FUNCTIONAL IMPLICATIONS OF IGF2-MEDIATED POST-TRANSLATIONAL MODIFICATIONS OF REST



Mehak Mumtaz

St. Edmund Hall

&

Sir William Dunn School of Pathology

Oxford University

This dissertation is submitted for the degree of Doctor of Philosophy

April 2017

ABSTRACT

Repressor element 1 silencing transcription factor/neuron-restrictive silencing factor (REST/NRSF) is a transcriptional silencer of neuronal gene expression. Experimental evidence supports its vital role in neuronal differentiation, repression of neuronal genes in non-neuronal cells and more recently, in the regulation of the cell cycle. Aberrant REST activity can lead to oncogenesis and indeed, REST can act as both a tumour suppressor and an oncogene, depending on the cellular and molecular context.

Preliminary data, using microarray analysis on mouse embryos with differential IGF2 dose supply, identified REST as one of the key transcriptional hub genes that define the signalling output of IGF2. REST was stabilised within 6 hours of exogenous exposure to IGF2, in a proteasome and PI3K-dependent manner. IGF2 is a mitogenic peptide that is crucial for normal fetal growth. IGF2 dosage control has been shown to be essential in regulating the different neurological niches of the adult brain as well as in growth and development. The hypothesis tested in this thesis was whether IGF2 stimulation induces REST protein stabilisation by a post-translational modification preventing its ubiquitination and subsequent proteasomal degradation. Here, we evaluate the mechanism of this stabilisation.

Results showed that IGF2 stabilises REST by de-phosphorylation, preventing its ubiquitination and degradation by the proteasome, leading to an increase in global REST levels, which, in turn, drives a secondary increase in the nuclear and potentially functional form of REST. The IGF2-mediated REST stability was found to be dependent on the cell lines used, the timing of the IGF2 exposure and feedback signalling. I attempted to characterise the REST phosphosite that might be responsive to IGF2 by using an in silico approach followed by mass spectrometry of exogenously expressed tagged-REST. Multiple high molecular weight (MW) REST forms were identified by Western blotting, which initiated an attempt to characterise the different MW forms of REST systematically. The 220 kDa high MW REST form was found to be phosphorylated, O-glycosylated, nuclear in localisation, and hence potentially the functional REST form. A dimerisation mechanism, mediated by a covalent linkage, was hypothesised to explain the increase in MW. CRISPR/Cas9 technology was next utilised to achieve a REST knock out for functional studies. An efficiency of 38-54% in indel generation was achieved which resulted in heterozygous REST knock out cells and a complete failure to obtain REST null homozygotes, indicating knockout-related cell lethality or differentiation phenotype.

This work suggests a role for IGF2 in stabilising and maintaining levels of REST, possibly during development, when neuronal precursor cells undergo lineage commitment and terminal differentiation, as the brain and nervous system are forming. By stabilising REST and repressing neuronal differentiation, IGF2 might redirect cells towards proliferation. Hence, we have identified a key regulatory signalling pathway, which might regulate a balance between growth and neuronal differentiation.

ACKNOWLEDGEMENTS

Firstly, I am grateful to Professor Bass Hassan for providing me with an incredible opportunity to work in his lab, and for his continuous encouragement throughout these tumultuous four years. He has a knack for turning around the most depressive of moments and showing the light at the end of the tunnel. It has been brilliant working with him and under his guidance; he has been a great mentor and a true inspiration.

Besides my supervisor, I would like to thank the rest of the Hassan lab, especially David Barnes for his unfaltering support and patience in dealing with my endless queries, and Susana Frago and Sabina Sanegre, for all their help and guidance, in lab matters and beyond.

This journey would not have been possible without funding, and I would like to thank the Higher Education Commission in Pakistan and Cancer Research UK for their support.

I would also like to acknowledge the Dunn School Proteomics Facility, especially Svenja Hester and Monika Stegmann, for their support with Mass Spectrometry, and the Genomics Facility, especially Andrew Bassett, for help with CRISPR.

I am extremely fortunate to have a supportive family, and I extend my deepest appreciation to my parents, who have always believed in me and encouraged me through all my decisions.

Lastly, I would like to thank my lifeline in Oxford; all of my friends who have made the past four years such a joyful time.

CONTENTS

1 Introduction 1

1.1 RE1-Silencing Transcription Factor 1

1.1.1 Structure of REST 2

1.1.2 REST Binding Site - RE1/NRSE 2

1.1.3 REST Isoforms 11

1.1.4 REST localization 14

1.1.5 REST in Animal Models 15

1.1.6 Gene Regulation by REST 16

1.1.7 Regulation of Neurogenesis 17

1.1.8 Regulation of REST 22

1.1.9 REST and Pathogenesis of Disease 26

1.1.10 Conclusion 30

1.2 Insulin Like Growth Factor 2 (IGF2) 31

1.2.1 IGF2 Signalling 32

1.2.2 IGF2 Function 35

1.2.3 Regulation of IGF2 35

1.2.4 IGF2 in Neurogenesis 36

1.2.5 IGF2 in Cancer 38

1.3 Aims and Rationale 40

2 Materials and methods 45

2.1 Materials 45

2.1.1 Cell culture reagents 45

2.1.2 Primary and Secondary Antibodies 45

2.1.3 Plasmids 45

2.2 Methods 45

2.2.1 General Cloning Methods 45
2.2.2 Cell Culture, Treatments and Reporter Assays 52
2.2.3 Protein Analysis 56
2.2.4 Lambda phosphatase treatment 57
2.2.5 Immunoprecipitation 57
2.2.6 Immunofluorescence 58
2.2.7 Imaging of Native Gels 59
2.2.8 Mass Spectrometry 59
2.2.9 CRISPR/Cas9 62
2.2.10 Statistical Analysis 65

3 Characterisation of the Large Molecular Weight forms of REST 67

3.1 Introduction 67

3.2 Results 74

3.2.1 Endogenous REST has multiple isoforms in HEK293T cells 74
3.2.2 One REST Splice Variant in expressed in HEK293T cell line 75
3.2.3 Higher MW REST isoforms are commonly found in cell lines 78
3.2.4 Exogenous REST is predominantly present in the higher isoforms 80
3.2.5 220kDa HMW REST band is identified as REST by Mass Spectrometry 82
3.2.6 HMW REST form is resistant to Urea 82
3.2.7 HMW REST form does not contain disulphide bonds 83
3.2.8 HMW REST is not a hetero-oligomer 86
3.2.9 HMW REST form is not N-glycosylated 90
3.2.10 HMW REST form is O-glycosylated 90
3.2.11 HMW REST form is Phosphorylated 95
3.2.12 HMW REST form is Sumoylated 95
3.2.13 Subcellular localisation of HMW and LMW forms 99

3.3 Discussion 101

3.3.1 Phosphorylation 102
3.3.2 Glycosylation 103
3.3.3 Sumoylation 104
3.3.4 Dimerisation 105

4 IGF2 alters the post-translational modification of REST 110

4.1 Introduction 110

4.2 Aims 114

4.3 Results 115

4.3.1 IGF2 stabilises REST in NIH3T3 cells in a time dependent manner 115
4.3.2 IGF2 response in Ewing Cell lines is time and IGF2 dose dependent 117
4.3.3 IGF2 alters phosphorylation of REST 120
4.3.4 IGF2 increases REST half-life. 122
4.3.5 IGF2 increases the amount of nuclear REST 125

4.4 Discussion 129

5 Characterisation of Phosphorylation sites in HMW REST 135

5.1 Introduction 135

5.1.1 Cellular Protein Phosphorylation 135
5.1.2 Insulin-like Growth factor 2 signalling 138
5.1.3 Phosphorylation of REST 139
5.1.4 Using Mass Spectrometry to Identify Protein Phosphorylation 141

5.2 Aims 143

5.3 Results 144

5.3.1 Candidate post-translational modification sites in REST 144
5.3.2 Identification of phosphorylation sites by Mass Spectrometry 156
5.3.3 Identification of IGF2-dependent Phosphosite on REST 164

5.3.4 Adapting Flag-REST-GFP fusion peptides for identification of REST phosphorylation 167
5.4 Discussion 172
6 Targeted Deletion of REST Gene in Mammalian Cells by CRISPR/Cas9 System 176
6.1 Introduction 176
6.1.1 Genome Editing 177
6.1.2 REST knock down and knock out studies 179
6.2 Aims 181
6.3 Results 182
6.3.1 Design and construction of CRISPR/Cas9 system for REST editing 182
6.3.2 Generation of Generation of Heterozygous REST KO mutants 186
6.3.3 Validation of Heterozygous REST KO mutants 190
6.3.4 Usage of X5050, a putative REST inhibitor, for achieving partial REST knock down 193
6.3.5 Generation of Homozygous REST mutants 196
6.4 Discussion 199
7 Discussion 203
7.1 Summary and Significance 203
7.2 Future Directions 208
7.2.1 REST Truncation Mutants for Phosphorylation Studies 208
7.2.2 CRISPR/Cas9 Validation 209
7.2.3 Structure of REST 210
7.2.4 Ewing sarcoma- REST Dependency 211
8 References 214
9 Appendices 230

LIST OF TABLES

Table 2.1 List of Reagents 46

Table 2.2 List of Primary Antibodies Used in Experiments 47

Table 2.3 List of Secondary Antibodies used in Experiments 47

Table 2.4 List of Plasmids used in Experiments 48

Table 2.5 A Typical PCR reaction. 49

Table 2.6 Primers used for HRMA Analysis. 64

Table 2.7 Reaction volumes for HMR PCR 64

Table 2.8 PCR program used for HRM analysis. 64

Table 3.1 Summary of the High Molecular REST observations. 72

Table 3.2 List of proteins co-purified with REST 88

Table 3.3 Predicted O-GlcNAc sites by YinOYang 1.2. 94

Table 5.1 Phosphorylated residues of REST identified by previous publications. 140

Table 5.2 Phosphorylation predictions for serine residues of REST by NetPhos. 147

Table 5.3 Phosphorylation predictions for threonine and tyrosine residues of REST by NetPhos 148

Table 5.4 Mutations in Serine, Threonine and Tyrosine residues of REST found in Cancer. 152

Table 5.5 Phosphorylation sites in REST identified from Phosphositeplus. 154

Table 5.6 Phosphorylation sites detected by LC-MS/MS. 162

Table 5.7 Analysis of Phosphorylated residues identified by Mass Spectrometry. 163

LIST OF FIGURES

Figure 1.1 Domain Structure of REST protein. 4

Figure 1.2 An Alignment of REST Orthologues. 9

Figure 1.3 RE1/NRSE consensus sequence. 10

Figure 1.4 REST gene structure and isoforms. 13

Figure 1.5 Cell-dependent function of REST. 20

Figure 1.6 Control of Neurogenesis by titration of REST levels. 20

Figure 1.7 Regulation of Class I and Class II genes by REST-corepressor complex. 21

Figure 1.8 Control of REST by various mechanisms 25

Figure 1.9 Overview of the Insulin/IGF system. 34

Figure 1.10 Unbiased detection of the transcriptional signature of IGF2 signalling in defined genetic models. 42

Figure 1.11 Identification of REST as a downstream target of IGF2 signalling. 43

Figure 3.1 Scheme of identification of HMW forms of REST. 73

Figure 3.2 Characterisation of anti-REST antibodies available commercially. 76

Figure 3.3 REST splice variants found in HEK293T cells. 77

Figure 3.4 Higher MW REST forms are ubiquitously observed in cell lines. 79

Figure 3.5 Exogenous REST is predominantly present in the higher isoforms. 81

Figure 3.6 Exogenous REST can be immunoprecipitated and is identified as REST by Mass spectrometry. 84

Figure 3.7 Urea treatment and reduction of REST. 85

Figure 3.8 N- and O-linked Glycosylation of REST. 93

Figure 3.9 REST is phosphorylated. 97

Figure 3.10 REST is Sumoylated. 98

Figure 3.11 Subcellular localisation of HMW and LMW REST forms. 100

Figure 3.13 Model showing LMW and HMW REST with their post-translational modifications and functional consequence. 109

Figure 4.1 IGF2-dependent REST stabilisation in NIH3T3 cells is time-dependent. 116

Figure 4.2 IGF2-dependent REST stabilisation in Ewing Sarcoma cell lines is dose and time-dependent. 119
Figure 4.3 Modification of the HMW REST form by IGF2 121
Figure 4.4 Half-life of REST. 123
Figure 4.5 Effect of IGF2 on REST half-life. 124
Figure 4.6 Effect of IGF2 on the localisation of REST. 127
Figure 4.7 Effect of IGF2 on the localisation of REST. 128
Figure 4.8 Putative Nuclear export signal in REST. 133
Figure 4.9 Model for IGF2 stabilisation of REST. 134
Figure 5.1 Mutations of REST in Cancer. 151
Figure 5.2 Total phosphorylation sites identified from in silico approach. 155
Figure 5.3 Modification of exogenous REST by IGF2 157
Figure 5.4 Optimisation of Mass Spectrometry pipeline. 160
Figure 5.5 Purification of REST for LC-MS/MS. 165
Figure 5.6 Identification of IGF2 dependent phosphorylation sites. 166
Figure 5.7 Validation of Flag-REST-GFP peptide reporter constructs. 170
Figure 5.8 Generation and validation of Flag-GFP tagged-REST isoforms. 171
Figure 6.1 Targeting Strategy for CRISPR/Cas9 mediated REST knock out. 184
Figure 6.2 Validation of sgRNA for CRISPR. 185
Figure 6.3 Methodology for creating REST null cell lines 187
Figure 6.4 Generation of Heterozygous REST mutants 189
Figure 6.5 Validation of Heterozygous REST mutants. 192
Figure 6.6 Using X5050 to achieve transient REST knockdown in HEK293T. 195
Figure 6.7 Killer Strategy- development of a functional selection strategy. 198
Figure 7.1 Proposed Control of Cell Proliferation and Differentiation by IGF2. 207

List of Abbreviations and Acronyms

AKT	Protein kinase B
C	Celsius
CRISPR	Clustered regularly interspaced short palindromic repeats
DAPI	4',6-diamidino-2-phenylindole
DON	6-Diazo-5-oxo-L-norleucine
EGF	Epidermal growth factor
ESC	Embryonic stem cells
FACS	Fluorescence activated cell sorting
FCS	Foetal calf serum
HMW	High molecular weight
IGF	Insulin growth factor
IgG	Immunoglobulin G
IR	Insulin receptor
IRS	Insulin receptor substrate
LMW	Low molecular weight
LOH	Loss of heterozygosity
LOI	Loss of Imprinting
MAPK	Mitogen-activated protein kinase
MW	Molecular weight
NRSE	Neuron-restrictive silencer element
NSC	Neural stem cells
OGT	O-GlcNAc transferase
OSI-906	Linsitinib
Phosphosite	Phosphorylation site
PNGase F	N glycosidase F
PUGNAc	O-(2-Acetamido-2-deoxy-D-glucopyranosylidene)aminoN-phenyl carbamate
PI3K	Phosphoinositide 3-kinase
PTM	Post-translational modification
RE1	Repressor element 1
REST	Repressor element 1 silencing factor
SDS	Sodium dodecyl sulphate
sgRNA	Single guide RNA
shRNA	Short hairpin RNA
siRNA	Small interfering RNA
TBST	Tris buffer saline plus tween

LIST OF APPENDICES

Appendix 1 IGF2 stabilises REST protein in a mouse fibroblast NIH3T3 cell line. *231*

Appendix 2 Human cell line screen for REST response to IGF2 and MG132. ..... *232*

Appendix 3 Titration of cell lines with varying IGF2 doses. *233*

Appendix 4 REST exists as higher MW isoforms. *234*

Appendix 5 Higher MW bands show a concomitant change in levels along with 116 kDa REST form. *235*

Appendix 6 Common cancer-associated mutations do not correlate with REST stabilisation. *236*

Appendix 7 REST knock down causes de-repression of L1Cam. *237*

1 INTRODUCTION

1.1 RE1-Silencing Transcription Factor

Repressor Element 1 Silencing Factor (REST), also known as Neuron-Restrictive Silencing Factor (NRSF), was identified in 1995 as a master regulatory gene of the neuronal phenotype (Chong et al., 1995; Schoenherr and Anderson, 1995). REST was originally cloned as the transcription factor that bound the 21 bp repressor element 1 (RE1) or neuron-restrictive silencer element (NRSE) sequence found in the regulatory regions of many neuronal genes. REST is a multi-zinc finger protein, similar to the Drosophila melanogaster Krüppel repressor. It binds to RE1 sequence through its zinc fingers; that results in silencing of neuronal genes. RE1 sequences, initially identified for SCG10, a microtubule-destabilising protein, and the type II sodium channel, have subsequently been shown to be present in numerous neuronal genes, spanning different features of the neuronal phenotype (Chong et al., 1995; Schoenherr and Anderson, 1995). These genes include ion channels, neuropeptides, cytoskeletal components, neurotransmitter receptors and neural cell adhesion molecules (Bruce et al., 2004).

1.1.1 Structure of REST

REST is a 116 kDa protein with a DNA-binding domain consisting of eight zinc fingers (Chong et al., 1995; Schoenherr and Anderson, 1995). Each zinc finger domain is comprised of two conserved cysteine and histidine residues respectively, and binds both DNA and RNA. As shown in Figure 1.1, REST also has two repressor domains located on both N and C-termini (Thiel, Lietz and Cramer, 1998; Tapia-Ramirez et al., 1997). A ninth zinc finger is contained within the C-terminal repressor domain. The two repressor domains are thought to function independently, but together they allow for multifactorial control of gene expression (Ballas et al., 2005). In addition to these domains, REST contains an intrinsically disordered region, consisting of less conserved lysine and proline-rich regions (See Figure 1.2).

1.1.2 REST Binding Site - RE1/NRSE

REST binds to its target genes via an evolutionarily conserved canonical 21 bp RE1 site. Figure 1.3 shows the consensus sequence for this site. Multiple genome-wide association studies, including bioinformatics analysis and chromatin immunoprecipitation sequencing assays (ChIP-seq), have been performed to identify the complete set of RE1 sites and their associated genes (Bruce et al., 2004; Johnson et al., 2007; Otto et al., 2007). These studies demonstrate the presence of over 1900 RE1 sites within the human genome, the majority being located within intergenic regions (40%). Intronic RE1 occupy 24% of the RE1 sites and only 15% of RE1 sites lie within promoter sequences of target genes (Jothi et al., 2008). REST target genes include neuron-specific genes such as those encoding for ion channels, neurotransmitter receptors and synthesizing enzymes, transporters, neuronal cytoskeletal proteins, proteins involved in vesicular trafficking, and axonal guidance as well as genes involved in metabolic processes, immune response and

cardiovascular functions (Bruce et al., 2004). In addition to these genes, some REST-binding sites have also been found adjacent to non-coding RNA genes, including microRNA (miRNA), short hairpin RNA (shRNA) and long noncoding RNA (ncRNA) (Johnson and Buckley, 2009; Rossbach, 2011; Volvert et al., 2014). REST control of miRNA expression has been especially studied and experimentally validated. REST-dependent miRNA include mir-9, mir-124, mir-132, and mir-153, all of which have been shown to promote neuronal differentiation (Conaco et al., 2006).

The canonical 21bp RE1 sequence can be subdivided into two non-canonical half-sites, the left half-site and the right half-site, separated by a spacer sequence (See Figure 1.3). REST has been shown to bind to the two non-identical and non-palindromic sites independently, but the binding affinity is weaker and less functional than full-site binding (Jothi et al., 2008). These findings suggest a mechanism that facilitates active REST-mediated repression of canonical RE1 site-containing target genes, common to all cell types, whereas for the non-canonical motifs, these control target genes which are cell and tissue specific (Bruce et al., 2009).

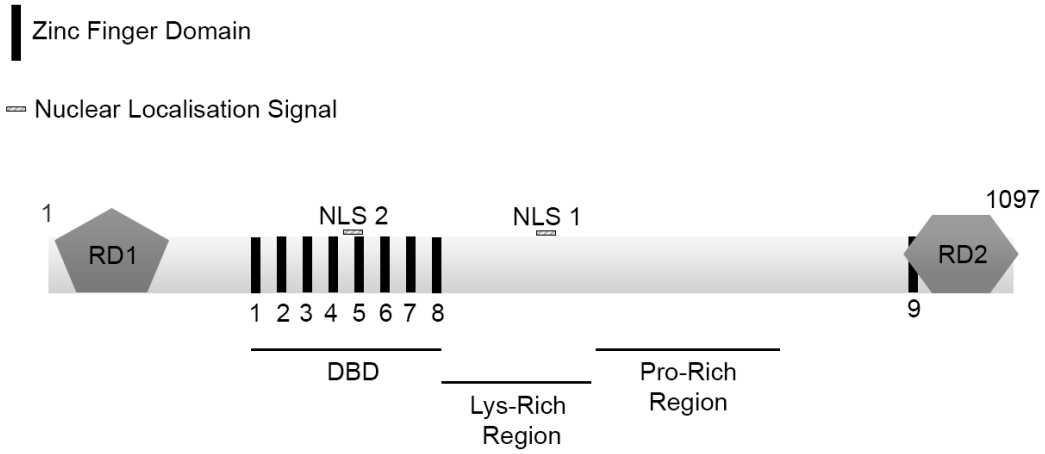
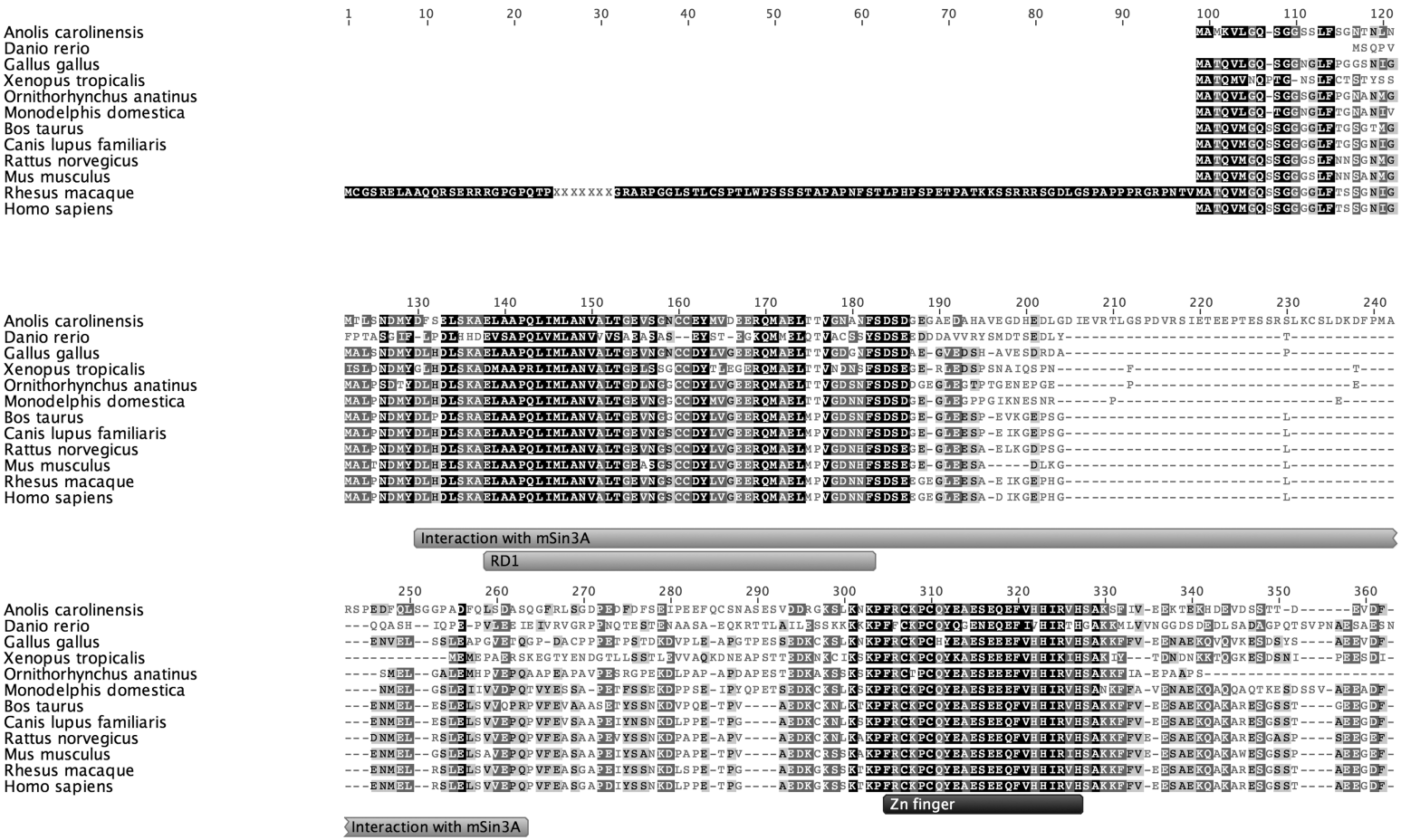
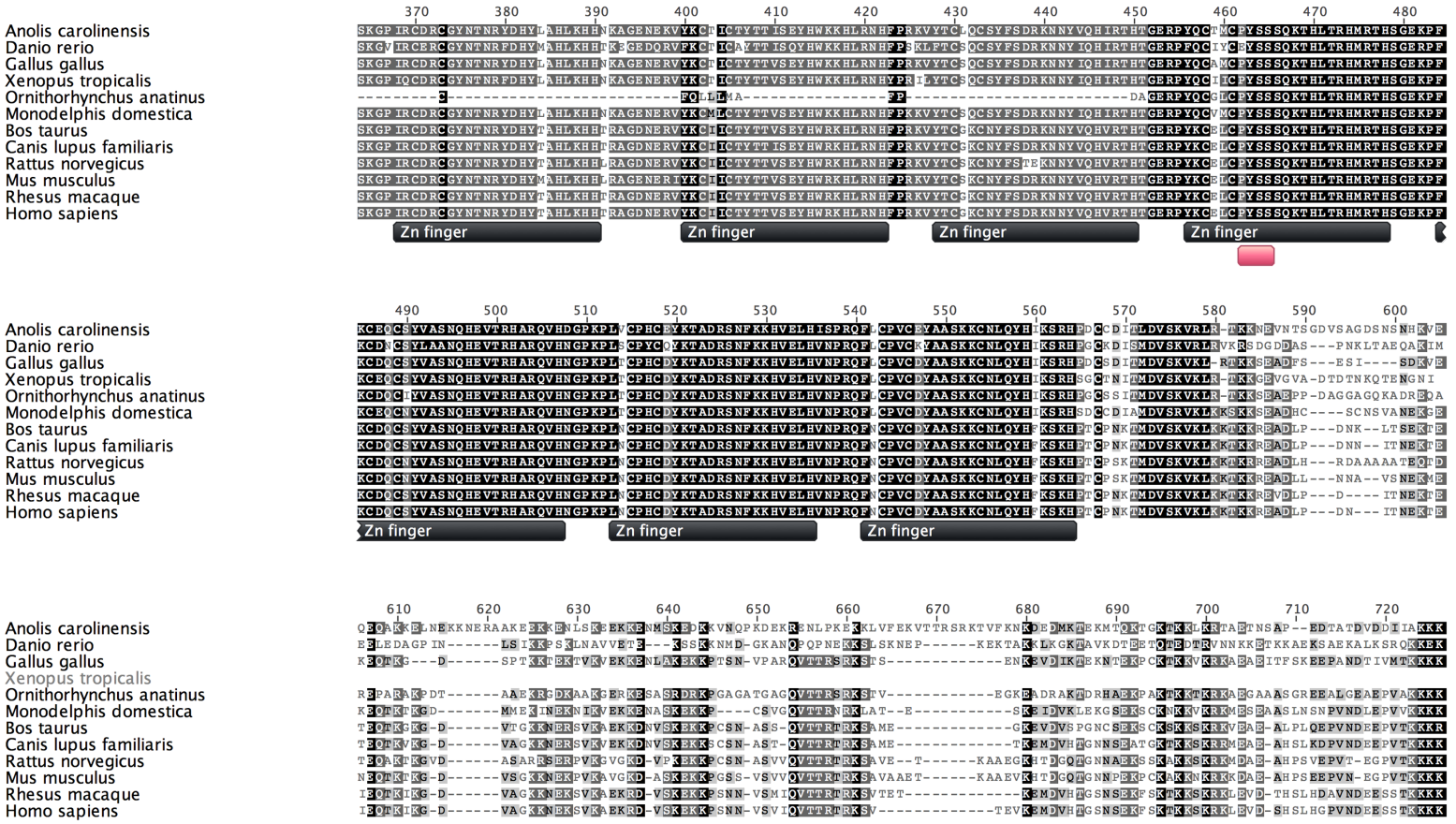
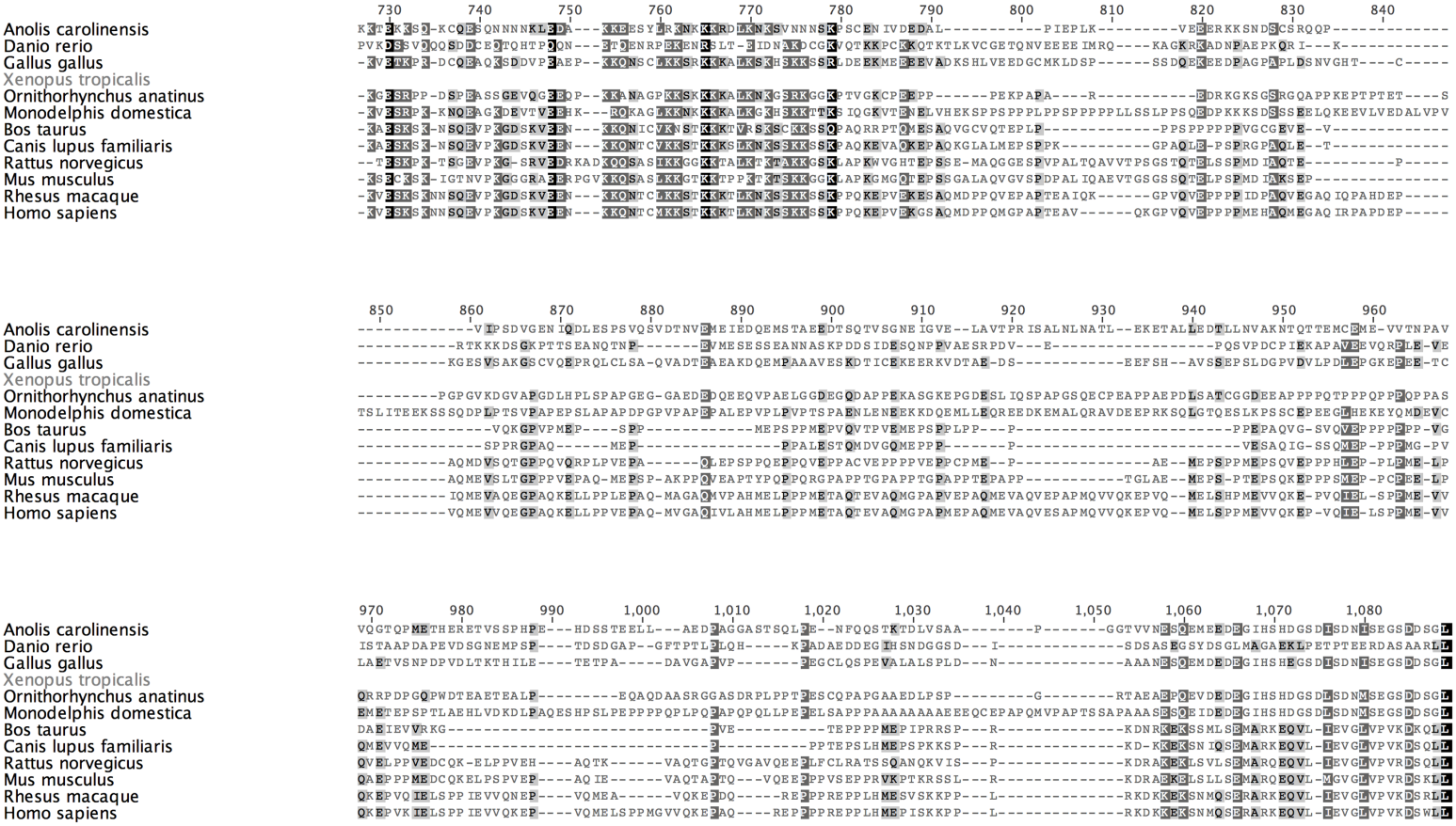


Figure 1.1 Domain Structure of REST protein.
A schematic of REST protein structure showing the DNA binding domain (DBD) consisting of 8 zinc fingers, C-terminal and N-terminal Repressor domains (RD1 and RD2), Lysine-rich and Proline-rich domains, and nuclear localisation signals (NLS).


Anolis carolinensis
Danio rerio
Gallus gallus
Xenopus tropicalis
Ornithorhynchus anatinus
Monodelphis domestica
Bos taurus
Canis lupus familiaris
Rattus norvegicus
Mus musculus
Rhesus macaque
Homo sapiens
Interaction with mSin3A
RD1
Zn finger
Interaction with mSin3A


Anolis carolinensis
Danio rerio
Gallus gallus
Xenopus tropicalis
Ornithorhynchus anatinus
Monodelphis domestica
Bos taurus
Canis lupus familiaris
Rattus norvegicus
Mus musculus
Rhesus macaque
Homo sapiens
Zn finger
Zn finger
Zn finger
Zn finger
Zn finger
Zn finger
Zn finger
Xenopus tropicalis


Anolis carolinensis
Danio rerio
Gallus gallus
Xenopus tropicalis
Ornithorhynchus anatinus
Monodelphis domestica
Bos taurus
Canis lupus familiaris
Rattus norvegicus
Mus musculus
Rhesus macaque
Homo sapiens

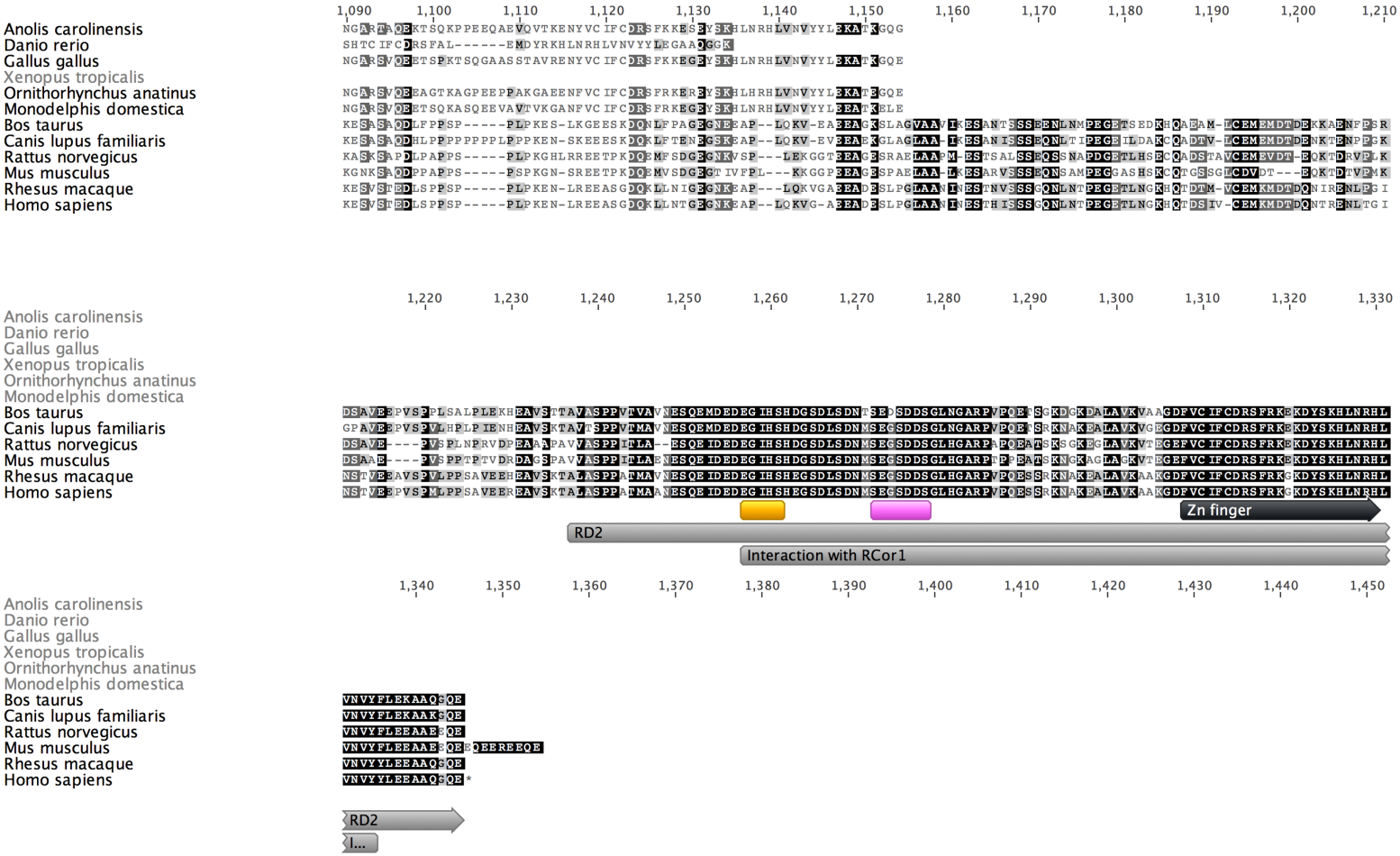

Anolis carolinensis
Danio rerio
Gallus gallus
Xenopus tropicalis
Ornithorhynchus anatinus
Monodelphis domestica
Bos taurus
Canis lupus familiaris
Rattus norvegicus
Mus musculus
Rhesus macaque
Homo sapiens
Zn finger
RD2
Interaction with RCor1
RD2

Figure 1.2 An Alignment of REST Orthologues.
An evolutionary alignment of REST orthologues, showing key features, in Homo sapiens (human), Macaca mulatta (Rhesus monkey), Mus musculus (Mouse), Rattus norvegicus (Rat), Canis lupus familiaris (Dog), Bos taurus (Cow), Monodelphis domestica (Opposum), Ornithorhynchus anatinus (Platipus), Xenopus tropicalis (Western clawed frog), Gallus gallus (Chicken), Danio rerio (Zebrafish) and Anolid carolinensis (Carolina Anole). Alignment was performed using Geneious software. The degree of alignment between sequences is represented by shades, darker shades indicating more conserved and lighter ones indicating less conserved residues. REST has a DNA binding domain consisting of 8 zinc fingers, N and C terminal repressor domains (RD1 and RD2) which assist the recruitment of various REST cofactors, 2 β-TrCP sites (annotated in pink and orange respectively) and a HAUSP binding site (annotated in pink).

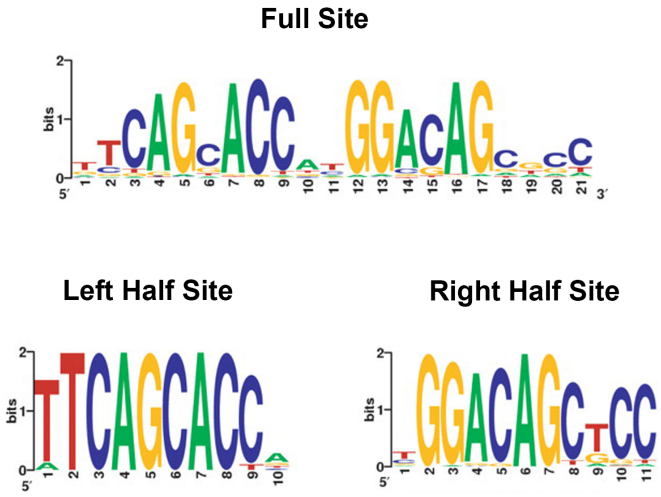


Figure 1.3 RE1/NRSE consensus sequence.
Above: the canonical 21bp sequence of RE1/NRSE is shown. Below: the canonical NRSE is split into two half-sites, the left-half site and right-half site. (Modified from Jothi et al., 2008)

1.1.3 REST Isoforms

Figure 1.4 shows the structural organisation of the human REST gene. REST gene consists of three coding exons, exon IV-VI, three non-coding exons, I-III, and an alternatively spliced exon N, situated between exon V and VI (Palm, Metsis and Timmusk, 1999). Exon IV encodes the translation initiation codon to the end of zinc finger 4, exon V encodes the spacer region between zinc finger 4 and 5 up till the end of zinc finger 5, and the remaining region is coded by exon VI. The exon and intron structure is conserved across human, mouse and rat.

