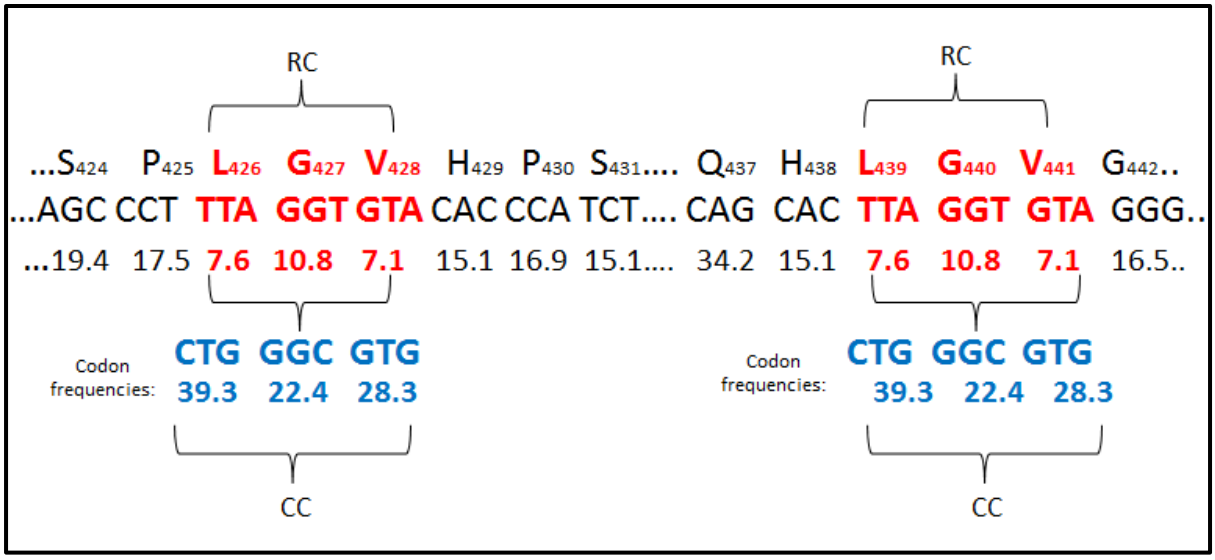


Figure 4.14: Location of RC clusters within the full length ZEB2 and it codon frequencies. Mutations of RC to CC sequences are shown.

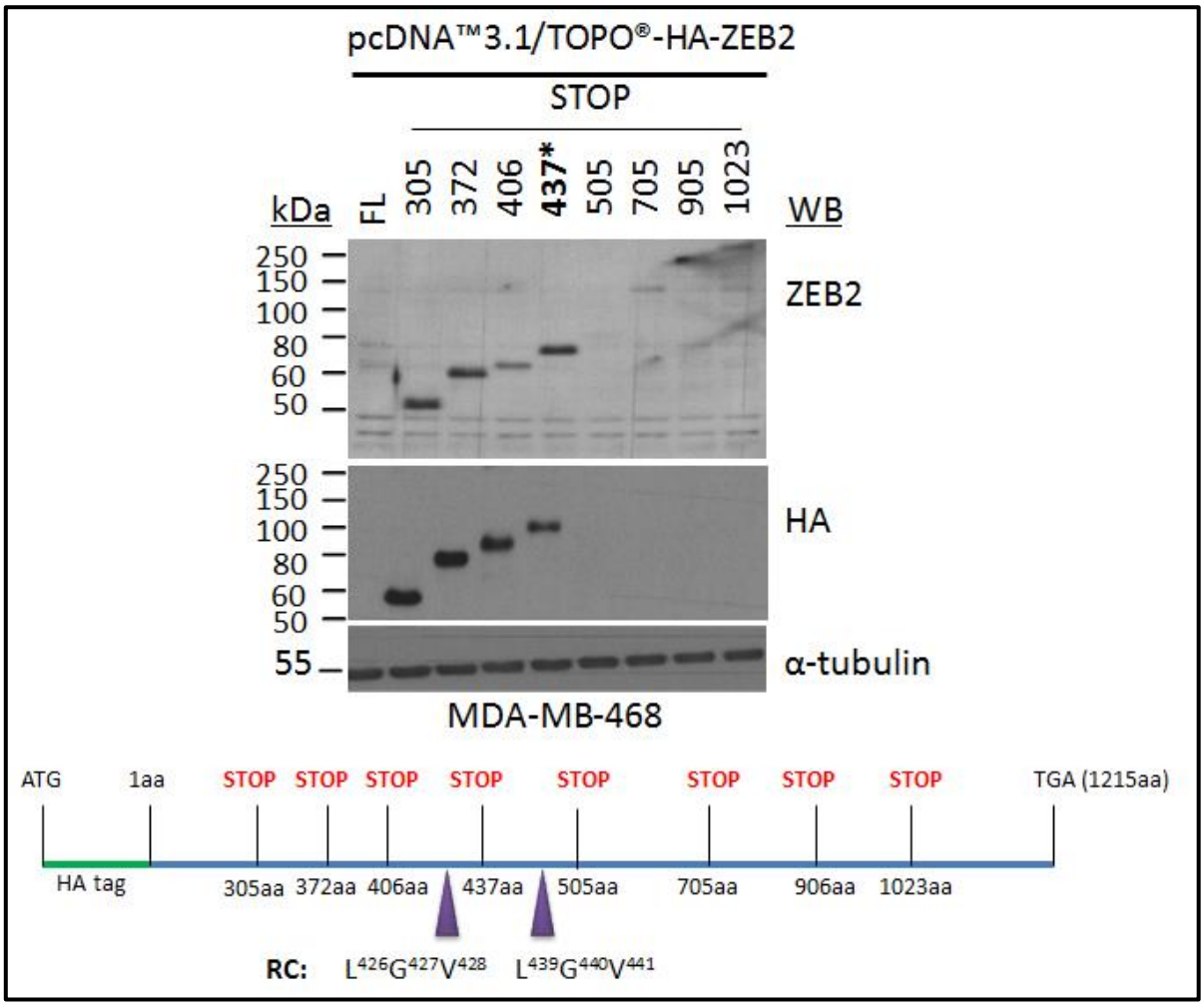


Figure 4.15: Western blot analysis of ZEB2 expression levels in cell line after MDA-MB-468 cells were transfected with the series of constructs expressing ZEB2 truncation mutants pcDNA™3.1/TOPO®-HA-ZEB2 (STOP). α-Tubulin staining was used as a protein loading control.

4.13 Expression of EGFP-ZEB2-(RC-CC) in cancer cell lines

The A431, MDA-MB-468 and SaOS-2 cell lines were co-transfected with 0.5 µg of pEGFP-C1 to control for efficiency and either EGFP-ZEB2 or EGFP-ZEB2-(RC-CC) constructs. 24 hours post-transfection cells were lysed and proteins were subjected to Western blotting. Data show that the simultaneous synonymous substitution of both LGV clusters with CC triplets enhanced ZEB2 expression level in A431, MDA-468 and SaOs-2 cell lines (Figure 4.16).

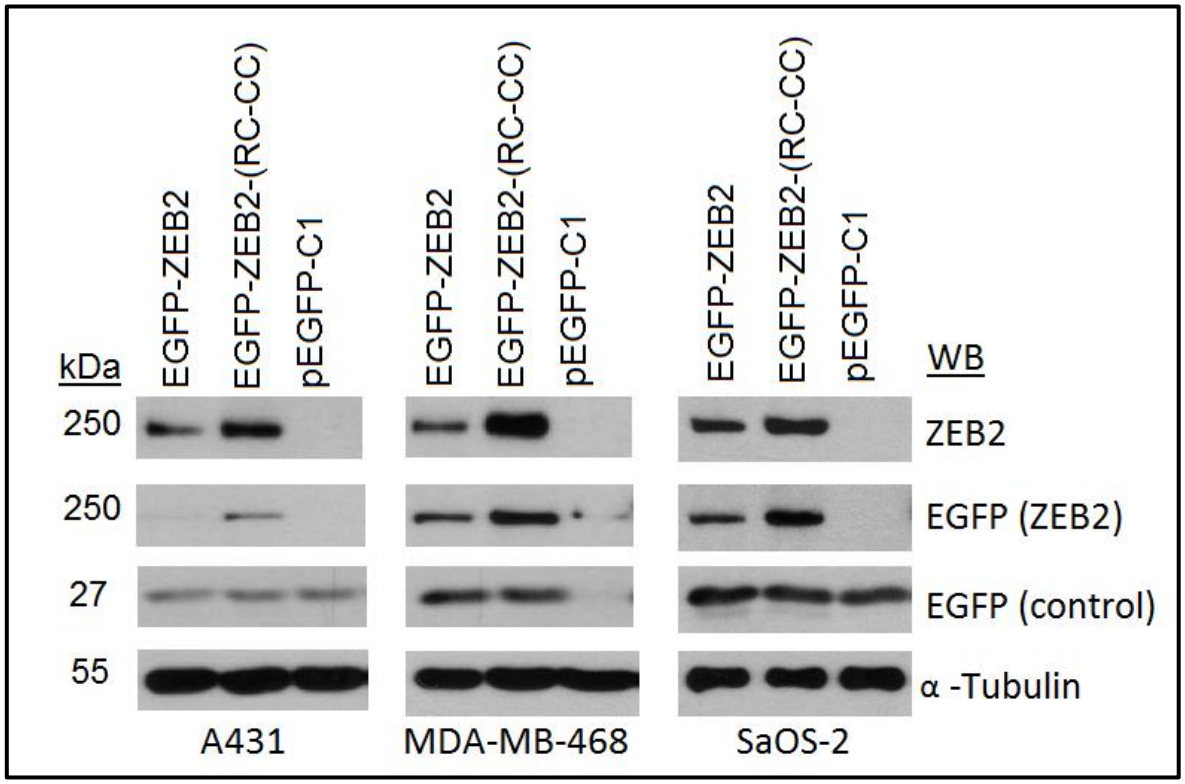


Figure 4.16: Western blot analysis demonstrated the level of EGFP and ZEB2 proteins expressions in A431, MDA-MB-468 and SaOS-2 cell lines after co-transfected with pEGFP-C1 and either EGFP-ZEB2 or EGFP-ZEB2-(RC-CC). α-Tubulin staining was used as a protein loading control.

4.14 Full length ZEB1 contains two high frequency CC clusters

We aimed to analyse whether RC clusters are functional in different protein context. We compared the region containing LGV RC clusters in ZEB2 with homologous area in ZEB1 and

found that they correspond to $L_{367}Q_{368}A_{369}$ and $V_{377}Q_{378}A_{379}$ triplets. According to the analysis performed using the Rare codons' search program, these triplets are composed of codons with relatively high frequencies (Figure 4.17). Therefore, we decided to substitute the LQA and VQA codons in context of the full length ZEB1 with LGV RC codons clusters derived from ZEB2 (Figure 4.17) (see section 2.2.10.5). The aim of this experiment was to analyse if these substitutions affect ZEB1 expression levels in cancer cell lines.

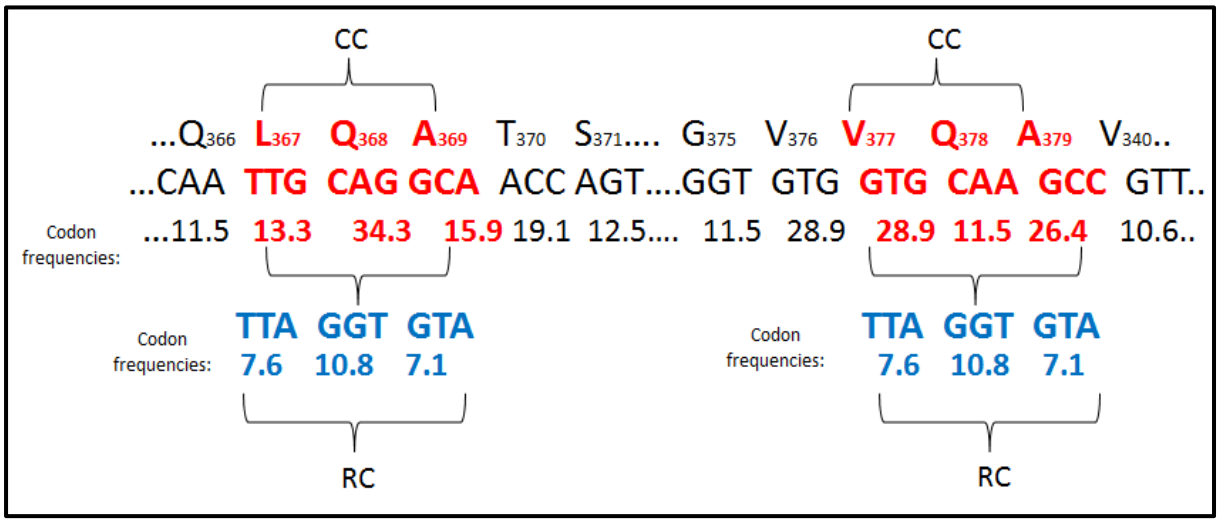


Figure 4.17: The location of LQA and VQA clusters (CC clusters) within ZEB1 sequence. Codon frequencies are indicated.

4.15 Expression of EGFP-ZEB1-(CC-RC) in cancer cell lines

The A431, MDA-MB-468 and SaOS-2 cell lines were co-transfected with 0.5 µg of pEGFP-C1 and either 4 µg of EGFP-ZEB2 or EGFP-ZEB2-(RC-CC) expression plasmids. 24 hours post-transfection cells were lysed and analysed by Western blotting. Importantly, the substitution of LQA and VQA codons with LGV RC clusters produced an effect on EGFP-ZEB1 expression in neither of cell lines (Figure 4.18).

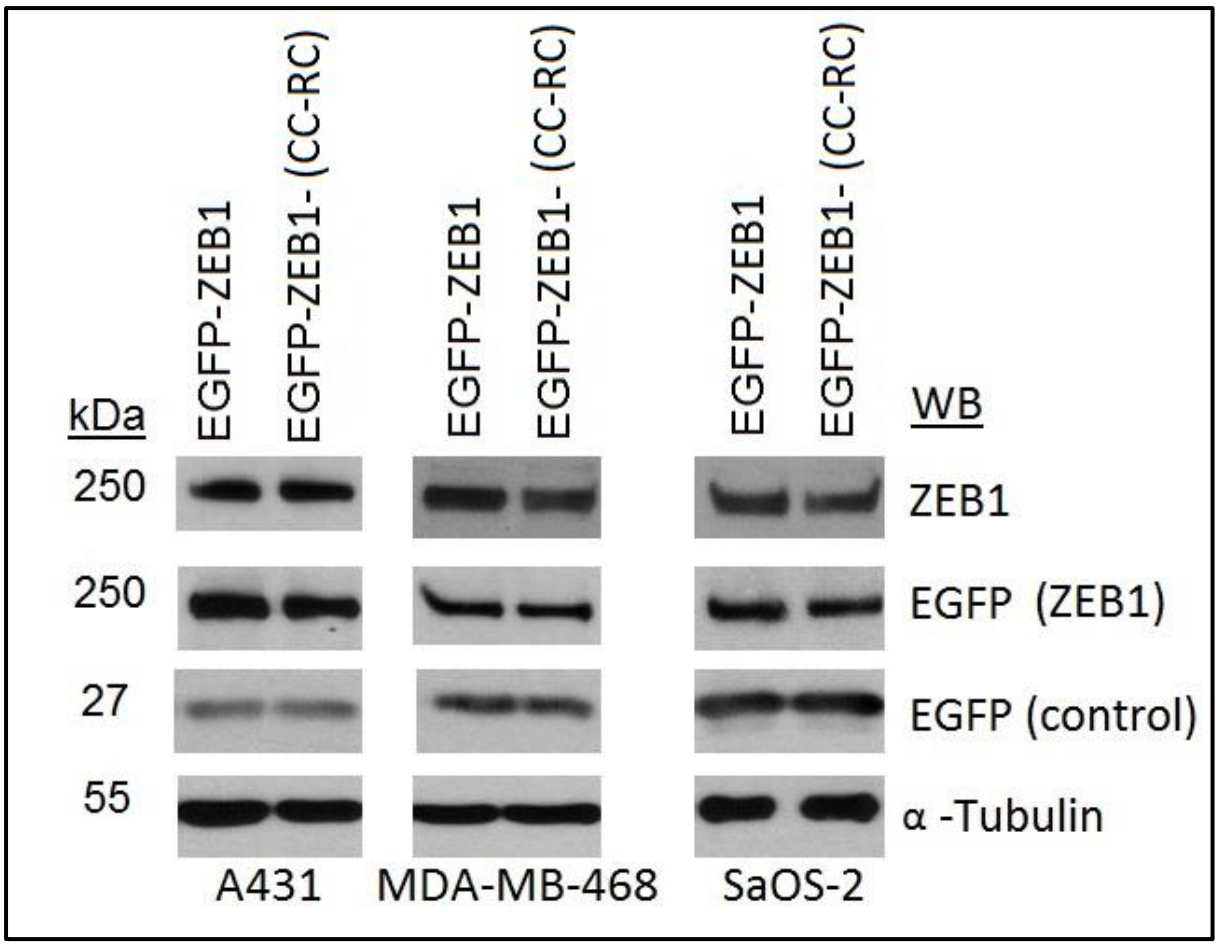


Figure 4.18: Western blot analysis demonstrates the level of EGFP and ZEB1 proteins expression in A431, MDA-MB-468 or SaOS-2 cell lines after transfection with pEGFP-C1 and either EGFP-ZEB1 or EGFP-ZEB1-(CC-RC). α-Tubulin staining was used as a protein loading control.

CHAPTER 5: DISCUSSION, CONCLUSION AND FUTURE STUDIES

EMT programs are regulated by transcription factors which are downstream effectors of several growth factors acting in coordinated fashion (Miyazono, 2009). These transcription factors are considered to be master regulators of EMT; and they act either as direct or indirect transcriptional repressors of the E-cadherin gene (*CDH1*) (Thiery et al., 2009). The list of these repressors includes a family of zinc finger transcription factors, ZEB1 and ZEB2. ZEB1 and ZEB2 proteins have been implicated in metastatic progression of several cancer types, including breast, bladder and pancreatic carcinoma, colorectal and lung cancer (Vandewalle et al., 2009, Sayan et al., 2009). ZEB1 and ZEB2 are similar in their genomic organization and structure; however their regulation in cancer cells remains unclear.

The control of ZEB1 and ZEB2 genes on the level of transcriptional initiation is insufficiently studied. However, several observations do suggest that this level of regulation contributes to the differential expression of ZEB genes. Indeed, homozygous knockout ZEB2 mice express significantly more ZEB1 in some embryonic tissues, such as neural folds and paraxial mesoderm than their wild-type counterparts. These data suggest that that ZEB1 gene is ZEB2-regulated (Miyoshi et al., 2006). Likewise, in carcinoma cells, hierarchical reciprocal regulation of ZEB1 and ZEB2 genes has been reported in both carcinoma and malignant melanoma cells (Caramel et al., 2013). Organization of ZEB1 and ZEB2 genes is extraordinary complex. Both genes seems to be controlled by multiple transcriptional enhancers, and currently described enhancers regulating ZEB2 gene expression are localised at the distance of 1.2 Mb and 60 kb away from the coding parts of the gene (El-Kasti et al., 2012). Obviously, this impedes attempts to delineate whether mutual regulation of ZEB genes is direct or not.

Previous studies have linked the ZEB proteins with several species of miRNA. It was found that repression of ZEB1 and ZEB2 by the miR-205 and miR-200 family increased expression of E-cadherin and promoted an epithelial phenotype (Gregory et al., 2008a). However, tight control over ZEB1 and ZEB2 expression by miR-205 and miR-200 family members cannot explain why in some cancer types, one or the other ZEB protein predominates. MiR-192 specifically targets ZEB2, but not ZEB1 in mesangial cells (Kato et al., 2007). This or possibly other miRNA species may contribute to differential expression of ZEB genes in different genetic backgrounds.

By studying the expression of ZEB1 and ZEB2 in cancer cell lines, we confirmed that ZEB1 and *ZEB2* mRNA were present in all E-cadherin-negative carcinoma cell lines; however, their protein levels were significantly different. Whereas ZEB1 protein was expressed in all the mesenchymal carcinoma cell lines, ZEB2 protein was mildly expressed in carcinoma cells in culture. Importantly, transient transfection of different cell cultures with identical expression vectors, pZ/EG-HA-ZEB1 and pZ/EG-HA-ZEB2 showed that ZEB1 protein was expressed at high level compared to ZEB2. This difference was primarily evident in epithelial carcinoma cell lines. We interpreted these data as an indication that ZEB1 and ZEB2 are differentially regulated at post-transcriptional level. Clearly, this level of regulation may be at least in part responsible for limited expression of ZEB2 in cancer cell lines and for differential expression of these proteins in cancer.

Our study on post-transcriptional regulation of ZEB2 in cancer cell lines revealed that ZEB2 ORF contains rare codon clusters that influence ZEB2 translation. The codon adaptation index (CAI) is a coefficient characterising the representation of rare codons in a gene relative to an ideal gene comprised of exclusively optimal codons. For decades, it has been known that genes with higher CAIs are highly expressed in both prokaryotes and eukaryotes (Grantham et al., 1981, Bennetzen and Hall, 1982, Gouy and Gautier, 1982). However, the data on the role of rare codons in the regulation of individual genes is very limited. A recent study addressed this issue in relation to RAS gene family (Lampson et al., 2013). It was found that the difference in oncogenic potentials of HRAS and KRAS relies on overrepresentation of rare codons in the latter resulting in its low expression as compared with more tumourigenic HRAS. In addition, this study identified 60 pairs of homologous proteins which differed in the density of rare codons and whose differential expression was affected by codon bias. Our data support general importance of codon bias for the regulation of gene expression in higher eukaryotes. However, we conclude that not just the total number of rare codons but also their distribution throughout the ORF is a determinant of inefficient protein synthesis. Indeed, whereas CAIs for HRAS and KRAS equal to 0.87 and 0.69 respectively, for ZEB1 and ZEB2, these figures are 0.70 and 0.75. Therefore, ZEB2 gene structure is supposed to be more adjusted for efficient translation than ZEB1. Contrary to this consideration, however, ZEB2 translation is less efficient than that of ZEB1 due to triplets of rare codons, LGV, localised to the border of Smad binding domain. Importantly,

these triplets are non-functional when introduced in homologous region of ZEB1 indicating the importance of protein context.

Ubiquitination is a post-translational modification that targets cellular proteins for degradation. Almost all cellular processes including EMT (Voutsadakis, 2012a) are regulated by the ubiquitin proteasome system. Knowledge about a link between ubiquitination and ZEB1 is very limited. ZEB1 regulation by cullin7 E3 ligase has been reported in the trophoblast cell lineage (Fu et al., 2010). However, the relationship between ZEB2 and the ubiquitination machinery is still unknown.

In this study, we showed that fusion with the 372-437 aa ZEB2-derived sequence confers instability on EGFP, and expression levels of EGFP-ZEB2 is strongly enhanced in cells treated with the proteasome inhibitor MG132. Proteasomes are co-localised with elongating ribosomes where they assure degradation of improperly folded nascent peptides (Sha et al., 2009). Therefore, we suggest that proteasomal degradation has a role in control of ZEB2 expression, but only on the stage of its synthesis (Figure 5.1). The prerequisite of proteasomal degradation is ubiquitination of a substrate at lysine residues, which enables the recognition of the protein by proteasome. However, we were not able to detect an ubiquitinated EGFP-ZEB2 fusion in MG132-treated cells by mass spectrometry. 372-437 ZEB2 fragment contains three lysine residues. Therefore, we propose to substitute these amino acids by structurally similar arginine residues, which cannot be ubiquitinated. The effect of these mutations on the expression level of EGFP-ZEB2 fusions would clarify whether ubiquitination is involved in the control of ZEB2 translation.

The interpretation of ribosome profiling data presented in Figure 4.11a is not simple. In this study we found that the molecular weight of peptides associated with polysomes shows that just one species was detected. It is possible that other peptides containing longer ZEB2 sequences are immediately degraded. The degradation is possible more efficient than in cytosol, because of close association of proteasomes with translational machinery. According to the secondary structure prediction program (JPred), ~383-437 region of ZEB2 represents a beta-strand, and LGV motif is required to complete the formation of a secondary structure of the domain. This observation is consistent with a proposed model, in

which LGV clusters impose translation pauses preventing protein folding. This permits degradation of unfolded peptides and abrogates ZEB2 protein synthesis (Figure 5.1).

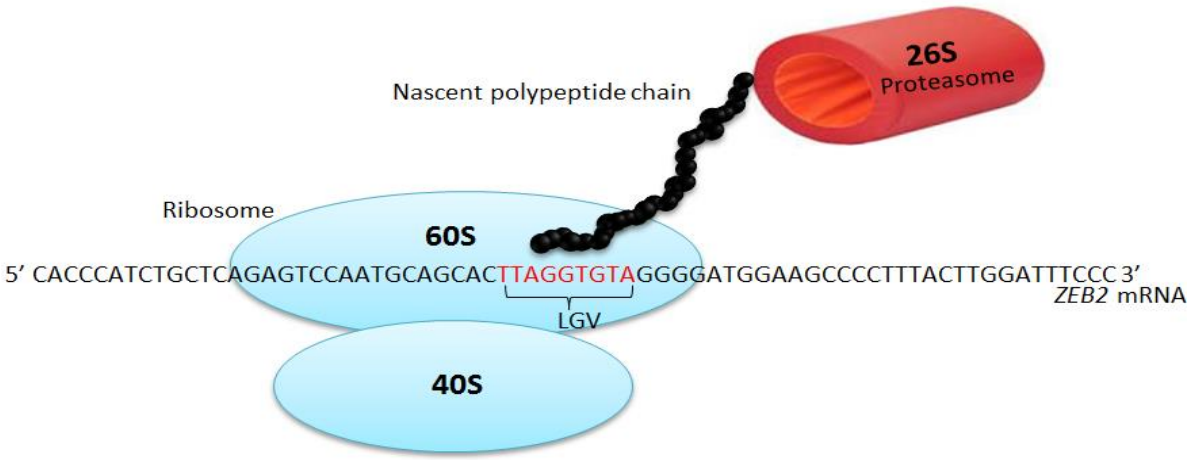


Figure 5.1: An illustration of the ribosome stalling during translation elongation of *ZEB2* mRNA.

The physiological consequences of ZEB1- or ZEB2-driven EMT programs are often different. ZEB2 possesses some features of a tumour suppressor by maintaining cell differentiation or inducing cellular senescence and cell cycle (Mejlvang et al., 2007, Ozturk et al., 2006, Caramel et al., 2013). However, these features have not been attributed to ZEB1 and ZEB2-induced EMT programs contribute to malignant transformation (Caramel et al., 2013). From this study we conclude that the expression of ZEB2 but not ZEB1 is controlled by yet another mechanism operating at translational level. This finding indicates that presently uncharacterised mechanisms may exist which select between ZEB1 and ZEB2 expression and therefore determine the outcome of EMT programs.

Data presented here demonstrate that the contribution of LGV triplets to the translational control of ZEB2 protein is different in different cell lines. Indeed, in MDA-468 cells the expression of EGFP-ZEB2 fusion is hardly detectable, and rare-to-common codon mutations strongly increase protein level. In contrast, this difference is moderate in mesenchymal cell lines, SaOs-2 or H1299 (Figure 4.13). Currently it remains not clear which factors are responsible for these differences. We propose that the expression level and the availability of transfer RNA (tRNA) corresponding to rare codons may represent such a factor. Transfer RNA genes are transcribed by RNAPolIII, and the regulation of their expression in different

cell backgrounds remains poorly studied. In particular, this is due to the fact that individual tRNA isoacceptors are encoded by several (up to 32) genes. In different study, we found that expression of rare tRNA for L and V are expressed at lower levels in epithelial carcinoma of different origins (A431 and MDA-MB-468 cell lines) compared to mesenchymal cells (H1299 and SaOS-2 cell lines). In addition, the nucleotide sequence analyses of rare isoacceptor valine tRNA has shown that the same genes are active in non-related SaOs-2 and MDA-MB-468 cells. In human genome, tRNAV-UAC is encoded by five genes, three of which remain inactive in both cell lines (personal communication with Dr E. Tulchinsky). Differential transcriptional activity of tRNA genes is in line with the ChIP seq data indicating that only 40%-50% of all tRNA gene promoters are bound by PolIII *in vivo* (Oler et al., 2010, Canella et al., 2012). Consistent with the differences in tRNAL-UAA and tRNAV-UAC levels in epithelial and mesenchymal cell lines, there are tissue- and cell line-specific differences in the occupancies of tRNA gene promoters by PolIII (Oler et al., 2010). Though mechanisms of transcriptional control of individual tRNA genes remain poorly understood, it has been shown that PolIII-occupied tRNA gene promoters are co-localised with genes actively transcribed by PolII in open chromatin regions (Barski et al., 2010). Given that there are global differences in gene expression patterns between epithelial and mesenchymal cells, one can propose that activation of EMT programs executed by ZEB1 or other EMT transcription factors may generate conditions permissive for ZEB2 expression. Involvement of ZEB2 in the EMT may alter the configuration of the program by shifting cells undergoing EMT towards cell cycle arrest and/or senescence.

Further studies addressing the genomic localisation of the genes coding for rare tRNA isoacceptors and linking EMT with tRNA field are required to test this hypothesis. Additionally, the dynamic measurements of ZEB2 translation and degradation rates can also be used to investigate fundamental principles of ZEB2 protein expression. Overall, better understanding of the configuration of EMT networks may open new perspective for targeting cancer metastases and development of new cancer therapies.

APPENDICES

Appendix 1

Appendix 1: List of chemicals, reagents and items used in this study

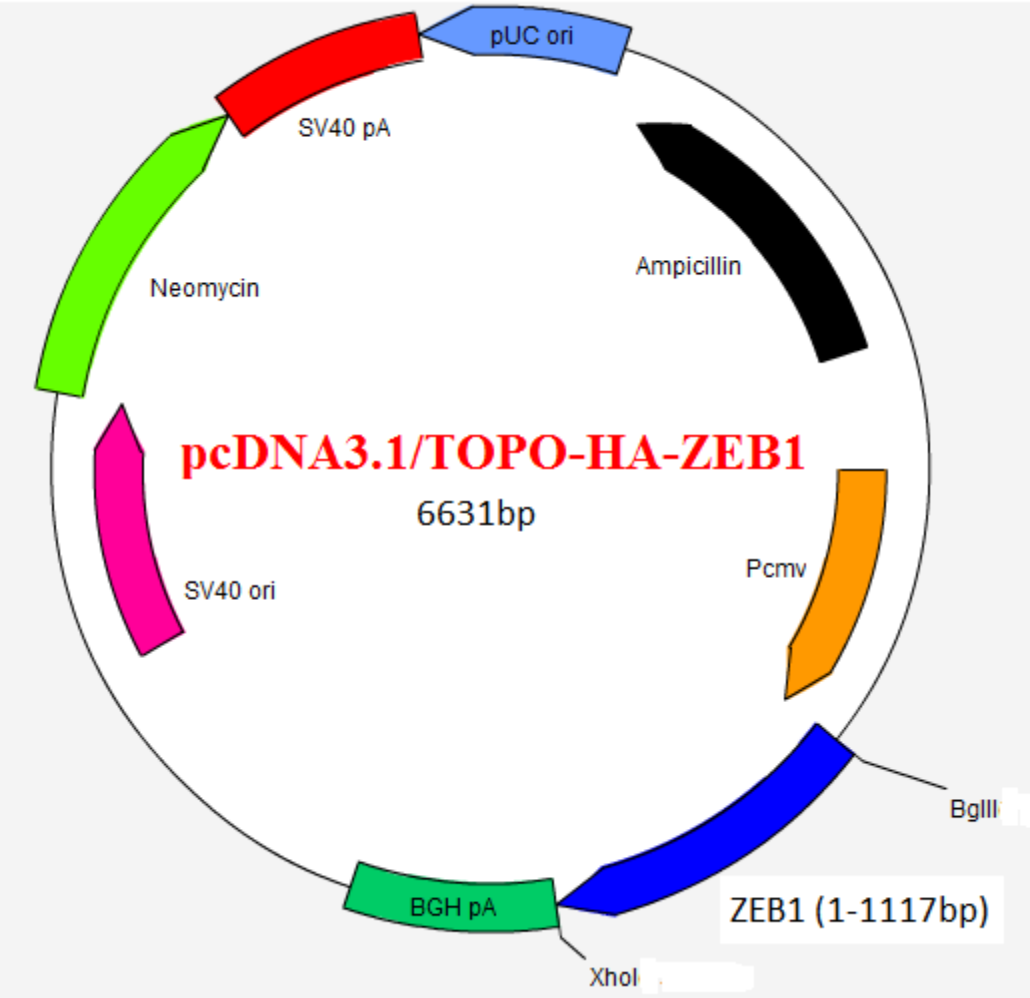
Reagents/supplements/items	Company	Catalogue number
Cell culture		
Dulbecco's modified eagle medium (DMEM) high glucose with L-glutamine	PAA, Laboratories, Inc., UK	E15-843
Roswell Park Memorial Institute (RPMI) 1640 with L-glutamine	PAA, Laboratories, Inc., UK	E15-040
McCoy's 5A (McCoy 5A) with L-glutamine	PAA, Laboratories, Inc., UK	E15-823
Heat inactivated foetal bovine serum (FBS)	PAA, Laboratories, Inc., UK	A11-151
100× Penicillin-Streptomycin (PS)	PAA, Laboratories, Inc., UK	P11-010
10× Trypsin EDTA (TE) UV inactivated	PAA, Laboratories, Inc., UK	L11-659
100× MEM Non-essential amino acid (NEAA) without L-glutamine	Invitrogen, UK	11140-035
Tissue culture flasks 25 cm^2	PAA, Laboratories, Inc., UK	PAA70025x
Tissue culture flasks 75 cm^2	PAA, Laboratories, Inc., UK	PAA70075x
Tissue culture flasks 175 cm^2	PAA, Laboratories, Inc., UK	PAA71175x
6-well plates	PAA, Laboratories, Inc., UK	30006x
96-well plates	PAA, Laboratories, Inc., UK	6555180
15 ml centrifuge tubes	Fisher Scientific, UK	CFT-900-031F
50 ml centrifuge tubes	Fisher Scientific, UK	CFT-420-077H
6 cm cell culture dishes	Greiner, UK	628960
10 cm cell culture dishes	Greiner, UK	664160
Freezing-cryotubes	Greiner, UK	123263
Cryo-tube container	Thermo Scientific, UK	CRY-120-010T
Coulter Counter vial	Greiner Bio One, UK	668102
Isoton®II diluent	Beckman Coulter Inc., UK	8448011
Electroporation cuvette (4mm)	Geneflow, UK	E6-0076
Ingenio® electroporation solution	Geneflow, UK	E7-0516
JetPRIMETM DNA transfection reagent	Source BioScience, UK	114-01
Cycloheximide (CHX)	Sigma-Aldrich, Dorset, UK	C1988-1G
Chloroquine	Sigma-Aldrich, Dorset, UK	C6628-25G
Dimethyl sulphoxide (DMSO)	Sigma-Aldrich, Dorset, UK	D5879
MG-132	Calbiochem, UK	474790
Phosphate Buffered Saline (PBS) pH7.3 tablets	Oxoid Ltd, UK	BR0014
Protein analysis		
Ammonium persulphate	Sigma-Aldrich, Dorset, UK	A3678-100G
N,N,N',N'-tetramethyletylenediamine (TEMED)	Sigma-Aldrich, Dorset, UK	T9281
2-mercaptoethanol	Sigma-Aldrich, Dorset, UK	M7154
Bromophenol blue	Sigma-Aldrich, Dorset, UK	B-8026

30% Protogel Acrylamide	Geneflow, UK	A2-0072
20% solution Sodium Dodecyl Sulphate (SDS)	Geneflow, UK	B9-0038
PageRuler™ Plus Prestained Protein Ladder	Thermo Scientific, UK	SM1811
Pierce BCA protein assay kit	Thermo Scientific, UK	23227
Ponceau S solution	Sigma-Aldrich, Dorset, UK	78376-10G-F
Tween-20	Sigma-Aldrich, Dorset, UK	P5927
CL xposure film (180mmx240mm)	Thermo Scientific, UK	34089
Pierce ECL Western Blotting Substrate	Thermo Scientific, UK	32106
Immobilon-PVDF (polyvinylidenedifloride) membrane (26.5cm×3.75m)	Millipore, Bedford MA, USA	IPVH 00010
Filter paper	Amersham Bioscience, UK	SE1141
GFP-Trap®_A	ChromoTek, UK	GTAK-20
Brilliant Blue G Acros Organics	Fisher Scientific, UK	19148-0050
Coomassie (Bradford) Protein Assay Kit	Thermo Scientific, UK	23200
Cloning, DNA and RNA analysis		
Ethidium bromide (EtBr)	Sigma-Aldrich, UK	46067-50ML-F
Chloroform:Isoamyl alcohol (24:1)	Sigma-Aldrich, UK	C-0549
S.O.C media	Sigma-Aldrich, UK	S1797
Ampicillin Na salt	Sigma-Aldrich, UK	A9518-5G
Agarose	Geneflow, UK	A4-0700
Low ranger 100 bp DNA ladder	Geneflow, UK	11500
High ranger 1 kb DNA ladder	Geneflow, UK	11900
TRIzol® reagent	Invitrogen, UK	15596-026
Luria Broth base (Miller's LB Broth Base)	Invitrogen, UK	12795-027
LB agar (Lennox L agar)	Invitrogen, UK	22700-025
Kanamycin	Invitrogen, UK	11815-024
QuikChange II site directed mutagenesis kit	Agilent Technologies, UK	200523
QuikChange Lightning Multi site-directed mutagenesis kit	Agilent Technologies, UK	210515
KOD Hot Start Master Mix	Calbiochem, UK	71482-3
Wizard® SV Gel and PCR Clean-up System Kit	Promega, USA	A9281
RNeasy Plus Mini Kit	Qiagen, UK	74134
RevertAid™ H Minus First Strand cDNA Synthesis Kit	Fermentas, UK	K1631
TaqMan® Fast Univ PCR Master Mix (2×) No AmpEraser® UNG	Applied Biosystem, UK	4352042
α-select chemically competent cells (Bronze)	Bioline, UK	Bio-85025
BamHI restriction enzyme (with NEBuffer 3+ BSA)	New England Biolabs, UK	R0136S
BglII restriction enzyme (with NEBuffer 3)	New England Biolabs, UK	R0144S
HindIII restriction enzyme (with NEBuffer 2)	New England Biolabs, UK	R0104S
KpnI restriction enzyme (with NEBuffer 1+ BSA)	New England Biolabs, UK	R0142S
SalI restriction enzyme (with NEBuffer3+ BSA)	New England Biolabs, UK	R0138S
NdeI restriction enzyme (with NEBuffer 4)	New England Biolabs, UK	R0111S