Several human splice variants of REST have been reported in literature (Palm, Metsis and Timmusk, 1999). The hREST-N62 transcript is generated by the usage of a 3' splice site resulting in an insertion of the 62 bp alternative exon N. The second transcript, hREST-N4, is due to a cryptic 5' splice site within exon N62. This splice site is located 4 bp downstream of the N62 3' splice site and results in a 4 bp addition of exon N. Both of these human splice variants encode truncated proteins with five zinc fingers, similar to the two REST isoforms (REST4 and REST5) expressed in mature neurons of adult rat brain. A third splice variant, hREST-5FΔ, is a result of exon 5 and exon N skipping and results in a protein with a missing fifth zinc finger (Figure 1.4). Another transcript, hREST1 contains the first exon only, resulting in a four-zinc finger protein.

The truncated REST isoform has been identified in human SCLC cells (where it is called sNRSF) (Coulson et al., 2000) as well as in neuroblastoma cells (where it is called REST4) (Palm, Metsis and Timmusk, 1999). This isoform has five zinc fingers, a missing C-terminal repressor domain and has been proposed to antagonize REST-mediated gene repression, hence acting in a dominant-negative fashion. In neurons, REST is frequently expressed in its truncated form, which competes with full-length REST for target gene binding. Shimojo et al. showed that the rodent REST4 inhibited

REST binding to RE1 of the cholinergic gene, causing de-repression (Shimojo et al., 1999). Additionally, REST4 was found to stimulate gene expression in non-neuronal cells stimulated by glucocorticoids (Abramovitz et al., 2008).

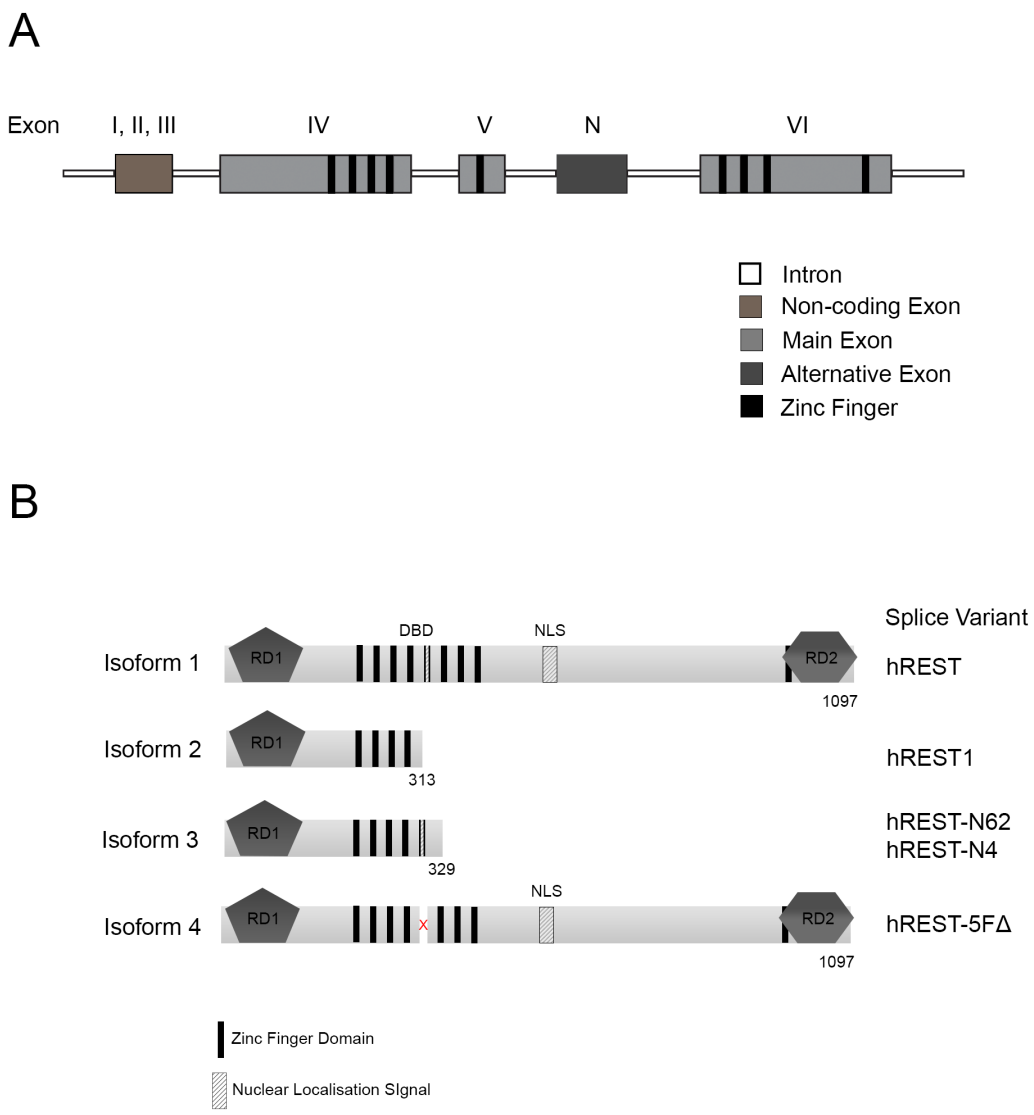


Figure 1.4 REST gene structure and isoforms.
A) A schematic illustrating REST gene structure including exons, introns and zinc fingers. B) A schematic illustrating REST full-length protein (Isoform 1) and alternative isoforms.

1.1.4 REST localization

A nuclear localisation signal (NLS) of REST was initially shown to lie between amino acids 512 and 522 (NLS1 in Figure 1.1) (Grimes et al., 2000). It was later observed that the truncation mutant, REST4, localises to the nucleus despite the absence of the NLS. This observation led to further investigation using deletion mutagenesis and point mutations, which concluded with zinc finger 2-5 of REST4 being essential for localisation (Shimojo, Lee and Hersh, 2001). Zinc finger domains 2-5, but not 2-4, were found to target β-galactosidase to the nucleus, demonstrating the importance of zinc finger 5. Furthermore, the 4-zinc finger isoform of REST, which is similar to REST4 but lacks zinc finger 5, was only found to be cytoplasmic in localisation. However, point mutations in zinc finger 5 do not prevent nuclear targeting. These findings suggest that the nuclear targeting signal is present in the region proximal and involving zinc finger 5 (amino acids 296-328) and that the signal is probably independent of the zinc finger domain structure (Shimojo, Lee and Hersh, 2001).

The translocation of REST into the nucleus appears tightly regulated. This regulation is proposed to be mediated by a complex formed by REST-interacting LIM domain protein (RILP) - a nuclear translocation receptor - the Huntingtin protein, and dynactin p150Glued (Shimojo, 2008). REST directly binds to RILP via its LIM domains and is translocated into the nucleus (Shimojo and Hersh, 2003). Knockdown of RILP using siRNA is sufficient to re-localise REST to the cytosol. Another protein, the huntingtin-associated protein-1 (HAP1), which binds to huntingtin protein, prevents the transport complex from entering the nucleus (Shimojo, 2008). As HAP1 is expressed predominantly in neuronal cells, it might play an important role in controlling the cellular localization of REST in these cells. In addition, REST4 acts as a dominant-negative regulator of REST nuclear translocation by competing with full-length for the REST binding site on RILP (Shimojo and Hersh, 2003).

1.1.5 REST in Animal Models

REST plays an essential role during embryogenesis. Loss of REST leads to embryonic lethality as demonstrated by homozygous REST knockout mice (Chen, Paquette and Anderson, 1998). Mice appear healthy until embryonic day 9.5, when all three germ layers have been formed, indicating that neuronal progenitor development can progress in the absence of REST. After this time, however, widespread apoptosis follows, resulting in malformations in the brain telencephalon and death by E11.5 (Chen, Paquette and Anderson, 1998). Loss of REST led to de-repression of neuronal tubulin, a REST target gene. It was observed that the lack of REST resulted in expression of very few REST target genes, suggesting the requirement of other positive activators, along with REST absence, for the expression of multiple neuronal differentiation genes. This result was also observed in chicken embryos, where downregulation of REST using a dominant-negative REST form, resulted in expression of neuronal tubulin and some other neuronal genes, in both non-neuronal tissues and neuronal progenitor cells (Chen, Paquette and Anderson, 1998).

Aoki et al. used REST conditional knock out (KO) mice to show that mice lacking REST in the brain appear normal, suggesting that REST is not required for brain development (Aoki et al., 2012). Furthermore, genetic ablation of REST in the brain does not appear to lead to altered target gene expression, in contrast to *in vitro* findings in cultured neuronal progenitor cells. This indicates the importance of the brain-specific environment in neuronal gene repression in the absence of REST, making REST dispensable for embryonic neurogenesis *in vivo*. On the other hand, REST target gene expression was found to be upregulated in non-neuronal cells upon loss of REST, suggesting active repression of neuronal genes by REST in non-neuronal cells. Conversely, Nechiporuk et al. show that switching off the *REST* gene in neural progenitor cells resulted in significantly smaller brains with fewer neurons,

accompanied with high levels of DNA damage during the S-phase of the cell cycle (Nechiporuk et al., 2016). DNA damage was demonstrated to cause an accumulation of chromosomal abnormalities in the cells, leading to p53-triggered apoptosis, which resulted in microcephaly. Furthermore, loss of p53 and REST together was shown to result in glioblastoma. The contradictory results with Aoki et al. were explained by the presence of a C-terminal REST peptide that is able to bind certain REST target genes and recruit corepressors (Nechiporuk et al., 2016).

1.1.6 Gene Regulation by REST

REST carries out its silencing function by acting as a platform, which recruits numerous cofactors. This is mediated via two repressor domains located at the N and C termini, which recruit the two major co-repressors, mSin3 (Grimes et al., 2000) at the N-terminus and CoREST (Andres et al., 1999) at the C-terminus. mSin3 further engages proteins such as the methyl-DNA-binding protein MeCP2 and histone deacetylases (HDACs), whereas Co-REST interacts with histone H3K9 methyltransferases, HDACs, SWI/SNF chromatin remodelling complex via brg1, among others (Huang, Myers and Dingledine, 1999).

The recruitment of numerous cofactors and formation of the repressor complex turns ‘off’ target gene expression by employing chromatin modification. For instance, HDACs remove acetyl groups from histone proteins, resulting in a condensed chromatin state that impedes transcription (Bingham et al., 2007). The histone demethylase LSD1 demethylates H3K4, which is a marker of gene activation (Shi et al., 2004). G9a dimethylates H3K9, resulting in H3K9me2, a marker for gene repression (Roopra et al., 2004). CDYL plays an essential role in this process by physically bridging REST and G9a. H3K9 dimethylation recruits HP1, which causes further condensation of chromatin (Roopra et al., 2004). Recruitment of BRG1 leads

to chromatin remodeling allowing REST better access to target genes (Battaglioli et al., 2002). In addition to this, REST also plays a direct role in repressing transcription by interacting with the basal transcription machinery. REST has been shown to interact with TATA-binding protein (TBP), inhibiting the formation of the preinitiation complex (PIC), as well as with small CTD phosphatases bound to RNA Pol II, inhibiting Pol II activity (Yeo et al., 2005; Murai et al., 2004). Hence, REST functions as an epigenetic regulator of gene expression and turns off target gene expression indirectly by histone deacetylation, methylation, and chromatin remodeling and directly by inactivation of RNA Pol II. Consistent with this, RE1 sites are not only found upstream of genes but also within gene introns as shown by CHIPseq experiments (Johnson et al., 2007). Binding of these cofactors also appears cell type-specific and context dependent, and therefore, represents an elegant mechanism for differential RE1-dependent gene regulation.

1.1.7 Regulation of Neurogenesis

REST acts to regulate a large number of genes, with regulation fundamentally dependent on the affinity of REST repressor complex to its target genes as well as the cellular context and so the total amount of REST protein and localisation in the cell (Figure 1.5). REST plays a crucial role in ensuring repression of neuronal genes in non-neuronal tissues, supported by predominant expression of REST in non-neuronal tissues (Jones and Meech, 1999). Additionally, by preventing neuronal differentiation, through keeping the untimely expression of neuronal genes in neural precursor cells (partially differentiated cells) in check, it plays an important role in contributing to and maintaining a stem cell-like phenotype.

Several mechanisms have been shown to regulate REST activity in the nervous system. Ballas et al. have shown that the transition from embryonic stem (ES) cells to

neural stem cells (NSC) is also accompanied by proteasomal degradation of REST (Ballas et al., 2005). Lower REST levels relieve gene repression in NSCs, poising them for neuronal gene transcription. As NSCs differentiate into mature neurons, they undergo inhibition of REST transcription. This may be mediated by the retinoic acid receptor element in REST promoter, which binds the unliganded retinoic acid receptor-repressor complex (Ballas et al., 2005). Absence of REST transcription diminishes REST protein levels, de-repressing neuronal genes (Figure 1.6).

An epigenetic mechanism for REST control was discovered by Kuwabara et al. and involves a small non-coding dsRNA with sequence homology to RE1 (Kuwabara et al., 2004). These dsRNA molecules are predominantly expressed in the early differentiation stages of adult neuronal progenitor cells. This dsRNA recognizes and interacts with the REST repressor complex to trigger a conformational change, which prevents its association with corepressor proteins. Hence, by binding dsRNA, REST is converted from a transcriptional repressor to a transcriptional activator. Introduction of REST dsRNA is enough to trigger expression of REST target genes and induce neuronal differentiation. A third mechanism for REST control within the CNS involves the REST isoform, REST4, which acts as a dominant negative and prevents REST binding to its target site (Zuccato et al., 2003).

Two types of neuronal genes have been classified, class I and class II, which have different mechanisms of regulation (Figure 1.7). For class I genes, removal of REST from RE1 sites is enough to cause expression. Class II genes have an additional regulatory control step in the form of CoREST/HDAC and MeCP2/mSin3/HDAC complexes occupying methylated CpG sites in the promoter region. These complexes remain intact, repressing gene expression, even after removal of REST upon transition to mature cortical neurons. Application of a stimulus, such as membrane

depolarization, can remove the MeCP2 complex allowing gene expression and providing neuronal plasticity (Ballas et al., 2005). Overall, this provides for two modes of action of REST, which can result in either active repression or gene silencing, and emphasises the importance of genetic context in determining the effect of REST on gene silencing.

In ES cells, high REST levels, in concert with pluripotency factors, Oct4, Sox2, and Nanog, function to repress a large number of genes (Jorgensen et al., 2009). However, the role of REST in maintaining ESC pluripotency is currently contentious. Singh et al. found evidence that REST is a key regulator of pluripotency and controls the switch between stem cells and neuronal cells (Singh et al., 2008), but this was refuted by other studies (Buckley et al., 2009; Jorgensen et al., 2009) which concluded that REST is not required for maintaining ES cell pluripotency. Two recent studies provided further evidence for the latter claim. The first study used zinc finger nucleases to disrupt REST gene expression in zebrafish and reveals normal neurogenesis for the most part in mutants (Kok et al., 2012). The second study, which employs REST conditional knockout mice, demonstrates that mice lacking REST in the brain grow into normal adults and concluded that REST is expendable for embryonic neurogenesis (Aoki et al., 2012). This indicates that besides genetic context, the other important determinant of REST function is the cellular/tissue context.

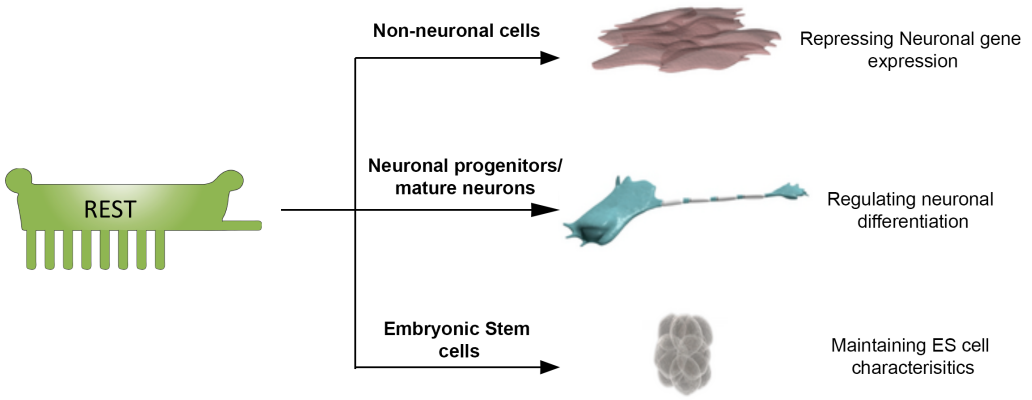


Figure 1.5 Cell-dependent function of REST.
REST performs different functions depending on the cellular context. In non-neuronal cells, it represses neuronal genes, in neuronal progenitors and mature neurons, it has a role in regulating neuronal differentiation and in ES cells, it functions to maintain ES cell characteristics. Adapted from (Hermanson, 2008).

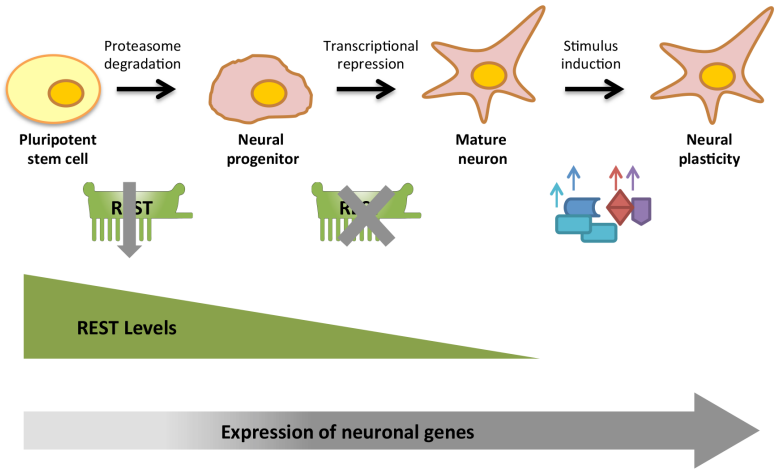


Figure 1.6 Control of Neurogenesis by titration of REST levels.
See text for detail. Adapted from(Coulson, 2005).

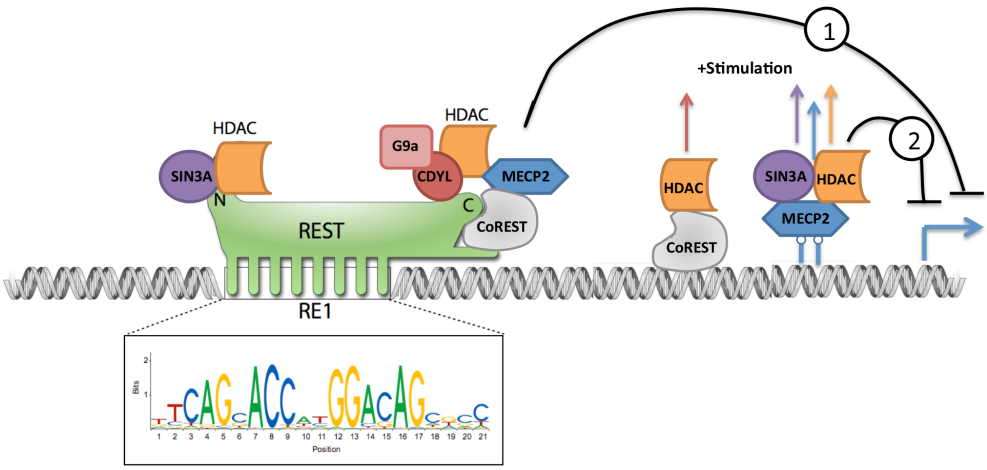


Figure 1.7 Regulation of Class I and Class II genes by REST-corepressor complex.
REST binds to RE1 elements on target genes and recruits multiple co-repressors to epigenetically silence target genes. Class I genes are repressed by bound REST-corepressor complex whereas Class II genes have an additional control step consisting of CoREST/HDAC and MeCP2/mSin3/HDAC complexes that occupy methylated CpG sites in promoter regions of target genes and repress gene expression, even after removal of REST. Application of a stimulus, such as membrane depolarization, results in removal of these complexes allowing gene expression and providing neuronal plasticity. Adapted from (Coulson, 2005).

1.1.8 Regulation of REST

Endogenous REST levels appear considerably different in different tissues and different cell types, suggesting differential regulation of the amount and activity of REST (Weissman, 2008). Recent studies have revealed various mechanisms by which cells achieve this (See Figure 1.8 for a summary).

One of the most critical mechanisms for the regulation of REST appears to be mediated by ubiquitination and deubiquitination, described as a 'Ying-Yang' control system by Huang et al. (Huang et al., 2011). Westbrook et al. showed that the E3 ubiquitin ligase, β-TrCP, recognises a conserved phosphodegron (S1024, S1027, S1030) at the C-terminus of REST (highlighted in Figure 1.2), binds and ubiquitinates it, thus targeting REST for proteasomal degradation (Westbrook et al., 2008). Degron-mutant REST fails to bind β-TRCP leading to increased REST stability. Neural differentiation is accompanied by degradation of REST (Westbrook et al., 2008). β-TrCP overexpression is frequent in human epithelial cancers, and ectopic expression can transform human mammary epithelial cells (HMEC) in a REST-dependant manner (Westbrook et al., 2008). A second adjacent phosphodegron (E1009, S1013, highlighted in Figure 1.2) was identified by Guardavaccaro et al., who demonstrated a drop in REST levels during the G2 phase of the cell cycle, contingent on β-TrCP presence and the phosphodegron (Guardavaccaro et al., 2008a). Conversely, a deubiquitinase, HAUSP (herpesvirus- associated ubiquitin-specific protease) acts to counterbalance β-TrCP activity via a consensus site on REST (310-PYSS-313, highlighted in Figure 1.2), hence thwarting neural differentiation(Huang et al., 2011). HAUSP and REST protein levels fall concurrently upon differentiation of neural progenitors whereas the opposite is true for β-TrCP levels. HAUSP knockdown in neuronal progenitor cells decreases REST levels and brings about neuronal differentiation, whilst HAUSP overexpression results in increased REST stability by

counteracting β-TrCP-mediated ubiquitination (Huang et al., 2011). Hence, β-TrCP and HAUSP form a very efficient post-translational mechanism wherein one player stabilizes REST protein by deubiquitination, and the other degrades it through ubiquitination. The cell's decision to either differentiate or maintain its stem-like state appears to be guided by REST protein levels, which are in turn determined by the balance between these two activities. Casein Kinase 1 (CK1) was identified as an upstream effector of both the C-terminal degrons (Kaneko et al., 2014). CK1 phosphorylates both degrons, the phosphorylation being crucial for β-TrCP association and proteasomal degradation of REST. The regulatory factors acting upstream of HAUSP, which control its activity, are yet to be identified.

A role for the telomere shelterin protein, TRF2, has also been implicated in the stabilization of REST (Zhang et al., 2008). TRF2 colocalises with REST, and these TRF2-REST complexes are shown to be localised in PML-nuclear bodies by co-immunoprecipitation. RNAi knockdown and dominant-negative inhibition of TRF2, separately, have been shown to cause dissociation of the complex, resulting in degradation of REST via the proteasome. Consequently, REST target genes are de-repressed, and cells begin to acquire a neuronal phenotype. However, selective telomere damage without affecting TRF2 levels does not phenocopy cell differentiation (Zhang et al., 2008).

Two signalling pathways have been described to have an effect on REST gene expression. Bone morphogenetic protein (BMP) signalling was shown to be directly involved in the expression of REST, as neuronal precursor cells (NPCs) differentiate to astrocytes, a type of brain cells (Kohyama et al., 2010). Astrocyte formation requires a restriction of neuronal gene expression, and this is achieved by the presence of REST. Treatment of NPCs with BMP2 causes sustained expression of REST

through the binding of Smad1, Smad5, or Smad8, which are downstream components of the signalling pathway, to Smad-binding elements (SBEs) in REST regulatory regions. REST subsequently associates with neuronal genes and maintains their repression. Wnt signalling is another pathway implicated in control of REST. The Wnt pathway is involved in growth and patterning of neural tissues (Klaus and Birchmeier, 2008). Overexpression of Wnt and β-catenin, a major downstream effector of the pathway, causes an upregulation of REST in chick spinal cord (Nishihara, Tsuda and Ogura, 2003) as well as in embryonic carcinoma cells (Willert et al., 2002). Moreover, a binding site for TCF, a transcription factor that interacts with β-catenin to promote gene expression, has also been identified in the REST promoter (Willert et al., 2002).

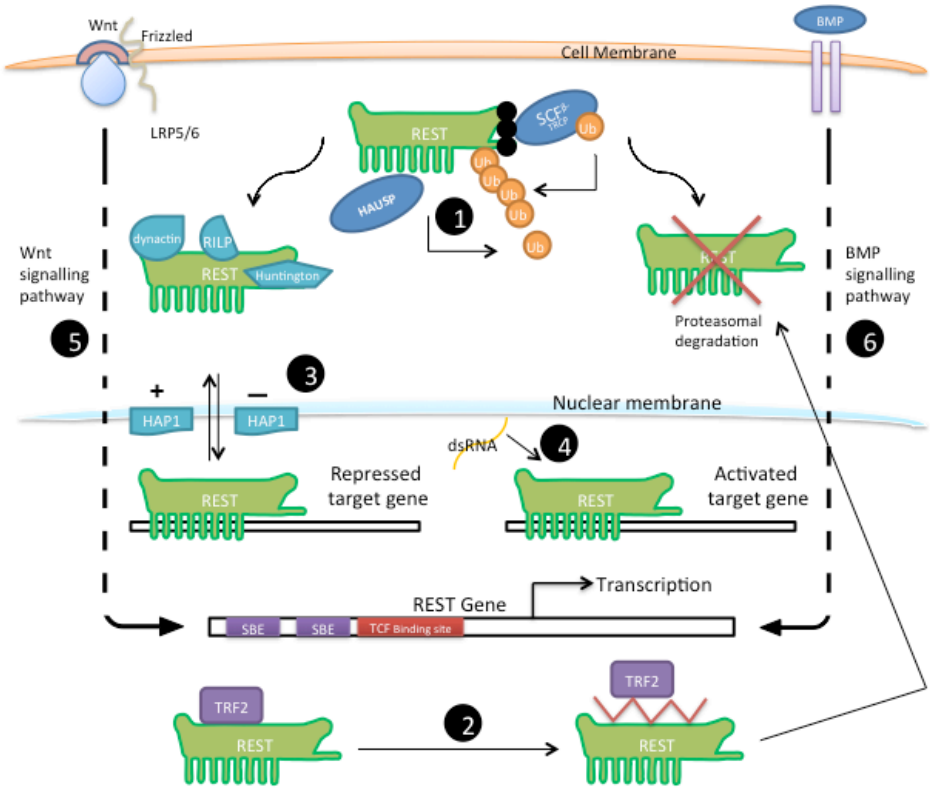


Figure 1.8 Control of REST by various mechanisms.
1) Post-translational control of REST by ubiquitination and proteasomal degradation through binding of SCF $^{\beta\text{-TrCP}}$ and deubiquitination and stabilization by binding of HAUSP. 2) Stabilised REST binds to TRF2, REST-TRF2 complexes are localised to PML-nuclear bodies, and removal of TRF2 directs protein for degradation. 3) RILP, dynactin and Huntington bind to REST and enable translocation into the nucleus while HAP1 binds to Huntington and prevents translocation. 4) A dsRNA molecule binds to REST and converts it from a transcriptional repressor to a transcriptional activator. 5 and 6) The Wnt and BMP signalling pathway cause an increase in transcription of REST.

1.1.9 REST and Pathogenesis of Disease

A direct corollary of REST's crucial role in gene regulation is the link between aberrant REST activity and various pathological diseases, including neurological diseases and cancer.

1.1.9.1 Neurological Diseases

REST and its target genes have been implicated in the pathogenesis of various neurological diseases. As mentioned earlier, in neuronal cells, REST is usually sequestered in the cytoplasm by interaction with the huntingtin protein. In Huntington's disease, mutant huntingtin protein is unable to bind to REST, leading to translocation of REST into the nucleus, increased binding of REST to target genes and their repression. Brain-derived neurotrophic factor (BDNF) is one such REST target gene, repression of which is believed to cause neuronal degeneration (Zuccato et al., 2003). Similarly, in Down syndrome, aberrant REST expression and resultant dysregulation of its target genes contributes to the disease phenotype. These alterations in REST levels are mediated by a REST regulator called dual specificity tyrosine phosphorylation-regulated kinase 1A (DYRK1A). Trisomic levels of DYRK1A in Down syndrome result in high REST levels in adult neurons and low levels in embryonic neurons (Canzonetta et al., 2008; Lepagnol-Bestel et al., 2009). Rest has also been linked to X-linked mental retardation and epilepsy. REST repressor complex associates with the histone H3K4 demethylase SMCX, a protein that is implicated in mental retardation and epilepsy. RNAi knockdown of SMCX de-represses REST target genes, suggesting that impaired SMCX activity might interfere with REST-dependent neuronal gene repression and contribute to disease pathogenesis (Tahiliani et al., 2007). Furthermore, genes that are regulated by REST

as well as SMCX, such as SCG10, BDNF, and SCN2A, have been shown to be mutated and dysregulated in mental retardation and epilepsy (Bahn et al., 2002).

REST has also been linked to aging. Aging neurons normally undergo an upregulation of REST, associated with repression of genes promoting cell death and Alzheimer disease, as well as protection from oxidative stress and toxicity by amyloid β-protein (Lu et al., 2014). These features are lost in Alzheimer's disease and dementia as REST is removed from the nucleus and directed to autophagosomes, along with other misfolded proteins(Lu et al., 2014).

1.1.9.2 Cancer

In recent years, an extensive amount of data has accumulated which maintains REST deregulation to be a key factor in many different forms of cancers. Loss of REST activity, overexpression, as well as copy number changes are all observed in numerous cancer types leading to a very fascinating deduction – REST can act as both a tumour suppressor and an oncogene, depending on the cellular and molecular context.

1.1.9.2.1 Tumour suppressor role of REST

REST was first identified as a tumour suppressor in epithelial cells from an RNAi-based genetic screen by Westbrook et al. (Westbrook et al., 2005). REST knockdown caused anchorage-independent growth in epithelial cells, which was countered by REST overexpression. The same study also found one-third of colon cancers to contain deletions in at least one of their REST alleles, resulting in haploinsufficiency. Truncated forms of REST can act as dominant-negative inhibitors of wild-type REST and abrogate its function. A premature translation termination mutant, REST- FS, was also found in a colon cancer cell line (Westbrook et al., 2005) and a splice variant identified as a clinical marker in small cell lung cancer (SCLC) (Coulson et al., 2000).

Diminished REST activity has also been documented in breast cancer and prostate cancer LNCaP cells. Wagoner et al. identified a RESTless subtype of breast cancer, using a 24-gene signature of REST targets, which expressed a truncated splice variant of REST and is associated with poor patient prognosis (Wagoner et al., 2010). REST degradation by β-TrCP overexpression has been shown to transform HMEC cells (Westbrook et al., 2008). Loss of REST is accompanied by aberrant expression of many neuroendocrine genes which are REST targets, including the glycine receptor α1 subunit (Gurrola-Diaz et al., 2003) and vasopressin (Coulson, Stanley and Woll, 1999) in SCLC cancer cell lines, hence giving rise to a neuroendocrine tumour phenotype.

The mechanistic details of REST's tumour suppressor function are slowly coming into light. REST has been shown to work with its corepressor, CDYL, which facilitates the recruitment of the histone methylase G9a to target genes, resulting in their repression (Mulligan et al., 2008). One such target gene is the oncogene, TrkC. RNAi knockdown of REST, DYL and G9a depress TrkC and can transform cells *in vitro* (Mulligan et al., 2008). REST may also have a role in regulating cell proliferation and apoptosis by regulating growth pathways like the PI3K signalling pathway. Loss of REST can lead to enhanced signalling and ultimately in transformation. For example, in lung and breast cancers, where there is REST loss of function, Akt phosphorylation and BCL-2 expression may be amplified several-fold (Lv et al., 2010). Another potential mechanism for tumorigenesis was recently identified in RESTless breast tumours where loss of REST was accompanied by an upregulation of LIN28A (Gunsalus et al., 2012). LIN28A, a direct REST target, is an RNA-binding protein, which plays a role in maintaining the self-renewal capacity of stem cells. Overexpression of LIN28A has been shown to drive tumour growth in xenografts in

nude mice, as well as metastasis in a mouse model of breast cancer (Gunsalus et al., 2012). Hence, aberrant expression of LIN28A in RESTless tumours may promote the growth of breast tumours. However, REST loss of function may be context-dependent, as other genetic and epigenetic alterations might also be needed for transformation. This is confirmed by a recent study in which conditional deletion of REST alone was insufficient for the development of colon tumours in mice (Hatano et al., 2011).

1.1.9.2.2 Oncogenic Role of REST

In contrast to the role of REST in epithelial tumours, REST has been shown to possess oncogenic properties specifically in neuronal tumours. This observation was first described in medulloblastoma, a rare paediatric brain tumour, where REST appeared to be aberrantly overexpressed, accompanied by a failure of cells to express REST target genes and to differentiate (Lawinger et al., 2000). Expression of REST-VP16, a recombinant form of REST which has repressor domains replaced by activator domains of a viral activator VP16, causes activation of REST target genes and halts tumorigenesis, demonstrating the importance of REST's repressor function in tumorigenesis (Fuller et al., 2005). Moreover, instead of promoting cell differentiation, REST-VP16 activated the apoptotic pathway, the mechanism of which remains unclear.

A mechanism for the oncogenic role of REST involves REST cooperation with the Myc oncogene to promote tumorigenesis. This was supported by the finding that mice with neuron-specific constitutive expression of REST do not form tumours (Su et al., 2006). However, REST overexpression was accompanied by Myc overexpression in many medulloblastoma samples. Thus, co-expression, and not expression alone, of REST and c-Myc in neuronal stem cells results in transformed cells which form

tumours, similar to human medulloblastoma tumours, in immunodeficient mice (Su et al., 2006). The abnormal expression of REST in these cells may prevent the cells from undergoing differentiation, forcing a stem cell or progenitor- like state, ultimately resulting in a transformed phenotype.

How is REST expression deregulated in these cancers? One possibility is the abnormal activity of various developmental pathways, which control REST gene expression. The Wnt pathway, for example, stimulates REST gene expression by stabilizing the β-catenin protein, which has been implicated in medulloblastoma pathogenesis (Nishihara, Tsuda and Ogura, 2003). In addition, mutations in the promoter or enhancer regions of REST can also cause constitutive REST expression. In the case where NSCs have managed to deregulate expression of both Myc and REST, rapid proliferation caused by Myc, and a block on neuronal differentiation due to presence of REST, causes the cells to proliferate abnormally, instigating medulloblastoma formation. An oncogenic role of REST has also been proposed for non-neuronal cells. Oncogenic REST, which fails to degrade in the G2/M transition of the cell cycle, affects continuous repression of Mad2, a spindle checkpoint gene, thus promoting genomic instability (Guardavaccaro et al., 2008a).

1.1.10 Conclusion

Hence, REST appears to be a master regulator of neuronal differentiation, with multiple mechanisms of upstream and downstream regulation. When the function of REST is compromised, such as loss of function, disease pathogenesis can result. In spite of these recent revelations, a few crucial questions remain to be answered. It remains unclear exactly which kinases act upstream of REST and set proteasomal degradation in motion and what is the relative importance of these various mechanisms in the control of REST. How do cells integrate these regulatory

mechanisms to maintain a steady state, and if there is any crosstalk between pathways? Furthermore, no structural studies of full-length REST have yet been performed. Finally, REST may be a possible target for therapeutic intervention against cancer and other neurological disorders, but whether this is achievable and by what means, remains to be defined.

1.2 Insulin Like Growth Factor 2 (IGF2)

IGF2 is a 7.5 kDa peptide hormone produced by many non-neuronal cell types, where it can act in an autocrine or paracrine manner. IGF2 is part of a complex growth-controlling system and possibly a differentiation regulator, acting potentially via REST. Other components of the growth-controlling system include insulin, another hormone IGF1, cell surface receptors and IGF binding proteins. There is considerable crosstalk between the IGF and insulin systems. IGF1 and IGF2 both signal through the type I IGF receptor (IGF1R). IGF2 can also signal through the Insulin receptor isoform A (IR-A), as well as through the classical Insulin receptor B (IR-B) and functional heterodimers of IGF1R and IR isoforms (Belfiore, 2007). This ability to function through differentially expressed receptors allows for tissue-specific effects of insulin and IGF2 and provides a broader range of functions than IGF1, which is limited to IGF1R binding (White, 2006).

In addition to the above, IGF2 can also bind to the IGF2R receptor with high affinity. IGF2R targets IGF2 for degradation (Scott and Firth, 2004). Like IGF2, IGF2R gene is also imprinted but unlike IGF2, it is expressed maternally (Reik et al., 2000). IGFs are found in blood and serum complexed with a family of proteins called IGF binding proteins (IGFBP). IGFBP regulate the bioavailability of IGFs. 75% of circulating IGFs exist in a ternary complex with IGFBP3 and acid-labile subunit (ALS), which cannot cross the endothelium, the rest exists as binary complexes with IGFBPs, which

can cross the endothelium, and only less than 1% of IGF is free and considered to be bioactive (Juul, 2003).

1.2.1 IGF2 Signalling

Most of the biological consequences of IGF2 signalling are mediated through IGF1R. IGF2 binds to IGF1R and activates it through autophosphorylation, leading to recruitment and subsequent phosphorylation of insulin receptor substrates (IRS) 1-4 (Vincent and Feldman, 2002). Phosphatidylinositol 3-kinase (PI3-K) binds to phosphorylated IRS, resulting in its activation (Zha and Lackner, 2010). Activated PI3-K activates AKT/Protein kinase B (PKB) through phosphorylation. AKT is a serine/threonine kinase that has numerous effects all mediated via phosphorylation of target proteins. This includes inhibition of apoptosis by inactivation of BAD (BCL-2 antagonist of cell death), activation of protein synthesis by phosphorylation of tuberous sclerosis complex (TSC1/2) and subsequent activation of mammalian target of rapamycin (mTOR) and ribosomal protein synthesis (See Figure 1.9). This constitutes the PI3K/AKT leg of IGF2 signalling and results in prevention of cell death (Kennedy et al., 1997).

The other parallel leg of the pathway is called the Ras/MAPK pathway which acts via recruiting different adapter proteins, SHC and growth factor receptor-bound protein 2 (Grb2), which form a complex with son-of-sevenless (Zha and Lackner, 2010). This complex activates Ras, a GTPase, which in turn leads to activation of Raf, triggering a kinase cascade, resulting in activation of mitogen-activated protein kinases (MAPKs), ERK1 and ERK2. ERK1/2 subsequently phosphorylate and activate several target genes including transcription factors, such as ELK1, triggering gene expression and cell proliferation (Zha and Lackner, 2010; Pouyssegur, Volmat and Lenormand, 2002).