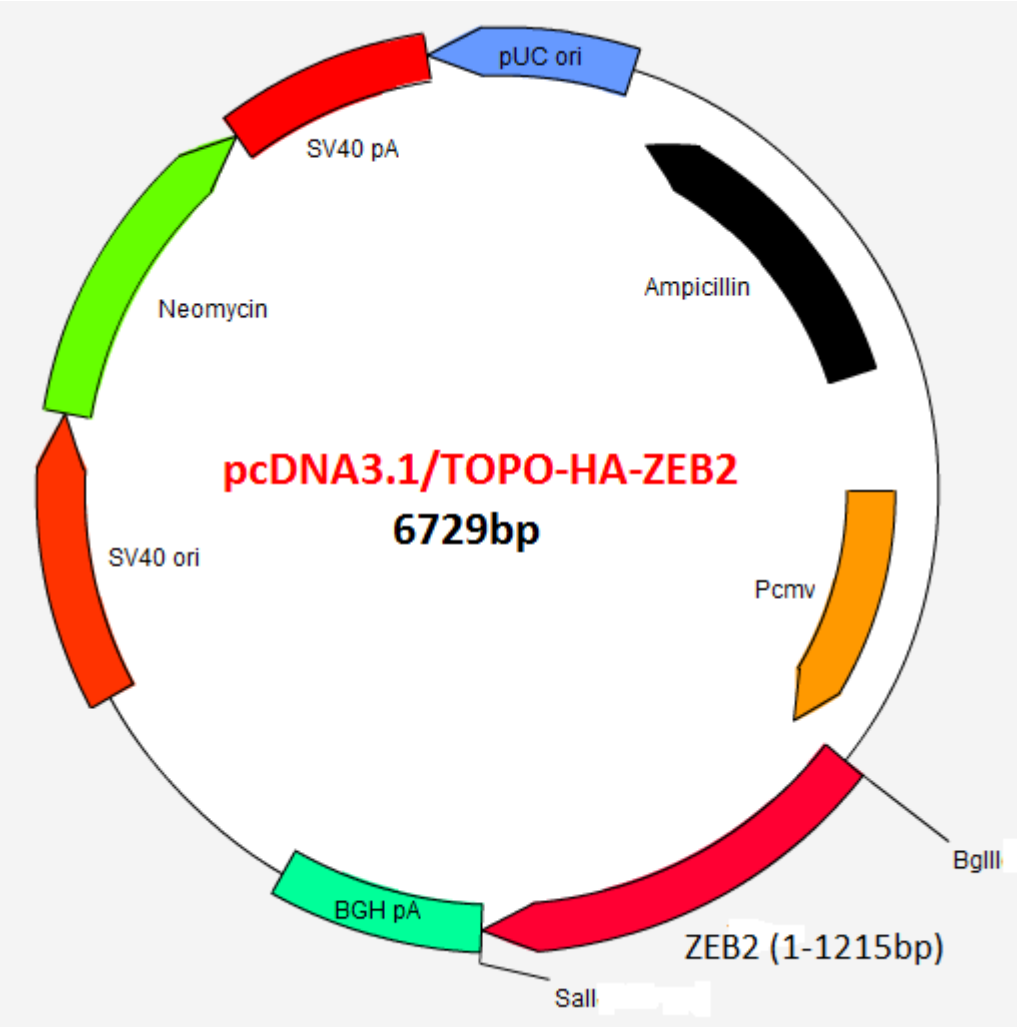
AgeI restriction enzyme (with NEBuffer 1)	New England Biolabs, UK	R0552S
NgoMIV restriction enzyme (with NEBuffer 4)	New England Biolabs, UK	R0564S
XhoI restriction enzyme (with NEBuffer 4+ BSA)	New England Biolabs, UK	R0146S
T4 DNA Ligase	Invitrogen, UK	15224-041
Polysome profiling		
Heparin Sodium Salt from porcine Intestinal Mucosa	Sigma-Aldrich, UK	H3393-500KU
Sucrose	Sigma-Aldrich, UK	84097-5KG
Trichloroacetic acid (TCA)	Sigma-Aldrich, UK	T6399-250G
Sorvall Polyallomer (PA) thin-walled 12 ml tubes	Seton Scientific Inc., USA	03699

Appendix 2

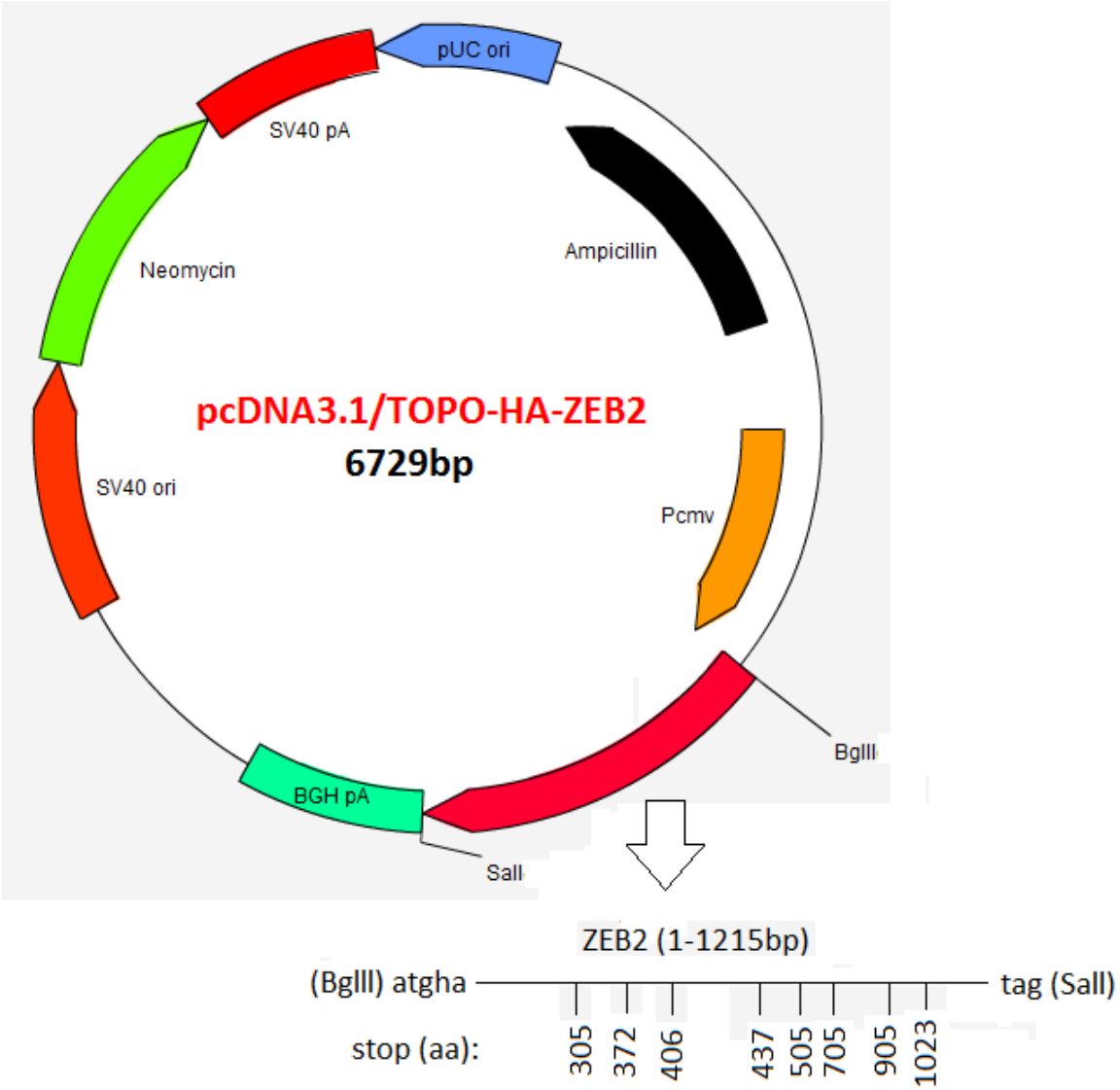
(A) Plasmid map for pcDNA™3.1/TOPO®-HA-ZEB1. Plasmid map produced using EZ Plasmid Map V1.3 (University of Texas, US).



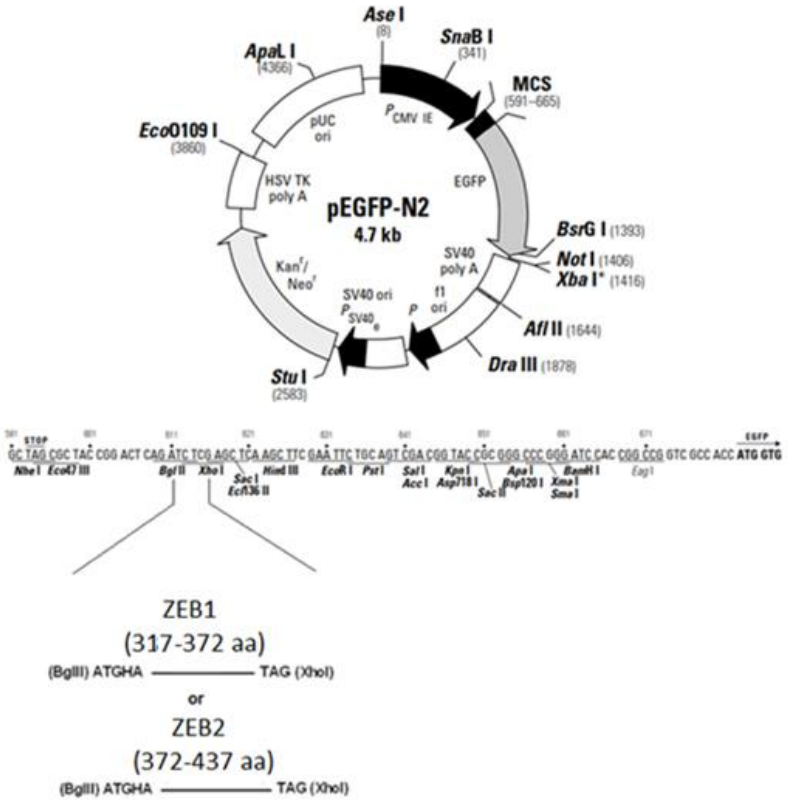
(B) Plasmid map for pcDNA™3.1/TOPO®-HA-ZEB2. Plasmid map produced using EZ Plasmid Map V1.3 (University of Texas, US).



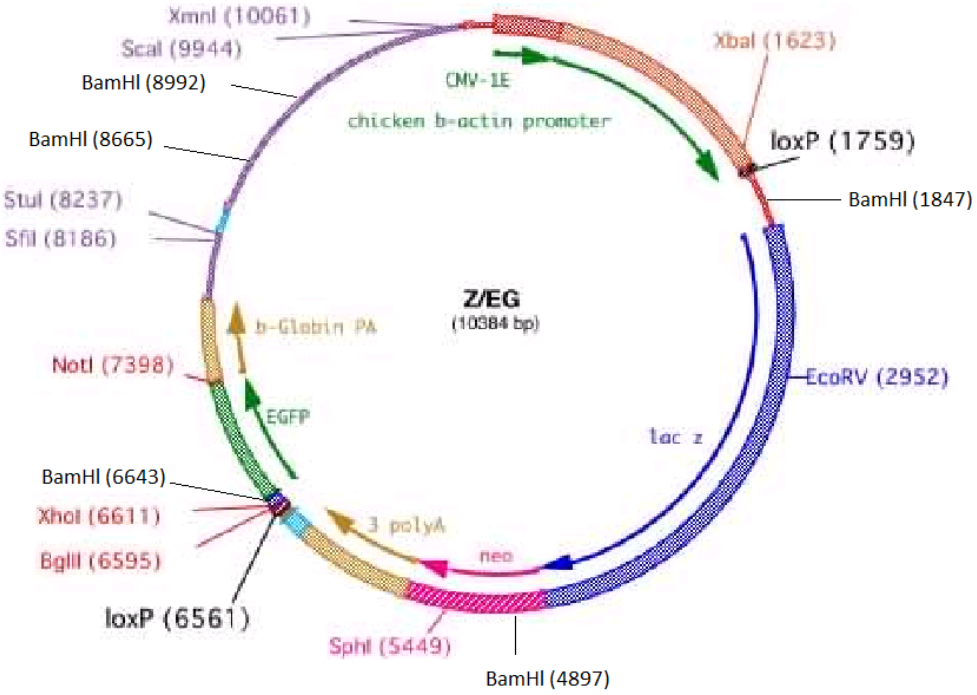
(C) Plasmid map for pcDNA™3.1/TOPO®-HA-ZEB2 (STOP 305 or 372 or 406 or 437 or 505 or 705 or 904 or 1023aa). Plasmid map produced using EZ Plasmid Map V1.3 (University of Texas, US).



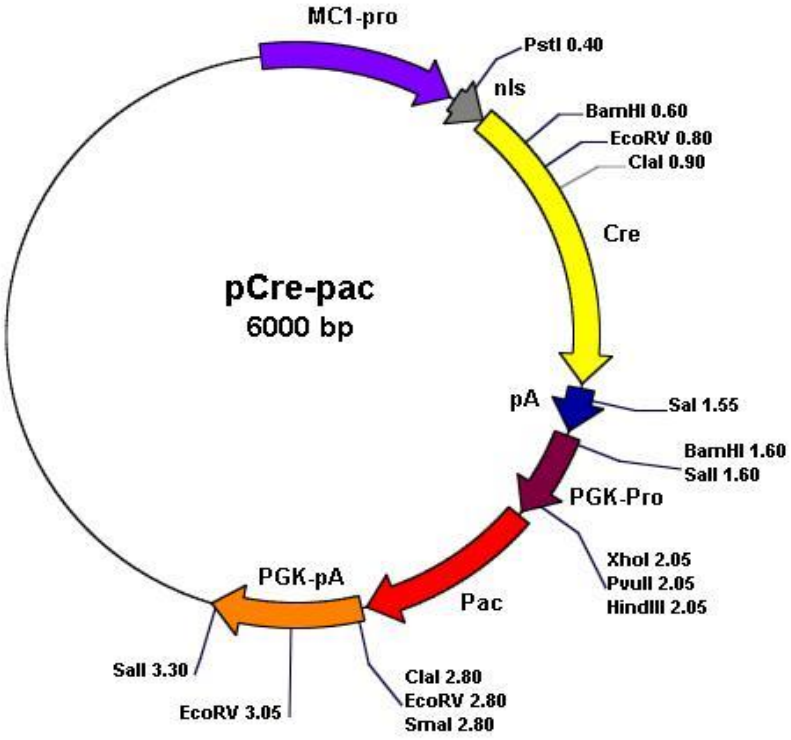
(D) Plasmid map for EGFP-ZEB1 (317-372aa) or EGFP-ZEB2 (372-437aa). Plasmid map modified from Clontech laboratories, UK.



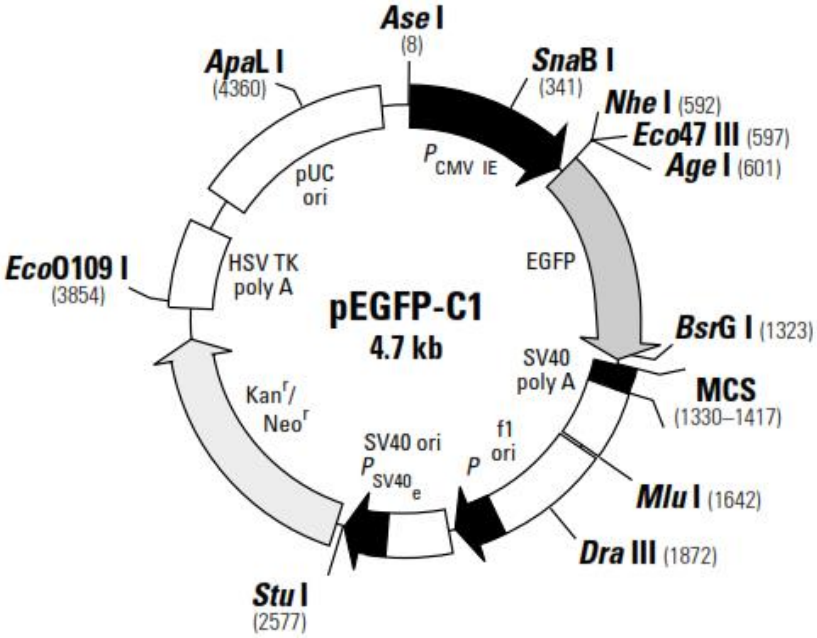
(E) Plasmid map for pZ/EG. Plasmid map produced using Bio-log Plasm (Bio-Log, Toulouse, France).



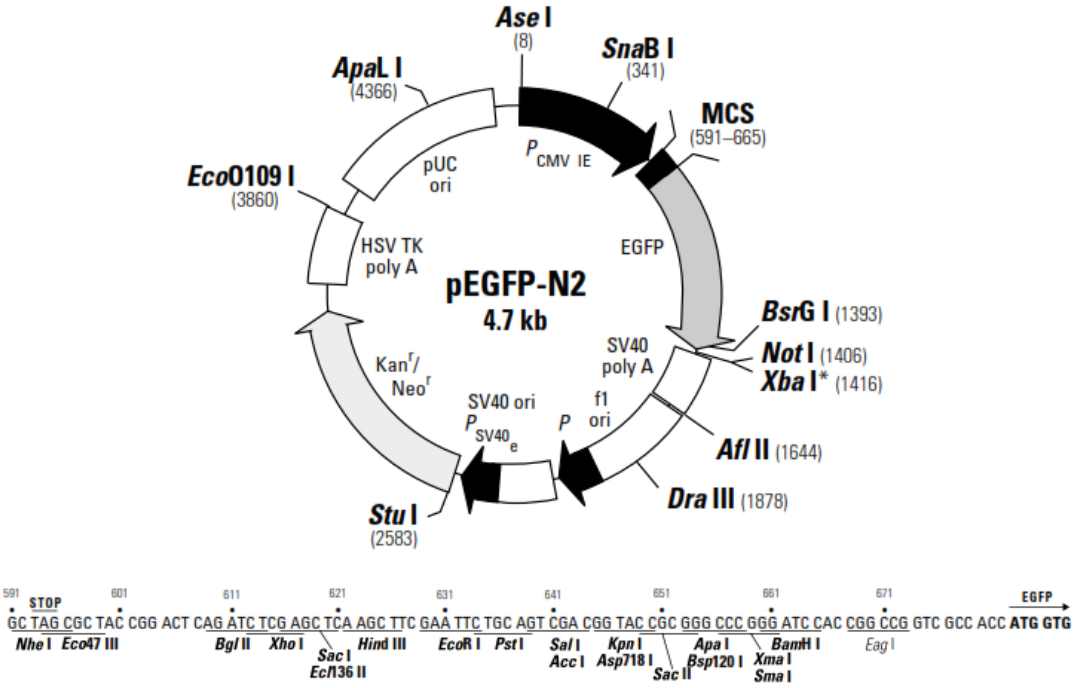
(F) Plasmid map for pCre-pac expression vector. Plasmid map produced using Bio-log Plasm (Bio-Log, Toulouse, France).



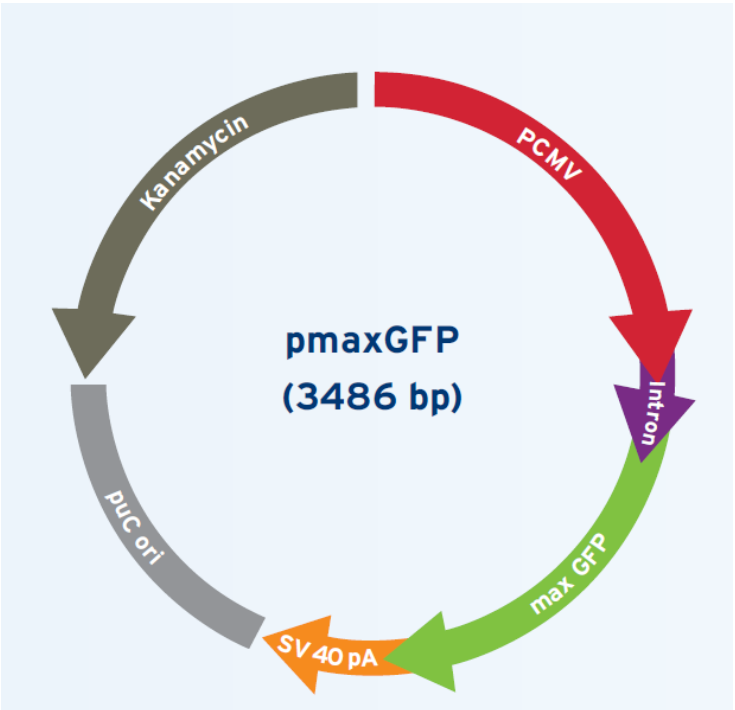
(G) Plasmid map for pEGFP-C1. Plasmid map taken from Clontech laboratories, UK.



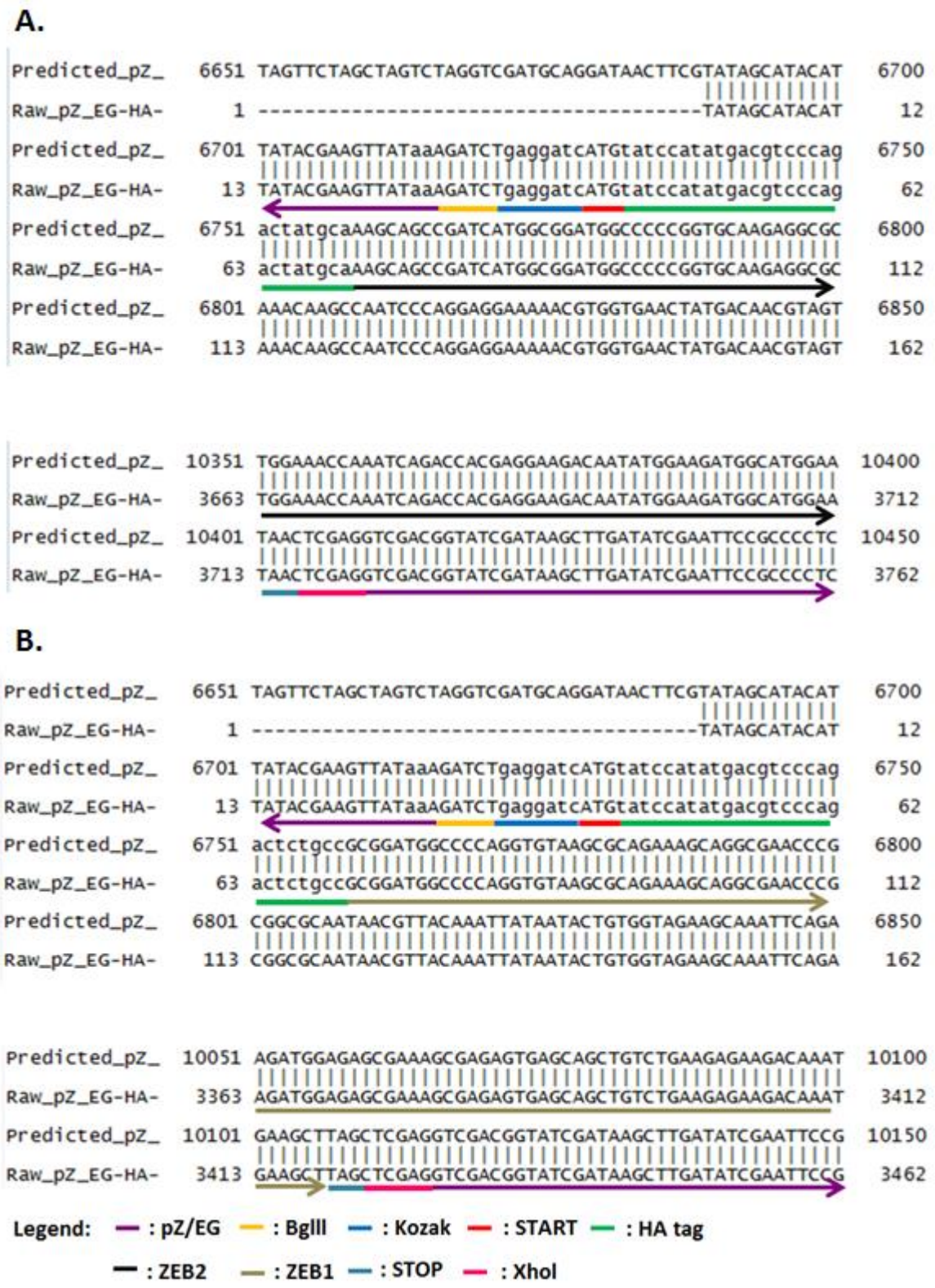
(H) Plasmid map for pEGFP-N2. Plasmid map taken from Clontech laboratories, UK.



(I) Plasmid map for pmaxGFP. Plasmid map taken from Lonza, UK.

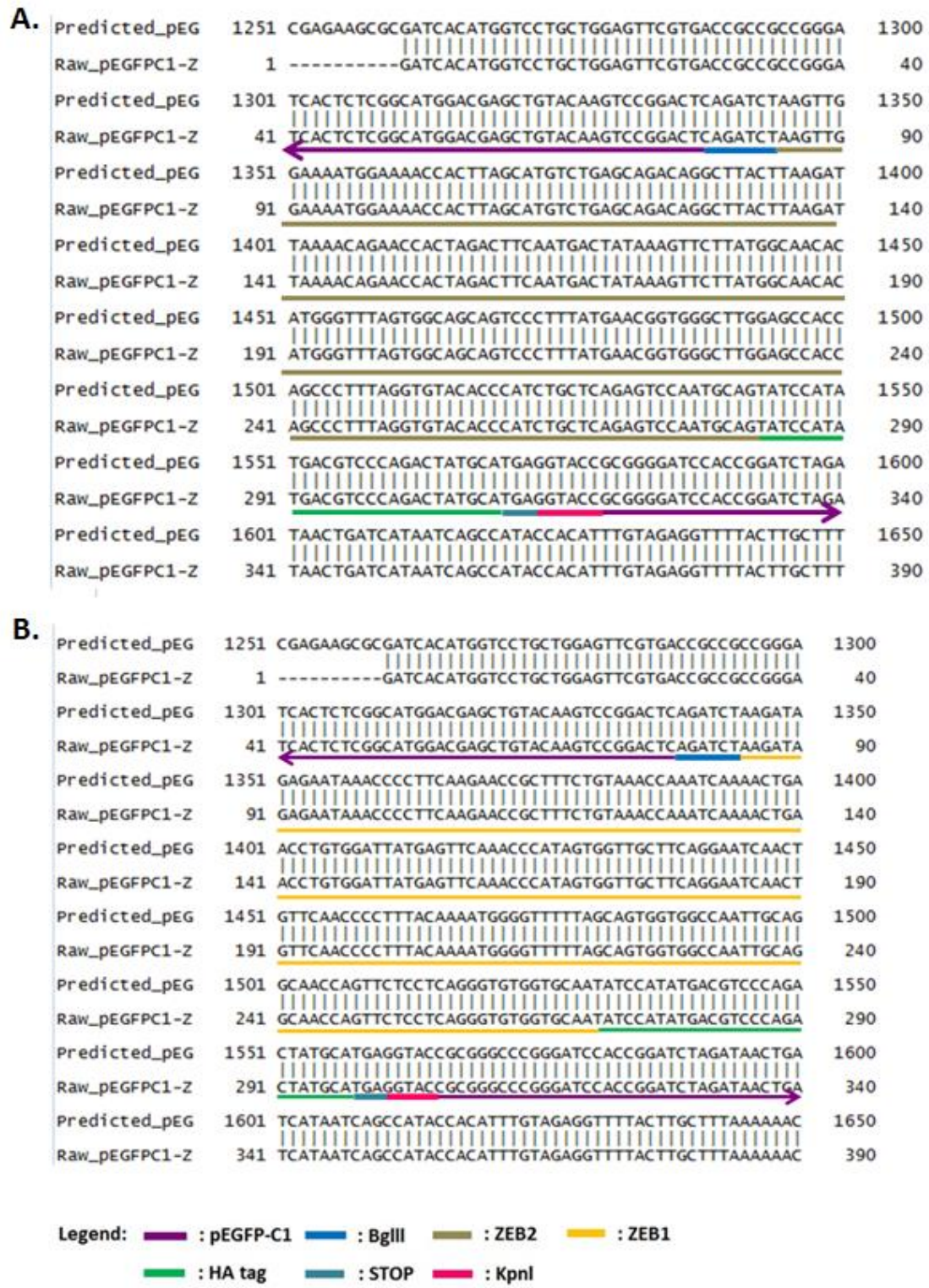


Appendix 3



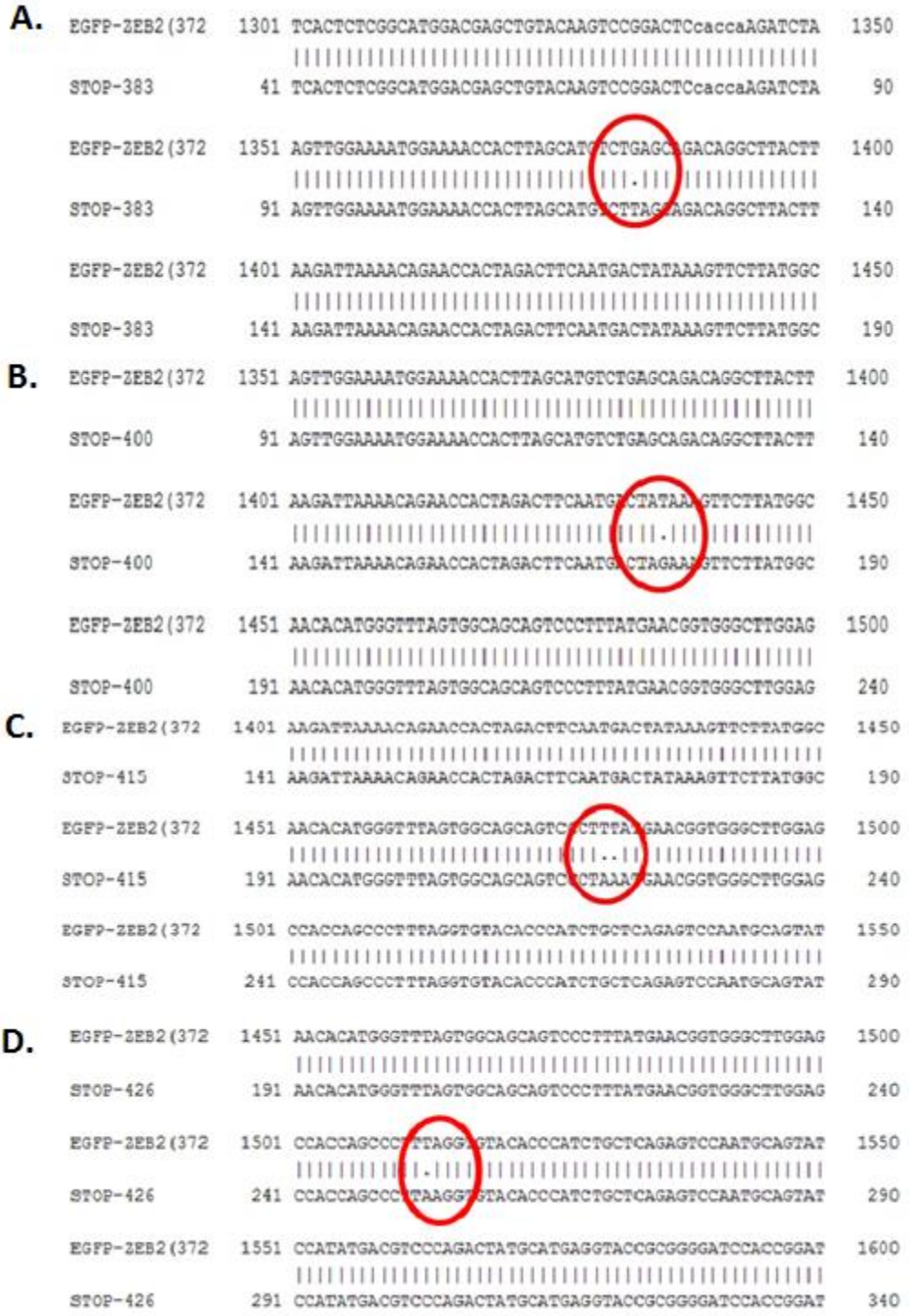
Appendix 3: Sequence alignment of vector-insert junctions in (A) pZ/EG-HA-ZEB2 and (B) pZ/EG-HA-ZEB1 plasmids. All sequencing data were analysed using the EMBOSS Pairwise Alignment Algorithm, EMBL-EBI Database. For complete sequence alignment of pZ/EG-HA-ZEB2 and pZ/EG-HA-ZEB1, see Attachment 1 and 2 respectively in the attached CD.

Appendix 4



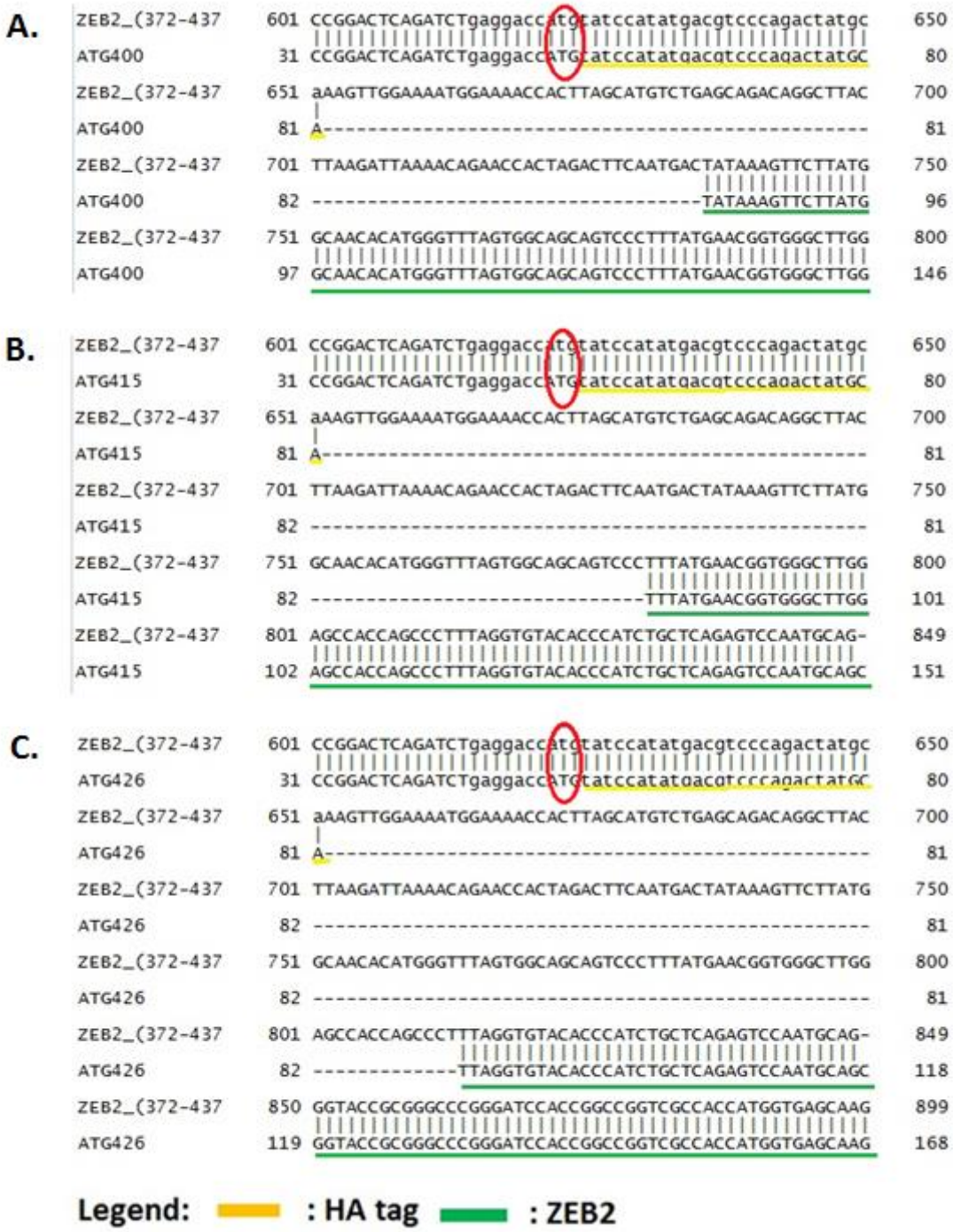
Appendix 4: Sequence alignment of vector-insert junctions in (A) pEGFP-C1-ZEB2 (372-437aa)-HA and (B) pEGFP-C1-ZEB1 (317-372aa)-HA. All sequencing data were analysed using the EMBOSS Pairwise Alignment Algorithm, EMBL-EBI Database. For complete sequence alignment of pEGFP-C1-ZEB2 (372-437aa)-HA and pEGFP-C1-ZEB1 (317-372aa)-HA, see Attachment 3 and 4 respectively in the attached CD.

Appendix 5



Appendix 5: Sequencing alignment of (A) EGFP-ZEB2 (372-437aa)-STOP383 (B) EGFP-ZEB2 (372-437aa)-STOP400 (C) EGFP-ZEB2 (372-437aa)-STOP415 and (D) EGFP-ZEB2 (372-437aa)-STOP426 (EMBOSS Pairwise Alignment Algorithm, EMBL-EBI Database). The base substitutions are indicated with the red circles. For complete sequence alignment of EGFP-ZEB2 (372-437aa)-STOP383, EGFP-ZEB2 (372-437aa)-STOP400, EGFP-ZEB2 (372-437aa)-STOP415 and EGFP-ZEB2 (372-437aa)-STOP426, see Attachment 5, 6, 7 and 8 respectively in the attached CD.

Appendix 6

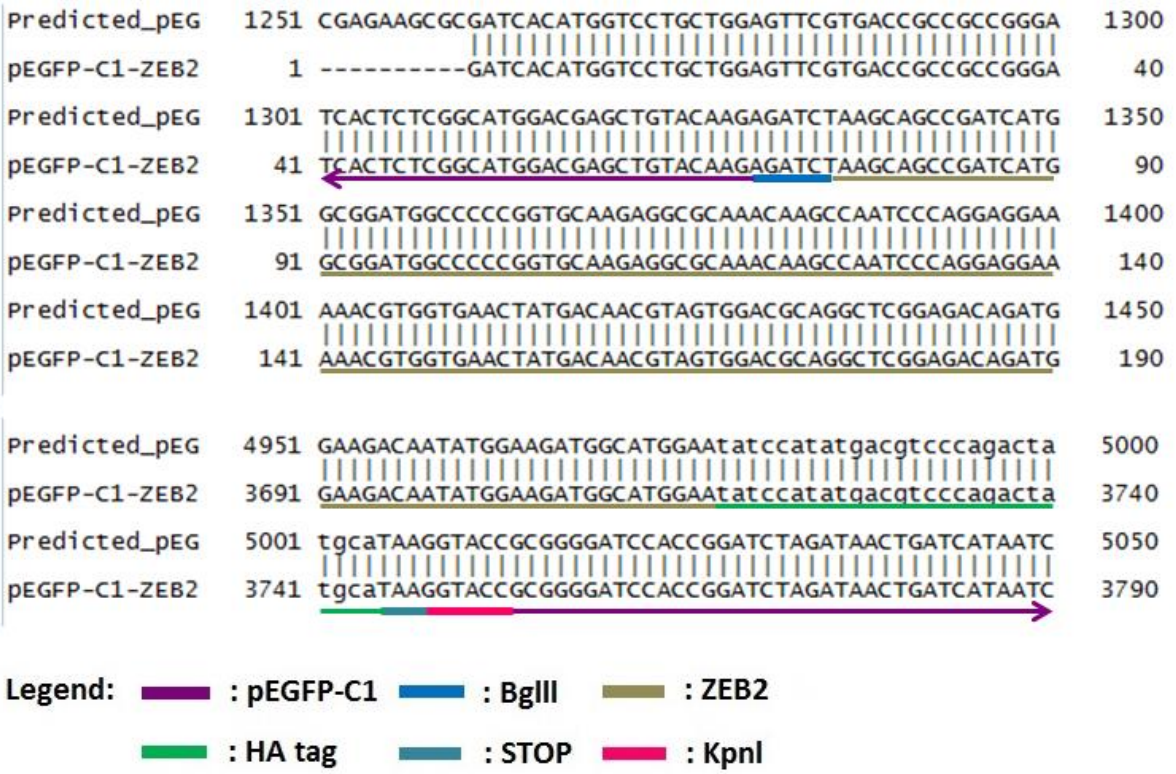


Appendix 6: Sequencing alignment of (A) ZEB2 (372-437aa)-ATG400-EGFP (B) ZEB2 (372-437aa)-ATG415-EGFP and (C) ZEB2 (372-437aa)-ATG426-EGFP (EMBOSS Pairwise Alignment Algorithm, EMBL-EBI Database). The base substitutions are indicated with the red circles. For complete sequence alignment of ZEB2 (372-437aa)-ATG400-EGFP, ZEB2 (372-437aa)-ATG415-EGFP and ZEB2 (372-437aa)-ATG426-EGFP see Attachment 9, 10 and 11 respectively in the attached CD.