Actions of the PI3K/AKT and Ras/MAPK pathways are opposed by phosphatases, notably the tumour suppressor phosphatase and tensin analogue (PTEN). PTEN dephosphorylates PI3K as well as SHC and mTOR, hence inhibiting both legs of the pathway (Soroceanu et al., 2007). PTEN expression has been shown to be increased by IGF2 signalling, forming a feedback loop, which serves to keep the signalling pathway in check (Moorehead et al., 2003; Church et al., 2012). Another feedback loop involving mTOR has also been described. mTOR targets include S6K1, a ribosomal kinase, which can inhibit IRS proteins by feedback signalling. This appears to be mediated through direct phosphorylation of IRS-1 on S302, which is suggested to interfere with IRS-1 ability to associate with the insulin receptor (Harrington et al., 2004). Inhibition of either mTOR or S6K1 triggers a negative feedback loop leading to an increase in AKT phosphorylation (Wan et al., 2007).

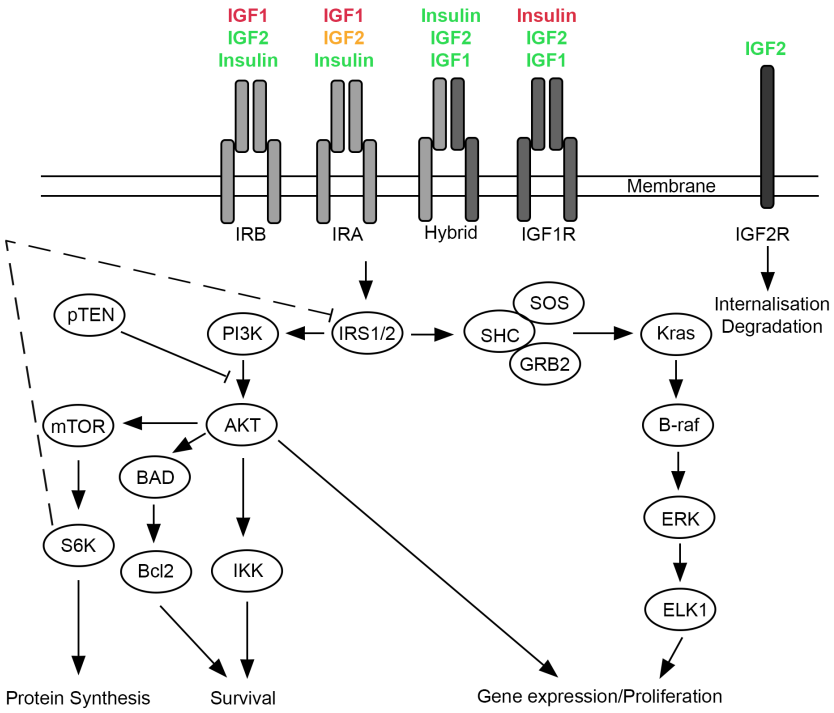


Figure 1.9 Overview of the Insulin/IGF system.
IGF1, IGF2 and Insulin mediate their cellular responses via IR-A, IR-B and IGF1R. IR-A and IR-B can exist as heterodimers and the affinity of the various receptors for their ligand is indicated by colour; green indicates strong affinity, orange indicates intermediate affinity and red indicates weak affinity. Binding of IGF2 to its cognate receptors leads to activation of two major signalling pathways, the PI3K/AKT and the Ras/Raf/MEK pathway, leading to protein synthesis, cellular proliferation and survival. IGF2 binding to IGF2R causes degradation of IGF2.
Adapted from (Chao and D'Amore, 2008).

1.2.2 IGF2 Function

IGF2 is expressed during embryogenesis and is important in promoting growth. Mice with disrupted paternal Igf2 alleles are 60% the size of their normal counterparts. The difference in cell number and proliferation occurs between embryonic days 9.5 to 11, where growth seems to be proportional to IGF2 bioavailability (Burns and Hassan, 2001). Conversely, overexpression of IGF2 results in fetal overgrowth and malformations resembling Beckwith-Wiedemann syndrome (Morison and Reeve, 1998).

IGF2 has a well-established role in muscle development and differentiation as well as in promoting proliferation and differentiation of bone cells (Chao and D'Amore, 2008). Inhibition of IGF2 results in a reduction in myoblast determination protein (MyoD) target gene expression, indicating that IGF2 potentiates MyoD effect for myocyte differentiation (Wilson and Rotwein, 2006). Reduction in bone mass due to cortisol use is also likely to be due to down-regulation of IGF2 (Minuto et al., 2005).

1.2.3 Regulation of IGF2

IGF2 supply is tightly regulated by multiple mechanisms. This multifactorial control allows fine-tuning of IGF2 response and prevents over or under-expression, which might lead to disease. IGF2 was the first gene found to be imprinted (DeChiara, Efstratiadis and Robertson, 1990). The imprinting mechanism is established in the germ line where an imprinting control region silences the maternal allele. This is achieved by the unmethylated differentially methylated domain (DMD) of the maternal allele recruiting the CCCTC-binding factor (CTCF), which restricts access of the IGF2 promoters to downstream enhancers (Thorvaldsen, Duran and Bartolomei, 1998; Ohlsson et al., 1994). IGF2 is transcribed from four tissue-specific promoters (P1-P4). P2-P4 are used during embryogenesis and cause monoallelic

expression whereas P1 is exclusively utilised in the liver during adulthood and results in biallelic IGF2 expression (Vu and Hoffman, 1994).

Murine IGF2 is widely expressed in the embryo, and its levels decline after birth (DeChiara, Robertson and Efstratiadis, 1991a). Human IGF2, however, is expressed during adulthood from the adult-specific P1 promoter in the liver (de Pagter-Holthuizen et al., 1987). Adult mice only show expression of *Igf2* transcripts in the choroid plexus and leptomeninges, which express IGF2 biallelically (DeChiara, Robertson and Efstratiadis, 1991a). Human IGF2 is also expressed biallelically in the choroid plexus, leptomeninges as well as the developing retina (Ohlsson et al., 1994).

In addition to imprinting, IGF2 supply is regulated by mRNA translation, pre-pro-peptide processing and IGF-binding proteins, which bind to IGF2 in the extra-cellular space reducing its bioavailability in a protease-dependent manner (Clemmons, 1997). IGF2 is internalized and degraded upon binding to the Mannose-6-phosphate/IGF2 receptor (Scott and Firth, 2004). The IGF2R gene is also imprinted but unlike IGF2, it is expressed from the maternal allele, the reciprocal imprinting is thought to regulate the abundance of both proteins (Reik et al., 2000). Changes in receptor expression also govern IGF2 effects. Increased IRA expression and increased IRA:IRB ratio is seen during embryogenesis and de-differentiation respectively, modulating IGF2 effects (White, 2006).

1.2.4 IGF2 in Neurogenesis

Whereas IGF2 role in cell proliferation, growth and survival has been well studied, only recently have studies started to shed light on its role in neurogenesis. As mentioned above, IGF2 levels drop postnatally when IGF signalling is dominated by IGF1 (Benarroch, 2012). In humans, apart from the liver, brain is the only organ that continues to express IGF2 (Benarroch, 2012). IGF2 is biallelically expressed in

human and mouse choroid plexus and leptomeninges (DeChiara, Robertson and Efstratiadis, 1991a). This is achieved via a regulatory mechanism distinct from other tissues that involves a centrally conserved domain enhancer region, located 5' of the DMD and unaffected by DMD control (Charalambous et al., 2004).

Adult neurogenesis can occur in either the subventricular zone (SVZ) or the subgranular zone (SGZ) located in the dentate gyrus of the hippocampus of the brain (Gage et al., 1998; Doetsch et al., 1999). SVZ neurogenesis is essential for generating new replacement neurons whereas that in SGZ is essential for learning and memory (Lois and Alvarez-Buylla, 1994; Bruel-Jungerman, Rampon and Laroche, 2007). Recent studies have revealed a role of IGF2 in both SVZ and SGZ by promoting expansion and self-renewal of NSCs (Bracko et al., 2012; Ziegler et al., 2012).

Differentiation assays, quantitative–PCR gene profiling and in vivo transplantation assays have been used to show that replacement of insulin with IGF2 in culture media has a profound effect on production of neurospheres from postnatal SVZ (Ziegler et al., 2012). Cerebrospinal (CSF) IGF2 can stimulate the growth of SVZ NSCs by binding to the primary cilium of NSCs that protrude into the lateral ventricles (Lehtinen et al., 2011; Bracko et al., 2012). The SGZ in the hippocampus are also responsive to IGF2. Microarray gene analysis has been used to show that *IGF2* gene is enriched in these cells and RNAi knockdown of Igf2 decreases proliferation of neural precursors (Bracko et al., 2012). The SVZ cells, however, do not produce IGF2 but use IGF2 produced in the choroid plexus. Hence, SVZ cells are unaffected by *IGF2* knockdown. Biallelic expression of IGF2 in the choroid plexus might serve to maintain the high levels of IGF2 in the CSF in order to supply SVZ, which are not adjacent to the CP, with IGF2. SGZ, on the other hand, are dependent on lower autocrine expression of IGF2, hence the monoallelic expression (Ferron et al., 2015).

Neurogenesis is thus only disturbed by deletion of the paternal IGF2 allele. Therefore, genomic imprinting results in efficient dosage control of IGF2 in order to regulate the different neurogenic niches of the adult brain. The implication of IGF2 deregulation can be seen in glioblastoma which displays elevated CSF-IGF2 levels, emphasizing the importance of IGF2 for CNS development and neurogenesis (Soroceanu et al., 2007).

1.2.5 IGF2 in Cancer

Deregulation of mammalian IGF2 has been shown to promote tumorigenesis. Loss of imprinting (LOI), a consequence of age-dependent dysregulation of DMD on the maternal chromosome is one of most frequent mechanism of IGF2 overexpression (Fu et al., 2004). Excessive IGF2 promotes growth and inhibits apoptosis in an autocrine manner (Gallagher and LeRoith, 2010). Experimental overexpression of IGF2 and IGF1R in murine tumour models leads to enhanced tumour growth and invasion (Hassan and Howell, 2000), while overexpression of IGF2 in primary mouse embryonic fibroblasts results in transformation and tumour formation of short latency (Hernandez et al., 2003). Altered IGF2 expression due to LOI occurs in paediatric cancers such as Wilms' tumours (Ravenel et al., 2001), rhabdomyosarcoma and common later onset tumours such as breast, lung, colorectal (where 11p15 is amplified) and hepatocellular carcinoma (McCann et al., 1996).

Loss of Imprinting (LOI), resulting in biallelic expression of IGF2 and increased ligand supply, is frequently the mechanism that causes overgrowth in Beckwith Weidemann syndrome (BWS). BWS is a rare congenital disorder arising from gene deletions resulting in LOI or uniparental disomy resulting in the inheritance of both paternal alleles. It is characterised by placental and postnatal overgrowth and a predisposition to tumours especially Wilms' tumour (Shapiro et al., 1982). An IGF2

over-expressing transgenic mouse model phenocopies the clinical features of BWS, confirming the link between IGF2 overexpression and BWS (Sun et al., 1997).

Increased IGF2 bioavailability due to mutations leading to LOH of IGF2R is another possible mechanism and has been reported in many cancers (De Souza et al., 1995; Hankins et al., 1996). Aberrant sonic hedgehog signalling in some tumours and decreased expression of the transcriptional repressor WT1 can also increase IGF2 expression (Ruiz i Altaba, Sanchez and Dahmane, 2002; Ward, 1997).

1.3 Aims and Rationale

The initial hypothesis that provided the impetus to study REST was the observation that IGF2 may regulate REST in mouse embryos. Previous work in the lab was undertaken to elucidate the key transcriptional hub genes that define the output of IGF2 signalling. Studying IGF2 signalling is complicated by autoregulatory signalling and feedback loops activated by exogenous ligands. Hence, two conditional models of IGF2 bioavailability were created. The first model consisted of mouse embryos (E9.5) that were either mono-allelic (wild-type), bi-allelic (H19-m/+p), or null (Igf2+m/-p) for expression of IGF2. Embryo harvesting was performed at E9.5, which is the time during development when growth is directly proportional to IGF2 ligand supply (Burns and Hassan, 2001). The second model consisted of a mouse fibroblast cell line, NIH3T3-ProIGF2, generated by retroviral transduction of NIH3T3 cells with a tetracycline-inducible form of the human IGF2 gene. Withdrawal of doxycycline causes accumulation of IGF2 in the culture media, hence modelling the *in vivo* paracrine action of this growth factor (Figure 1.10).

RNA extracted from both models was hybridised to Illumina mouse whole genome microarrays, and the resulting data were analysed in parallel, using Gene Set Enrichment Analysis (GSEA). Genes having a gene ontology annotation of 'transcriptional activity' were filtered from the leading-edge sets. Statistically over-represented transcription factor (TF) binding sites in the 1 kb of promoter sequence 5′ to the transcriptional start site of each gene were identified using PScan. This allowed transcriptional networks to be constructed from the leading-edge set TF genes and the putative TFs identified by promoter analysis. Common genes regulated by IGF2 in a time-dependent manner were identified from the embryo and NIH3T3-ProIGF2 networks. This process allowed a transcriptional network to be constructed with some

highly active 'hub' genes, with REST being one of the most prominent IGF2 dependent genes, in both embryos and fibroblasts. (Figure 1.11)

Induction of IGF2 did not affect REST mRNA, but REST stability was found to increase after 6 hours of incubation of NIH3T3 cells with increasing doses of IGF2 (Figure 1.11C) showing that regulation of REST is achieved post-translationally, most likely by a post-translational modification (PTM) on REST. This is consistent with the post-translational regulation of REST by ubiquitination and proteasomal degradation (Westbrook et al., 2008). The IGF2-dependent REST stability was abrogated upon treatment of cells with the PI3K inhibitor, LY294002, and rescued by treatment with MG-132, a proteasomal inhibitor, thus implicating the involvement of the IGF2 signalling pathway, via PI3K, in stabilising REST.

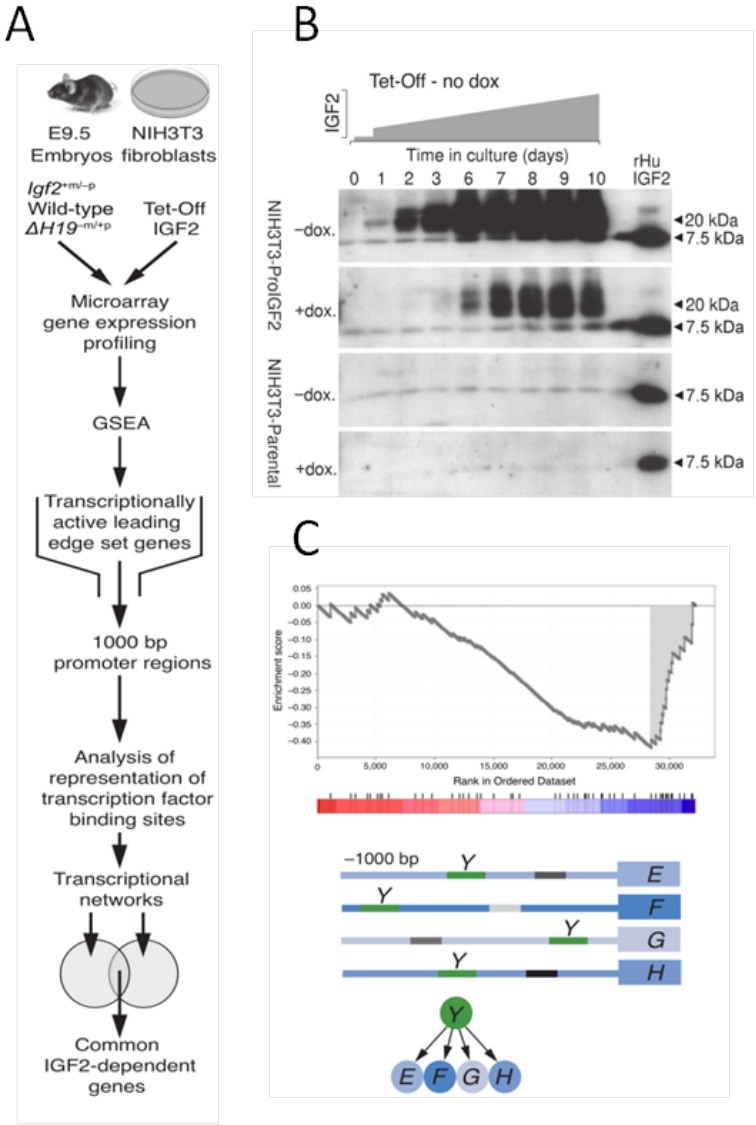


Figure 1.10 Unbiased detection of the transcriptional signature of IGF2 signalling in defined genetic models.

A) Microarray gene expression profiling was performed on two genetic models of IGF2 bioavailability: E9.5 mouse embryos monoallelic (wild-type), biallelic (ΔH19–m/+p) or null (Igf2+m/–p) for *Igf2* expression and NIH3T3 fibroblasts expressing a tetracycline-inducible form of the human *IGF2* gene. Data were analysed, in parallel, using Gene Set Enrichment Analysis (GSEA) and genes having a gene ontology annotation of 'transcriptional activity' were filtered from the leading-edge sets. Statistically over-represented transcription factor (TF) binding sites in the 1 kb of promoter sequence 5′ to the transcriptional start site of each gene were identified using PScan. Transcriptional networks were constructed from the leading edge set TF genes and the putative TFs identified by promoter analysis. Common genes regulated by IGF2 in a time-dependent manner were identified from the embryo and NIH3T3-ProIGF2 networks. B) A mouse fibroblast cell line, NIH3T3-ProIGF2 was generated by retroviral transduction of NIH3T3 with human *IGF2* cDNA under the control of a Tet-Off tetracycline responsive element (TRE). Withdrawal of doxycycline causes gradual accumulation of IGF2 in the culture media, modelling the in vivo paracrine action of this growth factor. C) Top - A typical GSEA profile for repressed genes. Middle - PScan was used to identify statistically over-represented TF binding sites, illustrated here as 'Y', in the proximal promoter regions (-1000 bp) of transcriptionally repressive genes, illustrated here as E, F, G and H. Bottom – Identification of Y as a transcriptional repressor of E, F, G and H allowed Y to be drawn as a 'node' connected by 'edges' to nodes E, F, G and H in a transcriptional network diagram. Data generated by Dr D J Barnes.

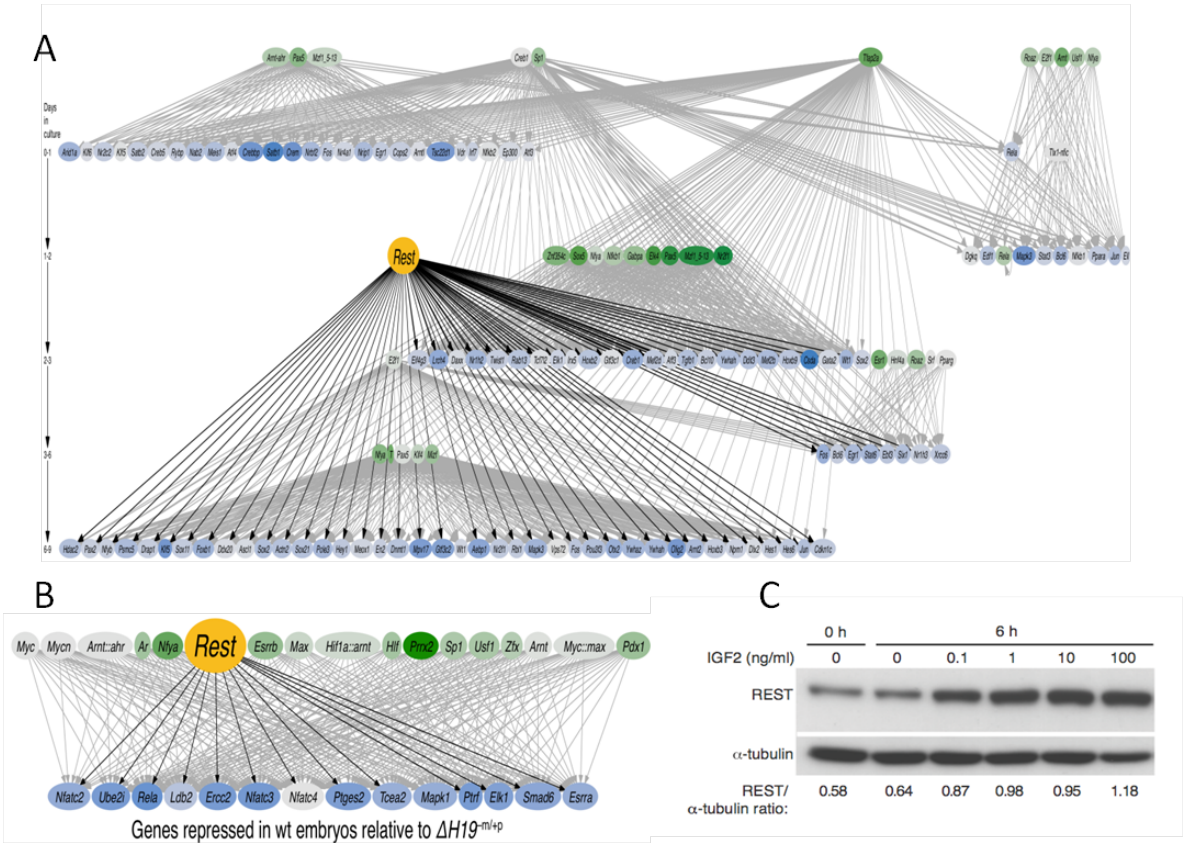


Figure 1.11 Identification of REST as a downstream target of IGF2 signalling.
(A) A transcriptional network constructed from gene expression data obtained from mouse embryos harvested on embryonic day 9.5 (E9.5), the time during development when growth and total cell number are determined by IGF2 bioavailability. This network consists of genes which are repressed in wild type embryos relative to H19-m/+p embryos (blue nodes) and the putative transcription factors responsible for this repression (green nodes). (B) A transcriptional network constructed from gene expression data obtained from NIH3T3-ProIGF2 cells. NIH3T3-ProIGF2 was used to study the time-dependent effects of IGF2 on gene expression. This network consists of genes, which undergo repression over a 10-day time course (blue nodes), and the putative transcription factors responsible for this repression (green nodes). RNA was extracted from E9.5 embryos and NIH3T3-ProIGF2 cells and hybridised to Illumina mouse whole-genome bead chip microarrays. Data was analysed using Gene Set Enrichment Analysis (GSEA) followed by a statistical analysis for over-represented transcription factor binding sites. Transcriptional networks were constructed from which common IGF2-regulated, and temporally specified genes were identified. Significantly, REST was identified as a key IGF2-dependent hub in networks constructed from both the NIH3T3-ProIGF2 data and the mouse embryo data. (C) Stabilisation of REST by IGF2 signalling. NIH3T3-ProIGF2 cells were serum-starved for 24 hours, washed and treated with increasing doses of IGF2 for 6 hours. Protein from cell lysates was immunoblotted with antibodies against REST and α-tubulin. Data generated by Dr D.J.Barnes.

Two phosophodegrons in the C-terminus of REST have been identified which are involved in binding the E3 ubiquitin ligase, β-TrCP, which targets REST for proteasomal degradation. Conversely, a binding site for a deubiquitinase lies in the fifth zinc finger of REST, which contributes to stabilising REST. Apart from this one site, no other site has yet been identified which contributes to stabilising REST, and neither have upstream players been defined in HAUSP control.

To test the hypothesis that IGF2 regulates neuronal differentiation through modification of REST, this thesis had the following aims:

1. To characterise the high molecular weight (MW) forms of REST (Chapter 3).
2. To evaluate the effects of IGF2 on phosphorylation, turnover and localisation of REST (Chapter 4).
3. To attempt to map phosphorylation sites of REST and identify the IGF2-dependent phosphosite if any (Chapter 5).
4. To evaluate *REST* gene knock out, both to define functional consequences and to provide a system for genetic complementation of mutant forms of REST (Chapter 6).

2 MATERIALS AND METHODS

2.1 Materials

2.1.1 Cell culture reagents

See Table 2.1 for a list of reagents used in this thesis.

2.1.2 Primary and Secondary Antibodies

See Table 2.2 and 2.3 for a list of primary and secondary antibodies used in this thesis.

2.1.3 Plasmids

See Table 2.4 for a list of plasmids used in this thesis.

2.2 Methods

2.2.1 General Cloning Methods

2.2.1.1 Primer design

Primers were designed using Geneious 5.1.7 software. The online tool NetPrimer was used to optimise primers for GC% content and secondary structures, and Thermo Fisher Scientific Tm Calculator tool was used to calculate the melting temperatures. BLASTN searches were performed to confirm the specificity of primers.

Item	Supplier	Catalogue Number
Dulbecco's modified eagles medium (DMEM)	Sigma	D6429
Eagle's Minimum Essential Medium	Sigma	M2279
RPMI-1640 Medium	Sigma	R8758
Foetal bovine serum (FBS)	Life Technologies	10500064
L-Glutamine 100X	Life Technologies	25030024
Non-Essential Amino Acids (NEAA)	Invitrogen	11140050
Puromycin	Sigma	P8833
Phosphate buffer saline (PBS)	Lonza	17-516F
TrypLE Express	Life Technologies	12604013
Penicillin/Streptomycin (P/S)	Life Technologies	15140122
Bovine serum albumin (BSA)	Sigma	A9306
Dimethyl sulfoxide (DMSO)	Sigma	67-68-5
IGF2	R&D	292-G2
EGF	R&D	291-G1
IGF1	R&D	236-EG
MG-132	Sigma	M7449
LY294002	Cell Signalling	9901
Halt Protease and Phosphatase inhibitors	Thermo Scientific	78440
Protease inhibitors	Roche	11697498001
2-mercaptoethanol	Sigma	M6250
ImmobilonP PVDF Membrane	Millipore	IPVH00010
Amersham ECL kit	GE Life Sciences	RPN 2232
Amersham Hyperfilm	GE Life Sciences	28-9068-35
Lumifilm	Roche	11666657001
CB-X assay	G-Biosciences	786-12X
Anti-Flag M2 Magnetic beads	Sigma	M8823
Prolong Gold Antifade	Invitrogen	P10144
Dual Glo Luciferase Assay	Promega	E2920
Urea	Sigma	U5378
PNGase F	NEB	P0704S
Lambda Phosphatase	NEB	P0753S
6-Diazo-50oxo-L-norleucine	Sigma	D2141
PUGNAc	Sigma	A722S9
Cycloheximide	Sigma	C7698
Poly-D-Lysine	Sigma	P6407
Iodoacetamide	Sigma	A3221
DAPI	Sigma	D9542

Table 2.1 List of Reagents

Primary Antibodies	Company	Cat. Number	Type	Dilution	
				IF	IB
REST (17-641)	Millipore	17-641	Polyclonal Rabbit	-	1:2000
REST (07-579)	Millipore	07-579	Polyclonal Rabbit	-	1:500
REST	Mandel	Gail Mandel	Polyclonal Rabbit	1:500	1:2000
REST (7D1.3)	Millipore	05-1477	Monoclonal Mouse	-	1:2000
REST (2436Y)	Abcam	Ab75785	Monoclonal Rabbit	-	1:1000
pAKT (Ser473)	Cell Signalling	4070	Monoclonal Rabbit	-	1:4000
α-tubulin	Cell Signalling	2144S	Polyclonal Rabbit	-	1:2000
Flag M2	Sigma	F1804	Monoclonal Mouse	1:300	1:2000
Lamin B	Santa Cruz	C20	Polyclonal Goat	-	1:500
AlexaFluor 647 Phalloidin	Invitrogen	A22287	-	1:50	
Phospho-GSK-3β	Cell Signalling	9336	Polyclonal Rabbit	-	1:2000
Sumo2/3	Cell Signalling	18H8	Monoclonal Rabbit		1:500
CD171 (L1CAM)-PE	Miltenyi Biotec	130-102-243	Monoclonal Mouse	1:10 (Cell surface staining	1:2000

Table 2.2 List of Primary Antibodies Used in Experiments.

IF=Immunofluorescence, IB: Immunoblotting, Cat. Number: Catalogue Number.

Secondary Antibodies	Company	Catalogue Number	Dilution
Goat anti-rabbit HPR conjugate	Dako	P0448	1:2000
Goat anti-mouse HPR conjugate	Dako	P0447	1:2000
Secondary goat anti-mouse Alexa Fluor 488	Invitrogen	A11001	1:500
Secondary goat anti-rabbit Alexa Fluor 647	Invitrogen	A21244	1:500
Secondary goat anti-rabbit Alexa Fluor 488	Invitrogen	A11008	1:500

Table 2.3 List of Secondary Antibodies used in Experiments

Plasmid Name	Description	Source
pEGFP-N1	Expression of GFP-fusion proteins	Gail Mandel
LPC-Flag-REST-WT	Flag-REST cloned into pLPC vector	Addgene
RE1 vector	pGL4.82- *Renilla* Luciferase Reporter Vector with PGK promoter and 12 RE1 sites	Anselme Perrier
RE1m Vector	pGL4.82- *Renilla* Luciferase Reporter Vector with PGK promoter and 12 mutated RE1 sites	Anselme Perrier
pGL4.10	*Firefly* Luciferase expressing reporter vector	Promega

Table 2.4 List of Plasmids used in Experiments

2.2.1.2 Polymerase Chain Reaction

Polymerase Chain Reactions were used to amplify DNA fragments using designed primers. PCR reactions were carried out in MJ Mini Personal Thermal Cycler (Bio-Rad) using Phusion Hot Start Polymerase (Thermo Scientific) according to manufacturer's instructions. A typical PCR reaction is shown in Table 2.5.

Reaction Component	Volume/µl
Template DNA (0.4ng/µl)	2.5
HF buffer (x10)	5
Forward Primer (10µM)	2.5
Reverse Primer (10µM)	2.5
dNTPS (10mM)	2
Phusion HotStart polymerase (2 units/µl)	0.5
DMSO	0.5
MilliQ H_2O	34.5

Table 2.5 A Typical PCR reaction.

2.2.1.3 Analysis of DNA using Agarose Gel-Electrophoresis

PCR reactions containing loading dye were run on 1-2% agarose gels, depending on amplified fragment size, at 80 mV. DNA was visualised using Gel Red Stain (Biotium). The size of the PCR product was determined by loading a DNA ladder (Hyper Ladder I or IV). Gels were analysed using Gene Genius Bioimaging System.

2.2.1.4 Purification of DNA

PCR reactions, which resulted in one band, were purified using DNA Clean & Concentrator™-5 kit (Zymo Research) according to manufacturer's instructions. PCR products that appeared as multiple bands were separated on agarose gels with Gel Red replaced by Gel Green Stain (Biotium). Bands of interest were excised using a Blue reader and purified using DNA & Gel Recovery Kit (Zymo Research). The purity and

concentration of DNA samples were analysed using a Nanodrop (ND-1000, Serial no. 780).

2.2.1.5 DNA ligation

DNA amount required for ligations was calculated using the ligation calculator at http://www.insilico.uni-duesseldorf.de/Lig_Input.html. In general, an initial insert: vector ratio of 3:1 was used, followed by 5:1 and 50 ng of the vector with the appropriate amount of insert was used for each ligation. Quick DNA Ligation kit (NEB) was used, and ligations performed according to manufacturer's instructions.

Briefly, insert and vector were made up to 10 µl with H2O and added to 10µl of 2x ligase buffer and 1µl of Quick DNA ligase. The reaction was incubated for 5 minutes at room temperature followed by incubation at 4^{o}C.

2.2.1.6 Transformation of E. coli

The ligation reactions were transformed into chemically competent DH5-α cells. Briefly, ligation or control reactions were added to 100 µl of competent cells and incubated at 4^{o}C for 30 minutes. Cells were then heat-shocked in a water bath for 45 seconds at 42^{o}C and then incubated at 4^{o}C for 2 minutes. Subsequently, 900 µl of super optimal broth with catabolite repression (SOC) media was added to cells and incubated at 37^{o}C on a shaker for 1 hour. 25-200 µl of the reaction was then plated on Agar plates supplemented with 100 µg/ml ampicillin or kanamycin, depending on the vector used. Plates were incubated at 37^{o}C overnight.

2.2.1.7 DNA mini-preparations

Single colonies from the transformation reactions were inoculated into 5 ml of lysogeny broth (LB) with 100 µg/ml ampicillin or kanamycin and cultured overnight on a shaker at 37^{o}C. QIAprep Spin Miniprep Kit (Qiagen) was used to isolate DNA from the resulting bacterial culture according to manufacturer's instructions.

2.2.1.8 DNA maxi-preparations

Single colonies from the transformation reaction were inoculated into 5 ml of LB with 100 μg/ml ampicillin or kanamycin for 8 hours. After 8 hours, this starter culture was diluted 1:500 into 250 μl LB supplemented with 100 μg/ml ampicillin or kanamycin and cultured overnight on a shaker at 37^{o}C. HighSpeed Plasmid Maxi kit (Qiagen) was used to isolate DNA according to manufacturer's instructions.

2.2.1.9 Sequencing

DNA sequencing was performed by Source Bioscience Life Sciences, Oxford. Samples were prepared according to instructions on the company website (www.lifesciences.sourcebioscience.com).

2.2.1.10 Cloning of REST Isoforms

pEGFP-N1 plasmid was provided by Gail Mandel. Full-length REST was cloned from LPC-Flag-REST plasmid into pEGFP-N1 plasmid using the following primers:

Forward_CP2_REST_Fw: TATTAAAGCTTACCATGGATTACAAGg

REST_CP2_REST_Rev: TGAATGGATCCTGTAATATACCTTACAAACTAA

To make Flag-REST1-GFP, the following primers were used to remove the STOP codon from the REST sequence.

Fw (REST_Fw_18.4.16): GGCTCTGGATCCATCGCCACCATGGTGAGCA

RC (REST_Rev_18.4.16): CTCCTGCCCTTGAGCTGCTTCTTC

Flag-REST2-GFP and Flag-REST3-GFP were cloned by deleting regions from Flag-REST1-GFP using the following primers:

Isoform 2

Fw (MMRESTiso2WOSTOP_Fw):

ATAAGTTCAGTTCTCTTTTCGGCTCTGGATCCATCGCCACC

GGCTCT

(MMRESTiso2_Rev):

GGACTAGAAAGCTTCTCTTACCTGTATGAGTTCTAACATGCT

Isoform 3

Fw (MMRESTiso3wostop_Fw): GTTGGGGGCTCTGGATCCATCGCCA

Rev complement (MMRESTiso3_Rev): TAGATCACAACCTGAATGAGTACGC

2.2.2 Cell Culture, Treatments and Reporter Assays

2.2.2.1 Cell Culture

MCF7, HEK293T, and NIH3T3 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) media supplemented with 10% (v/v) Fetal Calf Serum (FCS) and 1% (v/v) Penicillin/Streptomycin (P/S). Huh7, Hep3B and HT29 were cultured in Eagle's Minimum Essential Medium (EMEM) media supplemented with 10% (v/v) FCS, 1% (v/v) non-essential amino acids and 1% (v/v) Penicillin/Streptomycin (P/S). A673, CHP100, OSA, Hos, Hos143B were cultured in Roswell Park Memorial Institute 1640 (RPMI 1640) media supplemented with 10% (v/v) FCS and 1% (v/v) Penicillin/Streptomycin (P/S). Cells were grown in 175 cm^2 flasks and incubated at 37°C in a humidified atmosphere with 5% CO^2. Cell detachment was performed using TrypLE Express solution.

2.2.2.2 Transfections

Cells, at 60–70% confluence, were transiently transfected with plasmids using Lipofectamine 2000 transfection reagent for HEK293T cells, TransIT X2 reagent for SK-N-MC cells and TransIT 2020 reagent for MCF7 cells using manufacturer's instructions.

2.2.2.3 Growth Factor Experiments

For growth factor experiment, cells were grown on 6 well plates with $4x10^5$ cells per well. After 24 hours on adaptation, cells were washed three times with 2 ml of PBS

per well and serum starved by replacing media with FCS-free media for 24 hours. Cells were then replaced with fresh FCS free media after washing once with 2ml of PBS and treated with IGF2 doses (ranging from 0.1 to 100 ng/ml), MG-132 (0.75 μM) and LY294002 (50 μM) for 6 hours at 37^{o}C. Cells were subsequently washed with PBS and lysed in RIPA buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 0.5% (v/v) Sodium deoxycholate, 0.1% (v/v) SDS, 1% (v/v) Triton X-100) supplemented with protease and phosphatase inhibitors.

For experiments in HEK293T cells, which required attachment to plates prior to washing, poly-D-lysine-coated plates were used. Coating was performed by adding 300 μl of 0.1 mg/ml poly-D-lysine (PDL) per well, incubation of plates at 37^{o}C for 2 hours, followed by washing with 2 ml of sterile H_2O three times.

2.2.2.4 Treatment with O-Glycosylation Inhibitors

For O-glycosylation inhibition experiment, HEK293T cells, grown in complete media with FCS, were treated with DON and PUGNAc at a final concentration of 0-100 μM for 24 hours, followed by preparation of whole cell lysates.

2.2.2.5 Treatment with Urea

HEK293T ($1x10^7$) cells were transfected with Flag-REST and whole cell lysates were prepared 48 hours post-transfection. Lysates were incubated with anti-Flag M2 magnetic beads overnight with agitation at 4^{o}C. Following this, beads were washed three times with TBS, divided into four tubes and treated with increasing concentrations of urea (0-8 M) in TBS for 15 minutes. Subsequently, Flag-REST was eluted by addition of Laemmli buffer and boiling for ten minutes, followed by isolation of purified protein using a magnetic separator. The eluate was immunoblotted for REST using 07-579 antibody.

2.2.2.6 Reduction of Whole Cell Lysate

HEK293T whole cell lysate (in RIPA buffer) was treated with *tris*(2-carboxyethyl) phosphine.hydrochloride (TCEP.HCl) at a final concentration of 5 mM for 20 minutes at room temperature. This was followed by addition of iodoacetamide (10 mM final concentration) and incubation for 15 minutes at room temperature in the dark. Subsequently, treated cell lysates were boiled for ten minutes after addition of Laemmli buffer and immunoblotted for REST using 07-579 antibody.

2.2.2.7 Treatment with N-glycosidase F

Whole cell lysate was prepared from HEK293T cells expressing Flag-REST and treated with PNGase F enzyme for 1 hour at 37^{o}C according to manufacturer's protocol. Purified Fetuin was treated in parallel as a control (Green et al., 1988).

2.2.2.8 Treatment with Cycloheximide

HEK293T cells cultured in complete media with FCS were treated with cycloheximide alone or in the presence of IGF2 (100 ng/ml) or MG-132 (0.75 μM) for the indicated time points, followed by preparation of whole cell lysate.

2.2.2.9 Fluorescence Activated Cell Staining

L1CAM (L1 Cell Adhesion Molecule) was analysed using Fluorescence activated cell sorting. HEK293T or SK-N-MC cells were detached from 6 well plates using TrypLE, washed twice in fluorescence-activated cell sorting (FACS) buffer (PBS, 2% (v/v) FBS) and re-suspended in L1Cam conjugated to Phycoerythrin (L1Cam-PE) in FACS buffer at 1:5 ratio (10^6 cells in 50 μl total volume) for 15 minutes at 4^{o}C. Following incubation, cells were washed twice with FACS buffer and re-suspended in 1 ml of FACS buffer. Analysis was performed using Cyan flow cytometer analyser (BD Biosciences) and FlowJo Software. Stained cell profiles were compared with

unstained cells or alternatively, stained profiles of wild type cells were compared with stained profiles REST heterozygous cells.