Appendix 7

Appendix 7: Sequence alignment of EGFP-ZEB2 (372-437aa)-(RC-CC) showing the RC to CC substitutions (red circle). For complete sequence alignment see Attachment 12.

Appendix 8

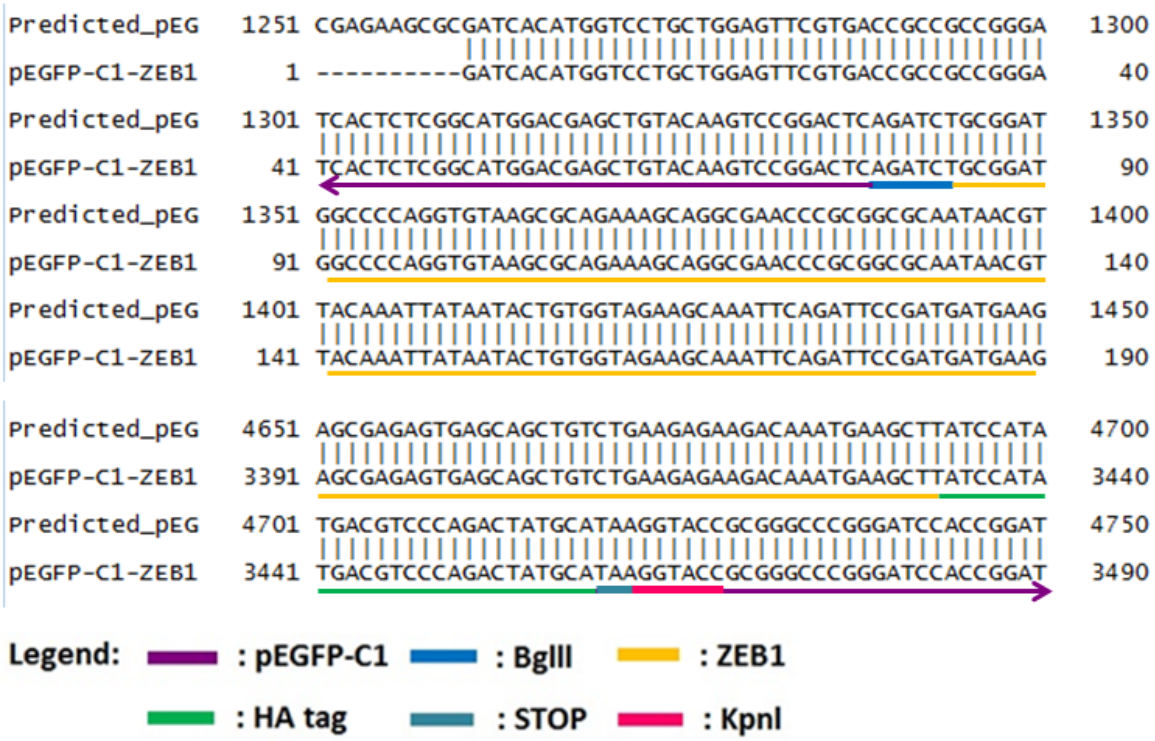


Appendix 8: Sequence alignment of vector-insert junctions in pEGFP-C1-ZEB2-HA plasmid. The sequencing data was analysed using the EMBOSS Pairwise Alignment Algorithm, EMBL-EBI Database. For complete sequence alignment see Attachment 13.

Appendix 9

Appendix 9: Sequence alignment of EGFP-ZEB2-(RC-CC) and EGFP-ZEB2 showing the RC to CC substitutions indicated with the red circles. For complete sequence alignment see Attachment 14.

Appendix 10



Appendix 10: Sequence alignment of vector-insert junctions in pEGFP-C1-ZEB1-HA plasmid. The sequencing data was analysed using the EMBOSS Pairwise Alignment Algorithm, EMBL-EBI Database. For complete sequence alignment of pEGFP-C1-ZEB1-HA, see Attachment 15.

Appendix 11

Appendix 11: Sequence alignment of EGFP-ZEB1-(CC-RC) showing the CC to RC substitutions indicated with the red circles. For complete sequence alignment see Attachment 16.

LIST OF ATTACHMENTS (See attached CD)

Attachment 1: Sequence alignment of vector-insert junctions in pZ/EG-ZEB1-HA

Attachment 2: Sequence alignment of vector-insert junctions in pZ/EG-ZEB2-HA

Attachment 3: Sequence alignment of vector-insert junctions in pEGFP-C1-ZEB2 (372-437aa)-HA

Attachment 4: Sequence alignment of vector-insert junctions in pEGFP-C1-ZEB1 (317-372aa)-HA

Attachment 5: Sequence alignment of EGFP-ZEB2 (372-437aa)-STOP383

Attachment 6: Sequence alignment of EGFP-ZEB2 (372-437aa)-STOP400

Attachment 7: Sequence alignment of EGFP-ZEB2 (372-437aa)-STOP415

Attachment 8: Sequence alignment of EGFP-ZEB2 (372-437aa)-STOP426

Attachment 9: Sequence alignment of ZEB2 (372-437aa)-ATG400-EGFP

Attachment 10: Sequence alignment of ZEB2 (372-437aa)-ATG415-EGFP

Attachment 11: Sequence alignment of ZEB2 (372-437aa)-ATG426-EGFP

Attachment 12: Sequence alignment of EGFP-ZEB2 (372-437aa)-(RC-CC) showing the RC to CC substitutions

Attachment 13: Sequence alignment of vector-insert junctions in pEGFP-C1-ZEB2-HA plasmid.

Attachment 14: Sequence alignment of EGFP-ZEB2-(RC-CC) and EGFP-ZEB2 showing the RC to CC substitutions indicated with the red circles.

Attachment 15: Sequence alignment of vector-insert junctions in pEGFP-C1-ZEB1-HA plasmid.

Attachment 16: Sequence alignment of EGFP-ZEB1-(CC-RC) showing the CC to RC substitutions indicated with the red circles.

Attachment 17: List of all components identified by mass spectrometric analysis of ZEB2 (372-437aa) and EGFP (control).

CD

REFERENCES

The Jackson Laboratory: Cre Expression Resource Web Site [Online]. Bar Harbor, Maine. Available: http://cre.jax.org/introduction.html [Accessed July 22 2013].

ABELL, A. N., JORDAN, N. V., HUANG, W., PRAT, A., MIDLAND, A. A., JOHNSON, N. L., GRANGER, D. A., MIECZKOWSKI, P. A., PEROU, C. M., GOMEZ, S. M., LI, L. & JOHNSON, G. L. 2011. MAP3K4/CBP-regulated H2B acetylation controls epithelial-mesenchymal transition in trophoblast stem cells. *Cell Stem Cell,* 8, 525-37.

ACUN, T., OZTAS, E., YAGCI, T. & YAKICIER, M. C. 2011. SIP1 is downregulated in hepatocellular carcinoma by promoter hypermethylation. *BMC Cancer,* 11, 223.

AIGNER, K., DAMPIER, B., DESCOVICH, L., MIKULA, M., SULTAN, A., SCHREIBER, M., MIKULITS, W., BRABLETZ, T., STRAND, D., OBRIST, P., SOMMERGRUBER, W., SCHWEIFER, N., WERNITZNIG, A., BEUG, H., FOISNER, R. & EGER, A. 2007. The transcription factor ZEB1 (deltaEF1) promotes tumour cell dedifferentiation by repressing master regulators of epithelial polarity. *Oncogene,* 26, 6979-88.

ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. 1990. Basic local alignment search tool. *J Mol Biol,* 215, 403-10.

ALVAREZ-DIAZ, S., VALLE, N., GARCIA, J. M., PENA, C., FREIJE, J. M., QUESADA, V., ASTUDILLO, A., BONILLA, F., LOPEZ-OTIN, C. & MUNOZ, A. 2009. Cystatin D is a candidate tumor suppressor gene induced by vitamin D in human colon cancer cells. *J Clin Invest,* 119, 2343-58.

ANAND, P. & STAMLER, J. S. 2012. Enzymatic mechanisms regulating protein S-nitrosylation: implications in health and disease. *J Mol Med (Berl),* 90, 233-44.

ANGOV, E. 2011. Codon usage: nature's roadmap to expression and folding of proteins. *Biotechnol J,* 6, 650-9.

ANSIEAU, S., BASTID, J., DOREAU, A., MOREL, A. P., BOUCHET, B. P., THOMAS, C., FAUVET, F., PUISIEUX, I., DOGLIONI, C., PICCININ, S., MAESTRO, R., VOELTZEL, T., SELMI, A., VALSESIA-WITTMANN, S., CARON DE FROMENTEL, C. & PUISIEUX, A. 2008. Induction of EMT by twist proteins as a collateral effect of tumor-promoting inactivation of premature senescence. *Cancer Cell,* 14, 79-89.

ARIMA, Y., INOUE, Y., SHIBATA, T., HAYASHI, H., NAGANO, O., SAYA, H. & TAYA, Y. 2008. Rb depletion results in deregulation of E-cadherin and induction of cellular phenotypic changes that are characteristic of the epithelial-to-mesenchymal transition. *Cancer Res,* 68, 5104-12.

BANNISTER, A. J. & MISKA, E. A. 2000. Regulation of gene expression by transcription factor acetylation. *Cell Mol Life Sci,* 57, 1184-92.

BANUMATHY, G. & CAIRNS, P. 2010. Signaling pathways in renal cell carcinoma. *Cancer Biol Ther,* 10, 658-64.

BARANWAL, S. & ALAHARI, S. K. 2009. Molecular mechanisms controlling E-cadherin expression in breast cancer. *Biochem Biophys Res Commun,* 384, 6-11.

BARITAKI, S., HUERTA-YEPEZ, S., SAHAKYAN, A., KARAGIANNIDES, I., BAKIRTZI, K., JAZIREHI, A. & BONAVIDA, B. 2010. Mechanisms of nitric oxide-mediated inhibition of EMT in cancer: inhibition of the metastasis-inducer Snail and induction of the metastasis-suppressor RKIP. *Cell Cycle,* 9, 4931-40.

BARSKI, A., CHEPELEV, I., LIKO, D., CUDDAPAH, S., FLEMING, A. B., BIRCH, J., CUI, K., WHITE, R. J. & ZHAO, K. 2010. Pol II and its associated epigenetic marks are present at Pol III-transcribed noncoding RNA genes. *Nat Struct Mol Biol,* 17, 629-34.

BATLLE, E., SANCHO, E., FRANCI, C., DOMINGUEZ, D., MONFAR, M., BAULIDA, J. & GARCIA DE HERREROS, A. 2000. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat Cell Biol,* 2, 84-9.

BECKER, K. F., ATKINSON, M. J., REICH, U., BECKER, I., NEKARDA, H., SIEWERT, J. R. & HOFLER, H. 1994. E-cadherin gene mutations provide clues to diffuse type gastric carcinomas. *Cancer Res,* 54, 3845-52.

BELLET, D., LAZAR, V., BIECHE, I., PARADIS, V., GIOVANGRANDI, Y., PATERLINI, P., LIDEREAU, R., BEDOSSA, P., BIDART, J. M. & VIDAUD, M. 1997. Malignant transformation of nontrophoblastic cells is associated with the expression of chorionic gonadotropin beta genes normally transcribed in trophoblastic cells. *Cancer Res,* 57, 516-23.

BENHAM, F., COTTELL, D. C., FRANKS, L. M. & WILSON, P. D. 1977. Alkaline phosphatase activity in human bladder tumor cell lines. *J Histochem Cytochem,* 25, 266-74.

BENNETZEN, J. L. & HALL, B. D. 1982. Codon selection in yeast. *J Biol Chem,* 257, 3026-31.

BERX, G., STAES, K., VAN HENGEL, J., MOLEMANS, F., BUSSEMAKERS, M. J., VAN BOKHOVEN, A. & VAN ROY, F. 1995. Cloning and characterization of the human invasion suppressor gene E-cadherin (CDH1). *Genomics,* 26, 281-9.

BERX, G. & VAN ROY, F. 2009. Involvement of members of the cadherin superfamily in cancer. *Cold Spring Harb Perspect Biol,* 1, a003129.

BINDELS, S., MESTDAGT, M., VANDEWALLE, C., JACOBS, N., VOLDERS, L., NOEL, A., VAN ROY, F., BERX, G., FOIDART, J. M. & GILLES, C. 2006. Regulation of vimentin by SIP1 in human epithelial breast tumor cells. *Oncogene,* 25, 4975-85.

BOLOS, V., PEINADO, H., PEREZ-MORENO, M. A., FRAGA, M. F., ESTELLER, M. & CANO, A. 2003. The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. *J Cell Sci,* 116, 499-511.

BRABANT, G., HOANG-VU, C., CETIN, Y., DRALLE, H., SCHEUMANN, G., MOLNE, J., HANSSON, G., JANSSON, S., ERICSON, L. E. & NILSSON, M. 1993. E-cadherin: a differentiation marker in thyroid malignancies. *Cancer Res,* 53, 4987-93.

BRABLETZ, S., BAJDAK, K., MEIDHOF, S., BURK, U., NIEDERMANN, G., FIRAT, E., WELLNER, U., DIMMLER, A., FALLER, G., SCHUBERT, J. & BRABLETZ, T. 2011. The ZEB1/miR-200 feedback loop controls Notch signalling in cancer cells. *EMBO J,* 30, 770-82.

BRABLETZ, S. & BRABLETZ, T. 2010. The ZEB/miR-200 feedback loop--a motor of cellular plasticity in development and cancer? *EMBO Rep,* 11, 670-7.

BRACKEN, C. P., GREGORY, P. A., KOLESNIKOFF, N., BERT, A. G., WANG, J., SHANNON, M. F. & GOODALL, G. J. 2008. A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition. *Cancer Res,* 68, 7846-54.

BRINKLEY, B. R., BEALL, P. T., WIBLE, L. J., MACE, M. L., TURNER, D. S. & CAILLEAU, R. M. 1980. Variations in cell form and cytoskeleton in human breast carcinoma cells in vitro. *Cancer Res,* 40, 3118-29.

BROWNE, G., SAYAN, A. E. & TULCHINSKY, E. 2010. ZEB proteins link cell motility with cell cycle control and cell survival in cancer. *Cell Cycle,* 9, 886-91.

CANELLA, D., BERNASCONI, D., GILARDI, F., LEMARTELOT, G., MIGLIAVACCA, E., PRAZ, V., COUSIN, P., DELORENZI, M. & HERNANDEZ, N. 2012. A multiplicity of factors contributes to selective RNA polymerase III occupancy of a subset of RNA polymerase III genes in mouse liver. *Genome Res,* 22, 666-80.

CANO, A., PEREZ-MORENO, M. A., RODRIGO, I., LOCASCIO, A., BLANCO, M. J., DEL BARRIO, M. G., PORTILLO, F. & NIETO, M. A. 2000. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat Cell Biol,* 2, 76-83.

CARAMEL, J., PAPADOGEORGAKIS, E., HILL, L., BROWNE, G. J., RICHARD, G., WIERINCKX, A., SALDANHA, G., OSBORNE, J., HUTCHINSON, P., TSE, G., LACHUER, J., PUISIEUX, A., PRINGLE, J. H., ANSIEAU, S. & TULCHINSKY, E. 2013. A Switch in the Expression of Embryonic EMT-Inducers Drives the Development of Malignant Melanoma. *Cancer Cell,* 24, 466-80.

CARVER, E. A., JIANG, R., LAN, Y., ORAM, K. F. & GRIDLEY, T. 2001. The mouse snail gene encodes a key regulator of the epithelial-mesenchymal transition. *Mol Cell Biol,* 21, 8184-8.

CHEN, T., YUAN, D., WEI, B., JIANG, J., KANG, J., LING, K., GU, Y., LI, J., XIAO, L. & PEI, G. 2010. E-cadherin-mediated cell-cell contact is critical for induced pluripotent stem cell generation. *Stem Cells,* 28, 1315-25.

CHESNOKOV, V., GONG, B., SUN, C. & ITAKURA, K. 2014. Anti-cancer activity of glucosamine through inhibition of N-linked glycosylation. *Cancer Cell Int,* 14, 45.

CHOMCZYNSKI, P. & SACCHI, N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem,* 162, 156-9.

CHRISTIANSEN, J. J. & RAJASEKARAN, A. K. 2006. Reassessing epithelial to mesenchymal transition as a prerequisite for carcinoma invasion and metastasis. *Cancer Res,* 66, 8319-26.

CHRISTOFFERSEN, N. R., SILAHTAROGLU, A., OROM, U. A., KAUPPINEN, S. & LUND, A. H. 2007. miR-200b mediates post-transcriptional repression of ZFHX1B. *RNA,* 13, 1172-8.

CHU, P. Y., HU, F. W., YU, C. C., TSAI, L. L., YU, C. H., WU, B. C., CHEN, Y. W., HUANG, P. I. & LO, W. L. 2013. Epithelial-mesenchymal transition transcription factor ZEB1/ZEB2 co-expression predicts poor prognosis and maintains tumor-initiating properties in head and neck cancer. *Oral Oncol,* 49, 34-41.

CHUA, H. L., BHAT-NAKSHATRI, P., CLARE, S. E., MORIMIYA, A., BADVE, S. & NAKSHATRI, H. 2007. NF-kappaB represses E-cadherin expression and enhances epithelial to mesenchymal transition of mammary epithelial cells: potential involvement of ZEB-1 and ZEB-2. *Oncogene,* 26, 711-24.

CHUNG, S. S., CHOI, H. H., CHO, Y. M., LEE, H. K. & PARK, K. S. 2006. Sp1 mediates repression of the resistin gene by PPARgamma agonists in 3T3-L1 adipocytes. *Biochem Biophys Res Commun,* 348, 253-8.

CIECHANOVER, A. 1998. The ubiquitin-proteasome pathway: on protein death and cell life. *EMBO J,* 17, 7151-60.

CLARHAUT, J., GEMMILL, R. M., POTIRON, V. A., AIT-SI-ALI, S., IMBERT, J., DRABKIN, H. A. & ROCHE, J. 2009. ZEB-1, a repressor of the semaphorin 3F tumor suppressor gene in lung cancer cells. *Neoplasia,* 11, 157-66.

COHEN, P. 2000. The regulation of protein function by multisite phosphorylation--a 25 year update. *Trends Biochem Sci,* 25, 596-601.

COMIJN, J., BERX, G., VERMASSEN, P., VERSCHUEREN, K., VAN GRUNSVEN, L., BRUYNEEL, E., MAREEL, M., HUYLEBROECK, D. & VAN ROY, F. 2001. The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Mol Cell,* 7, 1267-78.