2.2.2.10 Cell Viability Assay

CellTiter 96® Aqueous One Solution Cell Proliferation Assay (Promega) was used to analyse cell viability and proliferation according to manufacturer's instructions. 16000 cells per well were plated in 96 well plates and treated with REST inhibitor X5050 or puromycin respectively depending on the experiment. After 24 or 48 hours, 20 µl of MTS reagent was added per well, incubated for 2 hours at 37^{o}C and the absorbance read at 490nm on a plate reader (SpectraMax).

MTS Tetrazolium compound is bioreduced by cells into a formazan product, which is coloured and is soluble in culture media. Formazan concentration is directly proportional to the number of live cells and is used to estimate cell viability (Cory et al., 1991).

2.2.2.11 Luciferase Reporter Assay

Dual-Glo luciferase system (Promega) was used and assay performed according to manufacturer's protocol. Briefly, 16x10^{3} HEK293T or SK-N-MC cells were plated in a 96 well plate overnight and subsequently transfected with RE1m or RE1 vectors expressing Renilla luciferase with pGL4.10 (Firefly Luciferase expressing vector) in 50:1 ratios using Lipofectamine 2000 or TransIT X2 respectively. Firefly and Renilla luminescence was then measured using a plate reader (SpectraMax) at 48 hours after transfection. RE1 and RE1m signal was normalised to Firefly Luciferase luminescence.

2.2.3 Protein Analysis

2.2.3.1 Cell Lysis

Whole cell extracts were prepared by lysing cells in RIPA buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 0.5% Sodium deoxycholate, 0.1% (v/v) SDS, 1% (v/v) Triton X-100) supplemented with Halt protease and phosphatase inhibitors (Thermo Scientific). The lysate was cleared by centrifugation at 13000 rpm for 20 minutes, followed by isolation of the supernatant and storage at -80^{o}C.

2.2.3.2 Nuclear and Cytoplasmic Fractionation

Cytoplasmic fraction was isolated by lysing cells in cold NP-40 buffer (0.5% (v/v) NP-40, 25mM Tris ph 7.9, 100mM NaCl). After a quick centrifugation step (10 sec spin in an Eppendorf table top microfuge), the supernatant was isolated as the cytoplasmic fraction and the pellet was washed once in cold NP-40 buffer. The NP-40-insoluble pellet was then lysed by incubating it with RIPA buffer with 250 units/ml of Benzonase enzyme (to reduce viscosity of lysate) for 30 minutes at 4^{o}C with agitation. Following centrifugation, the supernatant was isolated as the nuclear fraction.

2.2.3.3 SDS-PAGE Electrophoresis and Immunoblotting

Protein concentration of each sample was determined by a CBX protein assay by G-Biosciences according to manufacturer's instructions. Lysate was mixed with 6x Laemmli sample buffer and boiled for 10 minutes. Equal amounts of total protein were resolved by SDS-PAGE gel electrophoresis. Protein was then transferred onto a PVDF membrane and the membrane was blocked with tris-buffered saline-0.1% (v/v) tween (TBS-T) with 5% (v/v) milk for 1 hour at room temperature with agitation. Membrane was then incubated with primary antibody overnight at 4^{o}C with agitation. After overnight incubation, membrane was washed 3 times with 0.1% TBST for 5

minutes each at room temperature followed by incubation with secondary antibody conjugated with HRP for 1 hour at room temperature with agitation. The antibody dilutions are shown in Table 2.2.

The membrane was once again washed 3 times with TBS-T for 5 minutes each at room temperature. Immunodetection was performed using Amersham ECL detection reagent and Amersham Hyperfilm according to manufacturer's instructions.

2.2.4 Lambda phosphatase treatment

HEK293T whole cell lysate was treated with 8 units of lambda phosphatase per µg of lysate or buffer alone for 1 hour at 30^{o}C. Treated lysates were immunoblotted for REST using anti-REST 07-579 and Mandel antibodies.

2.2.5 Immunoprecipitation

HEK293T cells were grown in 15 cm Corning plates and transfected with Flag-tagged REST using Lipofectamine 2000. Cells were lysed in 1 ml of RIPA buffer supplemented with protease and phosphatase inhibitors and 250 units/ml benzonase. Lysates were cleared by centrifugation at 13000 rpm for 20 minutes. Immunoprecipitation was performed using Anti-Flag M2 Magnetic beads (Sigma) according to manufacturer's protocol. Briefly, 1 ml of cleared lysate was incubated with 40 µl of M2 magnetic beads, equilibrated in TBS, overnight at 4^{o}C on a shaker. Flag-REST-bound beads were separated using a magnetic separator and washed three times with cold TBS for 3 minutes each to remove non-specifically bound proteins. Flag-REST was eluted by incubating beads with 20 µl of 2x Laemli buffer and boiled for 5 minutes. The supernatant was separated by a magnetic separator and stored at -20^{o}C.

The resulting eluate was analysed by resolving it on parallel lanes of an SDS-PAGE gel, which was then divided with one section being stained with Coomassie blue and the other section transferred to PVDF membrane and probed with an anti-REST antibody.

2.2.6 Immunofluorescence

HEK293T cells, untransfected or transiently transfected with Flag-REST, were plated on Poly-D-Lysine coated coverslips in 12 well plates. After 24 hours, cells were washed twice with 500 µl of cold TBS, fixed using PBS with 4% (v/v) paraformaldehyde (PFA) for 10 minutes at room temperature, followed by treatment with 100 mM glycine in PBS for 10 minutes. After two washes with 500 µl of TBS, cells were blocked with TBS/0.5% (v/v) Tween/10% (v/v) goat serum for 1 hour and incubated with the primary antibody at the described concentrations diluted in TBS/0.3% (v/v) Tween and 3% (v/v) goat serum for 2 hours or 4°C overnight. After two washes with TBS, cells were incubated with secondary goat antibody at the desired dilution for 1 hour in the dark, followed by two 3 minute washing steps with TBS. To probe for the actin cytoskeleton, cells were incubated with Phalloidin Alexa 488 (1:50) for 2 hours in the dark followed by two 3-minute washing steps. Nuclei were detected using the DNA stain 4',6-diamidino-2-phenylindole (DAPI) (1 µg/ml). Coverslips were mounted on slides using Prolong Gold Antifade reagent.

Image acquisition was performed using inverted Olympus Confocal microscope and Fluoview F1000 software. Images were taken using 60x oil objective with DAPI, alexa488 and alexa555nm laser. Images were taken using 1x binning and 4µs/pixels scan speed. Each image had a resolution of 2048 pixels and an aspect ratio of 1:1.

2.2.7 Imaging of Native Gels

HEK293T cells were transfected with Flag-REST-GFP constructs, serum-starved and treated with IGF2 (0.1-100 ng/ml) for 6 hours. Whole cell lysates were prepared and resolved on a 15% native acrylamide gel. In-gel GFP fluorescence was detected using the Biorad imaging system.

2.2.8 Mass Spectrometry

2.2.8.1 Preparation of Sample for MS

To obtain purified REST for MS analysis, HEK293T cells transiently expressing Flag-REST were grown in 5x150-mm plates and treated with 0.75 μM MG-132 for 6 hours before lysis. Cell lysate was prepared using RIPA buffer supplemented with Protease and Phosphatase inhibitors as described in 3.2.3.1. Flag-REST was immunoprecipitated from the lysate using Flag M2 magnetic beads as described in 3.2.4. Bound complexes were eluted with Laemmli sample buffer and eluted protein was resolved on NuPage 4-20% gradient SDS-Page gels. Gels were stained with Coomassie blue and the appropriate band was excised. Alternatively, bound complexes were eluted in an SDS-containing buffer.

For experiments in Section 5.4.2.2.1, lysate from 3x150mm HEK293T plates was used for immunoprecipitation of Flag-REST for gel-isolation or eluate. For experiments in Section 5.4.2.2.2 and Section 5.4.2.2.3, lysate from 5x150mm or 3x150mm HEK293T plates was used for immunoprecipitation of Flag-REST, respectively. For experiments in Section 5.4.3, lysate from 6x150mm plates was used for either IGF2 or IGF2+MG-132 samples.

2.2.8.2 Gel Digest

Flag-tagged REST was excised from Coomassie-stained gels. Gel pieces were cut into 1 mm pieces and destained with 150 μl 50% acetonitrile (ACN) and 50 mM

tetraethylammonium bicarbonate (TEAB) at 37°C. Protein was reduced using 10 mM tris(2-carboxyethyl)phosphine (TCEP) for 30 minutes at room temperature, followed by removal of the buffer and addition of 200 µl 100% ACN until gel pieces were dried out. Dried gel pieces were subsequently incubated with 50 mM 2-Chloroacetamide/ 100 mM TEAB in the dark for 30 minute at room temperature. Gel pieces were washed with 100% ACN once and incubated with 100% ACN until dried out. The proteins were then digested overnight with 0.5 µg of Trypsin in 50 mM TEAB at 37°C. Following Trypsin digest, the sample was divided into two; one-half was treated with 0.5 µg of GluC enzyme overnight at 37°C for the double digest. Peptides were extracted from the gel by rinsing gel with 10% formic acid (FA) and the supernatant transferred to a fresh tube.

2.2.8.3 Solution Digest

Solution digest was performed on the eluate resulting from Flag immunoprecipitation. Protein was denatured using 8 M Urea/100 mM TEAB for 30 minutes at room temperature. This was followed by reduction using 10 mM TCEP for 30 minutes at room temperature and subsequent alkylation with 50 mM 2-Chloroacetamide for 30 minutes. The proteins were digested overnight with 0.5 µg Trypsin, 1 M Urea and 50 mM TEAB at 37°C. Following digestion, FA was added to the solution up to 0.5%.

2.2.8.4 Phosphopeptide Enrichment

Phosphopeptide enrichment was performed as described in(Zhou et al., 2013). Briefly, a slurry of 10 mg/ml TiO2 beads (Sachtopore, Sachtleben Chemie, Germany) in 30% ACN/ 0.1% trifluoroacetic acid (TFA) was prepared and 40 µL of the slurry was packed into a GELoader column using a C8 plug by centrifugation (100g). The columns were conditioned with 50 µl of loading buffer (80% ACN, 6% TFA) by centrifugation (150g). The peptide sample that was digested in solution was

resuspended in loading buffer and loaded onto the column by centrifugation (100g). Washing of columns was performed by centrifugation (150g) using 50 µL of 50% CAN/ 0.5% TFA/200 mM NaCl followed by a second wash using 50 µL of 50% CAN/0.1% TFA. Elution of phosphopeptides was performed into 10% FA using 20 µl of 10% ammonia by centrifugation (100g) for 20 minutes, followed by a second elution using 2 µl of 80% ACN/2% FA. 3 µl of 100% FA was added to further acidify the samples.

2.2.8.5 LC-MS/MS

All MS experiments were performed on the Ultimate 3000 UHPLC system (Thermo Fischer Scientific). This was connected to a QExactive mass spectrometer (Thermo Fischer Scientific) equipped with an EASY-Spray nanoelectrospray ion source (Thermo Fischer Scientific). These experiments were performed by the proteomics facility at the Sir William Dunn School of Pathology.

The peptides were trapped on a C18 PepMap100 pre-column (300 µm inner diameter x 5 mm, 100Å, Thermo Fisher Scientific) using solvent A (0.1% FA in water) at a pressure of 500 bars. Peptide sample was separated on an in-house packed analytical column (75 µm inner diameter) packed with C18 beads (ReproSil-Pur 120 C18-AQ, 1.9 µm, 120 Å, Dr.Maisch GmbH) with a linear 0.5-hour gradient from 7% to 28% solvent B (0.1% formic acid in acetonitrile) and a flow rate of 200 nL/min.

The mass spectrometer was operated in the data-dependent mode to acquire raw data. Full scan MS spectra were acquired in the Orbitrap (scan range 350-2000 m/z, resolution 70000, AGC target 3e6, maximum injection time 50 ms). The ten most intense peptide ion peaks were selected and fragmented by higher-energy collisional dissociation (HCD) fragmentation using normalised collision energy of 30%. HCD

fragment ion spectra were recorded in the Orbitrap with a resolution 17500 (AGC target 5e4, maximum injection time 120 ms) with first fixed mass at 180 m/z.

2.2.8.6 Identification of Peptides and Proteins

Data analysis was performed using the Mascot software. Identification of proteins groups was carried out by searching peak lists against a human database (Swiss-Prot, version 04/13). Searches were also performed against a concatenated target/decoy human IPI database with an empirical false discovery rate (FDR) of 1%. Cysteine carbamidomethylation (C) was searched as a fixed modification, while Acetylation (Protein N-term), Oxidation (M) and Phosphorylation (S, T and Y) were searched as variable modifications. A maximum of two missed cleavages for Trypsin or GluC respectively were chosen. Peptide ion tolerance was set to 20 ppm and fragment mass tolerance was 0.02 Da. Peptides with an ion score < 20 were filtered out ($p<0.05$). Remaining peptides were then manually analysed to validate the identification of phosphorylation sites. For all phosphorylated peptides, unphosphorylated peptides were also identified.

2.2.9 CRISPR/Cas9

2.2.9.1 Construction of CRISPR/Cas9 Target Vector

Plasmids for CRSIPR/Cas9 editing were prepared by the Genome Editing facility at the Dunn School of Pathology. Briefly, guide RNAs were designed using the CRISPR guide tool at http://crispr.mit.edu and cloned into the pSpCas9n(BB)-2A-GFP (PX461) plasmid using restriction sites on either end. Guide RNA sequences are given below:

sgRNA 1A TGCAAGTCATACATGTCGTTAGG

sgRNA 1B CGCACCTCAGCTTATTATGCTGG

sgRNA 2A AGACTGCCCCATTACCTGGGTGG

sgRNA 2B GGAGGGCTGTTTACCAGCAGTGG

2.2.9.2 Cell Culture and Transfection

HEK293T, SK-N-MC, and MCF7 were cultured in recommended media as described in Section 2.2.2.1. HEK293T cells were transfected using Lipofectamine 2000 reagent, MCF7 using TransIT 2020 and SK-N-MC using TransIT X2 reagent according to manufacturer's protocols. Transfections were carried out in T175 flasks using 19 μg of each Cas9/sgRNA plasmid. Cell sorting for GFP-positive cells was carried out 24 hours after transfection using Beckman Coulter MoFlo XDP. GFP-positive cells were grown in recommended media supplemented with 20% FCS.

2.2.9.3 Genomic DNA Isolation and HRMA Analysis

Total genomic DNA was extracted from cells using the squish buffer method. Briefly, 100 μl of squishing buffer (10 mM Tris-HCl pH 8, 1 mM EDTA, 25 mM NaCl, 200 μg/ml Proteinase K) was added to a 24 well dish and incubated at 65°C for 30 min followed by inactivation of proteinase K at 95°C for 2 min. Extracted DNA was stored at 4°C. 1-2.5 μl of genomic DNA was used as a template for a 10 μl PCR reaction.

Genomic DNA was amplified using the Hotshot Diamond PCR mix (Clent Lifescience) with LCGreen Plus dye (BioFire Diagnostics) in RotorGeneQ followed by HRMA using a range of 70-98°C. Melt curves were analysed using the RotorGeneQ software and fluorescence was normalised before and after melting. Appropriate controls were used each time including mock-transfected cells and no template control. PCR reaction, primers used and PCR program are shown below.

Primer Name	Primer Sequence
MM_RES_HRMF	TGGTGCTCTTGTTTTGTTTTG
MM_RES_HRMR	TCACAGCAGCTGCCATTTAC
MM_RES_HRMF2	CGATGTGGTTTTAAGCCAGT
MM_RES_HRMR2	CTGTCTTTCTTCACCGACCAG

Table 2.6 Primers used for HRMA Analysis.

Reagent	Volume/ µl
2x Hotshot Diamond PCR mix	5
10 uM forward primer (200 nM final)	0.2
10 uM reverse primer (200 nM final)	0.2
A0x LC Green	1
ddH_2O	2.6
DNA prep	1
	10 (Total Volume)

Table 2.7 Reaction volumes for HMR PCR.

Temperature	Time
95°C	5 min
95°C	20 s
55-65°C	30 s
72°C	30 s (x45 cycles)
95°C	30 s
25°C	30 s
10°C	Store

Table 2.8 PCR program used for HRM analysis.

2.2.9.4 Target Sequencing and Sequence Alignment

PCR was performed in a 50 μl reaction using the conditions and program mentioned above, and run on an agarose gel. Resulting bands were gel-extracted using the Zymospin gel recovery kit according to manufacturer's protocol. Sequencing was performed using HRMF/R or HRMF2/R2 primers by Source Biosciences sequencing facility at the Department of Biochemistry, Oxford University and sequences aligned using the Geneious software.

2.2.9.5 Generation of Selection Strategy.

Killer and mutKiller plasmids were designed based on RE1 and RE1m plasmids obtained from Anselme Perrier (Charbord et al., 2013). These plasmids have 12 repeats of either RE1 sites (RE1 plasmid) or mutant RE1 sites (RE1m plasmid), cloned 5' of the PGK promoter and *hRluc* gene. A region of the plasmid containing SV40 promoter and hRluc was deleted using the following primers:

pgl4.82-crispr-FwMM primer: CACCATGACCGAGTACAAGCCT

pgl4.82-crispr-RevMM primer: GCTTTACCAACAGTACCGGATTGC

2.2.10 Statistical Analysis

Data were analysed using GraphPad Prism version 6.0. Statistical analysis was performed using paired t-test (groups of two) or the one-way ANOVA test (groups>2). $P \leq 0.05$ was considered to be significant.

3 CHARACTERISATION OF THE LARGE MOLECULAR WEIGHT FORMS OF REST

3.1 Introduction

REST is a transcription factor that has been implicated in the regulation of neurogenesis and more recently in embryogenesis. REST is a 1097 amino acid protein, with eight zinc fingers constituting the DNA binding domain, two repressor domains on either terminus and less conserved lysine and proline rich regions in the central region (Figure 1.1). The repressor domains bind a variety of co-repressors and associated complexes, which act in concert to repress gene expression of target genes containing specific 23-bp RE1 sequences. Given the central function of REST and the importance of its regulation, it is perhaps not surprising that protein levels appear under strict control and variations in these levels are crucial for its role in differentiation.

REST was discovered in 1995 initially as a 116 kDa protein (Chong et al., 1995; Schoenherr and Anderson, 1995). It is encoded by a large multiply spliced transcript

containing eight exons spanning a genomic contig of 70793 base pairs in man (Palm, Metsis and Timmusk, 1999). The two repressor domains located at the N and C terminus, recruit the two major co-repressors, mSin3 (Grimes et al., 2000) at the N-terminus and CoREST at the C-terminus, respectively (Andres et al., 1999). mSin3 further engages proteins such as the methyl-DNA- binding protein MeCP2 and histone deacetylases (HDACs) whereas Co-REST interacts with histone H3 K9 methyltransferases, HDACs, SWI/SNF chromatin remodelling complex via brg1, among others (Ooi et al., 2006; You et al., 2001; Huang, Myers and Dingledine, 1999; Ballas et al., 2001). REST also recruits numerous other regulatory cofactors. These cofactors employ various means to turn off target gene expression including histone deacetylation, methylation, chromatin remodelling and inactivation of RNA Pol II (Yeo et al., 2005). Hence, REST appears as an epigenetic regulator of gene expression, in a significant complex with other proteins.

There are discrepancies in the literature about the apparent molecular weight of REST as detected by immunoblotting. Some studies report REST to be 116 kDa whereas others describe a higher molecular weight (MW) from 180 to 220 kDa and some report both species. RNAi knockdown was shown to decrease levels of both lower molecular weight (LMW) and higher molecular weight (HMW) REST, showing that it is encoded by the same gene as LMW REST (Faronato et al., 2013).

Although recognised in the literature, efforts to characterise the HMW REST have not been exhaustive (see Table 4.1). Most studies fail to mention the size of the immune-reactive REST band being considered hence making it difficult to interpret and reproduce data.

A number of PTM may account for MW differences. O-linked glycosylation has been proposed to be a plausible explanation for the increased MW of REST. Inhibition of

O-linked N-acetylglucosamine (GlcNAc) transferase (OGT), the enzyme that is responsible for O-glycosylation, by siRNA knockdown or pharmacological intervention, has shown depleted levels of the 220kDa form of REST (Faronato et al., 2013).

Further evidence was provided by a study on a C-terminally truncated splice variant of REST, REST4 (Lee et al., 2000). REST 4 was found to have an apparent molecular weight of 53kDa as opposed to its predicted molecular weight of 37 kDa. This difference of 16 kDa is considerably less than that of about 100 kDa between that of calculated and observed molecular weight of full-length REST, indicating the presence of two regions of glycosylation, one in the N-terminus and another more heavily glycosylated region in the C-terminus. Furthermore, REST 4 and its deletion mutant REST4-N-Δ86 were retained on a wheat germ agglutinin-agarose column whereas deletion mutants of REST4, REST4-N-Δ152, REST4-N-Δ176, and REST4-N-Δ209 were not. These findings indicate the presence of *N*-acetylglucosamine or sialic acid on REST4 and narrows down the glycosylation site in the N-terminus to amino acid residues 87-152. Treatment with the enzyme N-Glycosidase F (PNGase), which cleaves N-linked glycans, did not alter the molecular weight of these proteins, hence suggesting O-linked glycosylation to be the predominant form of glycosylation.

The apparent molecular weight of REST appears almost twice that of the calculated weight leads to further interesting hypothesis to account for the HMW REST form. In a PhD thesis publication, Myc-tagged REST was immune-precipitated from a mixture of Flag and Myc-tagged REST using an anti-Flag antibody and vice versa, suggesting dimer formation (Vu, 2009). This interaction was proposed to be independent of zinc fingers as mutations in zinc fingers did not abrogate dimer formation. It has also been shown that REST4, a splice variant of REST, directly binds to REST. Shimojo et al.

used electrophoretic mobility shift assays to show that REST4 binds to REST and acts as a dominant-negative by interfering with REST binding to RE1 sequence of the cholinergic gene locus. However, REST4 itself cannot bind to the RE1 sequence. The molecular mechanism of this interaction remains unknown. These findings also point towards a binding mechanism that might be potentially employed by full-length REST molecules for dimerization (Shimojo et al., 1999). Despite these data, it is not yet clear if dimerisation also occurs *in vivo* and whether it is the direct mechanism that underlies the HMW REST band. The lack of either a crystal or solution structure of REST precludes further molecular assessment of the hypothesis.

Another clue for this puzzle might be found in the localisation of the two MW forms of REST. Cell fractionation studies suggest that the 220kDa form is the predominant MW of REST in the nucleus (Faronato et al., 2013; Kreisler et al., 2010). Hence, HMW REST might represent the functionally active pool of REST inside the nucleus, and the LMW form used as a transitory pool of REST in the cytoplasm, that potentially can be activated and directed towards the nucleus upon demand or stimulus.

Evaluating the effect of IGF2 on REST was complicated by the presence of multiple species of REST, additional to the 120kDa form (Appendix 4 and 5). The higher MW isoforms of REST appeared to replicate the behaviour of the smaller MW isoform and are either stabilised or destabilised by IGF2, respectively. Examining the 116 kDa band alone was found to be insufficient as other species forms of REST might also be functional. Hence, the initial impetus for this Chapter was to interpret my results with confidence and in the light of previous experiments (Appendix 1-6). This was especially important with regards to my data obtained with different antibodies. Hence, I attempted to deconvolute the identity of the HMW form. I first took a

systematic approach to try and rule out potential mechanisms. Data from previous studies was first combined to summarise the existing evidence about the large MW forms (Table. 3.1), followed by targeted experiments to assay for different PTMs, the scheme of which is shown in Figure 3.1.

Author	Cell lines	REST size/kDa	Antibody	Evidence/Info
Chong 1995, Cell (Mandel)	Rat Skeletal Muscle (Rat)	116	P73 Mandel (1-485 aa)	Minor diffused reactive band at a more slowly migrating position, which may be indicative of posttranslational modification.
Thiel 1998, JBC	COS	200	REST anti-sera	No explanation given
Shimojo et al. 1999, MCB	HEK293T	200	12C11-1	REST4 forms a hetero-oligomer with REST
Lee 2000, MBR	HEK293T	200	-	REST4 and presumably REST is O-glycosylated between 87 and 152. REST4 can form homo-oligomers, monomers bind to DNA
Grimes et al. 2000, JBC (Mandel)	HEK293T	200	P73 Mandel (1-485 aa)	-
Wood 2003, JMB	PC12, JTC19	120	R2174	-
Pance 2006, JNC	A35C	120, 200	07-579	-
Bruce et al. 2006, JNC	PC12, JTC19 (Rat lung fibroblast)	180	07-579	-
Guardavaccaro et al. 2008, Nature	Hela, HCT116	-	07-579, Mandel	Asynchronous cells – two bands Prometaphase – lower band only Slower migrating (upper) is phosphorylated and increased by inhibition of proteasome
Westbrook et al. 2009, Cell	SW620, SW1417 (colon cancer cell lines)	200	αRest-N, αRest-C (Mandel)	-
Vu 2009:23	HEK293T	-	07-579	REST dimerises independently of Zinc Fingers
Kreisler 2010, Oncogene	SCLC cell lines; HCC44, H128, H69, CRL2195, SW2	140, 200	07-579	200kDa predominantly nuclear
Faronato 2013	A549	-	07-579	180-220kDa REST form - nuclear and O-glycosylated. 220kDa is a double band – upper phosphorylated and degraded in mitosis, lower- unphosphorylated and accumulates in G1.

Table 3.1 Summary of the High Molecular REST observations.
Previous publications either addressing or mentioning alternative REST bands, the size of REST bands identified, cell lines and antibodies used and any additional information and characterisation.

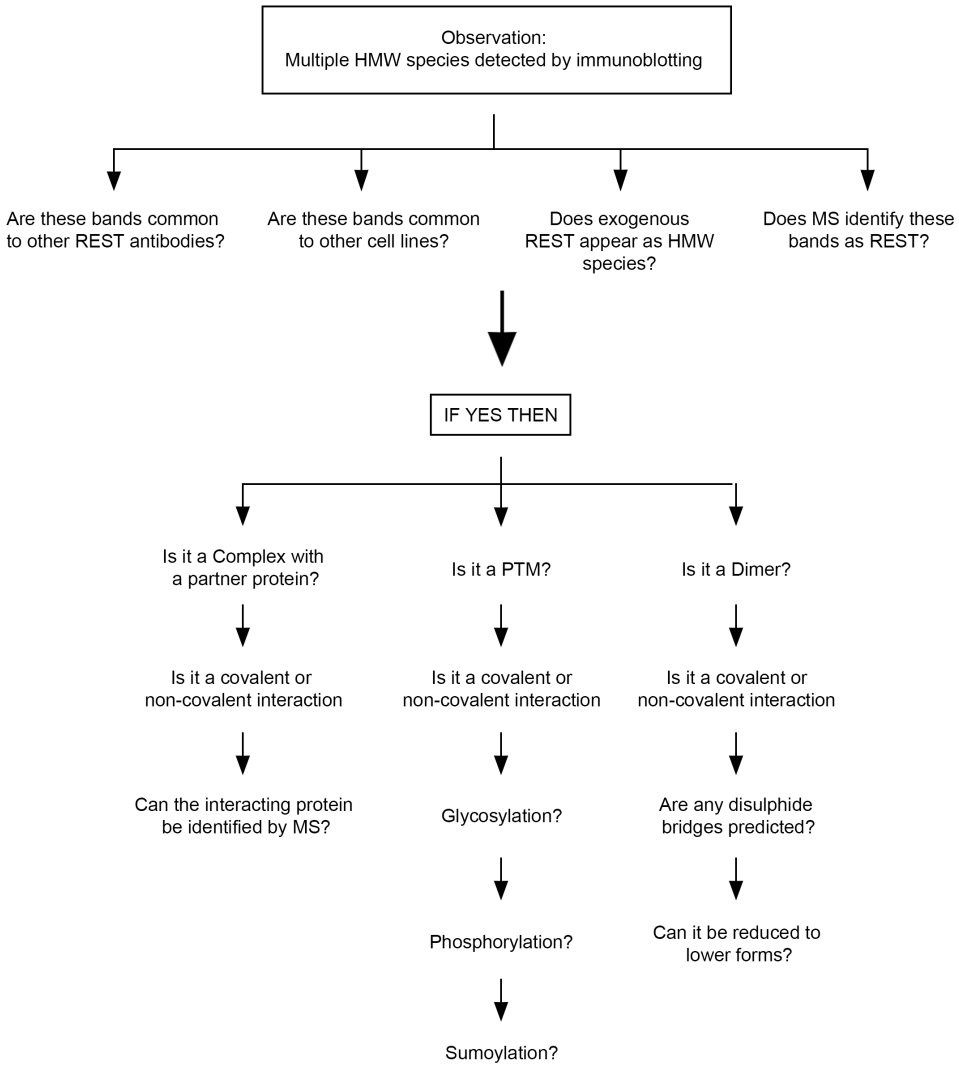


Figure 3.1 Scheme of identification of HMW forms of REST.

3.2 Results

3.2.1 Endogenous REST has multiple isoforms in HEK293T cells

Comparative immunoblotting on whole cell HEK293T extract using a REST antibody panel identified several high molecular weight (HMW) species in addition to 116 kDa REST. This was followed by an investigation into the antibody epitopes (Figure 3.2). Three of these antibodies, Mandel, 07-570 and 17-641, are polyclonal antibodies, all raised to C-terminus of REST, and two are monoclonal, 7D1.3 raised to the C-terminus and 2436Y to the N-terminus. The size of REST based on its peptide sequence is 116 kDa (Chong et al., 1995). A band around this size is recognised by two out of three polyclonal antibodies tested, 17-641 and 07-579, and one monoclonal antibody, the C-terminal 7D1.3. Furthermore, there is a consistency in the identification of the band close to 220kDa and 140kDa by all polyclonal antibodies. The 140 kDa band is additionally identified by the N-terminal monoclonal antibody, 2436Y. These data suggest multiple possibilities. One of which is that these bands are most likely to be post-translationally modified forms of REST with two main modifications, one smaller modification in the N-terminus, with a size of 20 kDa, and a bulkier modification in the C-terminal part lending an additional weight of 60 kDa. The N-terminal monoclonal 2436Y antibody preferentially identifies the N-terminal modified REST without the C-terminal modification, the C-terminal monoclonal preferentially picks up the unmodified REST and the polyclonal C-terminal antibodies, Mandel, 17-641 and 07-579 identify both the C-terminal and the N-terminal modified forms.

The 220 kDa form of REST is approximately a 100 kDa increase in molecular weight, suggesting that the existence of an interaction of REST with either itself or a partner protein, resulting in the 220 kDa band. The smaller observed bands (e.g. at 140 kDa

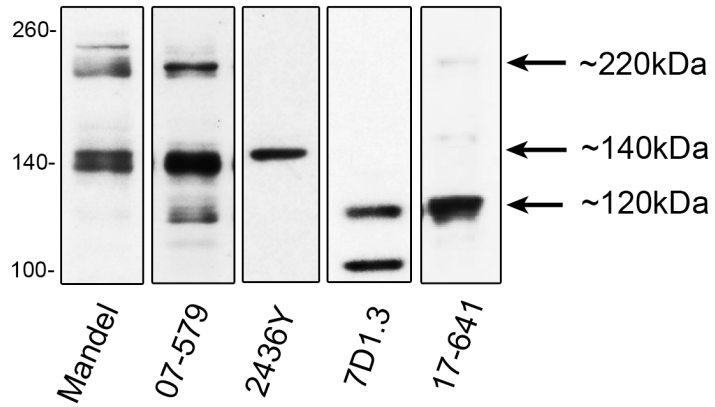
and 116 kDa) indicate either unbound protein or degradation products of the complex. A third possibility would be the existence of alternative REST splice variants, that, when combined with different PTMs, leads to multiple REST isoforms.

07-579 and Mandel antibody were found to generate the clearest bands across the various REST forms and were cited in most high impact papers (Table 3.1). Hence, subsequent experiments were largely performed using these antibodies.

3.2.2 One REST Splice Variant in expressed in HEK293T cell line

To determine if any of the bands identified by anti-REST antibodies were alternative splice variants of REST in HEK293T, whole transcriptome RNA sequencing data for HEK293T was analysed for alternative transcripts. Data was acquired from the GEO database (Gene Expression Omnibus) with the accession number GSE62591 for the control untreated RNAseq data for two HEK293T clones, A and B, and analysed using the cummerBund software (analysis was performed by David Barnes). Two transcript variants were identified in HEK293T cells (Figure 3.3). Variant 2, which is shorter than Variant 1 by 11 bp in the 5' UTR, was expressed at a higher level in both clones. Both of these transcripts are protein coding and code for the full-length REST (Isoform 1), as confirmed by their transcript ids in RefSeq. Hence, I disregarded the possibility of various REST isoforms expressed due to alternative transcript variants.

A



B

Antibody	Type	Species	Immunogen
17-641	Polyclonal	Rabbit	801-1097 of GST fusion protein
07-579	Polyclonal	Rabbit	801-1097 of GST fusion protein
Mandel	Polyclonal	Rabbit	C-terminus (exact epitope unknown)
7D1.3	Monoclonal	Mouse	801-1097 of GST-coupled recombinant protein
2436Y	Monoclonal	Rabbit	144-157 of synthetic peptide

Figure 3.2 Characterisation of anti-REST antibodies available commercially.
A) HEK293T whole cell lysate immunoblotted for REST using a panel of anti-REST antibodies. B) Table showing Anti-REST antibodies tested, their type, antibody species and the REST amino acid sequence of immunogen used to raise the antibody. ('Mandel' antibody was a generous gift from Gail Mandel).

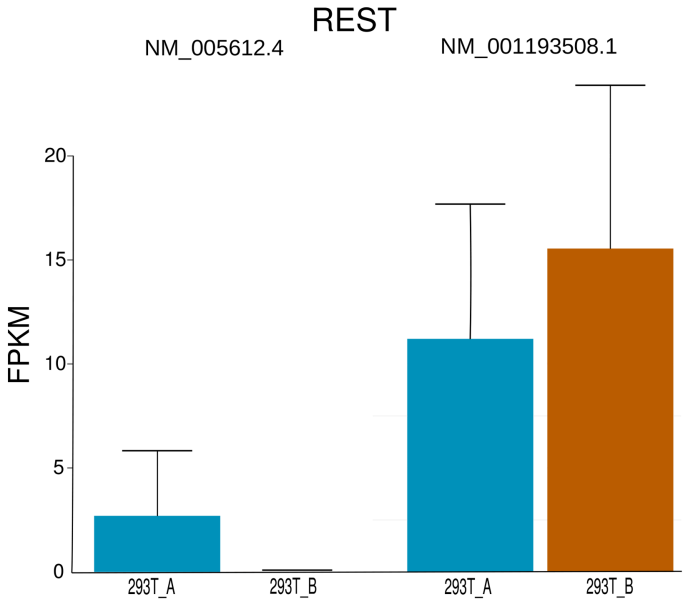


Figure 3.3 REST splice variants found in HEK293T cells.
Expression of REST isoforms was determined using cummerbund band and the output is shown here. RNAseq data for two HEK293T clones, A and B, was downloaded from GEO and expression of REST transcript variant determined using the cummerbund software. Normalised expression is indicated by fragment per kilobase per million mapped reads (FPKM). NM_005612.4 and NM_001193508.1 transcript variants differ in 11 bp at the 5' UTR and both code for full-length REST protein (Isoform 1).

3.2.3 Higher MW REST isoforms are commonly found in cell lines

To confirm whether the HMW forms are commonly found in cell lines, REST protein was examined in the panel of cancer cell lines used in section 3.2, using the 07-579 and Mandel antibodies. It can be observed that the HMW forms are found in all cell lines (Figure 3.4) and are mainly around 220 kDa or 140 kDa in size, in addition to the LMW form at 116 kDa. However, there are differences in the pattern and abundance of the HMW form. Both Mandel and 07-579 antibodies identify 140 kDa and 220 kDa forms. The 07-579 antibody also identifies the 116 kDa band, whereas the Mandel antibody only shows a faint band at 116 kDa.

Densitometry on all of the REST bands in each cell line was next performed, normalised to α-tubulin levels (Figure 3.4B). It can be seen that all cell lines show relatively similar levels of REST, with Hep3B showing the highest levels. REST levels in OSA are significantly lower than the other cell lines.

Both antibodies show similar levels of REST except for NIH3T3, where 07-579 antibody shows REST levels to be three times that of the REST levels shown by the Mandel antibody. When the abundance of the 116 kDa and 220kDa band was compared by performing densitometry on the REST bands identified by 07-579 antibody (Figure 3.4C), the difference was found to be because of the considerably higher expression of the 116 kDa band which is recognised by the 07-579 antibody.

One important observation here is that cancer cell lines of similar origin show comparable band pattern. For example, using the Mandel antibody, the Ewing sarcoma cell lines, SK-N-MC, CHP100 and A673, all show a double band at about 220 kDa of almost similar intensity, the hepatocellular carcinoma cell lines, Huh7 and Hep3B, show a double band with the top one being more intense than the lower one and the osteosarcoma cell lines, Hos and Hos143B show a triple band with the lower most band being the most intense and HT29 resembles this pattern.

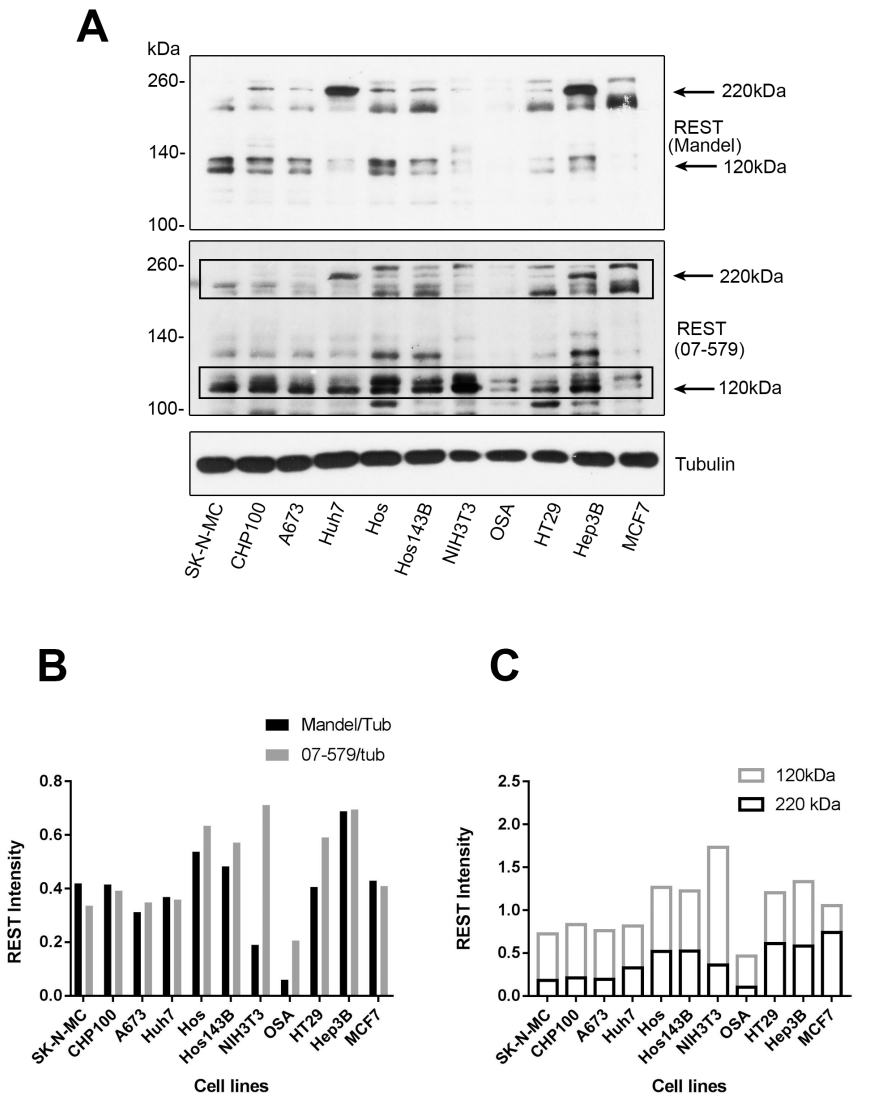


Figure 3.4 Higher MW REST forms are ubiquitously observed in cell lines.
A) A panel of cell lines was immunoblotted for α-tubulin and REST using Anti-REST (Mandel). The blot was stripped and re-probed for REST using 07-579 antibody. B) Densitometry was performed on total REST (all bands) identified by both antibodies, normalised to α-tubulin levels and plotted. C) Densitometry was performed on the 220kDa and 116 kDa band identified by 07-579 antibody (lower panel) for each cell line (bands outlined by a rectangle), normalised to α-tubulin and plotted.

3.2.4 Exogenous REST is predominantly present in the higher isoforms

Next, an attempt was made to identify if exogenously transfected REST was also present in these different MW isoforms. HEK293T cells were transfected with N-terminal Flag-tagged full-length REST and cell lysates prepared after 48 hours. Equal amounts of transfected and untransfected lysates were subsequently immunoblotted using various REST antibodies (Figure 3.5). It can be observed that there appears consistency in the increase in bands close to 220 kDa and 140 kDa as well as a slight increase in the 116 kDa band. The same lysate when probed using an Anti-Flag antibody gives a single band, which corresponds to the 220kDa REST band. When using the monoclonal antibodies, an important observation is that the 116 kDa band in the case of 7D1.3 or the 140kDa band in case of 2436Y show little increase but in both cases, a band close to 220 kDa appears. Since the Flag antibody does not recognise the smaller isoforms of REST, it is possible that those might be N-terminal degradation products with the removal of Flag tag.

This lysate was then immunoprecipitated using anti-Flag antibody and the resulting eluate immunoblotted using anti-REST and anti-Flag antibody. The eluate shows only one high intensity band around 220kDa, which is detected by both Flag and 07-579 REST antibodies (Figure 3.6A and B). There are additional bands detected by the antibody at 85 kDa and 116 kDa, which were not detected by the Flag antibody, hence they might either be C-terminal degradation product or non-specific bands. The eluate was also probed with the 17-641 and Mandel antibody with the same result, confirming the specificity of the polyclonal REST antibodies towards REST (Western blot not shown). Hence, exogenous REST can also exist in various MW isoforms, with the 220kDa isoform preferentially immunoprecipitated by Anti-Flag antibody.

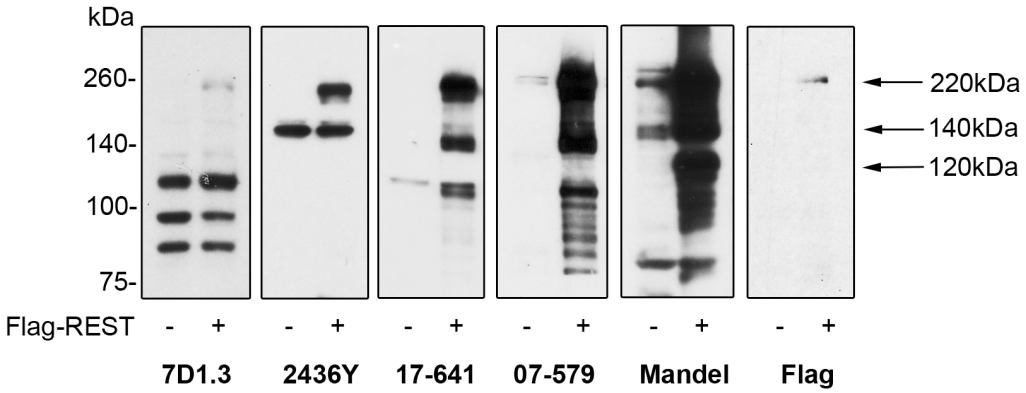


Figure 3.5 Exogenous REST is predominantly present in the higher isoforms.
Whole cell lysate from HEK293T transfected with Flag-REST or mock-transfected was immunoblotted for REST using a panel of anti-REST antibodies.

3.2.5 220kDa HMW REST band is identified as REST by Mass Spectrometry

To confirm the identity of the 220kDa band as an HMW form of REST, Flag-tagged REST was immunoprecipitated from whole cell lysates from HEK293T cells transiently expressing Flag-REST, by anti-Flag antibody and analysed by LC-MS/MS. Prior to mass spectrometry analysis, the immune-precipitated fraction was resolved on an SDS-page gel and the 220 kDa band, visualised by Coomassie staining, was cut out from the gel and subjected to tryptic digest. LC-MS/MS analysis recognised the 220kDa band to be REST, with a coverage of 44% (Figure. 3.6C), confirming that the 220kDa REST band represents a higher molecular weight species of REST.

3.2.6 HMW REST form is resistant to Urea

The size of HMW REST suggests that it is potentially a heavily modified form of LMW REST. This modification can be a result of either covalent interactions including peptide linkages with other proteins, intermolecular disulphide bonds, extensive phosphorylation, glycosylation or ubiquitination or non-covalent links including hydrogen or hydrophobic bonds leading to high molecular weight aggregates of REST.

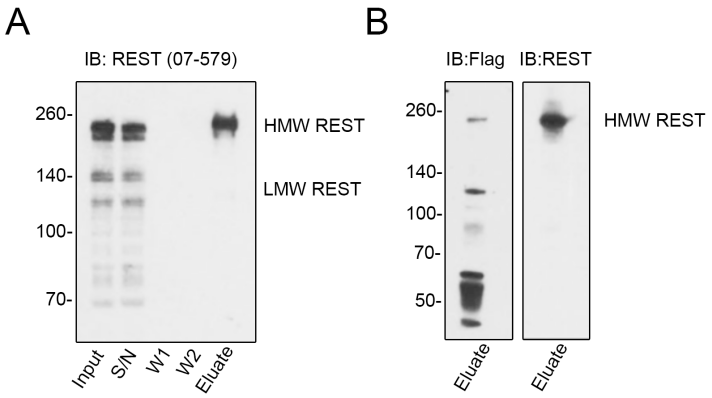
To investigate if the HMW REST was a covalently or non-covalently bonded species, a urea assay was employed. Urea is a chaotropic agent, which disrupts the three-dimensional structures of proteins by disrupting intramolecular interactions due to non-covalent forces like hydrogen bonding, hydrophobic interactions and van der Waals forces. Immunoprecipitated Flag-REST was treated with increasing concentrations of urea (0-8 M) and subsequently immunoblotted for REST using 07-579 antibody (Figure 3.7A). HMW REST form was found to be highly stable even

under conditions of 8M urea, suggesting that non-covalent bonds are unlikely to mediate interactions in the HMW form.

Amongst covalent bonds mediating a posttranslational modification, disulphide bonds, phosphorylation and glycosylation are additional possibilities, and these were subsequently explored.

3.2.7 HMW REST form does not contain disulphide bonds

The observed 220 kDa size of the HMW REST form is almost twice that of the 116 kDa predicted size. One possible explanation for this discrepancy is the presence of an intermolecular dimer. The dimer could be mediated through disulphide bonds between two cysteine residues in two separate monomers of REST. Handling of protein also sometimes favours the formation of aberrant intermolecular disulphide bonds not present endogenously. The peptide sequence of REST shows the presence of 24 Cysteines and sequence alignment of multiple REST orthologues indicates that 20 of these cysteine residues are highly conserved. To investigate the possibility of these cysteine residues forming intermolecular disulphide bonds, I treated HEK293T cell lysate with *tris*(2-carboxyethyl)phosphine.hydrochloride (TCEP.HCl), a potent reducing agent, followed by alkylation using iodoacetamide. This approach would be expected to break disulphide bonds in the protein, resulting in thiol groups, which are then transformed to the stable S-carboxyamidomethylcysteine, making the process irreversible. The resultant lysate was then immunoblotted for REST (Figure 3.7B). No change in either band size or intensity was observed between treated and untreated samples, suggesting that disulphide bonds are not present in the HMW REST form.



C

```
   1  MATQVMGQSS GGGGLFTSSG NIGMALPNDM YDLHDLSKAE LAAPQLIMLA
  51  NVALTGEVNG SCCDYLVGEE RQMAELMPVG DNNFSDSEEG EGLEESADIK
 101  GEPHGLENME LRSLELSVVE PQPVFEASGA PDIYSSNKDL PPETPGAEDK
 151  GKSSKTKPFR CKPCQYEAES EEQFVHHIRV HSAKKFFVEE SAEKQAKARE
 201  SGSSTAEEGD FSKGPIRCDR CGYNTNRYDH YTAHLKHHTR AGDNERVYKC
 251  IICTYTTVSE YHWRKHLRNH FPRKVYTCGK CNYFSDRKNN YVQHVRTHTG
 301  ERPYKCELCP YSSSQKTHLT RHMRTHSGEK PFKCDQCSYV ASNQHEVTRH
 351  ARQVHNGPKP LNCPHCDYKT ADRSNFKKHV ELHVNPRQFN CPVCDYAASK
 401  KCNLQYHFKS KHPTCPNKTM DVSKVKLKKT KKREADLPDN ITNEKTEIEQ
 451  TKIKGDVAGK KNEKSVKAEK RDVSKEKKPS NNVSVIQVTT RTRKSVTEVK
 501  EMDVHTGSNS EKFSKTKKSK RKLEVDSHSL HGPVNDEESS TKKKKKVESK
 551  SKNNSQEVPK GDSKVEENKK QNTCMKKSTK KKTLKNKSSK KSSKPPQKEP
 601  VEKGSAQMDP PQMGPAPTEA VQKGPVQVEP PPPMEHAQME GAQIRPAPDE
 651  PVQMEVVQEG PAQKELLPPV EPAQMVGAQI VLAHMELPPP METAQTEVAQ
 701  MGPAPMEPAQ MEVAQVESAP MQVVQKEPVQ MELSPPMEVV QKEPVQIELS
 751  PPMEVVQKEP VKIELSPPIE VVQKEPVQME LSPPMGVVQK EPAQREPPPP
 801  REPPLHMEPI SKKPPLRKDK KEKSNMQSER ARKEQVLIEV GLVPVKDSWL
 851  LKESVSTEDL SPPSPPLPKE NLREEASGDQ KLLNTGEGNK EAPLQKVGAE
 901  EADESLPGLA ANINESTHIS SSGQNLNTPE GETLNGKHQT DSIVCEMKMD
 951  TDQNTRENLT GINSTVEEPV SPMLPPSAVE EREAVSKTAL ASPPATMAAN
1001  ESQEIDEDEG IHSHEGSDLS DNMSEGSDDS GLHGARPVPQ ESSRKNAKEA
1051  LAVKAAKGDF VCIFCDRSFR KGKDYSKHLN RHLVNVYYLE EAAQGQE
```

Figure 3.6 Exogenous REST can be immunoprecipitated and is identified as REST by Mass spectrometry.

A) Flag-tagged REST transfected HEK293T lysate was immunoprecipitated using Anti-Flag M2 antibody. Input (IN), supernatant (S/N), wash 1 (W1), wash 2 (W2) and elution (Eluate) were run on an SDS-page gel and immunoblotted for REST using anti-REST 07-579 antibody. B) Eluate resulting from REST immunoprecipitation using anti-Flag M2 antibody was immunoblotted using anti-REST (07-579) or anti-Flag M2 antibodies. C) Coverage achieved by LC-MS/MS analysis of HMW REST band. A threshold of 20 was used to filter low scoring peptides. 44% of REST residues were covered. Covered residues are highlighted in red.

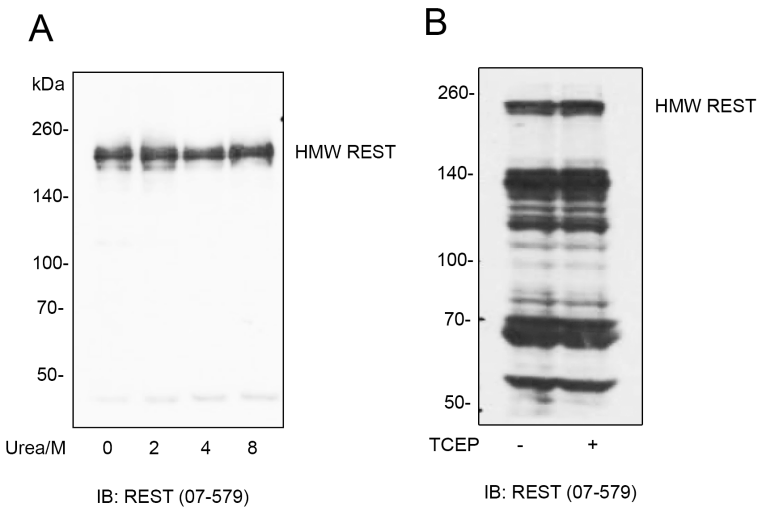


Figure 3.7 Urea treatment and reduction of REST.
A) Flag-REST immunoprecipitated from HEK293T cells was treated with 0-8 M Urea, the resulting eluate immunoblotted for REST using anti-REST (07-579) antibody. B) HEK293T lysate was reduced for 15 minutes with TCEP, followed by alkylation with IAA. Lysates were then immunoblotted for REST.

3.2.8 HMW REST is not a hetero-oligomer

To investigate the possibility of the 220kDa HMW form containing REST covalently interacting with other proteins, data obtained from LC-MS/MS of HMW REST band was analysed to identify a co-immunoprecipitating protein. There were in total 232 proteins co-purified with REST. Proteins having Mascot scores above 100 and at least 10 tryptic peptides were filtered to produce a final list of co-purified proteins (Table 3.2). An interacting protein co-purified with REST in the 220KDa band must have a molecular weight in the range of 80-116 kDa to result in an overall molecular mass of 220kDa. Alternatively, it could be two interacting proteins together adding up to 80-116 kDa. Furthermore, the interacting proteins would need to be present at the same abundance as REST. Peptide hits - the number of peptide-spectrum matches - can be used to describe the relative abundance of a protein, larger number corresponding to greater abundance. None of these identified proteins had the required abundance or molecular weight. Hence, they most likely co-purified due to interactions with the resin used or non-specifically with the anti-Flag antibody.

This was further confirmed by scrutinising CRAPome, an online database of negative-control affinity purification-mass spectrometry experiments. Raw mass spectrometry data files from numerous negative control experiments are processed using a uniform pipeline including strict quality control followed by association with detailed experimental protocols, establishing the base of the repository. Desired controls (a maximum of 30 if using the online workflow) can be selected using protocol details provided to derive the list of contaminant proteins for a particular set of experiment. Confidence in the occurrence of contaminants can be scored by the incidence of identification of a particular contaminant. Using the online workflow, background contaminants from 30 negative controls using Flag M2, magnetic agarose beads and whole cell HEK293T lysate, were analysed and searched for the list of co-purified

proteins with REST. Most of the proteins were found to be present in the list at a very high occurrence. The proteins not identified were Bloom Syndrome protein, Nuclear pore complex protein Nup153, Desmoglein-1 and Cornulin. The protein score of these proteins was however too low to be considered a covalent partner of REST. However, they might interact functionally with REST in the cell. Nup153, a component of the nuclear pore complex (NPC) localised to the nucleoplasmic side of the complex and implicated in import of nuclear proteins, especially is an interesting candidate, and might interact with REST to control nucleocytoplasmic transport (Walther et al., 2001)

Family	Member	Accession	Name	Score	Mass/Da	Num of matches	NUM_EXPT in CRAPome
1	1	P04264	Keratin, type II cytoskeletal 1	5120	66170	222	23
	2	P35908	Keratin, type II cytoskeletal 2	3412	65678	150	14
	3	P19013	Keratin, type II cytoskeletal 4	1390	57649	84	0
	4	P02538	Keratin, type II cytoskeletal 6A	954	60293	68	5
	5	P04259	Keratin, type II cytoskeletal 6B	886	60315	66	10
	6	P13647	Keratin, type II cytoskeletal 5	858	62568	65	5
	7	P12035	Keratin, type II cytoskeletal 3	850	64549	47	1
	8	Q8N1N4	Keratin, type II cytoskeletal 78	533	57629	24	0
	9	Q7Z794	Keratin, type II cytoskeletal 1b	469	62149	22	14
	10	Q86Y46-2	Isoform 2 of Keratin, type II cytoskeletal 73	399	42244	25	0
	12	K2C74	Keratin, type II cytoskeletal 74	248	58229	16	0
2	1	P35527	Keratin, type I cytoskeletal 9	3101	62255	141	12
	2	P13645	Keratin, type I cytoskeletal 10	2697	59020	122	15
	3	P13646	Keratin, type I cytoskeletal 13	1671	49900	77	6
	4	P02533	Keratin, type I cytoskeletal 14	1129	51872	67	8
	5	P08779	Keratin, type I cytoskeletal 16	891	51578	53	8
	6	Q04695	Keratin, type I cytoskeletal 17	431	48361	34	7
	7	P14923	Junction plakoglobin	254	82434	12	1
3	1	Q13127	RE1-silencing transcription factor	2452	123164	149	0
4	1	Q14152	Eukaryotic translation initiation factor 3 subunit A	842	166867	51	19
5	1	Q00610-2	Isoform 2 of Clathrin heavy chain 1	771	189538	35	20
6	1	P15924	Desmoplakin	753	334021	36	4
7	1	P07814	Bifunctional glutamate/proline--tRNA ligase	644	172080	36	14
8	1	Q9NTJ3	Structural maintenance of chromosomes protein 4	615	147775	25	7
9	1	P54132	Bloom syndrome protein	540	160611	22	0
10	1	Q13576	Ras GTPase-activating-like protein IQGAP2	469	181036	17	5
11	1	Q8TEQ6	Gem-associated protein 5	396	170821	14	2
12	1	P04083	Annexin A1 OS=Homo sapiens	365	38918	10	1
14	1	P49790	Nuclear pore complex protein Nup153	338	155440	11	0
16	1	Q02413	Desmoglein-1	268	114702	13	0
17	1	Q9UBG3	Cornulin	254	53730	10	0
21	1	O14744	Protein arginine N-methyltransferase 5	228	73322	11	25

Table 3.2 List of proteins co-purified with REST.

A hierarchical clustering algorithm groups proteins into families, with each family member identified by its accession number and name. Protein score corresponds to the summation of the individual peptide scores for each protein, Mass corresponds to the molecular weight of the proteins in kDa, num of matches indicates how many corresponding matches were identified for each protein and NUM_EXPT indicates the number of negative control experiments (out of 30)

the protein was identified in CRAPome. Proteins listed were MASCOT score above 100 and had at least more that 10 tryptic peptides identified.

3.2.9 HMW REST form is not N-glycosylated

A common PTM of proteins is N-linked glycosylation, which results from the attachment of a glycan to an asparagine side chain of a protein. Multiple linkages can result in a highly complex glycan, which can lead to large mass additions and play a major role in protein folding, transport and protein stability (Moremen, Tiemeyer and Nairn, 2012). To investigate the possibility of REST being N-glycosylated, N-glycosidase F (PNGase F) digestion was applied. PNGase F is an enzyme that cleaves N-linked oligosaccharides from proteins very efficiently, leaving the oligosaccharide intact but the asparagine, from which the oligosaccharide is removed, is deaminated to aspartic acid.

Flag-REST was transiently expressed in HEK293T cells to increase the amount of HMW REST and whole cell lysate treated with PNGase F enzyme for 1 hour at 37^{o}C. Similarly, untransfected HEK293T lysate was also treated with PNGase. Western blotting of treated and untreated extracts shows that there is no discernible difference in the migration of bands after treatment with PNGase whereas a control protein, Fetuin, is clearly deglycosylated and reduced in size (Figure 3.8A). This suggests the absence of N-linked glycosylation on REST.

3.2.10 HMW REST form is O-glycosylated

Another type of glycosylation is O-glycosylation, which consists of glycan attachment to serine or threonine side chains. Due to the heterogeneous nature of O-glycosylation and the lack of an enzyme that is as efficient as PNGase in removing O-glycosylation, I employed an O-linked glycosylation inhibitor, 6-Diazo-5-oxo-L-norleucine (DON), and an O-linked glycosylation agonist, O-(2-Acetamido-2-deoxy-D-glucopyranosylidene)amino *N*-phenyl carbamate (PUGNAc) to investigate the presence of O-linked glycosylation. HEK293T cells were treated with varying

concentration of DON and PUGNAc from 0 to 100μM for 24 hours after which they were lysed and immunoblotted for REST using anti-REST (17-641). This REST antibody was chosen for this experiment since it provides the clearest staining of the 120 kDa REST band along with the 220 kDa, in comparison with the other antibodies. Expression of HMW form was observed to be reduced with higher doses of DON (Figure 3.8B) with a small simultaneous increase in the LMW form at 116 kDa. PUGNAc increases expression of the HMW form slightly at lower doses, but the LMW form is unaltered. At higher doses, HMW REST starts going down, indicating a non-specific effect. Hence, I concluded the presence of O-linked glycosylation on the HMW form of REST.

To predict the locations of these O-glycosylated sites, a prediction program called YinOYang 1.2 was utilised. YinOYang uses an artificial neural network approach, trained by 40 experimentally identified O-GlcNAc sites, to analyse the sequence and surface accessibility of input sequences (Gupta and Brunak, 2002). This led to the correct prediction of 72.5% of glycosylated residues and 79.5% of unglycosylated residues in the tested proteins. The method can also identify putatively phosphorylated residues, giving sites, which might be reciprocally regulated by O-glycosylation and phosphorylation. REST residues predicted by YinOYang to be O-Glycosylated are given in Table 3.3, along with potential to be reciprocally regulated by phosphorylation.

Eleven REST residues were found to be putative O-glycosylation sites, with one of those sites having a higher probability than the others. Four sites were predicted to be YinOYang sites, reciprocally regulated by O-glycosylation and phosphorylation. Using this as an indicator, it was safe to predict that there are residues in REST protein that might be O-glycosylated. No putative residue was, however, identified in

the region (amino acids 87-152) that was narrowed down by Lee et al. to be potentially O-glycosylated (Lee et al., 2000).

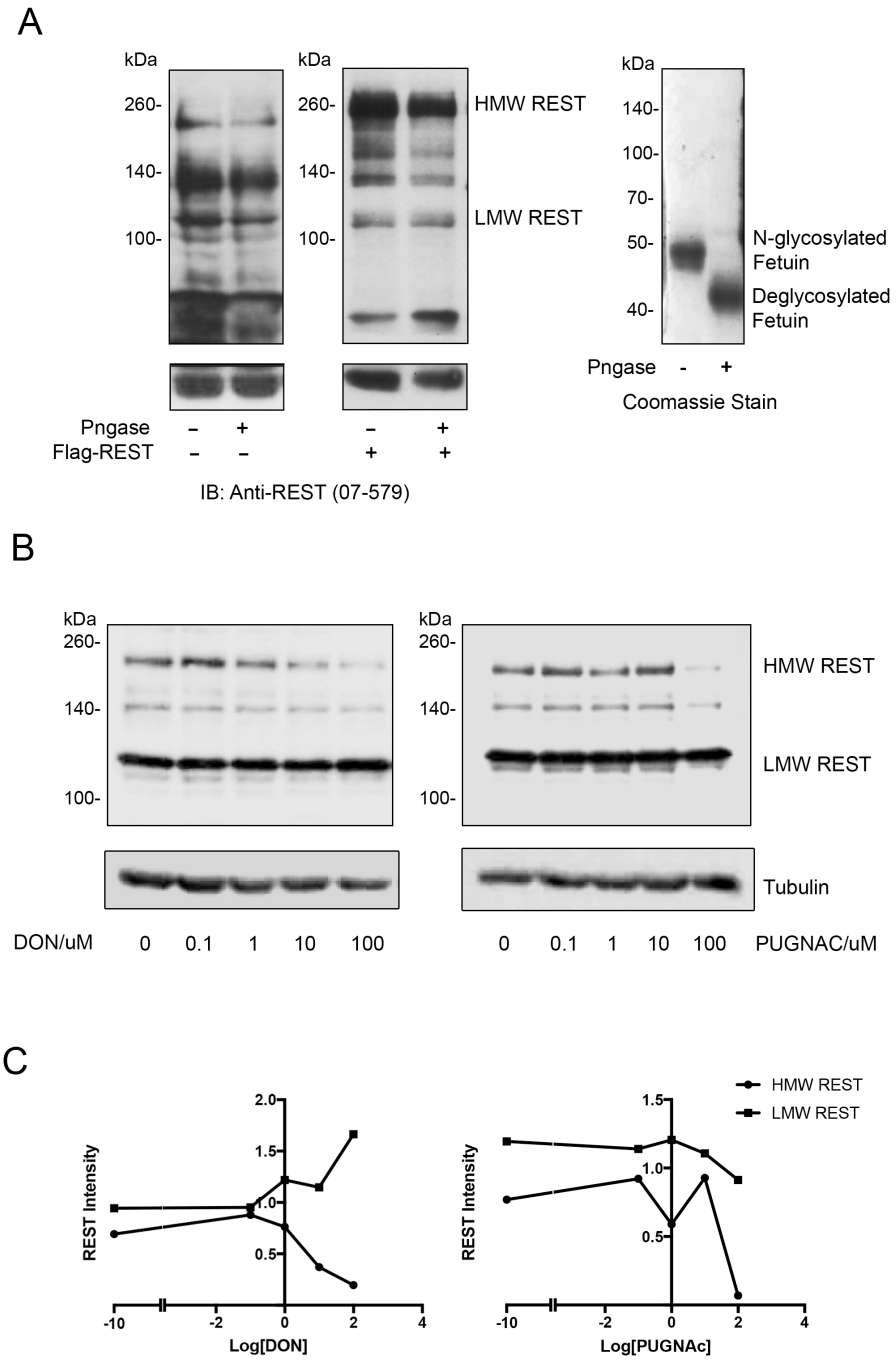


Figure 3.8 N- and O-linked Glycosylation of REST.
A) HEK293T whole cell lysates wild type or transfected with Flag-REST were incubated with PNGase F for 1 hour at 37°C, followed by immunoblotting for REST using anti-REST (07-579). Deglycosylation of Fetuin (left-hand panel) was used as a positive control. Treated or untreated Fetuin was run on an SDS page gel and stained with Coomassie stain. B) HEK293T cells were treated with varying doses of DON and PUGNAc for 24 hours, lysed and lysates immunoblotted for REST using anti-REST (17-641) antibody and α-tubulin. C) REST levels from A were quantified using Image J, normalised to α-tubulin levels and plotted.

Residue	O-GlcNAc Prediction	Conservation	Region	YinOYang
T3	+	Yes		
S312	+	Yes		
S313	+	Yes	Zinc Finger/ HAUSP binding	
S484	+	Partially	Zinc Finger/ HAUSP binding	
S588	+	Partially	Lys-rich	Yes
S589	+	Partially	Lys-rich	Yes
T618	+	No	Pro-rich	
T857	+	No		
S971	++	Yes		Yes
S986	+	Partially		Yes
T996	+	Yes	RD2	

Table 3.3 Predicted O-GlcNAc sites by YinOYang 1.2.
Residues are given using their sequence number in the REST sequence. Prediction confidence is indicated by + with ++ indicating higher probability. The residue conservation in vertebrate REST orthologues is indicated. YinOYang column indicates putative reciprocal modification by phosphorylation.

3.2.11 HMW REST form is Phosphorylated

The second possibility for covalently attached modification is phosphorylation. To test for this possibility, I treated HEK293T extract with lambda phosphatase. Lambda phosphatase is a Mn^{2+}-dependent protein phosphatase, which removes phosphate groups from serine, threonine and tyrosine residues. Treated lysate was immunoblotted for REST using anti-REST (07-579) and anti-REST (Mandel) (Figure. 3.9). Both antibodies show a decrease in the intensity of the top band of the HMW REST forms suggesting that it is phosphorylated. The LMW forms of REST are fairly resistant to Lambda phosphatase, hence suggesting that they are unphosphorylated bands.

3.2.12 HMW REST form is Sumoylated

A bioinformatics search for PTMs of REST resulted in the identification of multiple consensus sites for sumoylation (Figure 3.10). Sumoylation is a post-translational modification frequently found in nuclear proteins especially transcription factors and has been found to play important roles in their function. Sumo moieties can form long chains and add significantly to the mass of unmodified proteins. Hence, an attempt was made to check whether the HMW form is sumoylated. Whole cell lysate from HEK293T cells overexpressing Flag-REST was used to immunoprecipitate Flag-REST. Resulting eluate was immunoblotted for Sumo2/3 and REST using an Anti-Sumo2/3 antibody and anti-REST antibody respectively. As shown in Figure 3.10, there is a band detected by the Sumo 2/3 antibody in the eluate, which corresponds to the 140 kDa REST band. There is no immunoreactive band around 220kDa. This shows that 140kDa REST is potentially sumoylated whereas 220kDa HMW REST is unlikely to be sumoylated.

Experimental studies have identified a consensus motif Ψ-K-x-D/E for sumo sites where Ψ is a hydrophobic residue, K is the lysine to which SUMO is conjugated, x is any amino acid, and D/E is an acidic residue. 40% of sumo sites, however, do not follow this motif. SUMO can also interact with other proteins non-covalently by binding to SUMO-interacting-motifs (SIMs). There is no existent consensus sequence for SIMs, although approximately ten types of SIMs have been identified experimentally. An in silico prediction program called GPS-SUMO was used to predict potential SUMO sites and SIMs in REST (Zhao et al., 2014). GPS-SUMO uses a Group-based Prediction System (GPS) algorithm in addition to a matrix mutation (MaM) method to identify putative SUMO sites and SIMs. Analysis of REST identified 8 putative SUMO sites, of which 6 were consensus sites. 4 SIMs were identified. Of these, 6 SUMO binding sites were conserved, 2 SIMS were fully conserved, and 1 SIM was half-conserved, amongst human, mouse, rat, dog, cow and rhesus monkey. Hence, these are strong candidates for experimental verification.

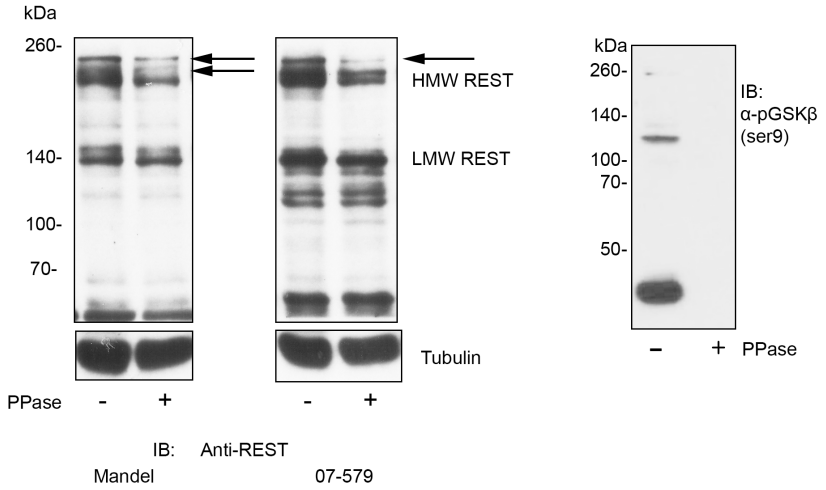
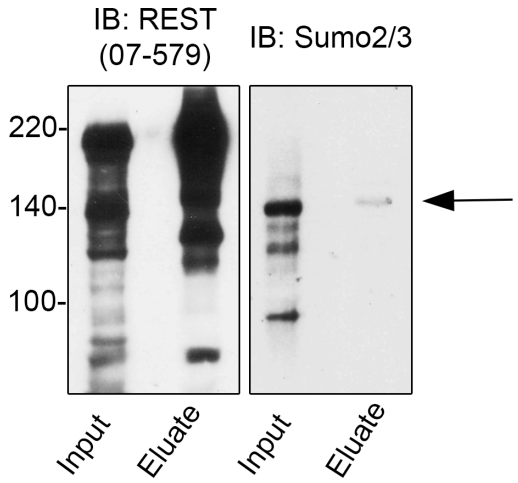


Figure 3.9 REST is phosphorylated.
HEK293T whole cell lysate was incubated with or without λ-Phosphatase for 1 hour at 37°c, after which lysates were immunoblotted for REST using Anti-REST (07-579) and Anti-REST (Mandel). Left-hand panel shows the disappearance of α-pGSKβ upon phosphatase treated, used as positive control.

A



B

Position	Peptide	Score	Cutoff	P-Value	Type	Annotation	Conserved
100	LEESADIKGEPHGLE	23.559	2.13	0.002	SUMO Consensus	Interaction with msin3A/rd1	Yes
431	KVKLKKTKKREADLP	3.676	3.32	0.061	SUMO Nonconsensus		Yes
432	VKLKKTKKREADLPD	3.808	3.32	0.048	SUMO Nonconsensus		Yes
454	EIEQTKIKGDVAGKK	7.623	2.13	0.021	SUMO Consensus		Yes
467	KKNEKSVKAEKRDVS	19.304	2.13	0.002	SUMO Consensus		Yes
485 - 489	KPSNNVSVIQVTTRTRKSV	34.156	29.92	0.117	SUMO Interaction		Half-site conserved
560	NNSQEVPKGDSKVEE	5.002	2.13	0.048	SUMO Consensus		Yes
762	VVQKEPVKIELSPPI	12.478	2.13	0.008	SUMO Consensus		No
769 - 773	KIELSPPIEVVQKEPVQME	46.645	29.92	0.031	SUMO Interaction		No
837 - 841	RARKEQVLIEVGLVPVKDS	37.417	29.92	0.093	SUMO Interaction		Yes
948	DSIVCEMKMDTDQNT	5.182	2.13	0.032	SUMO Consensus		No
1083-1087	SKHLNRHLVNVYYLEEAA Q	33.123	29.92	0.09	SUMO Interaction	Interaction with RCor1	Yes

Figure 3.10 REST is Sumoylated.
A: Flag-REST was immunoprecipitated from Flag-REST transfected whole cell lysate and immunoblotted for REST using 07-579 antibody and SUMO2/3 using SUMO2/3 antibody. The arrow indicates the REST band that is SUMOylated. B: Putative SUMOylated residues and SUMO interaction motifs in REST. A cut-off value of 2.13 for SUMO consensus motifs, 3.32 for SUMO Nonconsensus and 29.92 for SUMO interaction motifs was used as guided by the prediction server. The p-value for each prediction is given to score the confidence in each prediction along with conservation of the residue amongst REST orthologues; human mouse, rat, cow, dog and rhesus monkey. Residues marked by red indicate the conjugated lysine residue or the SIM.

3.2.13 Subcellular localisation of HMW and LMW forms

It is a possibility that the two forms of REST might be localised differently in sub-cellular compartments, hence representing two different pools of REST, with the functionally active pool of REST inside the nucleus. The PTMs might be serving to differentiate the two pools and allowing access to the nucleus. To explore this possibility, cytoplasmic and nuclear extracts from HEK293T, MCF7 and NIH3T3 cells were prepared and immunoblotted for REST (Figure 3.11). It can be observed that the HMW form is predominantly localised in the nucleus in both HEK293T and MCF7 cells, with a slightly faster migrating species of the 220kDa HMW form predominantly nuclear whereas the slower migrating species is also found in the cytoplasm. Similarly, in NIH3T3 cells, the faster migrating species are nuclear. However, it appears that REST is distributed more equally between the nucleus and cytoplasm. Immunofluorescence of REST using the same REST antibody (Mandel) confirms these observations, with REST in HEK293T and MCF7 being mostly nuclear whereas in NIH3T3 it appears to be more localised in the cytoplasm (Figure 3.11). Subsequently, the localisation of full length Flag-REST was confirmed in HEK293T cells and found to be exclusively nuclear.

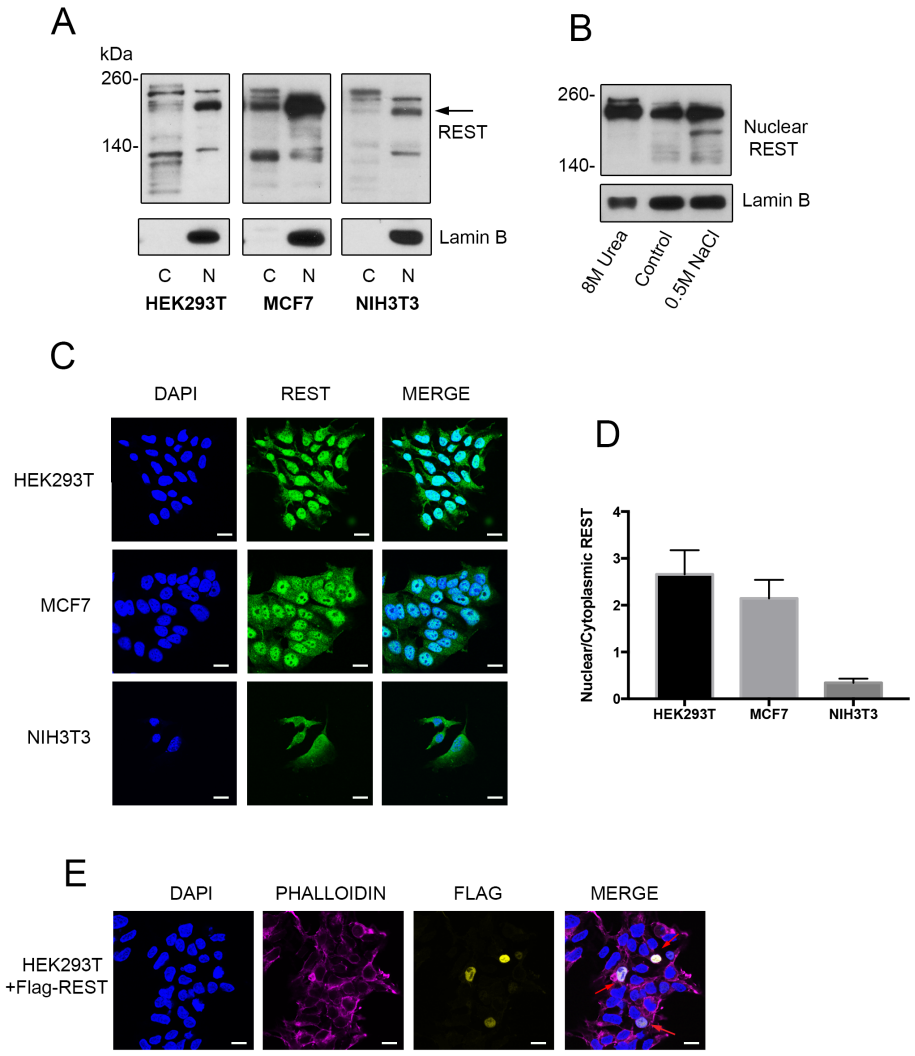


Figure 3.11 Subcellular localisation of HMW and LMW REST forms.
A) Nuclear and cytoplasmic extracts of HEK293T, MCF7 and NIH3T3 cells were prepared and immunoblotted for REST using Mandel antibody. B) Nuclear extract was treated with 8M urea or 0.5M NaCl and immunoblotted for REST using Mandel antibody. C) Immunolabelling of HEK293T, MCF7 and NIH3T3 cells for REST using Mandel antibody and DAPI stain for nuclei. Scale bar = 10 µm. D) Nuclear and cytoplasmic REST levels were quantified from 5 fields of three separate samples for each cell line, using Image J and the average nuclear/cytoplasmic intensity plotted. $P<0.0005$. Scale bar = 10 µm. E) Immunolabelling of HEK293T post 48 hours of Flag-REST transfection using Phalloidin, Flag antibody and DAPI for nuclear stain.