CONIDI, A., VAN DEN BERGHE, V. & HUYLEBROECK, D. 2013. Aptamers and Their Potential to Selectively Target Aspects of EGF, Wnt/beta-Catenin and TGFbeta-Smad Family Signaling. *Int J Mol Sci,* 14, 6690-719.

COOPER, G. M. 2000a. The Cell: A Molecular Approach. *Protein Folding and Processing.* 2nd edition ed. Sunderland (MA): Sinauer Associates.

COOPER, G. M. 2000b. The Cell: A Molecular Approach. *Regulation of Protein Function.* 2nd edition ed. Sunderland (MA): Sinauer Associates.

COPELAND, N. G., JENKINS, N. A., GILBERT, D. J., EPPIG, J. T., MALTAIS, L. J., MILLER, J. C., DIETRICH, W. F., WEAVER, A., LINCOLN, S. E., STEEN, R. G. & ET AL. 1993. A genetic linkage map of the mouse: current applications and future prospects. *Science,* 262, 57-66.

COSTANTINO, M. E., STEARMAN, R. P., SMITH, G. E. & DARLING, D. S. 2002. Cell-specific phosphorylation of Zfhep transcription factor. *Biochem Biophys Res Commun,* 296, 368-73.

DANG, C. V. 1999. c-Myc target genes involved in cell growth, apoptosis, and metabolism. *Mol Cell Biol,* 19, 1-11.

DAVID, A., NETZER, N., STRADER, M. B., DAS, S. R., CHEN, C. Y., GIBBS, J., PIERRE, P., BENNINK, J. R. & YEWDELL, J. W. 2011. RNA binding targets aminoacyl-tRNA synthetases to translating ribosomes. *J Biol Chem,* 286, 20688-700.

DEL PRETE, M. J., VERNAL, R., DOLZNIG, H., MULLNER, E. W. & GARCIA-SANZ, J. A. 2007. Isolation of polysome-bound mRNA from solid tissues amenable for RT-PCR and profiling experiments. *RNA,* 13, 414-21.

DEMONTIS, S., RIGO, C., PICCININ, S., MIZZAU, M., SONEGO, M., FABRIS, M., BRANCOLINI, C. & MAESTRO, R. 2006. Twist is substrate for caspase cleavage and proteasome-mediated degradation. *Cell Death Differ,* 13, 335-45.

DERIBE, Y. L., PAWSON, T. & DIKIC, I. 2010. Post-translational modifications in signal integration. *Nat Struct Mol Biol,* 17, 666-72.

DIKIC, I. & ROBERTSON, M. 2012. Ubiquitin ligases and beyond. *BMC Biol,* 10, 22.

DOHADWALA, M., YANG, S. C., LUO, J., SHARMA, S., BATRA, R. K., HUANG, M., LIN, Y., GOODGLICK, L., KRYSAN, K., FISHBEIN, M. C., HONG, L., LAI, C., CAMERON, R. B., GEMMILL, R. M., DRABKIN, H. A. & DUBINETT, S. M. 2006. Cyclooxygenase-2-dependent regulation of E-cadherin: prostaglandin E(2) induces transcriptional repressors ZEB1 and snail in non-small cell lung cancer. *Cancer Res,* 66, 5338-45.

DORUDI, S., SHEFFIELD, J. P., POULSOM, R., NORTHOVER, J. M. & HART, I. R. 1993. E-cadherin expression in colorectal cancer. An immunocytochemical and in situ hybridization study. *Am J Pathol,* 142, 981-6.

DRAKE, J. M., BARNES, J. M., MADSEN, J. M., DOMANN, F. E., STIPP, C. S. & HENRY, M. D. 2010. ZEB1 coordinately regulates laminin-332 and {beta}4 integrin expression altering the invasive phenotype of prostate cancer cells. *J Biol Chem,* 285, 33940-8.

DWEK, R. A., BUTTERS, T. D., PLATT, F. M. & ZITZMANN, N. 2002. Targeting glycosylation as a therapeutic approach. *Nat Rev Drug Discov,* 1, 65-75.

EGER, A., AIGNER, K., SONDEREGGER, S., DAMPIER, B., OEHLER, S., SCHREIBER, M., BERX, G., CANO, A., BEUG, H. & FOISNER, R. 2005. DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. *Oncogene,* 24, 2375-85.

EL-KASTI, M. M., WELLS, T. & CARTER, D. A. 2012. A novel long-range enhancer regulates postnatal expression of Zeb2: implications for Mowat-Wilson syndrome phenotypes. *Hum Mol Genet,* 21, 5429-42.

ELLIS, A. L., WANG, Z., YU, X. & MERTZ, J. E. 2010. Either ZEB1 or ZEB2/SIP1 can play a central role in regulating the Epstein-Barr virus latent-lytic switch in a cell-type-specific manner. *J Virol,* 84, 6139-52.

ELLOUL, S., ELSTRAND, M. B., NESLAND, J. M., TROPE, C. G., KVALHEIM, G., GOLDBERG, I., REICH, R. & DAVIDSON, B. 2005. Snail, Slug, and Smad-interacting protein 1 as novel parameters of disease aggressiveness in metastatic ovarian and breast carcinoma. *Cancer,* 103, 1631-43.

ELLOUL, S., SILINS, I., TROPE, C. G., BENSHUSHAN, A., DAVIDSON, B. & REICH, R. 2006. Expression of E-cadherin transcriptional regulators in ovarian carcinoma. *Virchows Arch,* 449, 520-8.

ELO, J. P., HARKONEN, P., KYLLONEN, A. P., LUKKARINEN, O., POUTANEN, M., VIHKO, R. & VIHKO, P. 1997. Loss of heterozygosity at 16q24.1-q24.2 is significantly associated with metastatic and aggressive behavior of prostate cancer. *Cancer Res,* 57, 3356-9.

ENGEL, L. W., YOUNG, N. A., TRALKA, T. S., LIPPMAN, M. E., O'BRIEN, S. J. & JOYCE, M. J. 1978. Establishment and characterization of three new continuous cell lines derived from human breast carcinomas. *Cancer Res,* 38, 3352-64.

EVANS, A. J., RUSSELL, R. C., ROCHE, O., BURRY, T. N., FISH, J. E., CHOW, V. W., KIM, W. Y., SARAVANAN, A., MAYNARD, M. A., GERVAIS, M. L., SUFAN, R. I., ROBERTS, A. M., WILSON, L. A., BETTEN, M., VANDEWALLE, C., BERX, G., MARSDEN, P. A., IRWIN, M. S., TEH, B. T., JEWETT, M. A. & OHH, M. 2007. VHL promotes E2 box-dependent E-cadherin transcription by HIF-mediated regulation of SIP1 and snail. *Mol Cell Biol,* 27, 157-69.

FANG, Y., WEI, J., CAO, J., ZHAO, H., LIAO, B., QIU, S., WANG, D., LUO, J. & CHEN, W. 2013. Protein expression of ZEB2 in renal cell carcinoma and its prognostic significance in patient survival. *PLoS One,* 8, e62558.

FIDLER, I. J. 2002. Critical determinants of metastasis. *Semin Cancer Biol,* 12, 89-96.

FOGH, J. & TREMPE, G. 1975. *New human tumor cell lines: Human Tumor Cell Lines in vitro*, New York: Plenum Press.

FRIXEN, U. H., BEHRENS, J., SACHS, M., EBERLE, G., VOSS, B., WARDA, A., LOCHNER, D. & BIRCHMEIER, W. 1991. E-cadherin-mediated cell-cell adhesion prevents invasiveness of human carcinoma cells. *J Cell Biol,* 113, 173-85.

FU, J., LV, X., LIN, H., WU, L., WANG, R., ZHOU, Z., ZHANG, B., WANG, Y. L., TSANG, B. K., ZHU, C. & WANG, H. 2010. Ubiquitin ligase cullin 7 induces epithelial-mesenchymal transition in human choriocarcinoma cells. *J Biol Chem,* 285, 10870-9.

FUNAHASHI, J., KAMACHI, Y., GOTO, K. & KONDOH, H. 1991. Identification of nuclear factor delta EF1 and its binding site essential for lens-specific activity of the delta 1-crystallin enhancer. *Nucleic Acids Res,* 19, 3543-7.

FUNAHASHI, J., SEKIDO, R., MURAI, K., KAMACHI, Y. & KONDOH, H. 1993. Delta-crystallin enhancer binding protein delta EF1 is a zinc finger-homeodomain protein implicated in postgastrulation embryogenesis. *Development,* 119, 433-46.

FURUKAWA, K. & KOBATA, A. 1992. Protein glycosylation. *Curr Opin Biotechnol,* 3, 554-9.

GARCIA-DOMINGUEZ, M. & REYES, J. C. 2009. SUMO association with repressor complexes, emerging routes for transcriptional control. *Biochim Biophys Acta,* 1789, 451-9.

GEOFFROY, M. C. & HAY, R. T. 2009. An additional role for SUMO in ubiquitin-mediated proteolysis. *Nat Rev Mol Cell Biol,* 10, 564-8.
GIACCONE, G., BATTEY, J., GAZDAR, A. F., OIE, H., DRAOUI, M. & MOODY, T. W. 1992. Neuromedin B is present in lung cancer cell lines. *Cancer Res. ,* 52, 2732-36.
GIARD, D. J., AARONSON, S. A., TODARO, G. J., ARNSTEIN, P., KERSEY, J. H., DOSIK, H. & PARKS, W. P. 1973. In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. *J Natl Cancer Inst,* 51, 1417-23.
GILL, G. 2004. SUMO and ubiquitin in the nucleus: different functions, similar mechanisms? *Genes Dev,* 18, 2046-59.
GLICKMAN, M. H. & CIECHANOVER, A. 2002. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol Rev,* 82, 373-428.
GOUY, M. & GAUTIER, C. 1982. Codon usage in bacteria: correlation with gene expressivity. *Nucleic Acids Res,* 10, 7055-74.
GRAHAM, F. L., SMILEY, J., RUSSELL, W. C. & NAIRN, R. 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J Gen Virol,* 36, 59-74.
GRANTHAM, R., GAUTIER, C., GOUY, M., JACOBZONE, M. & MERCIER, R. 1981. Codon catalog usage is a genome strategy modulated for gene expressivity. *Nucleic Acids Res,* 9, r43-74.
GREGORY, P. A., BERT, A. G., PATERSON, E. L., BARRY, S. C., TSYKIN, A., FARSHID, G., VADAS, M. A., KHEW-GOODALL, Y. & GOODALL, G. J. 2008a. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol,* 10, 593-601.
GREGORY, P. A., BRACKEN, C. P., BERT, A. G. & GOODALL, G. J. 2008b. MicroRNAs as regulators of epithelial-mesenchymal transition. *Cell Cycle,* 7, 3112-8.
GROETTRUP, M., PELZER, C., SCHMIDTKE, G. & HOFMANN, K. 2008. Activating the ubiquitin family: UBA6 challenges the field. *Trends Biochem Sci,* 33, 230-7.
GROSSMAN, H. B., WEDEMEYER, G., REN, L., WILSON, G. N. & COX, B. 1986. Improved growth of human urothelial carcinoma cell cultures. *J Urol,* 136, 953-9.
GUILFORD, P., HOPKINS, J., HARRAWAY, J., MCLEOD, M., MCLEOD, N., HARAWIRA, P., TAITE, H., SCOULAR, R., MILLER, A. & REEVE, A. E. 1998. E-cadherin germline mutations in familial gastric cancer. *Nature,* 392, 402-5.
GUO, H. B., JOHNSON, H., RANDOLPH, M. & PIERCE, M. 2009. Regulation of homotypic cell-cell adhesion by branched N-glycosylation of N-cadherin extracellular EC2 and EC3 domains. *J Biol Chem,* 284, 34986-97.
GUPTA, P. B., KUPERWASSER, C., BRUNET, J. P., RAMASWAMY, S., KUO, W. L., GRAY, J. W., NABER, S. P. & WEINBERG, R. A. 2005. The melanocyte differentiation program predisposes to metastasis after neoplastic transformation. *Nat Genet,* 37, 1047-54.
HAJRA, K. M., CHEN, D. Y. & FEARON, E. R. 2002. The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer Res,* 62, 1613-8.
HAN, I. & KUDLOW, J. E. 1997. Reduced O glycosylation of Sp1 is associated with increased proteasome susceptibility. *Mol Cell Biol,* 17, 2550-8.
HANAHAN, D. & WEINBERG, R. A. 2011. Hallmarks of cancer: the next generation. *Cell,* 144, 646-74.
HARDER, J. L., WHITEMAN, E. L., PIECZYNSKI, J. N., LIU, C. J. & MARGOLIS, B. 2012. Snail destabilizes cell surface Crumbs3a. *Traffic,* 13, 1170-85.
HATSELL, S., MEDINA, L., MEROLA, J., HALTIWANGER, R. & COWIN, P. 2003. Plakoglobin is O-glycosylated close to the N-terminal destruction box. *J Biol Chem,* 278, 37745-52.

HAY, R. T. 2005. SUMO: a history of modification. *Mol Cell,* 18, 1-12.
HENNIG, G., LOWRICK, O., BIRCHMEIER, W. & BEHRENS, J. 1996. Mechanisms identified in the transcriptional control of epithelial gene expression. *J Biol Chem,* 271, 595-602.
HERMISTON, M. L., WONG, M. H. & GORDON, J. I. 1996. Forced expression of E-cadherin in the mouse intestinal epithelium slows cell migration and provides evidence for nonautonomous regulation of cell fate in a self-renewing system. *Genes Dev,* 10, 985-96.
HILL, L., BROWNE, G. & TULCHINSKY, E. 2013. ZEB/miR-200 feedback loop: At the crossroads of signal transduction in cancer. *Int J Cancer,* 132, 745-54.
HONG, J., ZHOU, J., FU, J., HE, T., QIN, J., WANG, L., LIAO, L. & XU, J. 2011. Phosphorylation of serine 68 of Twist1 by MAPKs stabilizes Twist1 protein and promotes breast cancer cell invasiveness. *Cancer Res,* 71, 3980-90.
HU, P., BERKOWITZ, P., MADDEN, V. J. & RUBENSTEIN, D. S. 2006. Stabilization of plakoglobin and enhanced keratinocyte cell-cell adhesion by intracellular O-glycosylation. *J Biol Chem,* 281, 12786-91.
HUANG, X. & MILLER, W. 1991. A Time-efficient, Linear-Space Local Similarity Algorithm. *Advances in Applied Mathematics,* 12, 337-357.
HUNTZINGER, E. & IZAURRALDE, E. 2011. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet,* 12, 99-110.
HURT, E. M., SAYKALLY, J. N., ANOSE, B. M., KALLI, K. R. & SANDERS, M. M. 2008. Expression of the ZEB1 (deltaEF1) transcription factor in human: additional insights. *Mol Cell Biochem,* 318, 89-99.
HURTEAU, G. J., CARLSON, J. A., SPIVACK, S. D. & BROCK, G. J. 2007. Overexpression of the microRNA hsa-miR-200c leads to reduced expression of transcription factor 8 and increased expression of E-cadherin. *Cancer Res,* 67, 7972-6.
IMAMICHI, Y., KONIG, A., GRESS, T. & MENKE, A. 2007. Collagen type I-induced Smad-interacting protein 1 expression downregulates E-cadherin in pancreatic cancer. *Oncogene,* 26, 2381-5.
JACKSON, S. P. & TJIAN, R. 1988. O-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation. *Cell,* 55, 125-33.
JAISSER, F. 2000. Inducible gene expression and gene modification in transgenic mice. *J Am Soc Nephrol,* 11 Suppl 16, S95-S100.
JANDA, E., LEHMANN, K., KILLISCH, I., JECHLINGER, M., HERZIG, M., DOWNWARD, J., BEUG, H. & GRUNERT, S. 2002. Ras and TGF[beta] cooperatively regulate epithelial cell plasticity and metastasis: dissection of Ras signaling pathways. *J Cell Biol,* 156, 299-313.
KACZYNSKI, J., COOK, T. & URRUTIA, R. 2003. Sp1- and Kruppel-like transcription factors. *Genome Biol,* 4, 206.
KALLURI, R. & NEILSON, E. G. 2003. Epithelial-mesenchymal transition and its implications for fibrosis. *J Clin Invest,* 112, 1776-84.
KALLURI, R. & WEINBERG, R. A. 2009. The basics of epithelial-mesenchymal transition. *J Clin Invest,* 119, 1420-8.
KAMEMURA, K., HAYES, B. K., COMER, F. I. & HART, G. W. 2002. Dynamic interplay between O-glycosylation and O-phosphorylation of nucleocytoplasmic proteins: alternative glycosylation/phosphorylation of THR-58, a known mutational hot spot of c-Myc in lymphomas, is regulated by mitogens. *J Biol Chem,* 277, 19229-35.

KARRETH, F. A., TAY, Y., PERNA, D., ALA, U., TAN, S. M., RUST, A. G., DENICOLA, G., WEBSTER, K. A., WEISS, D., PEREZ-MANCERA, P. A., KRAUTHAMMER, M., HALABAN, R., PROVERO, P., ADAMS, D. J., TUVESON, D. A. & PANDOLFI, P. P. 2011. In vivo identification of tumor- suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma. *Cell,* 147, 382-95.

KATAGIRI, A., WATANABE, R. & TOMITA, Y. 1995. E-cadherin expression in renal cell cancer and its significance in metastasis and survival. *Br J Cancer,* 71, 376-9.

KATO, M., ZHANG, J., WANG, M., LANTING, L., YUAN, H., ROSSI, J. J. & NATARAJAN, R. 2007. MicroRNA-192 in diabetic kidney glomeruli and its function in TGF-beta-induced collagen expression via inhibition of E-box repressors. *Proc Natl Acad Sci U S A,* 104, 3432-7.

KAWANISHI, H., MATSUI, Y., ITO, M., WATANABE, J., TAKAHASHI, T., NISHIZAWA, K., NISHIYAMA, H., KAMOTO, T., MIKAMI, Y., TANAKA, Y., JUNG, G., AKIYAMA, H., NOBUMASA, H., GUILFORD, P., REEVE, A., OKUNO, Y., TSUJIMOTO, G., NAKAMURA, E. & OGAWA, O. 2008. Secreted CXCL1 is a potential mediator and marker of the tumor invasion of bladder cancer. *Clin Cancer Res,* 14, 2579-87.

KEYDAR, I., CHEN, L., KARBY, S., WEISS, F. R., DELAREA, J., RADU, M., CHAITCIK, S. & BRENNER, H. J. 1979. Establishment and characterization of a cell line of human breast carcinoma origin. *Eur J Cancer,* 15, 659-70.

KIM, K. I. & BAEK, S. H. 2006. SUMOylation code in cancer development and metastasis. *Mol Cells,* 22, 247-53.

KIRSCHMANN, D. A., SEFTOR, E. A., NIEVA, D. R., MARIANO, E. A. & HENDRIX, M. J. 1999. Differentially expressed genes associated with the metastatic phenotype in breast cancer. *Breast Cancer Res Treat,* 55, 127-36.

KO, H., SO, Y., JEON, H., JEONG, M. H., CHOI, H. K., RYU, S. H., LEE, S. W., YOON, H. G. & CHOI, K. C. 2013. TGF-beta1-induced epithelial-mesenchymal transition and acetylation of Smad2 and Smad3 are negatively regulated by EGCG in human A549 lung cancer cells. *Cancer Lett,* 335, 205-13.

KORPAL, M., LEE, E. S., HU, G. & KANG, Y. 2008. The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. *J Biol Chem,* 283, 14910-4.

KOS, C. H. 2004. Cre/loxP system for generating tissue-specific knockout mouse models. *Nutr Rev,* 62, 243-6.