3.3 Discussion

A difficulty emerging from my work was the interpretation of results with respect to various bands of REST and variability arising from usage of individual antibodies. The monoclonal 7D1.3 and polyclonal 17-641 antibodies were giving variable results, and there were multiple bands identified by the polyclonal antibody (Appendix 4). Hence, to correctly interpret my results, an attempt was made to thoroughly characterise these multiple bands using a multistep approach.

It is clear that these HMW REST forms are very consistent between cell lines and between antibodies. Both the monoclonal antibodies only recognise one band whereas the polyclonal antibodies give multiple bands of different intensity. These bands are unlikely to be splice variants of REST as HEK293T cells only express one splice variant. They are unlikely to be non-specific bands as they are detected by all polyclonal antibodies, and overexpressing full-length REST increases the amount of the HMW forms of REST. When over-expressed REST is immunoprecipitated, the 220kDa REST species is predominantly detected in the eluate. This species is recognised as REST by mass spectrometry. Hence, exogenous REST is preferentially directed into a pool of REST, which is altered to result in a higher molecular form. The higher form might form a pool for newly synthesised REST, allowing a way to differentiate between ‘new’ and ‘old’ REST protein pools.

HMW Rest was found to be resistant to urea indicating that covalent bonds would be causing the shift in mass. Amongst covalent bonds, REST can be post-translationally modified, phosphorylation, glycosylation and sumoylation being candidate covalent modifications. The two other possibilities are REST interacting with a partner protein covalently or covalently binding to itself resulting in a dimer. I was able to rule out an interaction with a partner protein since none was identified by our mass spectrometry

analysis. Examining the immunoblot bands resulting from different cell lines, a common pattern was identified between cell lines of similar origin. Since different cancer cell lines might have different PTMs, this fits in well with the PTM hypothesis. Amongst covalent PTMs, we explored phosphorylation, glycosylation and sumoylation.

3.3.1 Phosphorylation

Lambda phosphatase was found to effect one of the 220kDa HMW REST bands but did not eliminate both bands hence indicating that REST might be phosphorylated on several residues, but the HMW form is not due solely due to phosphorylation. This is consistent with published data where REST has been found to be phosphorylated at multiple sites. Two well-studied sites include the phosphodegrons involved in proteasome-dependent degradation (Westbrook et al., 2008). Faronato et al. also demonstrated the existence of two REST bands around 220kDa using A549 cells and the 07-579 antibody, the top one being phosphorylated, the relative intensities of these bands varying along the cell cycle (Faronato et al., 2013). My data in HEK293T cells using the 07-579 antibody is similar, in that the top band is phosphorylated. However, the bands are slightly different in size and relative intensity compared to those in A549 cells. It would be useful to repeat the experiment with other cell lines used in our panel to mark phosphorylated bands. Phosphorylated residues run differently on a gel due to the added mass and negative charge and the relative positioning is dependent on the balance between these two components. Some phosphorylated proteins run slower whereas others run faster than the unphosphorylated proteins depending on this interplay. The top band, which is phosphorylated, represents a heavier, slower-migrating species.

Most signalling pathways, including the IGF2 pathway, involve phosphorylation of proteins. Hence, either phosphorylation or conversely dephosphorylation might form an integral part of the mechanism of REST stability as observed in Chapter 3. The IGF2 response might be mediated either by the identified degrons or as yet undiscovered phosphorylated residues. This calls for a thorough analysis of REST phosphorylated residues.

As the size of a phosphorylation modification adds on an additional mass of approximately 80 Da, multiple phosphorylations are unlikely to explain the 100 kDa increase in mass seen on an SDS-page gel. Hence I explored other potential PTMs.

3.3.2 Glycosylation

Glycosylation is one of the most common PTMs as yet discovered with more than half of all proteins estimated to be glycosylated (Apweiler, Hermjakob and Sharon, 1999). N-linked and O-linked glycans are the most common type of linkages. I failed to identify the presence of N-linked glycans whereas the 220kDa HMW REST band was found to be reduced by DON, an inhibitor of the OGT enzyme that places O-linked glycans onto serine and threonine residues. This is consistent with previous experiments (Faronato et al., 2013; Lee et al., 2000). Faronato et al. used DON and PUGNAc to demonstrate that the 220kDa HMW band is increased and decreased in abundance, respectively, upon treatment. Furthermore, siRNA against OGT also eliminates the 220 kDa REST band.

Two types of O-glycosylation are known, one involving the addition of an N-acetylgalactosamine (GalNAc) to Ser/Thr residues, most commonly associated with mucin glycoproteins, and the other linking β-N-acetylglucosamine (O-GlcNAc) to Ser/Thr residues. Whereas the former modification is mainly restricted to the secretory channels, the latter can be found in nuclear and cytosolic proteins and hence

is a stronger candidate for REST PTM. Many transcription factors including p53 and FOXO1 (Yang et al., 2006; Housley et al., 2008) are O-GlcNAcylated, the modification playing an important role in stability, oligomerisation, nuclear targeting, DNA binding and cell signalling (Vaidyanathan, Durning and Wells, 2014). The latter is achieved by cross talk between O-GlcNAcylation and phosphorylation signalling cascades and might be used to regulate protein activity dynamically.

It is challenging to predict candidate sites, as there is no consensus sequence for O-glycosylation, unlike N-glycosylation. I employed YinOYang to predict candidate O-glycosylation sites and identified eleven REST, which were strongly predicted to be O-glycosylated. These serine and threonine residues can be mutated, transfected in cells and cell lysates immunoblotted using an antibody against O-GlcNAc to identify the O-glycosylated site.

3.3.3 Sumoylation

Sumoylation is another PTM with roles in transcription, chromatin structure and DNA repair. Sumo are ubiquitin-like proteins (UbLs) that can be covalently conjugated to a lysine residue on substrate proteins via a process which is homologous to ubiquitin conjugation. There are four isoforms of sumo in humans, sumo1, sumo2, sumo3 and sumo4. They are small proteins, around 12kDa in size but Sumo2/3, which show a high degree of similarity, can form polymeric chains, which can be terminated by sumo1, which lacks the internal sumo consensus sites. Sumo4 is used for modification of proteins under stress.

My experiments indicate that the 140kDa band might be sumoylated. Indeed, several SUMO consensus sites were identified using a sumo prediction program. SUMO modification plays an important role in protein localisation for example sumoylated RanGAP interacts with RanBP2, a component of the nuclear pore complex (Pichler

and Melchior, 2002) and sumoylated NEMO, the IKB kinase regulator, localises to the nucleus (Huang et al., 2003). Sumoylation has also been shown to affect transcription factor activity, mostly repressing transcription, through altered interactions with DNA, chromatin or adaptor proteins. Common examples include Sp3 and Elk1 (Ross et al., 2002; Yang and Sharrocks, 2004).

Sumoylation of several substrate proteins has been shown to be regulated by phosphorylation. For example, SUMO modification of p53 and PML was negatively regulated by phosphorylation whereas that of HSF1 was positively regulated by phosphorylation (Muller et al., 2000; Hietakangas et al., 2003). Increased phosphorylation due to activation of the MAPK pathway was found to reduce sumoylation of Elk-1, hence de-repressing its activity (Yang et al., 2003). Hence, there is significant cross-talk between sumoylation and phosphorylation pathways, and this might form an elegant mechanism of control of REST by IGF2 signalling.

It remains to be seen which SUMO isoform is conjugated to REST, whether it creates SUMO polymers and the identity of the lysine residue in question. These questions can be addressed by co-transfecting REST with different tagged SUMO isoforms, immunoprecipitating REST and probing the eluate for the tag. Once the identity of the isoform is determined, mutation analyses of the predicted lysine residues would identify the modified lysine. The shift in mass upon mutating modified lysine would also help determine the number of SUMO monomers attached to REST.

3.3.4 Dimerisation

My characterisation of HMW REST identified phosphorylated, glycosylated and sumoylated species. However, these modifications alone fail to account for the approximately 100 kDa mass change of the 220kDa HMW REST form. Multiple PTMs together might be a possibility, but the 220kDa HMW REST bands are too

precise and consistent to be a result of combined PTMs. Furthermore, the increase in mass to almost twice the size of unmodified 116 kDa REST strongly points towards oligomerisation.

Dimerisation has essential roles in virtually all signalling pathways, from the cell surface to the nucleus. One of the most common dimerisation events is the dimerisation of transcription factors. This allows more DNA-protein contacts to be made, hence increasing sequence specificity. The nuclear hormone receptor family is one such example which binds to bipartite response elements as homo or heterodimers via a DNA-binding domain further consisting of two subdomains which bind zinc (Tsai and O'Malley, 1994).

Tyrosine phosphorylation of STATs by JAKs also results in dimerisation by the interaction of SH2 domains of STATs and phospho-tyrosine residues. The dimerisation of STATs is imperative for DNA binding and might also control nuclear localisation (Shuai et al., 1994). Other common examples include basic helix-loop-helix proteins and the bzip family proteins (Klemm, Schreiber and Crabtree, 1998). Hence there is plenty of evidence of transcription factor dimerisation in order to facilitate DNA binding and nuclear translocation. Since the HMW REST form and exogenous Flag-REST is predominantly nuclear, dimerization might form a mechanism for nuclear translocation. Furthermore, since RE1 sequence consists of two half-sites, REST dimerization might facilitate binding to RE1, each monomer interacting with one half-site.

A common mechanism for oligomerisation is via disulphide linkages. However, REST was found to be resistant to reduction suggesting the absence of any such linkages. Since HMW REST is also resistant to urea, it is proposed that REST might dimerise via an interaction strong enough to resist urea treatment. This might include

a peptide link such as in the case of human superoxide dismutase 1 which dimerises via a ditryptophan cross-link (Medinas et al., 2010). Sumoylation might also mediate a link between two or more REST monomers via two sumo motifs or one sumo and one sumo-interacting motif. Another possibility involves dimerization either directly or indirectly mediated by zinc fingers of REST. RAG1 recombinase dimerises via a domain containing a RING finger and a C2H2 finger, the C2H2 finger indirectly participating in dimerization by providing a stable scaffold (Bellon et al., 1997). On the other hand, Ikaros, a protein expressed in hematopoietic cells serving to silence and activate genes has a C-terminal domain consisting of two C2H2 fingers which are not directly involved in DNA binding but instead self-interact or interact with parallel zinc fingers of other Ikaros family members (Sun, Liu and Georgopoulos, 1996).

While writing up, I came across a PhD thesis, which provides evidence for REST dimerization (Vu, 2009). Myc and Flag-tagged REST was co-transfected in cells, immunoprecipitated using one tag and immunoblotted using the other to show a direct association with other REST molecules. Mutation of zinc finger 1,6 and 9 decreased association but did not completely abrogate it suggesting a zinc finger-independent mechanism for dimerisation. The next step would be to repeat this experiment with our cell lines and reagents to verify these results and then use truncation mutants to identify the dimerization domain and mechanism.

As mentioned before, HMW form of REST is O-glycosylated and removal of glycosylation leads to decrease in the HMW form and a concomitant increase in the LMW form. This might point towards a dimerization mechanism that is mediated by O-glycosylation. There is evidence to show that N-linked glycosylation plays an important role in oligomerisation of proteins, by stabilising the interaction between monomers or providing an interface for dimerization, for example the gonadotropin

hormone hCG and Transferrin Receptor 2 (Ulloa-Aguirre et al., 1999; Zhao and Enns, 2013). There is no such precedence in literature for O-glycosylation mediating dimerisation, but my data indicates the potential of a similar mechanism. Mutation of the O-glycosylation site should abrogate dimerization and validate this hypothesis.

Based on these results, a model can be envisaged to help explain the HMW REST forms (Figure 4.13). The model consists of two REST molecules coming together to form a functional REST dimer, which is further modified by phosphorylation and O-glycosylation. The dimer resides mostly in the nucleus, hence potentially being the functional REST form, where it is unphosphorylated, with phosphorylation acting as a signal for degradation. The 140kDa form might be either a degradation product of the dimer also sumoylated, or it might be the monomeric form, which is sumoylated and competes with the dimer form to alter transcriptional regulation. Further biochemical analyses including truncation and point mutations, as well as structural studies are required to fully test this model.

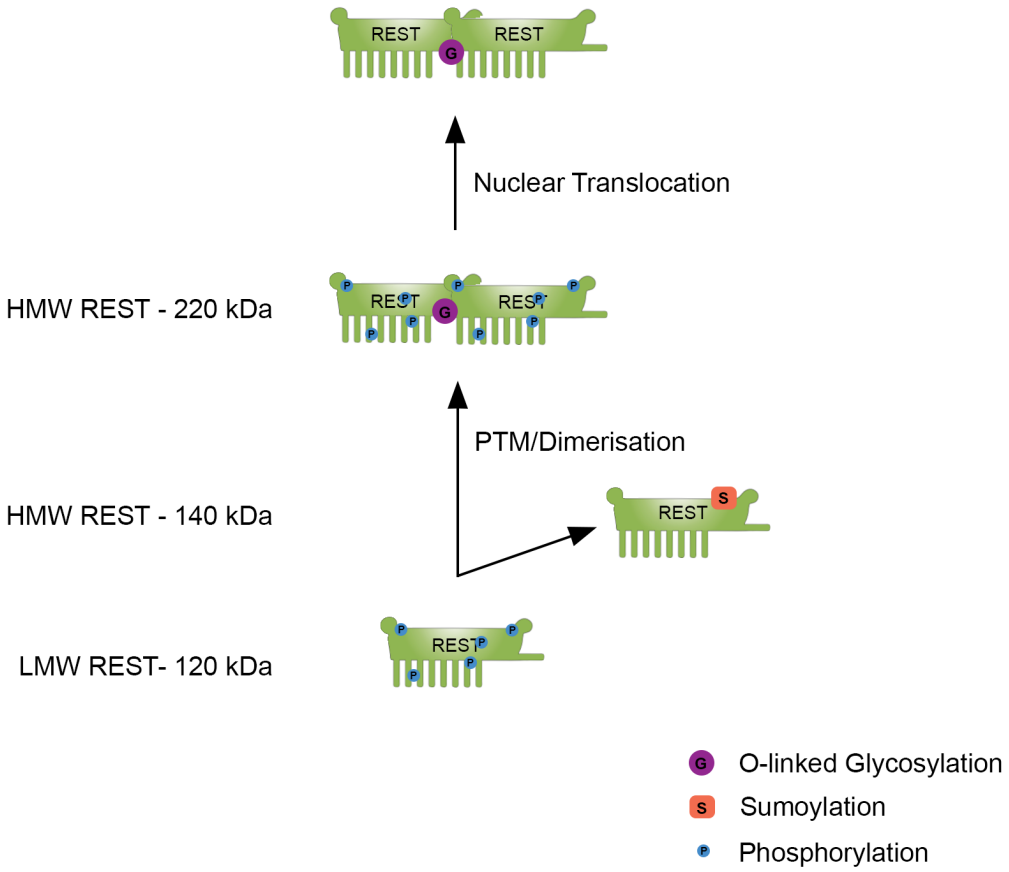


Figure 3.12 Model showing LMW and HMW REST with their post-translational modifications and functional consequence.

4 IGF2 ALTERS THE POST-TRANSLATIONAL MODIFICATION OF REST

4.1 Introduction

Previous work in the lab was undertaken to elucidate the key transcriptional hub genes that define the output of IGF2 signalling. As discussed in Chapter 1, this work resulted in the identification of transcriptional network with a small number of highly active hub genes, REST being the most prominent one. REST was identified as a negative regulator of neuronal differentiation and plays a crucial role in ensuring repression of neuronal genes in non-neuronal tissue (Jones and Meech, 1999). Additionally, it is implicated in regulating neuronal differentiation in NSCs (Ballas et al., 2005) and in the maintenance of ESCs by being a part of the core Oct4/Sox2/Nanog transcriptional network (Johnson et al., 2008). Furthermore, REST has an emerging role in cancer. It was first identified as a tumour suppressor in epithelial cells from an RNAi-based genetic screen and was found to be frequently deleted in colorectal cancer cell lines and primary tumour samples. REST has also

been shown to possess oncogenic properties specifically in neuronal tumours (Conti et al., 2012; Fuller et al., 2005). Bearing in mind IGF2's role in neurogenesis and tumorigenesis, REST appeared as a candidate that may shape IGF2 downstream signalling.

Induction of IGF2 did not affect REST mRNA, but REST protein stability was found to increase after 6 hours of incubation of NIH3T3 cells with increasing doses of IGF2, showing that regulation of REST is achieved post-translationally (Figure 1.11). This suggests a post-translational modification (PTM) on REST and is consistent with the post-translational regulation of REST by ubiquitination and proteasomal degradation. The IGF2-dependent REST stability was abrogated upon treatment of cells with the PI3K inhibitor, LY294002, and rescued by treatment with MG-132, a proteasomal inhibitor, thus implicating the involvement of the IGF2 signalling pathway, via PI3K, in stabilising REST.

As discussed in Chapter 1, REST appears to be under tight control in different neuronal and non-neuronal tissues. Numerous mechanisms that regulate REST activity have been identified, the most prominent of which seems to be mediated by ubiquitination and deubiquitination. Westbrook et al. showed that the E3 ubiquitin ligase, β-TrCP, recognises a conserved phosphodegron (S1024, S1027, S1030) at the C-terminus of REST, binds and ubiquitinates it, targeting REST for proteasomal degradation (Westbrook et al., 2008). Degron-mutant REST fails to bind β-TrCP leading to increased REST stability (Westbrook et al., 2008). A second adjacent phospho-degron (E1009, S1013) was identified by Guardavaccaro et al., that also binds β-TrCP and is involved in decreased REST levels during the G2 phase of cell cycle (Guardavaccaro et al., 2008a). Conversely, a deubiquitinase, HAUSP (herpesvirus-associated ubiquitin-specific protease) acts to counterbalance β-TrCP

activity via a consensus site on REST (310-PYSS-313) (Huang et al., 2011). HAUSP knockdown in neuronal progenitor cells decreases REST levels and results in neuronal differentiation (Huang et al., 2011). Together, β-TrCP and HAUSP form a very efficient post-translational mechanism.

There is evidence from previous studies to support the regulation of REST by PI3K/MAPK signalling. REST was reported to be upregulated by fetal calf serum as well as IGF-1 in neuroblastoma cells in a time-dependent manner (Di Toro, Baiula and Spampinato, 2005). This is accompanied by a repression of neuronal differentiation genes and maintenance of cells in an undifferentiated and proliferative state. Prolonged exposure of IGF1, however, decreases REST expression and de-represses neuronal genes (Di Toro, Baiula and Spampinato, 2005; Bedini, Baiula and Spampinato, 2008). Similarly, REST expression in HeLa cells was reported to be elevated in response to EGF, accompanied with a decrease in cell apoptosis (Baiula et al., 2012). Furthermore, Nesti et al. discovered a direct mechanism of REST regulation by EGF-Ras-ERK signalling via phosphorylation of S861 and S864 of REST by ERK2 (Nesti et al., 2014). Phosphorylation was shown to increase peptidyl prolyl cis/trans isomerase (Pin1) binding, in turn stimulating βTrCP-mediated proteasomal degradation of REST.

Conversely, a REST regulation loop on AKT has also been observed in other experiments. The tumour suppressor function in epithelial cells was found to be partly related to attenuation of P13K/AKT signalling (Westbrook et al., 2008). Knockdown of REST in mammary epithelial cells was shown to dysregulate P13K/AKT signalling. Similarly in SCLC lines, REST loss of function was shown to lead to EGF-mediated increased AKT2 phosphorylation, essential for cell proliferation and survival (Kreisler et al., 2010). These opposing observations either indicate a

feedback loop of PI3K/MAPK control of REST to maintain REST levels in the cell or indicate potential cell-line and context specific mechanisms.

4.2 Aims

Led by background preliminary data, I hypothesised that IGF2 signalling may induce a post-translational modification of REST, most likely via phosphorylation. This PTM is proposed to protect REST from ubiquitination and degradation in the proteasome.

The aims of this chapter are:

1. To characterise IGF2 induced REST stability in NIH3T3 cells.
2. To characterise the variability in REST response to IGF2 in different cell lines, using Ewing cell lines as an example.
3. To investigate whether the IGF2 effect on REST is via a post-translational modification.
4. To explore the effects of IGF2 on REST turnover in the cell.
5. To investigate if IGF2 modifies REST translocation into the nucleus.

4.3 Results

4.3.1 IGF2 stabilises REST in NIH3T3 cells in a time dependent manner

To evaluate whether IGF2 stabilises REST, NIH3T3 cells were serum starved and treated with IGF2 (100ng/ml), and REST levels assessed at different time points ranging from 0.5 to 12 hours (Figure 4.1). REST was quantified using immunoblotting and correlated to pAKT levels. It was observed that pAKT levels peak at approximately 1 hour after addition of IGF2 and then start falling rapidly to slightly below levels of unstarved cells. These observations are consistent with signalling autoregulation. REST levels were observed to increase steadily upon IGF2 addition, and peak at 6 hours after which they start falling, also potentially demonstrating a feedback loop. Therefore, REST levels appear to replicate effects of IGF2 on AKT phosphorylation with a lag of 5 hours and so treatment time is crucial for inferring any upregulation of REST. Moreover, pAKT and REST levels, although appear to correlate, do not imply direct causality, as there may be other mechanisms that result in REST feedback regulation.

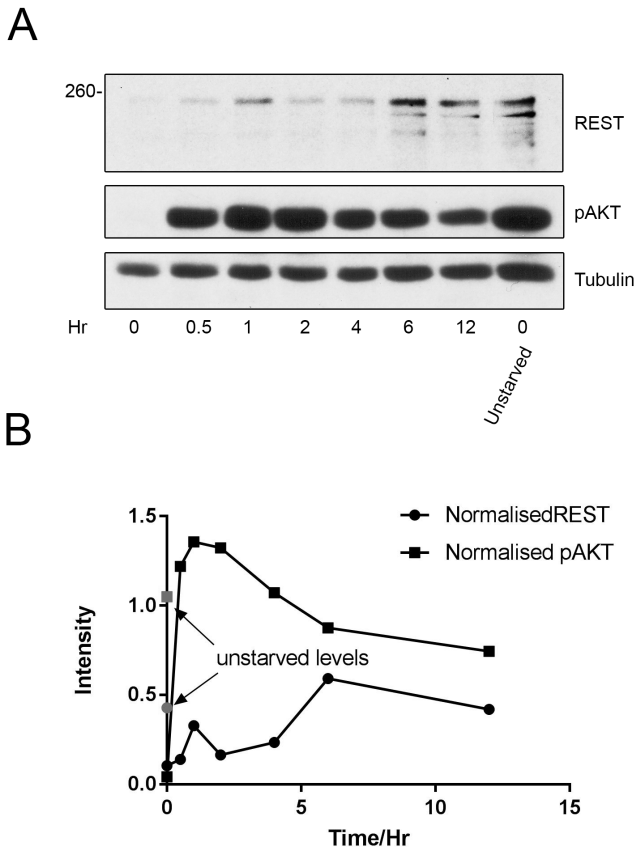


Figure 4.1 IGF2-dependent REST stabilisation in NIH3T3 cells is time-dependent.
A) NIH3T3 cells were serum-starved, washed and treated with IGF2 (100ng/ml). Cells were lysed at the indicated time points and immunoblotted for REST using anti-REST (Mandel), anti-pAKT and anti-α-tubulin. Representative Western blot from three independent experiments is shown. B) REST levels from (A) were quantified by Image J software, normalised to α-tubulin levels and plotted against time. Unstarved levels of pAKT and REST are indicated by arrows.

4.3.2 IGF2 response in Ewing Cell lines is time and IGF2 dose dependent

A concern in earlier results obtained (Appendix 1-6) was the variability of REST response in different cell lines. Once the IGF2 exposure time dependency of REST stability was established in NIH3T3, I further explored this effect, including downregulation of REST, observed in some cell lines (Appendix 5). REST expression in SKNMC and CHP100 was found to decrease upon IGF2 treatment as shown by the polyclonal 17-641 antibody, and the converse was observed in A673 cells. Hence, I chose these Ewing sarcoma cell lines to validate this observation further. SKNMC cells were serum-starved and treated with IGF2 for 1-6 hours (Figure. 4.2A). To evaluate the IGF2 feedback loop, a lower dose of IGF2 (10ng/ml) was tested. REST levels were found to increase with time, peaking at 6 hours. This increase in levels was considerably higher than that observed with a 100ng/ml IGF2 dose at 6 hours. Hence, not only the exposure time was found to be crucial for modifying REST levels, but also the IGF2 dose was seen to have a significant effect.

An IGF2 dosage experiment was performed on all three Ewing cell lines, keeping the time point fixed at 6 hours (Figure 4.2B). REST levels were found to peak in SKNMC cells at 10 ng/ml IGF2. Increasing IGF2 dose to 100ng/ml had no additional effect. Inhibition of the proteasome using MG-132 combined with stimulation of the IGF2 pathway (100 ng/ml IGF2) further increases REST levels, correlated with AKT phosphorylation. For CHP100, REST levels were stable but start to drop at 1 ng/ml IGF2. This decline in REST levels was rescued by MG-132, suggesting that the decrease was proteasome dependent. AKT phosphorylation in response to IGF2 was significantly lower in CHP100, perhaps suggesting deregulation of the pathway in the cell line, which might explain this opposing effect. In A673 cells, REST levels also increase with 10 ng/ml IGF2, the increase being less than that observed in SKNMC cells. At 100ng/ml, REST levels decreased to pre-treatment levels, the decrease being

rescued by proteasome inhibition. The levels of AKT phosphorylation (S473) in A673 are intermediate to the other two cell lines, correlating with the intermediate REST response.

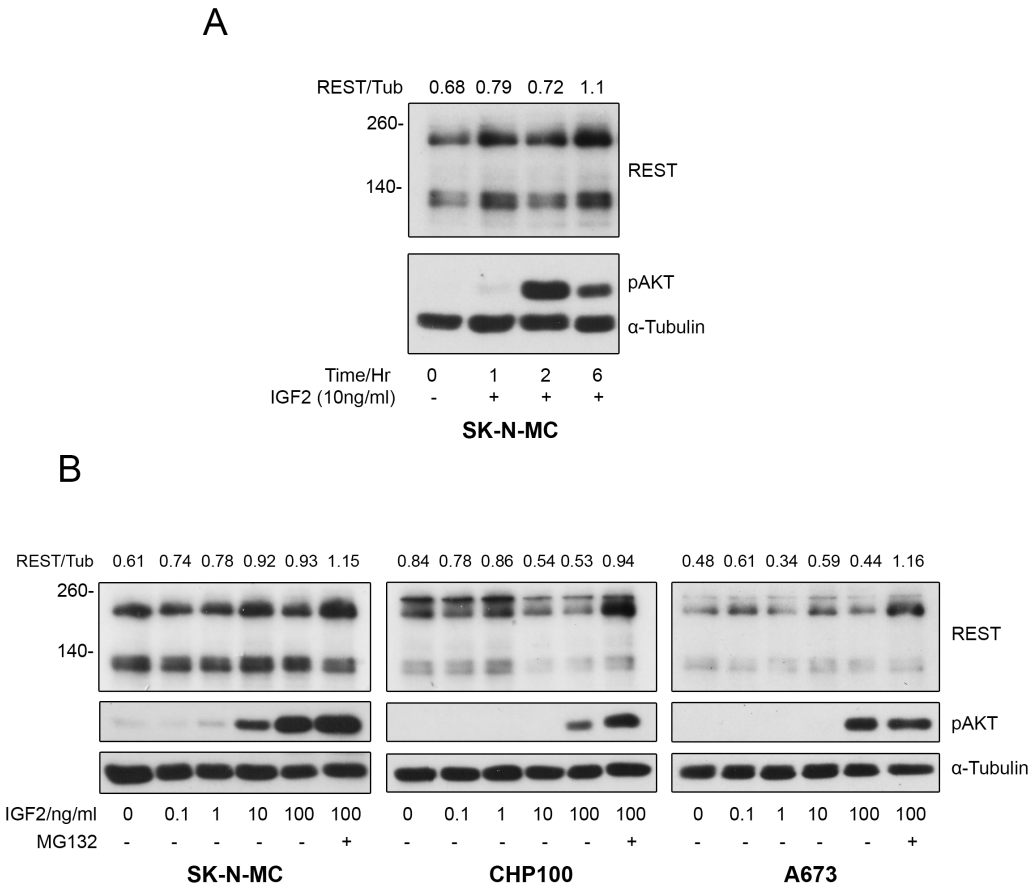


Figure 4.2 IGF2-dependent REST stabilisation in Ewing Sarcoma cell lines is dose and time-dependent.

A) SK-N-MC cells were serum-starved, washed and treated with IGF2 (10ng/ml). Cells were lysed at the indicated time points and immunoblotted for REST using anti-REST (Mandel), anti-pAKT and anti- α-tubulin. Representative Western blot from three independent experiments is shown. B) SK-N-MC, CHP100 and A673 cells were serum-starved, washed and treated with increasing doses of IGF2 and 0.75 μM MG-132 for 6 hours. Cell lysates were immunoblotted for REST using anti-REST (Mandel), anti-pAKT and anti- α-tubulin. Representative Western blot from two independent experiments is shown. REST levels, in the represented Western blot, were normalised to α-tubulin (quantified by Image J software) and indicated on the blots.

4.3.3 IGF2 alters phosphorylation of REST

IGF2 pathway like most other signalling pathways works predominantly via phosphorylation of intracellular proteins to transmit signals (Kennedy et al., 1997). As shown in Chapter 4, REST is phosphorylated in HEK293T cells. Furthermore, there have been two phosphodegrons identified in the C-terminal region of REST (Westbrook et al., 2008; Guardavaccaro et al., 2008b). Hence, it was proposed that REST stability in response to IGF2 might be mediated via phosphorylation of REST. To explore this, HEK293T cells were serum starved for 24 hours, treated with IGF2 and MG-132 for 6 hours and the effect on the 220kDa HMW form was noted by immunoblotting (Figure 4.3). It was observed that the top band, which is phosphorylated and responds to λ-Phosphatase treatment, was unresponsive to IGF2. However, levels of the unphosphorylated REST form were found to increase with IGF2 treatment. Both forms increase upon MG-132 treatment, with the unphosphorylated form being significantly more responsive. These data suggest that IGF2 selectively stabilises the dephosphorylated HMW form of REST.

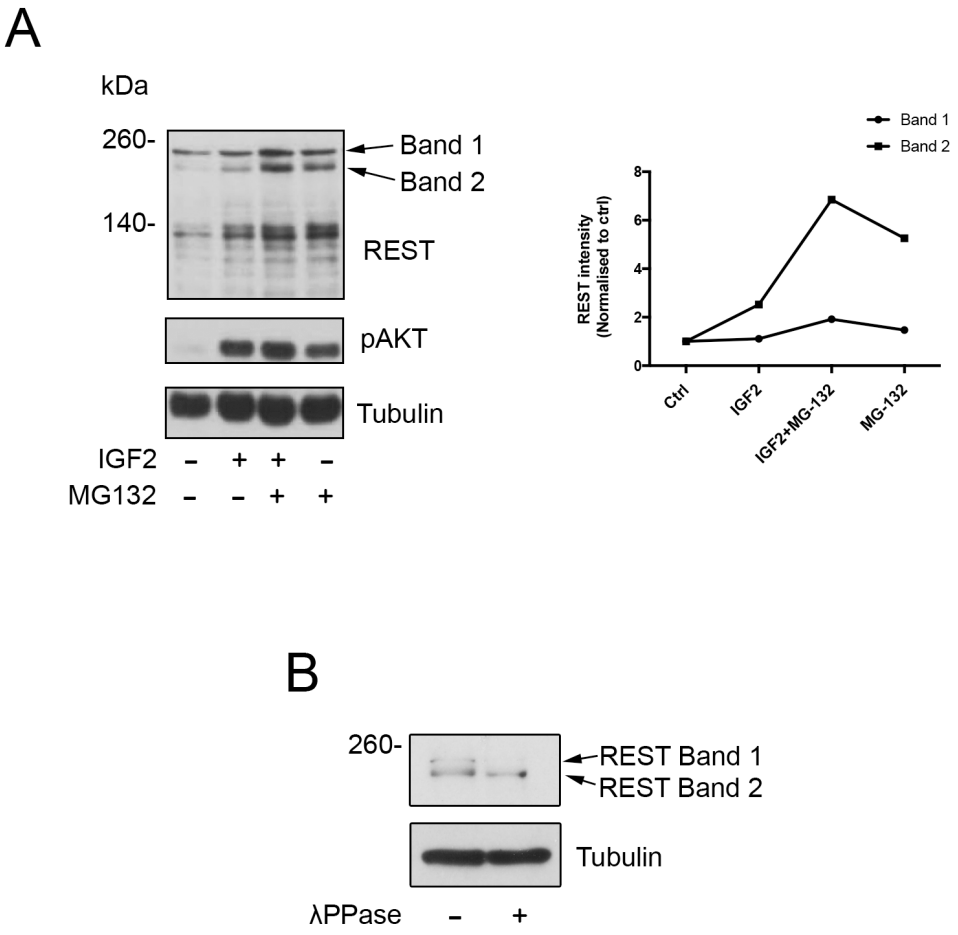


Figure 4.3 Modification of the HMW REST form by IGF2
A) HEK293T cells were serum-starved for 24 hours, washed and treated with 100ng/ml IGF2 and 0.75 μM MG-132 for 6 hours. Whole cell lysates were immunoblotted for REST (Mandel antibody), pAKT and α-tubulin. Representative Western blot from three independent experiments is shown. REST levels were quantified by Image J software, normalised to α-tubulin levels and the control and plotted. B) HEK293T cell lysate was treated with λ-Phosphatase at 30^{o}c for 1hr. Untreated and treated lysates were immunoblotted for REST (Mandel Antibody).

4.3.4 IGF2 increases REST half-life.

Next, I wanted to see whether IGF2 influences the half-life of REST. To characterise REST stability, a cycloheximide chase assay was performed. Cycloheximide inhibits eukaryotic protein synthesis by preventing translational elongation and allows estimation of protein turnover (Schneider-Poetsch et al., 2010). HEK293T cells were treated with 10 μg/ml of cycloheximide for different time points, and the cell lysates obtained immunoblotted for REST (Figure 4.4). The two MW forms of REST, both phosphorylated and unphosphorylated, were quantified separately, and the half-lives of both bands were estimated. The phosphorylated form of REST was found to be relatively stable with a half-life of slightly more than 9 hours. The unphosphorylated form was significantly labile, with a half-life of ~2.5 hours.

To determine the effect of IGF2 on the half-life of REST, HEK293T cells were then treated with either cycloheximide alone or cycloheximide in the presence of IGF2, and cells were lysed at different time points (Figure 4.5). It was observed that IGF2 increases the half-life of the unphosphorylated REST HMW form specifically, to more than 9 hours. The phosphorylated form of REST remains virtually unchanged. There was a peak at 6 hours of cycloheximide and IGF2 treatment, which is consistent with a maximal IGF2 effect on REST observed at 6 hours. This suggests that IGF2 stabilises the unphosphorylated band and prevents its turnover, the turnover being partially proteasome-dependent.

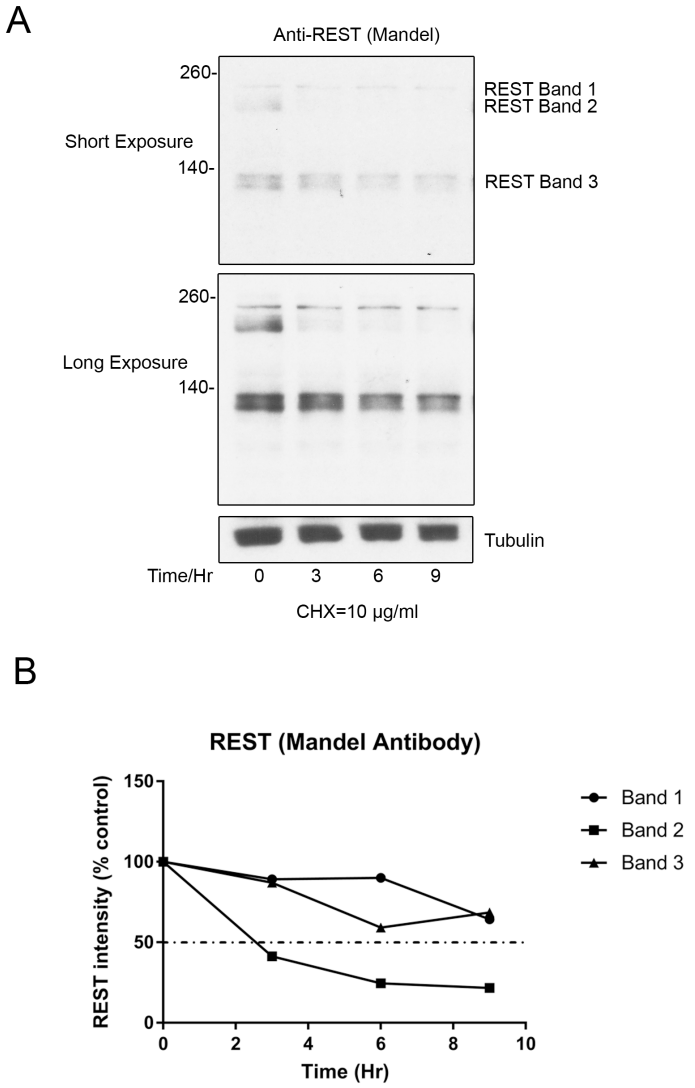


Figure 4.4 Half-life of REST.
A: HEK293T cells were treated with 10 µg/ml Cycloheximide for the indicated time points. Whole cell lysates prepared were immunoblotted for REST (Mandel antibody) and α-tubulin. Representative Western blot from three independent experiments is shown. B: Image J was used to quantify the indicated REST bands in A. Normalised REST levels, represented as a percentage of control (0 hr) are plotted against time.

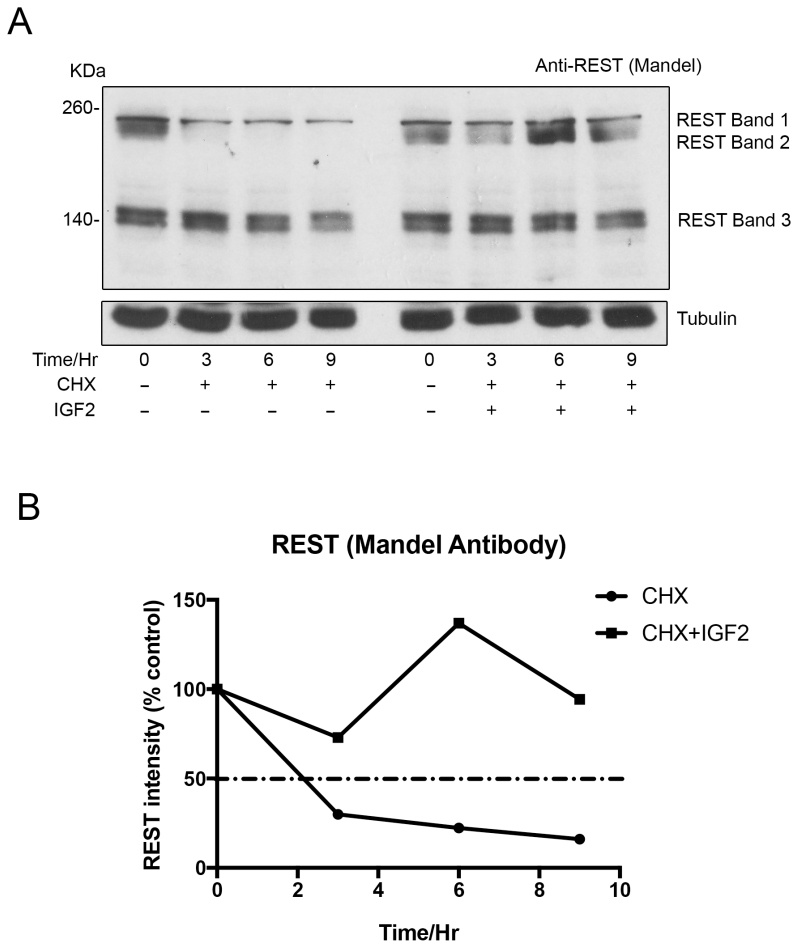


Figure 4.5 Effect of IGF2 on REST half-life.
A) HEK293T cells were treated with 10μg/ml Cycloheximide and/or IGF2 (100 ng/ml) for the indicated time points. Whole cell lysates prepared were immunoblotted for REST (Mandel antibody) and α-tubulin. Representative Western blot from three independent experiments is shown. B) Image J was used to quantify REST bands indicated in A. Normalised REST levels (Band 1+2), represented as a percentage of control (0 hr) are plotted against time.

4.3.5 IGF2 increases the amount of nuclear REST

`As REST is a transcription factor, functional REST must reside inside the nucleus, bound to DNA in complex (Ballas et al., 2005). Any change in REST level may cause a change in functional levels of REST in the nucleus. I wanted to find out if the IGF2-dependent increase in REST levels translated into an increase in REST nuclear levels. HEK293T cells were serum starved, treated with IGF2, MG-132 and the dual IR/IGF1R kinase inhibitor OSI-906. Cytoplasmic and nuclear extracts were prepared from cells. REST immunoblotting on these lysates shows that IGF2 treatment leads to an increase in both cytoplasmic and nuclear levels of REST (Figure 4.6). The increase in cytoplasmic REST is far more than that in nuclear REST, but when the proteasome is blocked together with IGF2 treatment, it leads to a significant increase in nuclear REST. Inhibition of the proteasome only increases cytoplasmic REST. The effect of OSI-906 on HEK293T cells is an increase in REST. This effect is ambiguous as only one dose and time point was used. Hence, the possibility of a feedback effect cannot be ignored, and may explain the effect. In addition, it was also observed that only the lower REST form migrated inside the nucleus. This indicates that dephosphorylated REST is found in the nucleus. It is plausible that IGF2 increases global levels of REST, causing an increase in the nuclear translocation of dephosphorylated REST.

Immunofluorescence was used to verify this hypothesis further (Figure 4.7). HEK293T cells, that were serum starved and treated with IGF2 and MG-132, were prepared for immunofluorescence using the REST (Mandel) antibody. REST was found to be predominantly nuclear in localisation. There was a visible increase in REST signal using the same microscope settings upon IGF2 and MG-132 treatment. Quantification of images also showed an increase in the REST protein localised in the

nucleus. This suggests that IGF2 increases global REST levels, leading to a secondary increase in nuclear levels of REST.

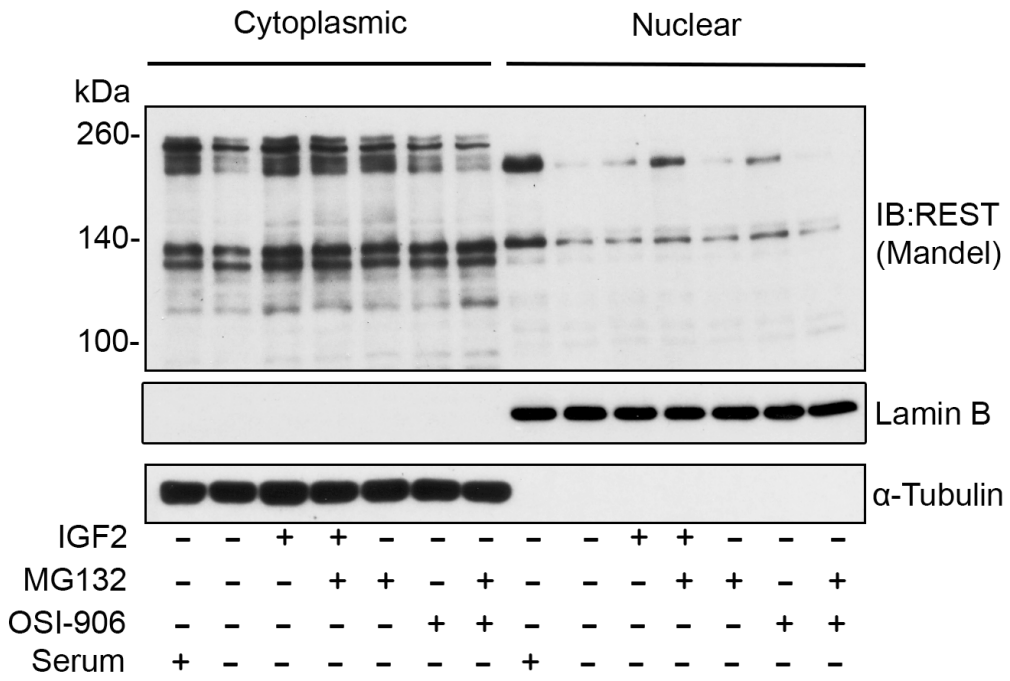


Figure 4.6 Effect of IGF2 on the localisation of REST.
HEK293T cells were serum-starved and treated with IGF2 and MG-132 for 6 hours. Nuclear and cytoplasmic extracts were prepared and immunoblotted for REST (Mandel Antibody), α-tubulin and lamin B. Representative Western blot from two independent experiments is shown.

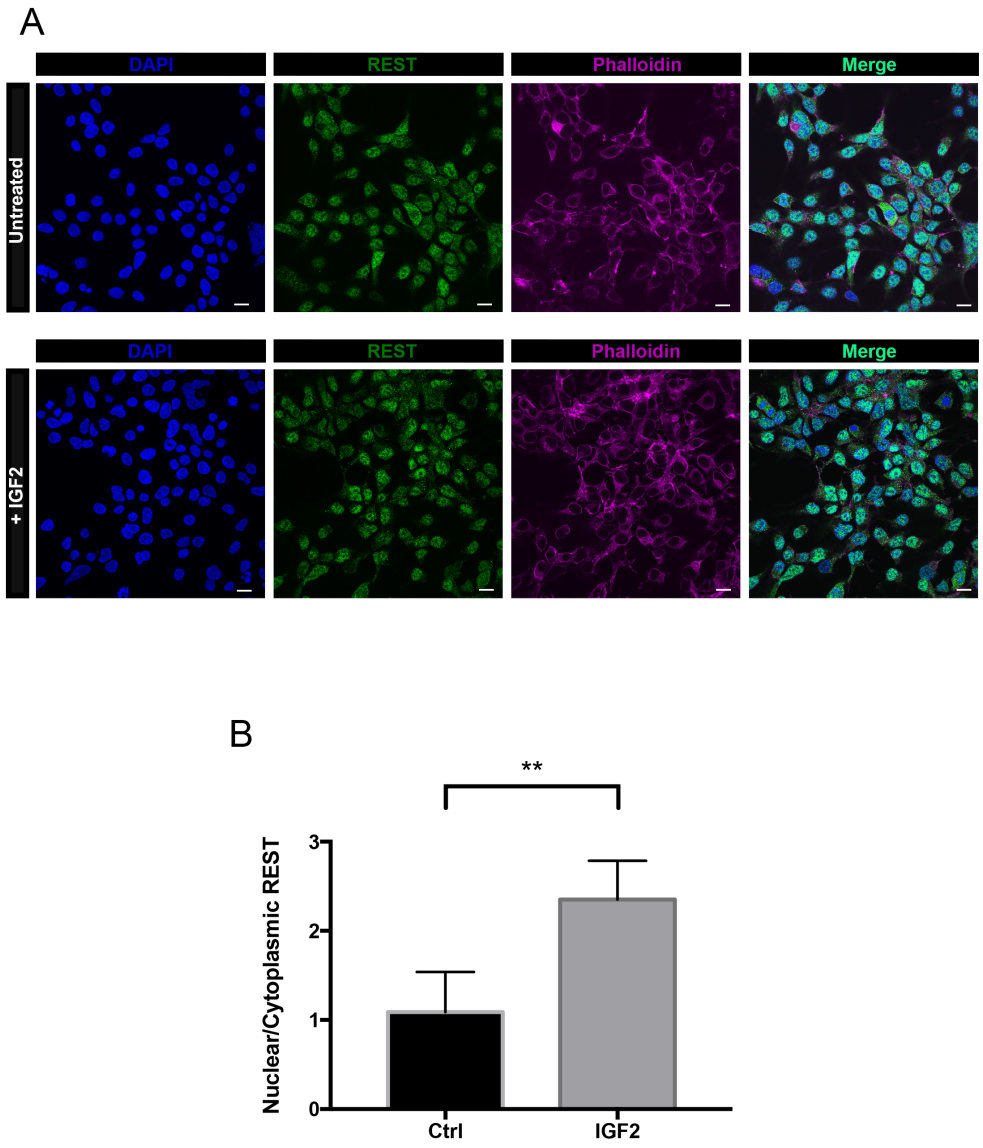


Figure 4.7 Effect of IGF2 on the localisation of REST.
HEK293T cells were plated on coverslips, serum-starved and treated with IGF2 (100 ng/ml) for 6 hours. Cells were subsequently immunolabelled for REST using Mandel antibody, DAPI stain for nuclei and Phalloidin for cytoplasm and imaged using an Olympus confocal microscope. B) Nuclear and cytoplasmic REST levels were quantified using Image J. 5 fields of three separate samples, and two biological repeats were used and the average nuclear/cytoplasmic intensity was plotted. $P<0.005$ by unpaired student's t test. Scale bar = 10 µm.

4.4 Discussion

The biological activity of REST is regulated by post-translational modifications. REST is dynamically turned over inside cells, and that provides a mechanism for efficient and immediate response to extracellular signals. Previous studies have demonstrated a dynamic turnover model, which is phosphorylation-dependent, the phosphorylation targeting REST to the proteasome (Westbrook et al., 2008). Conversely, deubiquitination by the deubiquitinase HAUSP protects REST from degradation (Huang et al., 2011). Another deubiquitinase, USP15, acts to stabilise newly synthesised REST, a process that aids in the recovery of REST levels after mitosis (Faronato et al., 2013). The data presented in this study complements that model by adding in a nuclear versus cytoplasmic component, and linking REST stability to IGF2 signalling.

In this chapter, it has been shown that IGF2 selectively stabilises the dephosphorylated REST protein. The phosphorylated levels of REST remain unchanged upon IGF2 treatment. Inhibition of the proteasome, however, increases both forms of REST, suggesting that both REST forms are targeted to the proteasome. However, the turnover of the dephosphorylated REST protein is much higher than that of the phosphorylated protein, and it is this highly turned-over form that is stabilised by IGF2. The lower band also resides in the nucleus and IGF2, in combination with proteasomal inhibition, dramatically increases the amount of REST protein in the nucleus. This suggests a model of IGF2 stabilisation of REST, which is phosphorylation and localisation-dependent. According to this model, cytoplasmic REST may exist in two pools, a phosphorylated and a dephosphorylated form. IGF2 increases the dephosphorylated form by either removing the phosphorylation or stabilising the dephosphorylated form by association with stabilising factors, possibly

HAUSP. The dephosphorylated pool is translocated to the nucleus where it acts to repress neuronal genes.

Dephosphorylated REST being the functional REST form in the nucleus is consistent with previous studies which show that phosphorylation acts as a means of inducing ubiquitination and protein degradation (Westbrook et al., 2008; Guardavaccaro et al., 2008b; Lenk and Sommer, 2000). Further verification of the model would involve, treating nuclear protein with lambda phosphatase to confirm if nuclear REST is indeed dephosphorylated. It is unclear whether IGF2 leads to either dephosphorylation of REST protein leading to increase in REST levels or induces its stability by acting on another target.

The different turnover rates of the two REST forms in the cell, the dephosphorylated form being faster turnover is contrary to previous studies, which show phosphorylation to be required for βTrCP binding and ubiquitination of REST. Since the dephosphorylated form is also in the nucleus, this would suggest a very efficient and constant mechanism of nuclear export. Although the nuclear localisation signals have been identified for REST, there is no evidence in the literature for a nuclear export signal (NES). A precursory search for NES in REST using online prediction program called LocNES resulted in the identification of a putative leucine-rich NES (Figure 4.8). Mutation analysis would confirm if this motif is functional. On the other hand, if the dephosphorylated and phosphorylated REST is indeed partitioned between the nucleus and cytoplasm with different turnover rates, this can imply a different mechanism of degradation of the nuclear protein, possibly nuclear proteolysis. The faster turnover might be because of nuclear proteasomes having better access to nuclear REST.

Many transcription factors proteins have been shown to be degraded by nuclear proteolysis. These include Matα2, a transcription factor which only degrades upon nuclear import, MyoD, as well as p53 which degrades in nuclear proteasomes in the case of downregulated stress response resulting from DNA damage (Shirangi, Zaika and Moll, 2002; Floyd et al., 2001; Lenk and Sommer, 2000). To investigate if REST is indeed degraded by nuclear proteasome, inhibition of nuclear export by leptomycin B should not prevent degradation of nuclear REST protein if nuclear degradation also occurs.

This chapter also helps to explain the original findings and variability in cell line data (Appendix 2-3). The IGF2-dependent REST stability was found to be reliant on treatment time as well as IGF2 dose. This indicates the presence of a feedback loop. Indeed it has been shown that IGF2 induces PTEN expression via PI3K/AKT pathway, which in turn acts to reduce AKT phosphorylation, modulating the physiological effects of IGF2 (Moorehead et al., 2003). Another feedback loop operates through mTOR and S6K1 activation, leading to feedback inhibition of the pathway through IRS-1 phosphorylation and inhibition (Harrington et al., 2004). Interestingly, REST was also shown to suppress PI3K signalling in HMECs expressing hTERT and SV40 LT (TLM-HMECs) (Westbrook et al., 2005). This might point towards a feedback loop involving REST moderating its own expression levels.

The findings in this chapter and possible mechanisms for REST stability and turnover are summarised in Figure 4.9. Due to the dynamic nature of REST turnover in the cell, pulse-chase experiments with live imaging would be the approach forward. In this way, it would be possible to pull apart the mechanistic details with the effects of

different pathway inhibitors, IGF2, and inhibition of proteasome clearly visualised on REST.

	102	116
Cow	EPSGLENMELESLELSVV	
Dog	EPSGLENMELESLELSVV	
Monkey	EPHGLENMELRSLELSVV	
Rat	DPSGLDNMELRSLELSVV	
Mouse	DLKGLENMELGSLELSAV	
Human	EPHGLENMELRSLELSVV	

Figure 4.8 Putative Nuclear export signal in REST.
LocNES was used to predict NES in REST sequence. The sequence conservation of the highest scoring NES is shown in Human, Mouse, Rat, Monkey, Dog and Cow.

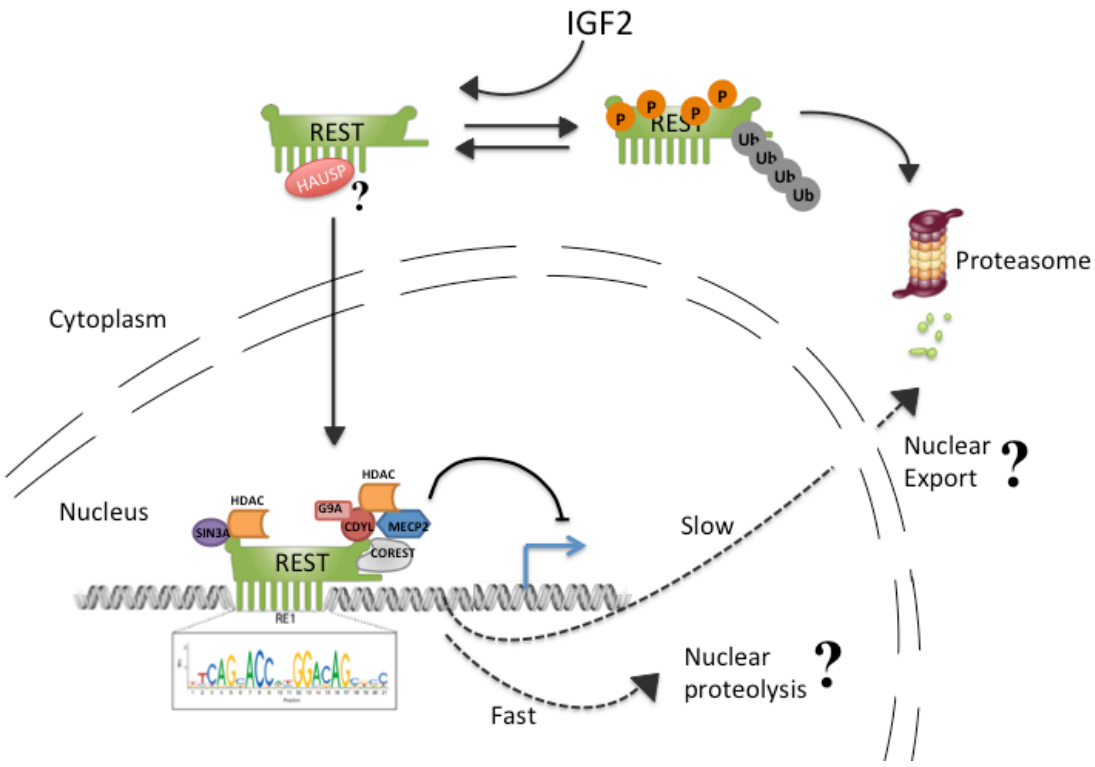


Figure 4.9 Model for IGF2 stabilisation of REST.
IGF2 acts to increase the dephosphorylated pool of REST. Dephosphorylated REST enters the nucleus and carries out silencing of neuronal genes. Phosphorylated REST is degraded by the proteasome in the cytoplasm. An additional possibility of nuclear degradation is indicated by a question mark.

5 CHARACTERISATION OF PHOSPHORYLATION SITES IN HMW REST

5.1 Introduction

5.1.1 Cellular Protein Phosphorylation

Protein phosphorylation is a reversible PTM controlling many cellular processes in all living cells (Humphrey, James and Mann, 2015). In eukaryotes, phosphorylation occurs at serine, threonine and tyrosine residues of proteins resulting from the opposing action of kinases and phosphatases. Histidine residues are also phosphorylated but they are known to be more labile, and hence, not readily detected by conventional proteomics. The discovery of phosphorylated proteins can be traced back to 1955, when the first phosphorylated protein, glycogen phosphorylase (GP), was discovered, the phosphorylation converting GP from inactive glycogen phosphorylase b to the active glycogen phosphorylase a (Fischer and Krebs, 1955; Krebs and Fischer, 1955). This was closely followed by the discovery of glycogen synthase phosphorylation and phosphorylation was relegated to control of glycogen

metabolism. Phosphorylation was shown to be a widespread phenomenon in 1968 with the isolation and purification of protein kinase A (PKA) (Walsh, Perkins and Krebs, 1968). This opened the doors to a new era in the research of protein regulation and paved the way to identification of hundreds of phosphorylated proteins and kinases/phosphatases, with far reaching implications in disease pathology and therapy (Johnson, 2009). Almost half of all human proteins are phosphorylated at any one point in time and thousands of phosphorylation sites exist (Olsen et al., 2006). Protein phosphorylation impacts all major biological processes including transcriptional and translational regulation, metabolism, protein degradation, cellular signalling and communication, cell survival, proliferation and apoptosis (Johnson, 2009). Moreover, dysregulation of protein phosphorylation is linked to pathogenic mechanisms in various physiological systems ranging from the cardiovascular system, to neurological systems, immune response and musculoskeletal regulation as well as cancer (Fabbro, Cowan-Jacob and Moebitz, 2015; Engelman, 2009).

Protein kinases are the enzymes that are responsible for protein phosphorylation. The whole human kinome was identified in 2002 following completion of the human genome sequencing project (Manning et al., 2002). A total of 518 protein kinases were identified and assigned to seven major eukaryotic protein kinase (ePK) groups based on their target sequences and an atypical protein kinase (aPK) group lacking sequence similarity to the ePK domains. Protein phosphatases evolved parallel to kinases and have been classified into three structurally and functionally diverse gene families. These include the phosphoprotein phosphatase (PPP) family and metallo-dependent protein phosphatase (PPM) family, which dephosphorylate serine and threonine residues, and the protein tyrosine phosphatase (PTP) family, a subfamily of which can dephosphorylate all residues (Moorhead et al., 2009). The catalytic

domains are highly conserved within each family and diversity of function is provided by regulatory domains. The dynamic interplay between these two enzymes regulates phosphorylation and dephosphorylation, both temporally and spatially, on specific target sites.

Phosphorylation of a target protein can alter protein function in multiple ways (Johnson and Lewis, 2001). Addition of a phosphoryl group, which is dianionic at physiological pH, can mediate strong interactions by hydrogen bonding with arginine side chains, or medium strength interactions with main-chain nitrogens situated at the beginning of α-helices. Other polar residues like lysine, histidine, tyrosine, serine and threonine can also mediate contacts. These strong interactions can stabilise a conformational change, thus activating enzyme activity, found to be essential for many protein kinases activated by upstream kinase activity. The converse is also true; the phosphate group can sterically block conformational changes, inhibiting enzyme activity. Other roles for phosphorylation include creation of recognition sites for other proteins, promoting conformational changes leading to protein association or disassociation, regulating cellular localization and stability of target proteins (Johnson, 2009). Owing to its transient and reversible nature, phosphorylation by kinases and phosphatases, serves as a rapid molecular switch, which can alter the behaviour of its target proteins directly or indirectly, in response to cellular or external stimuli (Humphrey, James and Mann, 2015). Deep phosphoproteomics studies have revealed that proteins can be phosphorylated on at least five different sites (Olsen et al., 2010; Sharma et al., 2014). These multiple residues are often targets of several different kinases; different kinases might even phosphorylate the same site depending on the cellular context (Humphrey, James and Mann, 2015). Phosphorylation on distinct sites might have shared or distinct effects on proteins function. This allows

proteins to adapt to multiple roles in the cells, depending on the phosphorylation site/sites.

5.1.2 Insulin-like Growth factor 2 signalling

The insulin-like growth factor (IGF) signalling axis is crucial for human growth and development. It is a classic example of a signalling pathway, which employs phosphorylation to mediate signals and regulate downstream proteins.

As mentioned in Chapter 1, most of the biological consequences of IGF2 signalling are mediated through IGF1R. IGF2 binding to IGF1R activates it through autophosphorylation, leading to recruitment and subsequent phosphorylation of insulin receptor substrates (IRS) 1-4 (Vincent and Feldman, 2002). Phosphorylated IRS binds and activates Phosphatidylinositol 3-kinase (PI3-K) which in turn activates the serine/threonine kinase AKT/PKB (Zha and Lackner, 2010). Active AKT has numerous phosphorylation-mediated effects on the cell, including inhibition of apoptosis and activation of protein synthesis, resulting in proliferation and prevention of cell death (Kennedy et al., 1997). Ras/MAPK signalling constitutes the other parallel leg of the pathway via recruitment of different adapter proteins, SHC and growth factor receptor-bound protein 2 (Grb2) which form a complex with son-of-sevenless (Zha and Lackner, 2010). The complex activates Ras, a GTPase, which in turn leads to activation of Raf, triggering a kinase cascade, resulting in activation of mitogen-activated protein kinases (MAPKs), ERK1 and ERK2. ERK1/2 subsequently phosphorylate and activate several target genes including transcription factors, for example ELK1, triggering gene expression and cell proliferation (Zha and Lackner, 2010; Pouyssegur, Volmat and Lenormand, 2002).

Actions of the PI3K/AKT and Ras/MAPK pathways are opposed by phosphatases, notable the tumour suppressor phosphatase and tensin analogue (PTEN). PTEN

dephosphorylates both PI3K as well as SHC and mTOR, hence inhibiting both legs of the pathway (Soroceanu et al., 2007). PTEN expression has been shown to be increased by IGF2 signalling, forming a feedback loop, which serves to keep the signalling pathway in check (Moorehead et al., 2003).

5.1.3 Phosphorylation of REST

Previous studies have identified two phosphodegrons in REST, which regulate REST stability. Westbrook et al identified a conserved phosphodegron (S1024, S1027, S1030) at the C-terminus of REST and Guardavaccaro et al identified a second adjacent degron (S1013, S1014) (Westbrook et al., 2008; Guardavaccaro et al., 2008). Phosphorylation of these sites recruits the E3 ubiquitin ligase, β-TrCP which then targets it for proteasomal degradation. The first degron was suggested to regulate degradation of REST during neural differentiation, and the second was demonstrated to be involved in cell cycle–dependent REST degradation during G2.

Casein Kinase 1 (CK1) was identified as an upstream effector of both the C-terminal degrons (Kaneko et al., 2014). CK1 phosphorylates both degrons, the phosphorylation being crucial for β-TrCP association and proteasomal degradation of REST. CK1 levels can be modulated by global ischaemia in rats *in vivo*, resulting in increase in REST levels in hippocampal CA1 neurons (Kaneko et al., 2014). Hence, CK1 plays a crucial role in controlling REST stability in adult neurons.

CK1 and β-TrCP are both present in stem cells, which demonstrate high REST levels. Phosphorylation of these two degrons by CK1 fails to explain the high REST levels and corresponding decrease in levels accompanied by neurogenesis. Nesti et al. identified a further phosphorylation motif at serines 861 and 864 that regulates REST stability (Nesti et al., 2014). This motif was found to bind the peptidyl prolyl cis/trans isomerase, Pin1, which stimulates βTrCP binding to the C-terminal degrons

and subsequent proteasomal targeting. Inhibition of Pin1 activity also inhibits the binding of βTrCP (Nesti et al., 2014). Furthermore, EGF-Ras-ERK signalling, which has been implicated in both embryonic and adult neurogenesis (Li et al., 2014), was demonstrated to target S861 and S864, with ERK2 found to directly phosphorylate REST. Conversely, the phosphatase small CTD phosphatase 1 (CTDSP1) was shown to dephosphorylate S861 and S864 (Nesti et al., 2014). REST has earlier been shown to colocalise with CTDSP1 on neuronal genes (Yeo et al., 2005). Nesti et al. hence propose a model for mammalian neurogenesis in which CTDSP1 protects REST from degradation in the expansion phase of neural progenitors, repressing expression of neuronal genes. In the later neurogenic stages, increased ERK signalling and Pin1 levels causes βTrCP-dependent REST degradation, de-repressing neuronal genes. Fittingly, CTDSP1 expression is dramatically reduced upon differentiation of neural progenitor cells to mature neurons (Yeo et al., 2005).

Residue	Kinase	PPase	Role	Publication
S861	ERK2	CTDSP1	Pin1 binding, indirectly increasing β-TrCP binding	Nesti et al., 2014
S864	ERK2	CTDSP1	Pin1 binding, indirectly increasing β-TrCP binding	Nesti et al., 2014
S1013	CK1	-	β-TrCP binding	Guardavaccro et al., 2008
S1024	CK1	-	β-TrCP binding	Westbrook et al., 2008
S1027	CK1	-	β-TrCP binding	Westbrook et al., 2008
S1030	CK1	-	β-TrCP binding	Westbrook et al., 2008

Table 5.1 Phosphorylated residues of REST identified by previous publications.

5.1.4 Using Mass Spectrometry to Identify Protein Phosphorylation

A comprehensive functional characterisation of protein phosphorylation sites requires identification of all the phosphorylation sites of the protein. High-resolution Mass Spectrometers are commonly used to identify protein phosphorylation sites, owing to their speed, sensitivity, and throughput (Yates, Ruse and Nakorchevsky, 2009).

Analysis by MS is preceded by digestion of purified protein using a protease that cuts at known sites, producing small peptides (Dephoure et al., 2013). Reverse-phase high-performance liquid chromatography is used to separate resulting peptides, which then enter the mass spectrometer and are ionised commonly by Electrospray Ionisation (ESI) or Matrix Assisted Laser Desorption Ionisation (MALDI). The resulting ions are then separated according to their mass-to-charge ratio (m/z), their m/z ratios and relative abundance are recorded and presented in the format of an m/z spectrum. Tandem mass spectrometers are capable of multiple rounds of mass spectrometry and enable fragmentation of selected sample ions. A tandem mass spectrometer records two basic types of information, the mass of intact peptide ions in an MS scan, and masses of fragment ions in an MS/MS scan. These fragment ions are produced by isolating peptides one at a time and fragmenting them by collision with an inert mass. Peptides are identified by matching experimental data to theoretical spectra, which are obtained from sequences databases, using online algorithms. The position of the phosphorylated residue can be determined from the MS/MS scan depending on the coverage of fragment ions (Dephoure et al., 2013).

Successful analysis requires purified protein in sufficient amounts, usually micrograms. This need is due to the low stoichiometry of phosphorylated proteins compared to unphosphorylated ones in the cell. Either an epitope tag is employed, or protein is immunoprecipitated using a suitable antibody. Proteins are subsequently purified by SDS-PAGE, proteins visualised by staining, the protein band excised,

proteolytically digested and subjected to LC-MS/MS (Aebersold and Mann, 2003). Alternatively, proteolytic digestion is performed in solution if proteins are eluted in low salt and detergent-free buffers. Phosphorylated proteins can be subsequently enriched by using metal ion affinity chromatography or titanium oxide (TiO_2) beads (Posewitz and Tempst, 1999; Pinkse et al., 2004). Digested and enriched phosphopeptides are subjected to microcapillary reversed-phase liquid chromatography/tandem mass spectrometry (LC-MS/MS) (Breitkopf and Asara, 2012).

The fragmentation data obtained is processed using a database search engine software, for example, Sequest or Mascot, and queried against protein databases like Swiss-Prot. MS uses mass/charge ratio (m/z ratio) to detect a +80 Da mass increase for singly phosphorylated peptides (HPO_3) compared to unmodified peptides. Site localisation on the peptide is attained by the fragmentation pattern of the ions and changes due to the neutral loss of a phosphate group (-98 Da) from serine and threonine phosphopeptides and a loss of metaphosphoric acid (-80 Da) for tyrosine phosphopeptides, ensuing gas-phase fragmentation (Hunter and Games, 1994).

A limitation of phosphorylation site identification is obtaining sufficient coverage of the protein sequence. Ideal MS-friendly peptides range from ~8-20 amino acids. However, not all phosphorylated sites lie in regions that will result in peptides of this size after proteolytic cleavage with a particular enzyme. Hence, it is important to use additional enzymes after examining the protein sequence that will give peptides of the ideal length (Dephoure et al., 2013).

5.2 Aims

The aim of this chapter is to attempt to map phosphorylation sites on REST and to identify the IGF2-dependent phosphosite on REST if any. The phosphorylation site mapping was achieved by two approaches:

1) An in silico approach to identify previously determined and predicted phosphorylation candidates.
2) A direct approach to identify REST phosphorylation sites by Mass Spectrometry.
3) Comparison of phosphosites on engineered isoforms of REST from IGF2 treated and untreated cells.

5.3 Results

5.3.1 Candidate post-translational modification sites in REST

An in silico approach was initially employed for detecting phosphorylation sites on REST. This approach consisted of using a prediction server to predict putative phosphorylation sites, identifying which residues were conserved amongst REST orthologues, identification of residues which have been found to be mutated in cancer, and hence display disease significance and finally searching any published data online for REST phosphorylation sites identified previously.

5.3.1.1 Prediction of Phosphorylation Sites using NetPhos

There are several phosphorylation prediction programs available online including NetPhos, ScanSite, and GPred. NetPhos 2.0 employs an artificial neural network method to predict phosphorylation sites in eukaryotic protein sequences (Blom, Gammeltoft and Brunak, 1999). The sensitivity of the method has been found to be in the range of 69% to 96%. In simple terms, the program uses common features of phosphorylated residues and their surrounding sequences to identify novel sites for phosphorylation. The output score is given in the range of 0.000 – 1.000 with a threshold of 0.5000. Sites above or below the threshold score are predicted to be phosphorylated and unphosphorylated respectively. A higher score indicates greater confidence in the prediction and higher similarity to one or more phosphorylation sites used in training of the program. Table 5.2 and 5.3 show the sites in REST with a score above 0.500 and hence strong candidate phosphorylation sites.

For REST, 58 Serine residues were identified as putative phosphorylation sites of which 36 Serine residues had a score above 0.9. I also checked the accuracy of the server by checking for sites validated by experimental methods, S861 and S864 (Binding site for Pin1 and MAPK), S1013, S1024 and S1027 (Binding site for

βTrCP). All except for S1027 were correctly identified as putative phosphorylation sites by NetPhos hence increasing our confidence in the program. Similar analysis for Threonine and Tyrosine residues identified nine putatively phosphorylated Threonine residues, with scores ranging from 0.565-0.871, and six putatively phosphorylated Tyrosine residues with scores ranging from 0.513-0.922.

These putative sites were found to be distributed across the sequence of REST. Five of the serine sites are within the region that interacts with the corepressor mSin3, one site is in the N-terminal repressor domain, nine sites lie within the DNA-binding domains, twenty-one within the lysine-rich region, two within the proline-rich region and nine within the C-terminal repressor domain. All putative threonine residues, except for one within the DBD domain, were found to lie in the lysine-rich region. Three tyrosine residues were in the DBD and one lied in the N-terminal repressor domain. The location for phosphosites in the disordered region is consistent with previous studies, which indicate that PTM in eukaryotic cell signalling proteins frequently occurs in disordered regions (Dunker et al., 2001; Collins et al., 2008). Prominent examples include c-src tyrosine kinase which is activated by a tyrosine phosphorylation in a flexible and disordered loop region, p53 which has its majority of phosphorylation sites in the putatively disordered N-terminal transactivation domain and C-terminal basic tail, and Bcl-2 which is inactivated by phosphorylation in its unstructured loop (Ojala et al., 2000; Ayed et al., 2001; Xu et al., 1999). The ability of unstructured regions to fold around the surface contours of enzymes that place the modifications is suggested to make them more accessible and hence, amenable to modification (Dunker et al., 2001).

5.3.1.2 Site Conservation of Putative Phosphorylation Sites

Phosphorylation sites with important function are likely to be conserved through evolution. Hence, I sought to filter phosphorylation candidates by checking for sites that were conserved amongst REST orthologues. Geneious Software was used to perform sequence alignment of predicted phosphorylation regions among vertebrate species including *H. sapiens*, *M. musculus, R. norvegicus, G. gallus, B. Taurus, M. mulatta* and *C. familiaris* (last column in Table 6.2 and 6.3). I found that amongst the 58 Serine residues identified to be putatively phosphorylated, 45 were conserved amongst these species with 20 being completely conserved and 25 semi-conserved amongst species. Of the Threonine residues identified, three were completely conserved and five semi-conserved, and all Tyrosine residues identified were conserved with five being completely conserved amongst all species. Of the phosphodegron that binds MAPK, S864 was semi-conserved, S861 was not found to be conserved, and S1020 and S1024 were fully conserved.

Position	Score	Conservation	Domain
S37	0.886	*	mSin3 Int.
S61	0.901	:	RD1, mSin3 Int.
S85	0.947	*	mSin3 Int.
S87	0.981	*	mSin3 Int.
S117	0.978	:	mSin3 Int.
S136	0.996	:	
S153	0.985	:	
S154	0.972	:	
S170	0.978	*	DBD
S182	0.982	*	DBD
S191	0.803	:	DBD
S201	0.996	*	DBD
S203	0.868	:	DBD
S204	0.992	:	DBD
S212	0.85	*	DBD
S259	0.87	*	DBD
S327	0.991	*	DBD
S423	0.525	*	Lys-Rich
S465	0.995		Lys-Rich
S474	0.997	:	Lys-Rich
S480	0.533		Lys-Rich
S495	0.998	*	Lys-Rich
S508	0.775		Lys-Rich
S510	0.972		Lys-Rich
S514	0.553		Lys-Rich
S519	0.995	:	Lys-Rich
S527	0.527		Lys-Rich
S529	0.93		Lys-Rich
S540	0.975		Lys-Rich

Position	Score	Conservation	Domain
S549	0.922	:	Lys-Rich
S551	0.924	:	Lys-Rich
S555	0.993	:	Lys-Rich
S563	0.768	:	Lys-Rich
S578	0.973	:	Lys-Rich
S588	0.995	:	Lys-Rich
S589	0.99	:	Lys-Rich
S592	0.995	:	Lys-Rich
S593	0.984	:	Lys-Rich
S750	0.552		Pro-Rich
S811	0.818		Pro-Rich
S824	0.954	:	
S828	0.705	:	
S856	0.997	*	
S861	0.987		MAPK
S864	0.91	:	MAPK
S964	0.605	:	
S971	0.815	*	
S977	0.969		
S986	0.818	:	
S992	0.817	*	RD2
S1002	0.886	*	RD2
S1013	0.994	*	RD2, Degron 1
S1020	0.694	*	RD2
S1024	0.926	*	RD2, Degron 2
S1042	0.916		RD2
S1043	0.959	:	RD2
S1068	0.844	*	RD2
S1076	0.52	*	RD2

Table 5.2 Phosphorylation predictions for serine residues of REST by NetPhos.
Column 1 shows the position of the residue, column 2 the output score in the range of 0.001 to 1.000 and column 3 shows the conservation of the residue with '*' to mean completely conserved and ':' to mean semi-conserved, and column 4 shows the domain in which the residue occurs (DBD: DNA binding domain, Lys-Rich: Lysine rich domain, Pro-Rich; Proline rich domain, RD1: Repressor Domain 1, RD2: Repressor Domain 2, mSin3 Int.: mSin3 interaction domain. Residues with score >0.9 are underlined.

Position	Score	Conservation	Domain
T144	0.784	:	
T257	0.657	*	DBD
T430	0.713	*	Lys-Rich
T451	0.565	:	Lys-Rich
T492	0.785	:	Lys-Rich
T516	0.701	:	Lys-Rich
T541	0.837	*	Lys-Rich
T579	0.871	:	Lys-Rich
T696	0.625		Pro-Rich

Position	Score	Conservation	Domain
Y31	0.835	*	
Y65	0.513	*	RD1, mSin3 Int.
Y134	0.922	:	
Y231	0.582	*	DBD
Y291	0.906	*	DBD
Y339	0.773	*	DBD

Table 5.3 Phosphorylation predictions for threonine and tyrosine residues of REST by NetPhos.