KRAFCHAK, C. M., PAWAR, H., MOROI, S. E., SUGAR, A., LICHTER, P. R., MACKEY, D. A., MIAN, S., NAIRUS, T., ELNER, V., SCHTEINGART, M. T., DOWNS, C. A., KIJEK, T. G., JOHNSON, J. M., TRAGER, E. H., ROZSA, F. W., MANDAL, M. N., EPSTEIN, M. P., VOLLRATH, D., AYYAGARI, R., BOEHNKE, M. & RICHARDS, J. E. 2005. Mutations in TCF8 cause posterior polymorphous corneal dystrophy and ectopic expression of COL4A3 by corneal endothelial cells. *Am J Hum Genet,* 77, 694-708.

KRISHNAMACHARY, B., ZAGZAG, D., NAGASAWA, H., RAINEY, K., OKUYAMA, H., BAEK, J. H. & SEMENZA, G. L. 2006. Hypoxia-inducible factor-1-dependent repression of E-cadherin in von Hippel-Lindau tumor suppressor-null renal cell carcinoma mediated by TCF3, ZFHX1A, and ZFHX1B. *Cancer Res,* 66, 2725-31.

KUHN, R. & TORRES, R. M. 2002. Cre/loxP recombination system and gene targeting. *Methods Mol Biol,* 180, 175-204.

LAMPSON, B. L., PERSHING, N. L., PRINZ, J. A., LACSINA, J. R., MARZLUFF, W. F., NICCHITTA, C. V., MACALPINE, D. M. & COUNTER, C. M. 2013. Rare codons regulate KRas oncogenesis. *Curr Biol,* 23, 70-5.

LARUE, L., OHSUGI, M., HIRCHENHAIN, J. & KEMLER, R. 1994. E-cadherin null mutant embryos fail to form a trophectoderm epithelium. *Proc Natl Acad Sci U S A,* 91, 8263-7.

LATIL, A., CUSSENOT, O., FOURNIER, G., DRIOUCH, K. & LIDEREAU, R. 1997. Loss of heterozygosity at chromosome 16q in prostate adenocarcinoma: identification of three independent regions. *Cancer Res,* 57, 1058-62.

LECKER, S. H., GOLDBERG, A. L. & MITCH, W. E. 2006. Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. *J Am Soc Nephrol,* 17, 1807-19.

LI, A., OMURA, N., HONG, S. M., VINCENT, A., WALTER, K., GRIFFITH, M., BORGES, M. & GOGGINS, M. 2010. Pancreatic cancers epigenetically silence SIP1 and hypomethylate and overexpress miR-200a/200b in association with elevated circulating miR-200a and miR-200b levels. *Cancer Res,* 70, 5226-37.

LI, X., HIRANO, R., TAGAMI, H. & AIBA, H. 2006. Protein tagging at rare codons is caused by tmRNA action at the 3' end of nonstop mRNA generated in response to ribosome stalling. *RNA,* 12, 248-55.

LIN, S. Y. & ELLEDGE, S. J. 2003. Multiple tumor suppressor pathways negatively regulate telomerase. *Cell,* 113, 881-9.

LIU, J., ZHENG, H., TANG, M., RYU, Y. C. & WANG, X. 2008a. A therapeutic dose of doxorubicin activates ubiquitin-proteasome system-mediated proteolysis by acting on both the ubiquitination apparatus and proteasome. *Am J Physiol Heart Circ Physiol,* 295, H2541-50.

LIU, X. & FENG, R. 2010. Inhibition of epithelial to mesenchymal transition in metastatic breast carcinoma cells by c-Src suppression. *Acta Biochim Biophys Sin (Shanghai),* 42, 496-501.

LIU, Y., EL-NAGGAR, S., DARLING, D. S., HIGASHI, Y. & DEAN, D. C. 2008b. Zeb1 links epithelial-mesenchymal transition and cellular senescence. *Development,* 135, 579-88.

LIU, Y., PENG, X., TAN, J., DARLING, D. S., KAPLAN, H. J. & DEAN, D. C. 2008c. Zeb1 mutant mice as a model of posterior corneal dystrophy. *Invest Ophthalmol Vis Sci,* 49, 1843-9.

LIU, Y., SANCHEZ-TILLO, E., LU, X., HUANG, L., CLEM, B., TELANG, S., JENSON, A. B., CUATRECASAS, M., CHESNEY, J., POSTIGO, A. & DEAN, D. C. 2013. Sequential inductions of the ZEB1 transcription factor caused by mutation of Rb and then Ras proteins are required for tumor initiation and progression. *J Biol Chem,* 288, 11572-80.

LONG, J., ZUO, D. & PARK, M. 2005. Pc2-mediated sumoylation of Smad-interacting protein 1 attenuates transcriptional repression of E-cadherin. *J Biol Chem,* 280, 35477-89.

LOWRY, W. E. 2011. E-cadherin, a new mixer in the Yamanaka cocktail. *EMBO Rep,* 12, 613-4.

LUO, J., LUBAROFF, D. M. & HENDRIX, M. J. 1999. Suppression of prostate cancer invasive potential and matrix metalloproteinase activity by E-cadherin transfection. *Cancer Res,* 59, 3552-6.

MACHADO, J. C., OLIVEIRA, C., CARVALHO, R., SOARES, P., BERX, G., CALDAS, C., SERUCA, R., CARNEIRO, F. & SOBRINHO-SIMOES, M. 2001. E-cadherin gene (CDH1) promoter

methylation as the second hit in sporadic diffuse gastric carcinoma. *Oncogene,* 20, 1525-8.
MAEDA, G., CHIBA, T., OKAZAKI, M., SATOH, T., TAYA, Y., AOBA, T., KATO, K., KAWASHIRI, S. & IMAI, K. 2005. Expression of SIP1 in oral squamous cell carcinomas: implications for E-cadherin expression and tumor progression. *Int J Oncol,* 27, 1535-41.
MANI, S. A., GUO, W., LIAO, M. J., EATON, E. N., AYYANAN, A., ZHOU, A. Y., BROOKS, M., REINHARD, F., ZHANG, C. C., SHIPITSIN, M., CAMPBELL, L. L., POLYAK, K., BRISKEN, C., YANG, J. & WEINBERG, R. A. 2008. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell,* 133, 704-15.
MARSHALL, C. J., FRANKS, L. M. & CARBONELL, A. W. 1977. Markers of neoplastic transformation in epithelial cell lines derived from human carcinomas. *J Natl Cancer Inst,* 58, 1743-51.
MARTIN, P. T. 2005. The dystroglycanopathies: the new disorders of O-linked glycosylation. *Semin Pediatr Neurol,* 12, 152-8.
MEARINI, G., SCHLOSSAREK, S., WILLIS, M. S. & CARRIER, L. 2008. The ubiquitin-proteasome system in cardiac dysfunction. *Biochim Biophys Acta,* 1782, 749-63.
MEJLVANG, J., KRIAJEVSKA, M., VANDEWALLE, C., CHERNOVA, T., SAYAN, A. E., BERX, G., MELLON, J. K. & TULCHINSKY, E. 2007. Direct repression of cyclin D1 by SIP1 attenuates cell cycle progression in cells undergoing an epithelial mesenchymal transition. *Mol Biol Cell,* 18, 4615-24.
METZGER, M. B., HRISTOVA, V. A. & WEISSMAN, A. M. 2012. HECT and RING finger families of E3 ubiquitin ligases at a glance. *J Cell Sci,* 125, 531-7.
MICHELLE, C., VOURC'H, P., MIGNON, L. & ANDRES, C. R. 2009. What was the set of ubiquitin and ubiquitin-like conjugating enzymes in the eukaryote common ancestor? *J Mol Evol,* 68, 616-28.
MIQUELAJAUREGUI, A., VAN DE PUTTE, T., POLYAKOV, A., NITYANANDAM, A., BOPPANA, S., SEUNTJENS, E., KARABINOS, A., HIGASHI, Y., HUYLEBROECK, D. & TARABYKIN, V. 2007. Smad-interacting protein-1 (Zfhx1b) acts upstream of Wnt signaling in the mouse hippocampus and controls its formation. *Proc Natl Acad Sci U S A,* 104, 12919-24.
MISRA, R. & REEVES, P. 1985. Intermediates in the synthesis of TolC protein include an incomplete peptide stalled at a rare Arg codon. *Eur J Biochem,* 152, 151-5.
MIYAZONO, K. 2009. Transforming growth factor-beta signaling in epithelial-mesenchymal transition and progression of cancer. *Proc Jpn Acad Ser B Phys Biol Sci,* 85, 314-23.
MIYOSHI, A., KITAJIMA, Y., SUMI, K., SATO, K., HAGIWARA, A., KOGA, Y. & MIYAZAKI, K. 2004. Snail and SIP1 increase cancer invasion by upregulating MMP family in hepatocellular carcinoma cells. *Br J Cancer,* 90, 1265-73.
MIYOSHI, T., MARUHASHI, M., VAN DE PUTTE, T., KONDOH, H., HUYLEBROECK, D. & HIGASHI, Y. 2006. Complementary expression pattern of Zfhx1 genes Sip1 and deltaEF1 in the mouse embryo and their genetic interaction revealed by compound mutants. *Dev Dyn,* 235, 1941-52.
MURRAY, D., PRECHT, P., BALAKIR, R. & HORTON, W. E., JR. 2000. The transcription factor deltaEF1 is inversely expressed with type II collagen mRNA and can repress Col2a1 promoter activity in transfected chondrocytes. *J Biol Chem,* 275, 3610-8.
MYEKU, N. & FIGUEIREDO-PEREIRA, M. E. 2011. Dynamics of the degradation of ubiquitinated proteins by proteasomes and autophagy: association with sequestosome 1/p62. *J Biol Chem,* 286, 22426-40.

NELLES, L., VAN DE PUTTE, T., VAN GRUNSVEN, L., HUYLEBROECK, D. & VERSCHUEREN, K. 2003. Organization of the mouse Zfhx1b gene encoding the two-handed zinc finger repressor Smad-interacting protein-1. *Genomics,* 82, 460-9.

NIESSEN, C. M. 2007. Tight junctions/adherens junctions: basic structure and function. *J Invest Dermatol,* 127, 2525-32.

NIETO, M. A. 2002. The snail superfamily of zinc-finger transcription factors. *Nat Rev Mol Cell Biol,* 3, 155-66.

NISHIMURA, G., MANABE, I., TSUSHIMA, K., FUJIU, K., OISHI, Y., IMAI, Y., MAEMURA, K., MIYAGISHI, M., HIGASHI, Y., KONDOH, H. & NAGAI, R. 2006. DeltaEF1 mediates TGF-beta signaling in vascular smooth muscle cell differentiation. *Dev Cell,* 11, 93-104.

NOVAK, A., GUO, C., YANG, W., NAGY, A. & LOBE, C. G. 2000. Z/EG, a double reporter mouse line that expresses enhanced green fluorescent protein upon Cre-mediated excision. *Genesis,* 28, 147-55.

O'TOOLE, C., PERLMANN, P., UNSGAARD, B., ALMGARD, L. E., JOHANSSON, B., MOBERGER, G. & EDSMYR, F. 1972. Cellular immunity to human urinary bladder carcinoma. II. EEffect of surgery and preoperative irradiation. *Int J Cancer,* 10, 92-8.

O'TOOLE, C., PRICE, Z. H., OHNUKI, Y. & UNSGAARD, B. 1978. Ultrastructure, karyology and immunology of a cell line originated from a human transitional-cell carcinoma. *Br J Cancer,* 38, 64-76.

OHIRA, T., GEMMILL, R. M., FERGUSON, K., KUSY, S., ROCHE, J., BRAMBILLA, E., ZENG, C., BARON, A., BEMIS, L., ERICKSON, P., WILDER, E., RUSTGI, A., KITAJEWSKI, J., GABRIELSON, E., BREMNES, R., FRANKLIN, W. & DRABKIN, H. A. 2003. WNT7a induces E-cadherin in lung cancer cells. *Proc Natl Acad Sci U S A,* 100, 10429-34.

OKA, H., SHIOZAKI, H., KOBAYASHI, K., INOUE, M., TAHARA, H., KOBAYASHI, T., TAKATSUKA, Y., MATSUYOSHI, N., HIRANO, S., TAKEICHI, M. & ET AL. 1993. Expression of E-cadherin cell adhesion molecules in human breast cancer tissues and its relationship to metastasis. *Cancer Res,* 53, 1696-701.

OLER, A. J., ALLA, R. K., ROBERTS, D. N., WONG, A., HOLLENHORST, P. C., CHANDLER, K. J., CASSIDAY, P. A., NELSON, C. A., HAGEDORN, C. H., GRAVES, B. J. & CAIRNS, B. R. 2010. Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. *Nat Struct Mol Biol,* 17, 620-8.

ORLOWSKI, R. Z. & BALDWIN, A. S., JR. 2002. NF-kappaB as a therapeutic target in cancer. *Trends Mol Med,* 8, 385-9.

OZTAS, E., AVCI, M. E., OZCAN, A., SAYAN, A. E., TULCHINSKY, E. & YAGCI, T. 2010. Novel monoclonal antibodies detect Smad-interacting protein 1 (SIP1) in the cytoplasm of human cells from multiple tumor tissue arrays. *Exp Mol Pathol,* 89, 182-9.

OZTURK, N., ERDAL, E., MUMCUOGLU, M., AKCALI, K. C., YALCIN, O., SENTURK, S., ARSLAN-ERGUL, A., GUR, B., YULUG, I., CETIN-ATALAY, R., YAKICIER, C., YAGCI, T., TEZ, M. & OZTURK, M. 2006. Reprogramming of replicative senescence in hepatocellular carcinoma-derived cells. *Proc Natl Acad Sci U S A,* 103, 2178-83.

PARK, S. M., GAUR, A. B., LENGYEL, E. & PETER, M. E. 2008. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev,* 22, 894-907.

PASSMORE, L. A. & BARFORD, D. 2004. Getting into position: the catalytic mechanisms of protein ubiquitylation. *Biochem J,* 379, 513-25.

PATTERSON, C., IKE, C., WILLIS, P. W. T., STOUFFER, G. A. & WILLIS, M. S. 2007. The bitter end: the ubiquitin-proteasome system and cardiac dysfunction. *Circulation,* 115, 1456-63.

PEINADO, H., OLMEDA, D. & CANO, A. 2007. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat Rev Cancer,* 7, 415-28.

PEINADO, H., PORTILLO, F. & CANO, A. 2004. Transcriptional regulation of cadherins during development and carcinogenesis. *Int J Dev Biol,* 48, 365-75.

PERL, A. K., WILGENBUS, P., DAHL, U., SEMB, H. & CHRISTOFORI, G. 1998. A causal role for E-cadherin in the transition from adenoma to carcinoma. *Nature,* 392, 190-3.

PETROSKI, M. D. & DESHAIES, R. J. 2003. Context of multiubiquitin chain attachment influences the rate of Sic1 degradation. *Mol Cell,* 11, 1435-44.

PETROSKI, M. D. & DESHAIES, R. J. 2005. Function and regulation of cullin-RING ubiquitin ligases. *Nat Rev Mol Cell Biol,* 6, 9-20.

PICKART, C. M. & EDDINS, M. J. 2004. Ubiquitin: structures, functions, mechanisms. *Biochim Biophys Acta,* 1695, 55-72.

PINHO, S. S., OSORIO, H., NITA-LAZAR, M., GOMES, J., LOPES, C., GARTNER, F. & REIS, C. A. 2009a. Role of E-cadherin N-glycosylation profile in a mammary tumor model. *Biochem Biophys Res Commun,* 379, 1091-6.

PINHO, S. S., REIS, C. A., PAREDES, J., MAGALHAES, A. M., FERREIRA, A. C., FIGUEIREDO, J., XIAOGANG, W., CARNEIRO, F., GARTNER, F. & SERUCA, R. 2009b. The role of N-acetylglucosaminyltransferase III and V in the post-transcriptional modifications of E-cadherin. *Hum Mol Genet,* 18, 2599-608.

PINHO, S. S., SERUCA, R., GARTNER, F., YAMAGUCHI, Y., GU, J., TANIGUCHI, N. & REIS, C. A. 2011. Modulation of E-cadherin function and dysfunction by N-glycosylation. *Cell Mol Life Sci,* 68, 1011-20.

PONTEN, J. & SAKSELA, E. 1967. Two established in vitro cell lines from human mesenchymal tumours. *Int J Cancer,* 2, 434-47.

POSTIGO, A. A., DEPP, J. L., TAYLOR, J. J. & KROLL, K. L. 2003. Regulation of Smad signaling through a differential recruitment of coactivators and corepressors by ZEB proteins. *EMBO J,* 22, 2453-62.

QI, S., SONG, Y., PENG, Y., WANG, H., LONG, H., YU, X., LI, Z., FANG, L., WU, A., LUO, W., ZHEN, Y., ZHOU, Y., CHEN, Y., MAI, C., LIU, Z. & FANG, W. 2012. ZEB2 mediates multiple pathways regulating cell proliferation, migration, invasion, and apoptosis in glioma. *PLoS One,* 7, e38842.

RAILE, K., HOFLICH, A., KESSLER, U., YANG, Y., PFUENDER, M., BLUM, W. F., KOLB, H., SCHWARZ, H. P. & KIESS, W. 1994. Human osteosarcoma (U-2 OS) cells express both insulin-like growth factor-I (IGF-I) receptors and insulin-like growth factor-II/mannose-6-phosphate (IGF-II/M6P) receptors and synthesize IGF-II: autocrine growth stimulation by IGF-II via the IGF-I receptor. *J Cell Physiol,* 159, 531-41.

REDMER, T., DIECKE, S., GRIGORYAN, T., QUIROGA-NEGREIRA, A., BIRCHMEIER, W. & BESSER, D. 2011. E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming. *EMBO Rep,* 12, 720-6.

REMACLE, J. E., KRAFT, H., LERCHNER, W., WUYTENS, G., COLLART, C., VERSCHUEREN, K., SMITH, J. C. & HUYLEBROECK, D. 1999. New mode of DNA binding of multi-zinc finger transcription factors: deltaEF1 family members bind with two hands to two target sites. *EMBO J,* 18, 5073-84.

RHIM, A. D., MIREK, E. T., AIELLO, N. M., MAITRA, A., BAILEY, J. M., MCALLISTER, F., REICHERT, M., BEATTY, G. L., RUSTGI, A. K., VONDERHEIDE, R. H., LEACH, S. D. & STANGER, B. Z. 2012. EMT and dissemination precede pancreatic tumor formation. *Cell,* 148, 349-61.

ROBINSON, M., LILLEY, R., LITTLE, S., EMTAGE, J. S., YARRANTON, G., STEPHENS, P., MILLICAN, A., EATON, M. & HUMPHREYS, G. 1984. Codon usage can affect efficiency of translation of genes in Escherichia coli. *Nucleic Acids Res,* 12, 6663-71.

ROCHE, J., NASARRE, P., GEMMILL, R., BALDYS, A., PONTIS, J., KORCH, C., GUILHOT, J., AIT-SI-ALI, S. & DRABKIN, H. 2013. Global Decrease of Histone H3K27 Acetylation in ZEB1-Induced Epithelial to Mesenchymal Transition in Lung Cancer Cells. *Cancers* 5, 334-356.

ROSIVATZ, E., BECKER, I., SPECHT, K., FRICKE, E., LUBER, B., BUSCH, R., HOFLER, H. & BECKER, K. F. 2002. Differential expression of the epithelial-mesenchymal transition regulators snail, SIP1, and twist in gastric cancer. *Am J Pathol,* 161, 1881-91.

ROTH, Z., YEHEZKEL, G. & KHALAILA, I. 2012. Identification and Quantification of Protein Glycosylation. *International Journal of Carbohydrate Chemistry,* 2012, 10.

ROYBAL, J. D., ZANG, Y., AHN, Y. H., YANG, Y., GIBBONS, D. L., BAIRD, B. N., ALVAREZ, C., THILAGANATHAN, N., LIU, D. D., SAINTIGNY, P., HEYMACH, J. V., CREIGHTON, C. J. & KURIE, J. M. 2011. miR-200 Inhibits lung adenocarcinoma cell invasion and metastasis by targeting Flt1/VEGFR1. *Mol Cancer Res,* 9, 25-35.

ROZEN, S. & SKALETSKY, H. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol,* 132, 365-86.

RUDDOCK, L. W. & MOLINARI, M. 2006. N-glycan processing in ER quality control. *J Cell Sci,* 119, 4373-80.

RUSHWORTH, S. A., MURRAY, M. Y., BARRERA, L. N., HEASMAN, S. A., ZAITSEVA, L. & MACEWAN, D. J. 2012. Understanding the role of miRNA in regulating NF-kappaB in blood cancer. *Am J Cancer Res,* 2, 65-74.

SANCHEZ-TILLO, E., DE BARRIOS, O., SILES, L., CUATRECASAS, M., CASTELLS, A. & POSTIGO, A. 2011. beta-catenin/TCF4 complex induces the epithelial-to-mesenchymal transition (EMT)-activator ZEB1 to regulate tumor invasiveness. *Proc Natl Acad Sci U S A,* 108, 19204-9.

SARIKAS, A., HARTMANN, T. & PAN, Z. Q. 2011. The cullin protein family. *Genome Biol,* 12, 220.

SASAKI, C. Y., LIN, H. & PASSANITI, A. 2000. Expression of E-cadherin reduces bcl-2 expression and increases sensitivity to etoposide-induced apoptosis. *Int J Cancer,* 86, 660-6.

SAUER, B. 1998. Inducible gene targeting in mice using the Cre/lox system. *Methods,* 14, 381-92.

SAVAGNER, P., YAMADA, K. M. & THIERY, J. P. 1997. The zinc-finger protein slug causes desmosome dissociation, an initial and necessary step for growth factor-induced epithelial-mesenchymal transition. *J Cell Biol,* 137, 1403-19.

SAYAN, A. E., GRIFFITHS, T. R., PAL, R., BROWNE, G. J., RUDDICK, A., YAGCI, T., EDWARDS, R., MAYER, N. J., QAZI, H., GOYAL, S., FERNANDEZ, S., STRAATMAN, K., JONES, G. D., BOWMAN, K. J., COLQUHOUN, A., MELLON, J. K., KRIAJEVSKA, M. & TULCHINSKY, E. 2009. SIP1 protein protects cells from DNA damage-induced apoptosis and has independent prognostic value in bladder cancer. *Proc Natl Acad Sci U S A,* 106, 14884-9.

SAYDAM, O., SHEN, Y., WURDINGER, T., SENOL, O., BOKE, E., JAMES, M. F., TANNOUS, B. A., STEMMER-RACHAMIMOV, A. O., YI, M., STEPHENS, R. M., FRAEFEL, C., GUSELLA, J. F., KRICHEVSKY, A. M. & BREAKEFIELD, X. O. 2009. Downregulated microRNA-200a in meningiomas promotes tumor growth by reducing E-cadherin and activating the Wnt/beta-catenin signaling pathway. *Mol Cell Biol,* 29, 5923-40.

SCHIPPER, J. H., FRIXEN, U. H., BEHRENS, J., UNGER, A., JAHNKE, K. & BIRCHMEIER, W. 1991. E-cadherin expression in squamous cell carcinomas of head and neck: inverse correlation with tumor dedifferentiation and lymph node metastasis. *Cancer Res,* 51, 6328-37.

SCHLESINGER, D. H., GOLDSTEIN, G. & NIALL, H. D. 1975. The complete amino acid sequence of ubiquitin, an adenylate cyclase stimulating polypeptide probably universal in living cells. *Biochemistry,* 14, 2214-8.

SEKIDO, R., TAKAGI, T., OKANAMI, M., MORIBE, H., YAMAMURA, M., HIGASHI, Y. & KONDOH, H. 1996. Organization of the gene encoding transcriptional repressor deltaEF1 and cross-species conservation of its domains. *Gene,* 173, 227-32.

SHA, Z., BRILL, L. M., CABRERA, R., KLEIFELD, O., SCHELIGA, J. S., GLICKMAN, M. H., CHANG, E. C. & WOLF, D. A. 2009. The eIF3 interactome reveals the translasome, a supercomplex linking protein synthesis and degradation machineries. *Mol Cell,* 36, 141-52.

SHASHIKANT, C. S. & RUDDLE, F. H. 2003. Impact of transgenic technologies on functional genomics. *Curr Issues Mol Biol,* 5, 75-98.

SHIH, J. Y., TSAI, M. F., CHANG, T. H., CHANG, Y. L., YUAN, A., YU, C. J., LIN, S. B., LIOU, G. Y., LEE, M. L., CHEN, J. J., HONG, T. M., YANG, S. C., SU, J. L., LEE, Y. C. & YANG, P. C. 2005. Transcription repressor slug promotes carcinoma invasion and predicts outcome of patients with lung adenocarcinoma. *Clin Cancer Res,* 11, 8070-8.

SHIN, S. & BLENIS, J. 2010. ERK2/Fra1/ZEB pathway induces epithelial-to-mesenchymal transition. *Cell Cycle,* 9, 2483-4.

SHIOZAKI, H., TAHARA, H., OKA, H., MIYATA, M., KOBAYASHI, K., TAMURA, S., IIHARA, K., DOKI, Y., HIRANO, S., TAKEICHI, M. & ET AL. 1991. Expression of immunoreactive E-cadherin adhesion molecules in human cancers. *Am J Pathol,* 139, 17-23.

SHIRAKIHARA, T., SAITOH, M. & MIYAZONO, K. 2007. Differential regulation of epithelial and mesenchymal markers by deltaEF1 proteins in epithelial mesenchymal transition induced by TGF-beta. *Mol Biol Cell,* 18, 3533-44.

SMIT, M. A. & PEEPER, D. S. 2010. Epithelial-mesenchymal transition and senescence: two cancer-related processes are crossing paths. *Aging (Albany NY),* 2, 735-41.

SMITH, A. J., DE SOUSA, M. A., KWABI-ADDO, B., HEPPELL-PARTON, A., IMPEY, H. & RABBITTS, P. 1995. A site-directed chromosomal translocation induced in embryonic stem cells by Cre-loxP recombination. *Nat Genet,* 9, 376-85.

SMITH, G. E. & DARLING, D. S. 2003. Combination of a zinc finger and homeodomain required for protein-interaction. *Mol Biol Rep,* 30, 199-206.

SONCIN, F. & WARD, C. M. 2011. The Function of E-Cadherin in Stem Cell Pluripotency and Self-Renewal. *Genes* 2, 229-259.

SOOY, K. & DEMAY, M. B. 2002. Transcriptional repression of the rat osteocalcin gene by deltaEF1. *Endocrinology,* 143, 3370-5.

SOULE, H. D., VAZGUEZ, J., LONG, A., ALBERT, S. & BRENNAN, M. 1973. A human cell line from a pleural effusion derived from a breast carcinoma. *J Natl Cancer Inst,* 51, 1409-16.

SPADERNA, S., SCHMALHOFER, O., HLUBEK, F., BERX, G., EGER, A., MERKEL, S., JUNG, A., KIRCHNER, T. & BRABLETZ, T. 2006. A transient, EMT-linked loss of basement membranes indicates metastasis and poor survival in colorectal cancer. *Gastroenterology,* 131, 830-40.

SPADERNA, S., SCHMALHOFER, O., WAHLBUHL, M., DIMMLER, A., BAUER, K., SULTAN, A., HLUBEK, F., JUNG, A., STRAND, D., EGER, A., KIRCHNER, T., BEHRENS, J. & BRABLETZ, T. 2008. The transcriptional repressor ZEB1 promotes metastasis and loss of cell polarity in cancer. *Cancer Res,* 68, 537-44.

SPANO, D., HECK, C., DE ANTONELLIS, P., CHRISTOFORI, G. & ZOLLO, M. 2012. Molecular networks that regulate cancer metastasis. *Semin Cancer Biol,* 22, 234-49.

SRIKANTH, B., VAIDYA, M. M. & KALRAIYA, R. D. 2010. O-GlcNAcylation determines the solubility, filament organization, and stability of keratins 8 and 18. *J Biol Chem,* 285, 34062-71.

STANLEY, P. 2011. Golgi glycosylation. *Cold Spring Harb Perspect Biol,* 3.

STRATHDEE, G. 2002. Epigenetic versus genetic alterations in the inactivation of E-cadherin. *Semin Cancer Biol,* 12, 373-9.

SWITZER, C. H., GLYNN, S. A., CHENG, R. Y., RIDNOUR, L. A., GREEN, J. E., AMBS, S. & WINK, D. A. 2012. S-nitrosylation of EGFR and Src activates an oncogenic signaling network in human basal-like breast cancer. *Mol Cancer Res,* 10, 1203-15.

TAKAGI, T., MORIBE, H., KONDOH, H. & HIGASHI, Y. 1998. DeltaEF1, a zinc finger and homeodomain transcription factor, is required for skeleton patterning in multiple lineages. *Development,* 125, 21-31.

TEPASS, U., TRUONG, K., GODT, D., IKURA, M. & PEIFER, M. 2000. Cadherins in embryonic and neural morphogenesis. *Nat Rev Mol Cell Biol,* 1, 91-100.

THIERY, J. P. 2002. Epithelial-mesenchymal transitions in tumour progression. *Nat Rev Cancer,* 2, 442-54.

THIERY, J. P., ACLOQUE, H., HUANG, R. Y. & NIETO, M. A. 2009. Epithelial-mesenchymal transitions in development and disease. *Cell,* 139, 871-90.

THIERY, J. P. & SLEEMAN, J. P. 2006. Complex networks orchestrate epithelial-mesenchymal transitions. *Nat Rev Mol Cell Biol,* 7, 131-42.

TSAI, C. J., SAUNA, Z. E., KIMCHI-SARFATY, C., AMBUDKAR, S. V., GOTTESMAN, M. M. & NUSSINOV, R. 2008. Synonymous mutations and ribosome stalling can lead to altered folding pathways and distinct minima. *J Mol Biol,* 383, 281-91.

UCHIKADO, Y., NATSUGOE, S., OKUMURA, H., SETOYAMA, T., MATSUMOTO, M., ISHIGAMI, S. & AIKOU, T. 2005. Slug Expression in the E-cadherin preserved tumors is related to prognosis in patients with esophageal squamous cell carcinoma. *Clin Cancer Res,* 11, 1174-80.

UMBAS, R., SCHALKEN, J. A., AALDERS, T. W., CARTER, B. S., KARTHAUS, H. F., SCHAAFSMA, H. E., DEBRUYNE, F. M. & ISAACS, W. B. 1992. Expression of the cellular adhesion molecule E-cadherin is reduced or absent in high-grade prostate cancer. *Cancer Res,* 52, 5104-9.

VAGIN, O., TOKHTAEVA, E., YAKUBOV, I., SHEVCHENKO, E. & SACHS, G. 2008. Inverse correlation between the extent of N-glycan branching and intercellular adhesion in epithelia. Contribution of the Na,K-ATPase beta1 subunit. *J Biol Chem,* 283, 2192-202.

VALASTYAN, S. & WEINBERG, R. A. 2011. Tumor metastasis: molecular insights and evolving paradigms. *Cell,* 147, 275-92.

VAN DE PUTTE, T., MARUHASHI, M., FRANCIS, A., NELLES, L., KONDOH, H., HUYLEBROECK, D. & HIGASHI, Y. 2003. Mice lacking ZFHX1B, the gene that codes for Smad-interacting protein-1, reveal a role for multiple neural crest cell defects in the etiology of Hirschsprung disease-mental retardation syndrome. *Am J Hum Genet,* 72, 465-70.

VAN GRUNSVEN, L. A., TAELMAN, V., MICHIELS, C., OPDECAMP, K., HUYLEBROECK, D. & BELLEFROID, E. J. 2006. deltaEF1 and SIP1 are differentially expressed and have overlapping activities during Xenopus embryogenesis. *Dev Dyn,* 235, 1491-500.

VAN ROY, F. & BERX, G. 2008. The cell-cell adhesion molecule E-cadherin. *Cell Mol Life Sci,* 65, 3756-88.

VANDEWALLE, C., COMIJN, J., DE CRAENE, B., VERMASSEN, P., BRUYNEEL, E., ANDERSEN, H., TULCHINSKY, E., VAN ROY, F. & BERX, G. 2005. SIP1/ZEB2 induces EMT by repressing genes of different epithelial cell-cell junctions. *Nucleic Acids Res,* 33, 6566-78.

VANDEWALLE, C., VAN ROY, F. & BERX, G. 2009. The role of the ZEB family of transcription factors in development and disease. *Cell Mol Life Sci,* 66, 773-87.

VARKI, A. 1993. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology,* 3, 97-130.

VEENA, M. S., QIN, M., ANDERSSON, A., SHARMA, S. & BATRA, R. K. 2009. CAR mediates efficient tumor engraftment of mesenchymal type lung cancer cells. *Lab Invest,* 89, 875-86.

VERSCHUEREN, K., REMACLE, J. E., COLLART, C., KRAFT, H., BAKER, B. S., TYLZANOWSKI, P., NELLES, L., WUYTENS, G., SU, M. T., BODMER, R., SMITH, J. C. & HUYLEBROECK, D. 1999. SIP1, a novel zinc finger/homeodomain repressor, interacts with Smad proteins and binds to 5'-CACCT sequences in candidate target genes. *J Biol Chem,* 274, 20489-98.

VICHALKOVSKI, A., GRESKO, E., HESS, D., RESTUCCIA, D. F. & HEMMINGS, B. A. 2010. PKB/AKT phosphorylation of the transcription factor Twist-1 at Ser42 inhibits p53 activity in response to DNA damage. *Oncogene,* 29, 3554-65.

VINAS-CASTELLS, R., BELTRAN, M., VALLS, G., GOMEZ, I., GARCIA, J. M., MONTSERRAT-SENTIS, B., BAULIDA, J., BONILLA, F., DE HERREROS, A. G. & DIAZ, V. M. 2010. The hypoxia-controlled FBXL14 ubiquitin ligase targets SNAIL1 for proteasome degradation. *J Biol Chem,* 285, 3794-805.

VLEMINCKX, K., VAKAET, L., JR., MAREEL, M., FIERS, W. & VAN ROY, F. 1991. Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. *Cell,* 66, 107-19.

VOUTSADAKIS, I. A. 2012a. The ubiquitin-proteasome system and signal transduction pathways regulating Epithelial Mesenchymal transition of cancer. *J Biomed Sci,* 19, 67.

VOUTSADAKIS, I. A. 2012b. Ubiquitination and the Ubiquitin-Proteasome System as regulators of transcription and transcription factors in epithelial mesenchymal transition of cancer. *Tumour Biol,* 33, 897-910.

WANG, F., SLOSS, C., ZHANG, X., LEE, S. W. & CUSACK, J. C. 2007. Membrane-bound heparin-binding epidermal growth factor like growth factor regulates E-cadherin expression in pancreatic carcinoma cells. *Cancer Res,* 67, 8486-93.

WANG, X., SU, H. & RANEK, M. J. 2008. Protein quality control and degradation in cardiomyocytes. *J Mol Cell Cardiol,* 45, 11-27.

WANG, Y. & DASSO, M. 2009. SUMOylation and deSUMOylation at a glance. *J Cell Sci,* 122, 4249-52.

WARE, J. H., ZHOU, Z., GUAN, J., KENNEDY, A. R. & KOPELOVICH, L. 2007. Establishment of human cancer cell clones with different characteristics: a model for screening chemopreventive agents. *Anticancer Res,* 27, 1-16.

WEI, P. L., CHANG, Y. J., HO, Y. S., LEE, C. H., YANG, Y. Y., AN, J. & LIN, S. Y. 2009. Tobacco-specific carcinogen enhances colon cancer cell migration through alpha7-nicotinic acetylcholine receptor. *Ann Surg,* 249, 978-85.

WEINBERG, R. A. 2008. Mechanisms of malignant progression. *Carcinogenesis,* 29, 1092-5.

WELLNER, U., SCHUBERT, J., BURK, U. C., SCHMALHOFER, O., ZHU, F., SONNTAG, A., WALDVOGEL, B., VANNIER, C., DARLING, D., ZUR HAUSEN, A., BRUNTON, V. G., MORTON, J., SANSOM, O., SCHULER, J., STEMMLER, M. P., HERZBERGER, C., HOPT, U., KECK, T., BRABLETZ, S. & BRABLETZ, T. 2009. The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nat Cell Biol,* 11, 1487-95.

WILKINSON, K. A. & HENLEY, J. M. 2010. Mechanisms, regulation and consequences of protein SUMOylation. *Biochem J,* 428, 133-45.

WU, Y. & ZHOU, B. P. 2008. New insights of epithelial-mesenchymal transition in cancer metastasis. *Acta Biochim Biophys Sin (Shanghai),* 40, 643-50.

YANG, J., MANI, S. A., DONAHER, J. L., RAMASWAMY, S., ITZYKSON, R. A., COME, C., SAVAGNER, P., GITELMAN, I., RICHARDSON, A. & WEINBERG, R. A. 2004. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell,* 117, 927-39.

YANG, J. & WEINBERG, R. A. 2008. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev Cell,* 14, 818-29.

YANG, X., LIU, S. & YAN, Q. 2013. Role of fucosyltransferase IV in epithelial-mesenchymal transition in breast cancer cells. *Cell Death Dis,* 4, e735.

YANG, X. J. & SETO, E. 2008. Lysine acetylation: codified crosstalk with other posttranslational modifications. *Mol Cell,* 31, 449-61.

YANG, Z., RAYALA, S., NGUYEN, D., VADLAMUDI, R. K., CHEN, S. & KUMAR, R. 2005. Pak1 phosphorylation of snail, a master regulator of epithelial-to-mesenchyme transition, modulates snail's subcellular localization and functions. *Cancer Res,* 65, 3179-84.

YOOK, J. I., LI, X. Y., OTA, I., HU, C., KIM, H. S., KIM, N. H., CHA, S. Y., RYU, J. K., CHOI, Y. J., KIM, J., FEARON, E. R. & WEISS, S. J. 2006. A Wnt-Axin2-GSK3beta cascade regulates Snail1 activity in breast cancer cells. *Nat Cell Biol,* 8, 1398-406.

YOSHIKAWA, M., HISHIKAWA, K., MARUMO, T. & FUJITA, T. 2007. Inhibition of histone deacetylase activity suppresses epithelial-to-mesenchymal transition induced by TGF-beta1 in human renal epithelial cells. *J Am Soc Nephrol,* 18, 58-65.

YOUNG, G. W., WANG, Y. & PING, P. 2008. Understanding proteasome assembly and regulation: importance to cardiovascular medicine. *Trends Cardiovasc Med,* 18, 93-8.

YUAN, B. Z., CHAPMAN, J. A. & REYNOLDS, S. H. 2008. Proteasome Inhibitor MG132 Induces Apoptosis and Inhibits Invasion of Human Malignant Pleural Mesothelioma Cells. *Transl Oncol,* 1, 129-40.

ZHANG, Z. & LUTZ, B. 2002. Cre recombinase-mediated inversion using lox66 and lox71: method to introduce conditional point mutations into the CREB-binding protein. *Nucleic Acids Res,* 30, e90.

ZHOU, B. P., DENG, J., XIA, W., XU, J., LI, Y. M., GUNDUZ, M. & HUNG, M. C. 2004. Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition. *Nat Cell Biol,* 6, 931-40.

ZHURINSKY, J., SHTUTMAN, M. & BEN-ZE'EV, A. 2000. Plakoglobin and beta-catenin: protein interactions, regulation and biological roles. *J Cell Sci,* 113 (Pt 18), 3127-39.