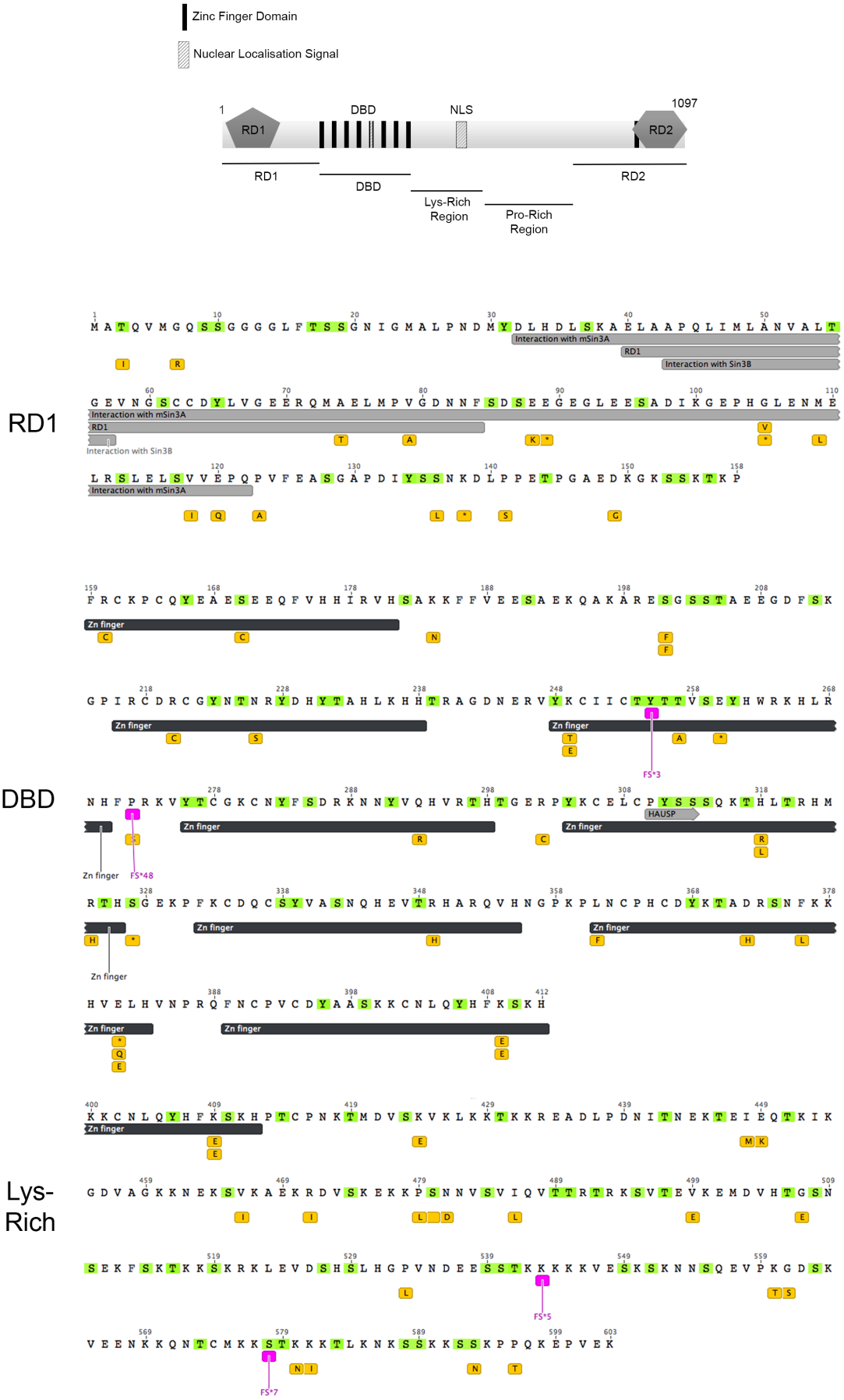
Putative phosphorylation sites of REST were mapped by NetPhos 2.0 server. Column 1 shows the position of the residue, column 2 the output score in the range of 0.001 to 1.000, column 3 shows the conservation of the residue with * to mean completely conserved and : to mean semi-conserved, and column 4 shows the domain in which the residue occurs (DBD: DNA binding domain, Lys-Rich: Lysine rich domain, Pro-Rich; Proline rich domain, RD1: Repressor Domain 1, RD2: Repressor Domain 2, mSin3 Int.: mSin3 interaction domain. Residues with score >0.9 are underlined.

5.3.1.3 Mutations of REST in Cancer

Next, I attempted to identify sites through functional relevance by searching for sites found to be mutated in Cancer, hence implying important function. This was performed by downloading mutation data for REST from Cosmic Database (Catalogue for Somatic Mutations in Cancer). As mentioned earlier, Cosmic is an online catalogue of genes and mutations that have been implicated in cancer (http://cancer.sanger.ac.uk/cosmic).

Downloaded mutations were mapped onto the REST sequence, which was divided into five regions, Repressor Domain 1 (RD1), DNA-binding Domain (DBD), Lysine-rich domain, Proline-rich domain and Repressor Doman 2 (RD2). There were in total 148 mutations identified, 143 point mutations of which 14 lead to premature stop codons, the rest leading to change in amino acid, four frameshift mutations, and one deletion. Figure 5.1 shows these sites mapped onto the REST sequence. It can be observed that these mutations are widespread and there are no particular regions of REST which are mutated more frequently than other. The frameshift mutations were found in the DBD and lysine-rich regions and resulted in precluding the C-terminal domain of REST. The deletion occurs in the proline-rich region and removes a 16 amino acid region containing one Serine residue.

The serine, threonine, and tyrosine residues found to be mutated are listed in Table 5.4. These were 20 in total and were spread across the REST sequence. None of them was mutated to a phosphomimetic, a glutamate or aspartate. Interestingly, S718 was mutated to a tyrosine and S750 was mutated to a threonine, maintaining phosphorylation potential.


Zinc Finger Domain
Nuclear Localisation Signal
1
DBD
NLS
1097
RD1
RD2
RD1
DBD
Lys-Rich Region
Pro-Rich Region
RD1
DBD
Lys-Rich
Interaction with mSin3A
Interaction with Sin3B
Zn finger
HAUSP
FS*3
FS*48
FS*5
FS*7

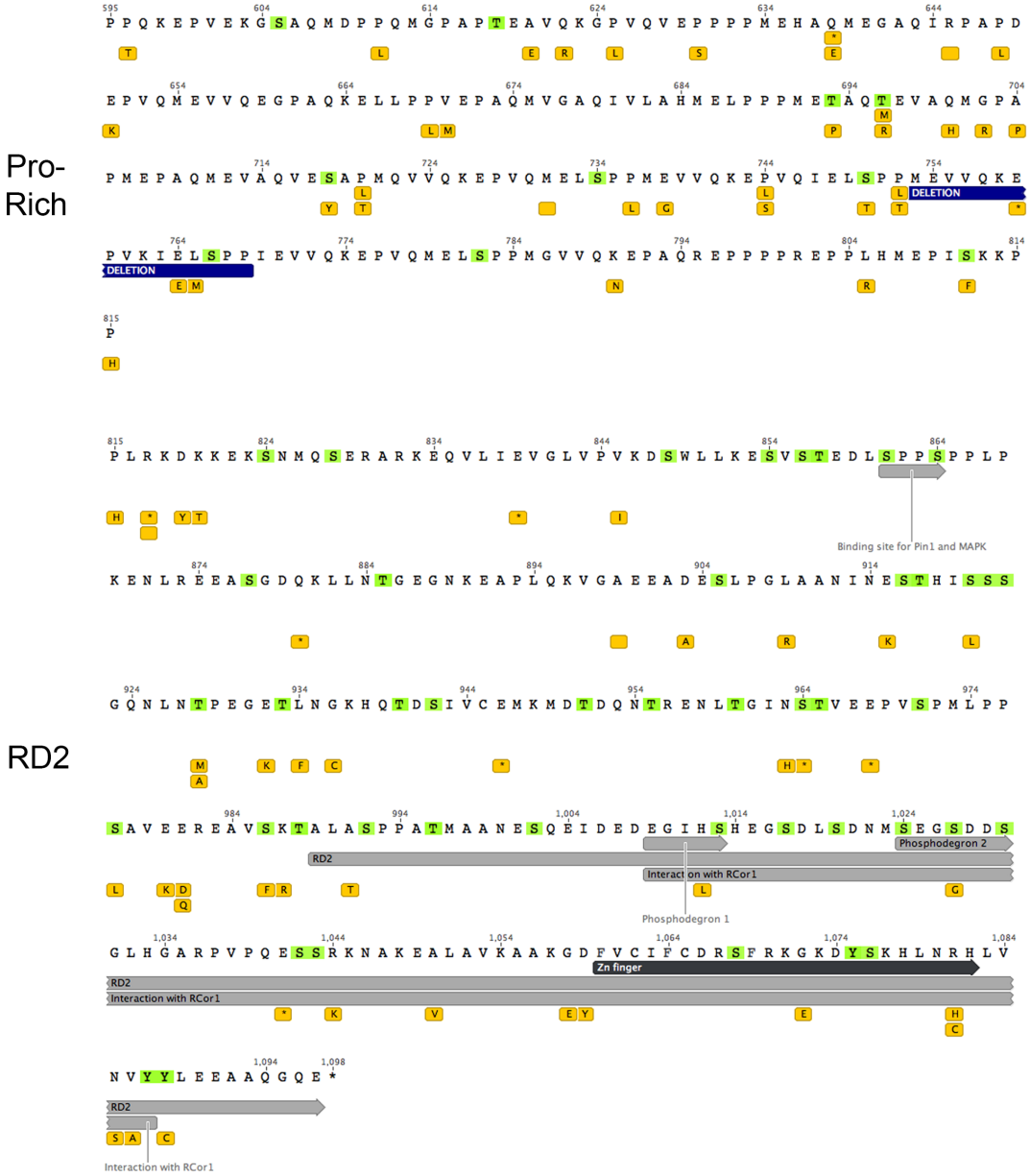


Figure 5.1 Mutations of REST in Cancer.
Top: A schematic of REST illustrating REST domains. Full-length REST consists of RD1; Repressor Domain 1, DBD; DNA-binding domain consisting of 8 zinc fingers, a Lysine-rich and a Proline-rich regions and RD2; Repressor Domain 2. Below: REST sequence was divided into five segments based on the schematic at the top and annotated to show interaction domains and zinc fingers. Ser/Thr/Tyr residues are depicted in green and cancer mutations are shown with a yellow box underneath, showing the mutated residue.

Position	WT Aminoacid	MUT Aminoacid	Occurrence
3	T	I	1
136	S	L	1
170	S	C	1
201	S	F	2
257	T	A	1
327	S	*	1
480	S	F	2
593	S	N	1
693	T	P	1
696	T	R,M	2
718	S	Y	1
750	S	T	1
811	S	F	1
920	S	L	1
928	T	M,A	2
964	S	*	1
977	S	L	1
986	S	F	1
1027	S	G	1
1088	Y	C	1

Table 5.4 Mutations in Serine, Threonine and Tyrosine residues of REST found in Cancer.

REST mutations in cancer obtained from the Cosmic database are shown with the amino acid position in REST sequence, the wild type amino acid, the mutated residue and the occurrence of the mutation in cancer samples.

5.3.1.4 Phosphorylated residues previously identified

Next, I scanned an online database of PTM for phosphorylations in REST. Phosphositeplus is a knowledgebase of protein PTMs obtained from mass spectrometry experiments as well as site-specific experiments (Hornbeck et al., 2015). It showed 20 phosphorylated sites, of which 18 were identified by mass spectrometry analysis and two sites were found by site-specific methods, namely S861 and S864. Three of the identified sites were AKT specific sites and three were ERK sites, one PKC and one CK2 site (Table 5.5).

Combining these three types of sites together, putatively phosphorylated and conserved, sites mutated in Cancer, and previously identified phosphosites, gives a plethora of potential sites (Figure 5.2). There is only one site that is common to all three, S986. However, it is not possible to narrow down the multitude of sites into selected sites for mutation analysis using an in silico approach without experimental validation.

Position	Phosphosite plus	
	Phosphorylated	Kinase
S203	Yes	AKT
S204	Yes	AKT
S212	Yes	AKT
S495	Yes	PKC
S527	Yes	CK2
S588	Yes	
S589	Yes	
S592	Yes	
S734	Yes	
S750	Yes	
S766	Yes	
S856	Yes	
S861	Yes	ERK1
S864	Yes	ERK1
S971	Yes	ERK1/2
S986	Yes	
S1020	Yes	
T317	Yes	
T960	Yes	
Y231	Yes	

Table 5.5 Phosphorylation sites in REST identified from Phosphositeplus.
The table shows residues identified to be phosphorylated and the predicted kinases (if any) responsible for the phosphorylation.

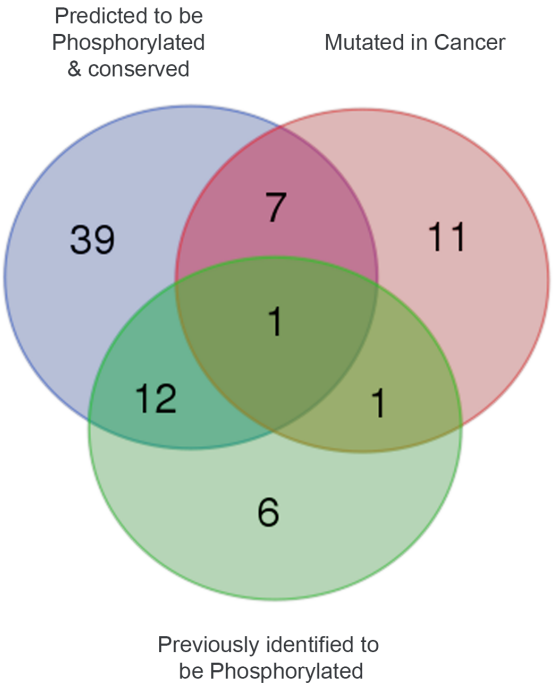


Figure 5.2 Total phosphorylation sites identified from in silico approach.
Venn diagram shows the number of sites predicted to be phosphorylated and conserved, mutated in cancer, and previously identified to be phosphorylated as well as the common site (S986) between them.

5.3.2 Identification of phosphorylation sites by Mass Spectrometry

Next, a mass spectrometry approach was used to identify actual phosphorylation sites on REST. This experiment consisted of selecting a cell line that demonstrated IGF2-mediated REST stability, producing sufficient amount of REST protein for Mass Spectrometry, followed by comparative phosphorylation identification of IGF2 treated and untreated samples.

5.3.2.1 IGF2 stabilised exogenous REST in HEK293T cells

NIH3T3 cells were unfortunately not found to be amenable to transfection, and since large amounts of protein are required for Mass Spectrometry, this precluded NIH3T3 cells from being used for Mass Spectrometry. HEK293T cells, owing to their ease of transfection and presence of a REST stability mechanism mediated by IGF2, were evaluated as an alternative.

Immunoprecipitation of endogenous REST using Anti-REST (17-641) failed to give enough material for MS analysis. That introduced the requirement for exogenous expressed REST with a tag to facilitate protein purification. The Flag-tagged REST that has been used in Chapter 4 was one possibility. Since the tag might interfere with proper folding or function of REST, it was necessary to check if Flag-tagged REST reflected the behaviour of endogenous REST. Flag-tagged REST was transfected into HEK293T cells, serum-starved for 24 hours and then treated with IGF2 and MG-132. Whole cell extracts were immunoblotted for REST using Flag, α-tubulin, and pAKT (Figure 5.3). Flag-REST was found to reproduce the effect of IGF2 and MG-132 on endogenous wild-type REST. As seen in Chapter 3, Flag-REST was found to be predominantly directed to the 220kDa HMW pool of REST. Hence, HEK293T cells and Flag-REST represent suitable tools for phosphorylation dissection of REST using Mass Spectrometry.

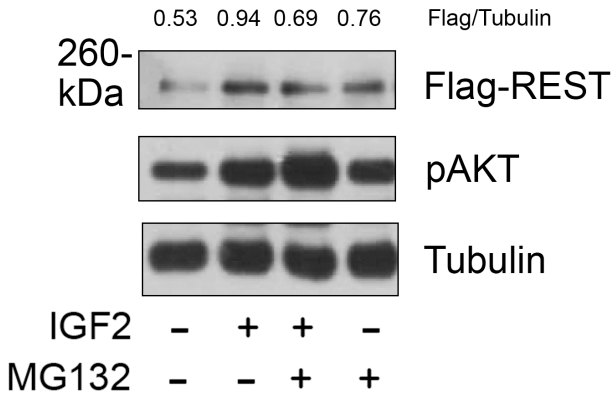


Figure 5.3 Modification of exogenous REST by IGF2.
HEK293T cells were transfected with Flag-REST, serum-starved and treated with IGF2 (100ng/ml) and MG-132 (0.75μM). Whole cell lysates were immunoblotted for Flag-REST using Flag antibody, α-tubulin and pAKT. Normalised Flag-REST levels quantified by Image J are shown above the blot. Western blot is representative of three independent experiments.

5.3.2.2 Optimisation of Mass Spectrometry pipeline

5.3.2.2.1 Gel band provides better coverage than eluate

A crucial issue in Mass spectrometry is dye contaminants, which are used in gel loading. To check whether gel-isolated protein or the eluate obtained from purification provided better coverage of REST peptides, HEK293T cells were transfected with Flag-REST construct. 48 hours post transfection, cells were treated with MG-132 for 6 hours to increase relative protein levels, lysed and whole cell extracts were prepared. REST was immunoprecipitated using Anti-Flag M2 antibody. After purification, fractions were run on an SDS-page gel, immunoblotted for REST using 07-579 antibody and confirmed for the presence of REST. The other half of the eluate was run on an SDS page gel and stained with Coomassie stain. A clean coomassie band was obtained at 220kDa, which was also identified as REST by immunoblotting. This band was cut out, exposed to in-gel tryptic digest and Mass Spectrometry was performed on the resulting peptides (Figure. 5.4A). Alternatively, bound protein was eluted in an SDS-containing buffer without bromophenol, subjected to trypsin in solution and resulting peptides subjected to LC-MS/MS.

Mass Spectrometry identified REST to be the top hit in the gel band with a coverage of 44%, whereas REST was seen as the 8th hit from the elution with a coverage of 20% (Figure 5.4B). Hence, gel band followed by in-gel tryptic digest was found to be the better option regarding peptide coverage.

5.3.2.2.2 Addition of GluC digest step improves coverage

The coverage obtained from the in-gel tryptic digest of the Coomassie-stained REST band was 44% with three major islands of non-coverage. Since candidate phosphorylation sites were widespread throughout the REST sequence, it was necessary to increase the coverage of REST to attempt to capture all potential phosphorylation sites. Westbrook et al. use an N1022R mutation to enable

identification of the tryptic peptides in the C-terminus of REST (Westbrook et al., 2008). An alternative way to improve coverage is to use alternative enzymes that cut at different amino acid residues. GluC is a serine protease, which cleaves peptides on the carboxyl side of Glutamic acid residues specifically. Cleavage by GluC would lead to alternative cleavage pattern allowing detection of peptides, which would otherwise remain undetected due to limitations in peptide size detection. In-gel Tryptic digestion was followed by digestion by GluC and subsequent detection with Mass Spectrometry. Coverage in one of the 'island' areas, in the C-terminus of REST, was improved by GluC (Figure 5.4C). Overall coverage after combining Trypsin digestion and Trypsin and GluC digestion improved from 44% to 53.1%.

5.3.2.2.3 Phosphopeptide enrichment using TiO2

In general, phosphopeptides do not ionise very efficiently in MS. That, along with their low stoichiometry, means that it is hard to detect them in tryptic digest pools of total fragments. Hence, a phosphopeptide enrichment step is often used to improve their detection. One such method uses TiO_2, which shows a high affinity towards phosphate ions in aqueous solutions. Tryptic peptides are loaded onto TiO_2 columns at highly acidic conditions to prevent binding of acidic peptides to the TiO_2 resin. Unphosphorylated peptides are washed away, and bound phosphorylated peptides eluted using alkaline conditions.

The eluate from Flag-REST immunoprecipitation was subjected to tryptic digest followed by phosphopeptide enrichment using TiO_2 and enriched phosphopeptides directed for LC-MS/MS. Coverage of 23% was obtained and only one previously unidentified peptide from amino acid 71-91, was discovered with two phosphorylation sites (Figure 5.4D). Hence, phosphopeptide enrichment did not provide a significant improvement in detection of phosphorylation.

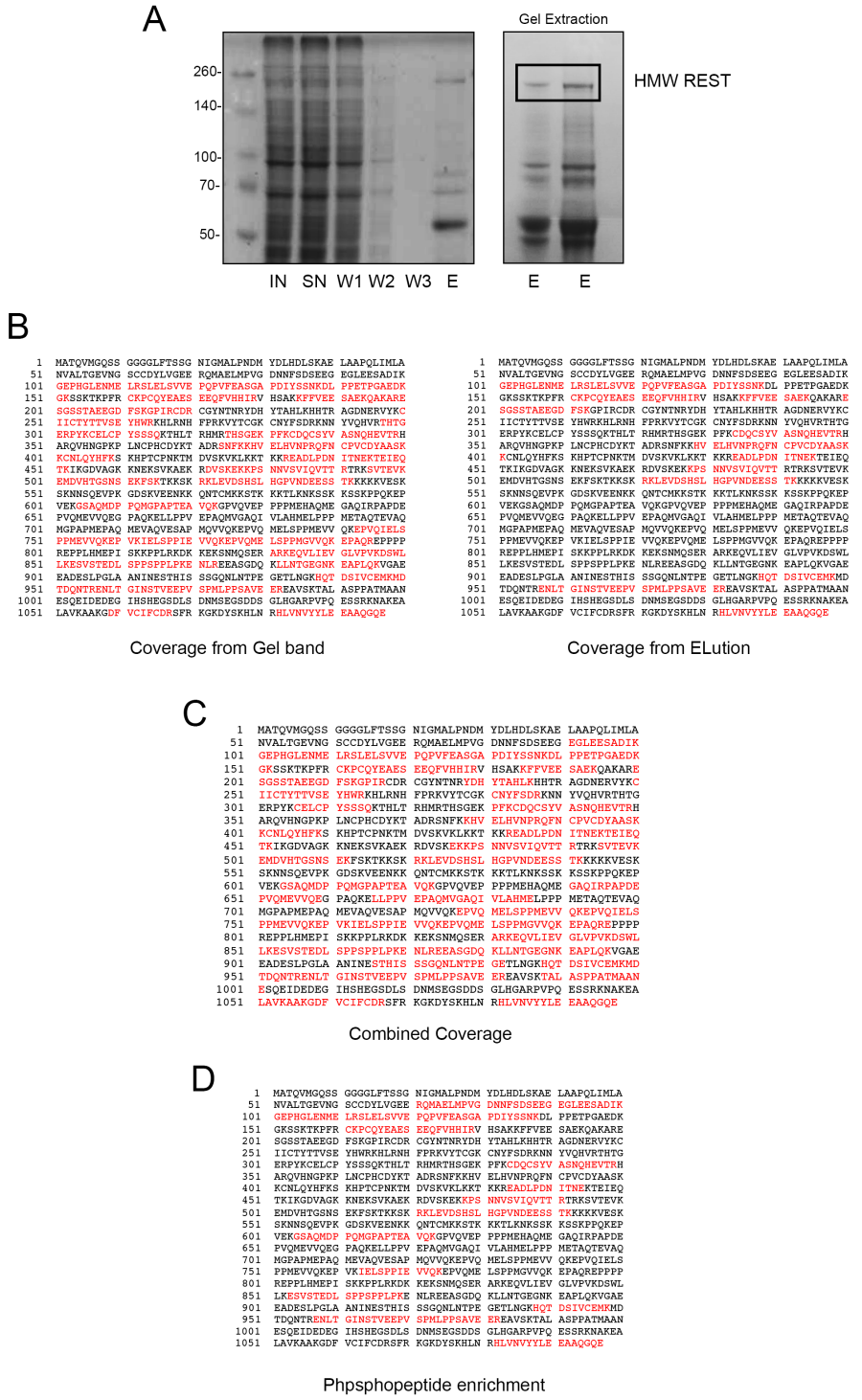


Figure 5.4 Optimisation of Mass Spectrometry pipeline.
A) Coomassie stained gel, showing purification of Flag-tagged REST. The right panel shows the bands that were cut for Mass spectrometry. B) Comparison of coverage achieved by gel band (left) and elution (right). Covered residues are given in red. C) Coverage obtained by GluC digest on peptides digested by trypsin combined with trypsin only digest. D) Coverage reached by trypsin digestion followed by phosphopeptide enrichment.

5.3.2.3 Mapping of Phosphorylation Sites by Mass Spectrometry

Data from separate MS runs was analysed for phosphorylated residues and pooled together. Total phosphorylated residues identified are given in Table 5.6. Only two of these residues, S87 and S96, were identified with phosphoenrichment. S861 and S864, the MAPK degron, were identified as phosphosites but S1013, S1024, and S1027, the βTrCP-binding degrons, were not covered by the experiments. Of the 17 phosphosites identified, eight were found to be novel sites. Of these eight sites, three were completely conserved and two were partially conserved amongst vertebrates (Table 5.7). Of these novel sites, one site, S327, was also found to be mutated in cancer. Upon comparing the identified sites with NetPhos predictions, it was noticed that nine sites out of seventeen were correctly predicted to be phosphorylated. The remaining eight sites were incorrectly predicted, giving a false negative rate of 47%.

Position	Peptide	Observed	Mr(expt)	Mr(calc)	ppm	Score	Runs
87 and 96	R.QMAELMPVGDNNFSDSpEEGEGLEESpADIKGEPHGLENMELR.S	1178.7349	4710.9107	4710.8962	3.07	24	1
203	R.ESGSpSTAEEGDFSKGPIR.C	645.2781	1932.8126	1932.8102	1.23	63	1
205	K.ARESGSSTpAEEGDFSKGPIR.C	720.9912	2159.9517	2159.9484	1.52	67	1
327	R.THSpGEKPFKCDQCSYVASNQHEVTR.H	609.866	3044.2939	3044.2906	1.06	32	1
474	R.DVSpKEKKPSNNVSVIQVTTR.T	770.4002	2308.1787	2308.1788	-0.013	97	1
734	K.EPVQMELSpPPMEVVQK.E	976.9437	1951.8729	1951.8672	2.92	47	1
750	K.EPVQIELSpPPMEVVQK.E	959.9695	1917.9245	1917.9159	4.52	39	1
766	K.IELSpPPIEVVQK.E	716.3802	1430.7458	1430.7422	2.55	46	1
782	K.IELSPPIEVVQKEPVQMELSpPPMGVVQK.E	1071.8868	3212.6387	3212.6114	8.5	55	2
766 and 782	K.IELSpPPIEVVQKEPVQMELSpPPMGVVQK.E	1098.5372	3292.5899	3292.5777	3.7	46	1
782	K.EPVQMELSpPPMGVVQK.E	940.9305	1879.8464	1879.8461	0.17	39	1
861	K.ESVSTEDLSpPPSPPLPK.E	930.4396	1858.8647	1858.8601	2.46	85	2
864	K.ESVSTEDLSPPSpPPLPK.E	930.4382	1858.8618	1858.8601	0.88	68	2
861 and 864	K.ESVSTEDLSpPPSpPPLPK.E	970.4198	1938.8251	1938.8264	-0.71	65	2
856	K.ESVSpTEDLSPPSpPPLPK.E	970.4216	1938.8287	1938.8264	1.17	21	1
857 and 864	K.ESVSTpEDLSPPSpPPLPK.E	970.4224	1938.8303	1938.8264	1.98	58	1
854 and 864	K.ESpVSTEDLSPPSpPPLPKENLR.E	818.0404	2451.0994	2451.0971	0.91	56	1
960	R.ENLTpGINSTVEEPVSPMLPPSAVEER.E	959.1205	2874.3397	2874.3358	1.38	33	1
971	R.ENLTGINSTVEEPVSpPMLPPSAVEER.E	964.4522	2890.3349	2890.3307	1.45	46	1

Table 5.6 Phosphorylation sites detected by LC-MS/MS.
Position; the position of amino acid in REST sequence. Peptide; the sequence of the peptide in 1-letter code. Phosphorylated residues are identified in red. Observed; experimental m/z value. Mr(expt); experimental m/z value converted to a relative molecular mass. Mr(calc); the relative molecular mass of REST peptides calculated from REST sequences digested by enzyme in silico. Ppm; the difference between experimental and calculated masses. Score; the ion score (set to a threshold of 20). Occurrence shows the number of times the peptides were identified in separate runs. Only the highest scoring peptides are given for those identified multiple times.

Position	Novel	Mutated in Cancer	Conserved	Phos prediction	Domain
S87	✓		✓✓	✓	mSin3A interaction
S96	✓		✓✓		mSin3A interaction
S203			✓	✓	DBD
T205	✓		✗		DBD
S327	✓	✓	✓✓	✓	DBD
S474	✓		✓	✓	Lys-rich
S734			✓		Pro-rich
S750		✓	✗	✓	Pro-rich
S766			✗		Pro-rich
S782	✓		✗		Pro-rich
S854	✓		✓		
S856			✓✓	✓	
T857	✓		✗		
S861			✗	✓	Pin1 binding
S864			✓	✓	Pin1 binding
T960			✗		
S971			✓✓	✓	

Table 5.7 Analysis of Phosphorylated residues identified by Mass Spectrometry.
Phosphosites identified in Table 6.6 were checked for previous identification, conservation, mutation in cancer, whether they are predicted to be phosphorylated by NetPhos, and the domain containing them.

5.3.3 Identification of IGF2-dependent Phosphosite on REST

To identify IGF2-dependent phosphosite on REST, HEK293T cells were transfected with Flag-REST, 48 hours post transfection serum starved for 24 hours and half of the plates subsequently treated with IGF2 and the other half treated with a combination of IGF2 and MG-132. Purification fractions were run on an SDS-page gel, and either stained with Coomassie or immunoblotted for REST using 07-579 antibody. Coomassie staining identified bands around 220kDa, which were also recognised to be REST by immunoblotting. Coomassie bands identified as REST were then excised, subjected to in-gel tryptic digest, followed by GluC digest and subsequent analysis by LC-MS/MS (Figure 5.5).

64.8% of REST residues from the IGF2+MG-132 treated sample were covered compared to 58.7% of residues obtained from MG-132 only treated sample. Four phosphosites were identified in the IGF2 treated sample namely S861, S864, S960 and S1002 and all except S960 were present in the MG-132 only treated sample, The peptide with S960 was not picked up in the MG-132 sample hence no deduction can be made about S960 being the differential site (Figure 5.6).

Whereas S861, SS864, and S960 have been identified in previous MS runs to be phosphorylated, S1002 was identified for the first time and was found to be a novel site, not yet reported in the literature. It was also found to be conserved amongst REST orthologues and predicted by NetPhos to be phosphorylated.

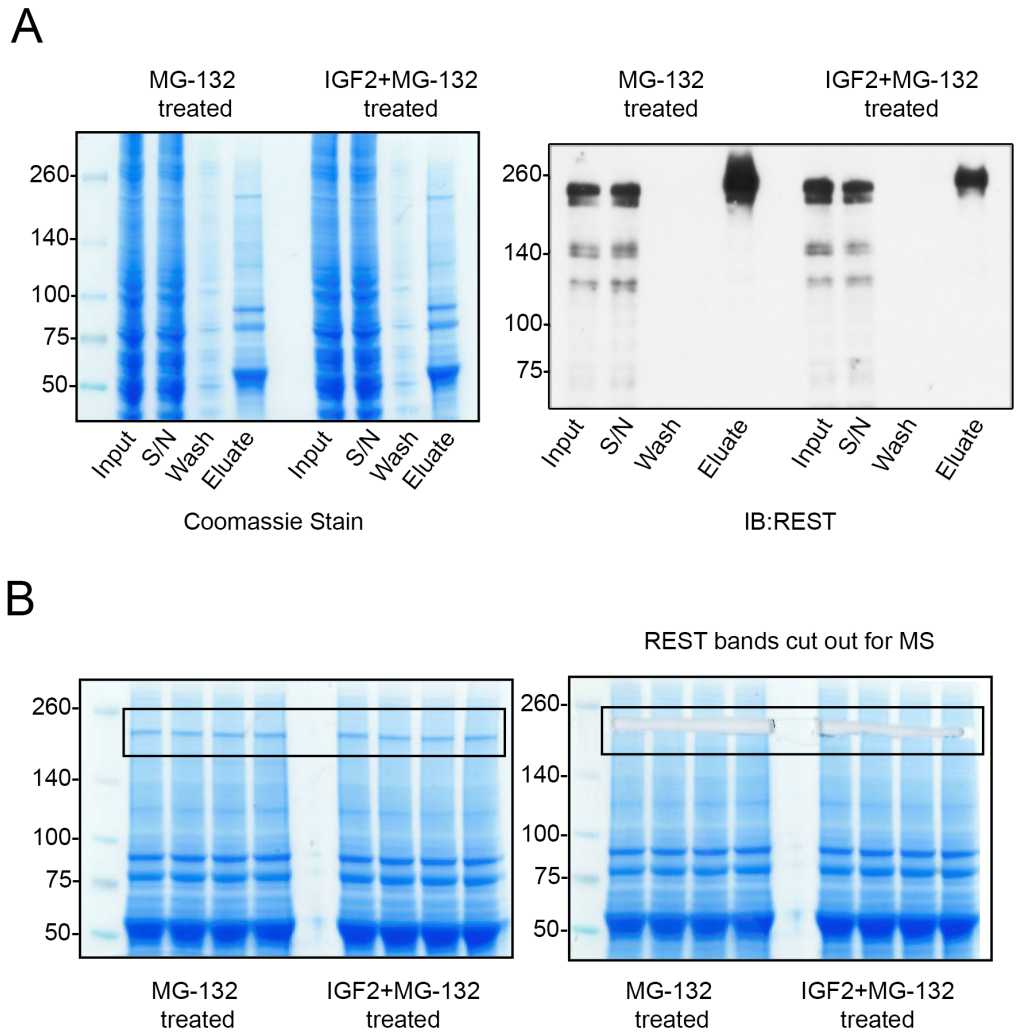


Figure 5.5 Purification of REST for LC-MS/MS.
A) HEK293T cells were transfected with Flag-REST, serum-starved and treated with MG-132 or MG-132+IGF2. Flag-REST was immunprecipitated and purification fractions were run on SDS-page and either stained for protein using Coomassie stain or immunoblotted for REST using 07-579 antibody. S/N: supernatant. B) The eluate obtained from Flag immunoprecipitation was separated on SDS-page gels, stained with Coomassie and the REST bands excised for MS. The right gel shows the gel bands cut out.

A



B

MG-132 + IGF2

Position	Peptide	Observed	Mr(expt)	Mr(calc)	ppm	Score	Covered in M sample
864	E.SVSTEDLSPP**Sp**PPLPK.E	865.9185	1729.8225	1729.8175	2.89	77	Yes
861 and 864	E.SVSTEDL**Sp**PP**Sp**PPLPK.E	905.9007	1809.7869	1809.7839	1.66	87	Yes
1002	K.TALA**Sp**PPATMAANE.S	720.8106	1439.6067	1439.6003	4.4	54	Yes
960	R.ENL**Tp**GINSTVEEPVSPMLPPS AVEEREAVSK.T	1135.8807	3404.6203	3404.6058	4.27	45	No

MG-132

Position	Peptide	Observed	Mr(expt)	Mr(calc)	ppm	Score	Covered in M+I sample
864	E.SVSTEDLSPP**Sp**PPLPK.E	865.9184	1729.8223	1729.8175	2.77	88	Yes
861 and 864	E.SVSTEDL**Sp**PP**Sp**PPLPK.E	905.8997	1809.7848	1809.7839	0.54	80	Yes
861	E.AVSKTALA**Sp**PPATMAANE.S	913.4243	1824.834	1824.8328	0.62	34	Yes
1002	K.TALA**Sp**PPATMAANE.S	720.8088	1439.603	1439.6003	1.83	51	Yes

Figure 5.6 Identification of IGF2 dependent phosphorylation sites.
A: Coverage achieved in MG-132+IGF2 and MG-132 only treated samples is shown. The covered peptides are given in red. B: Tables showing phosphopeptides identified from MS. Top table shows phosphopeptides identified from MG-132+IGF2 sample (M+I), the bottom table shows phosphopeptides identified from MG-132 treated sample (M). Observed; experimental m/z value. Mr(expt); experimental m/z value converted to a relative molecular mass. Mr(calc); the relative molecular mass of REST peptides calculated from REST sequences digested by enzyme in silico. Ppm; the difference between experimental and calculated masses. Score; the ion score (set to a threshold of 20). Peptides covered in the other sample are indicated with 'yes' in the last column.

5.3.4 Adapting Flag-REST-GFP fusion peptides for identification of REST phosphorylation

One reason for the limited interpretation of the mass spectrometry approach, to identify the IGF2-dependent sites, was the relatively large protein size and the resulting low coverage issue. The bioinformatics approach showed that REST has multiple putative phosphosites throughout its protein sequence. Hence, identification of the IGF2-dependent phosphosite would optimally require total coverage of the protein in both treated and untreated samples.

One alternative strategy was to isolate regions of protein to narrow down the putative phosphorylated region. Nesti et al. employed a similar approach to identify specific sites in REST protein (Nesti et al., 2014).

5.3.4.1 Validation of Flag-REST-GFP peptide reporter constructs

The first step was to validate the use of the constructs. We obtained two constructs from Nesti et al., one expressing a peptide encoding amino acids 810-910 and the other expressing amino acids 594-695. These peptides have a fused N-terminal Flag tag and a C-terminal eGFP tag (Figure 5.7). The increased charge of phosphorylated peptides causes differential migration of phosphorylated and unphosphorylated peptides in a native gel, allowing visualisation of consequences of the phosphorylation status. HEK293T cells were transfected with REST(810-910) and REST(594-695), serum starved and treated with different doses of IGF2 for 6 hours. Whole cell lysates obtained were resolved on a native gel. Peptides were detected by in-gel fluorescence (Figure 5.7). As reported by Nesti et al., REST(810-910) migrated as two distinct bands which were suggested to be phosphorylated peptides. These bands were converted to a single slow migrating species upon phosphatase treatment. Using mutational analysis, Nesti et al., report S861 and S864 to be the phosphorylated

residues in REST(810-910). REST (594-695) exists as one band resistant to phosphatase treatment (data not shown), hence, is non-phosphorylated.

Treatment of IGF2 does not change the relative band intensity of the two phosphorylated bands in REST(810-910). However, it decreases the intensity of both bands at 1ng/ml IGF2 dosage with no increase in the unphosphorylated band, suggesting that IGF2 potentially destabilises REST(810-910). This observation is consistent with ERK2 mediated degradation of REST (Nesti et al., 2014). REST(594-695) is unresponsive to IGF2 treatment suggesting the absence of an IGF2-dependent site. Hence, I established the utility of the system for narrowing down REST regions responsible for IGF2-dependent stabilisation.

5.3.4.2 Generation of REST isoforms

I then proceeded to generate constructs for narrowing down the IGF2-dependent site. I first decided to use the existing knowledge about two REST isoforms to create truncated mutants. These two isoforms have C-terminal truncations, leaving behind either 4 (Isoform 1) or 5 zinc fingers (Isoform 2). These two isoforms along with full-length REST were cloned into the pEGFP-N1 vector provided by Mandel. Constructs were expressed in HEK293T cells and detected by immunoblotting using a Flag antibody (Figure 5.8). Full-length REST (REST 1) was found to have a slightly higher molecular weight than the 220 kDa HMW REST indicative of the additional eGFP tag. REST 2 and REST 3 were about 75 and 80 kDa in size. The size of eGFP is 32.7 kDa, taking that into account the size of REST2 and REST3 comes down to about 42 and 47 kDa, which is only slightly higher than the 35 and 37 kDa MW of REST2 and REST3 predicted from their amino acid sequences.

HEK293T cells were then transfected with full-length REST and REST isoforms and 24 hours post-transfection treated with MG-132 for 6 hours. Whole cell lysates were

immunoblotted for REST using Flag antibody (Figure 5.8C). Both isoforms and full-length REST were found to be stabilised by MG-132 suggesting that all three tagged REST isoforms are being degraded by the proteasome. This implies the presence of other targeting signals independent of the C-terminal phosphodegrons that bind β-TrCP, resulting in proteasomal degradation. Furthermore, a cycloheximide chase assay showed that all three isoforms were being rapidly turned over in the cell (Figure 5.8D). Full-length REST had a half-life of about 2.7 hours, consistent with the half-life of endogenous REST observed in Chapter 4. Isoform 2 and 3, however, had much shorter half-lives of 1 hr and 1.25 hr respectively, suggesting that they are degraded more rapidly in the cell. This implies the presence of a C-terminal independent degradation mechanism that is being employed by the isoforms.

Hence, we show that these isoforms can be expressed, turnover in the cells and can be stabilised by proteasomal inhibition. Therefore, they can be used to investigate the identity of the IGF2-dependent phosphosite.

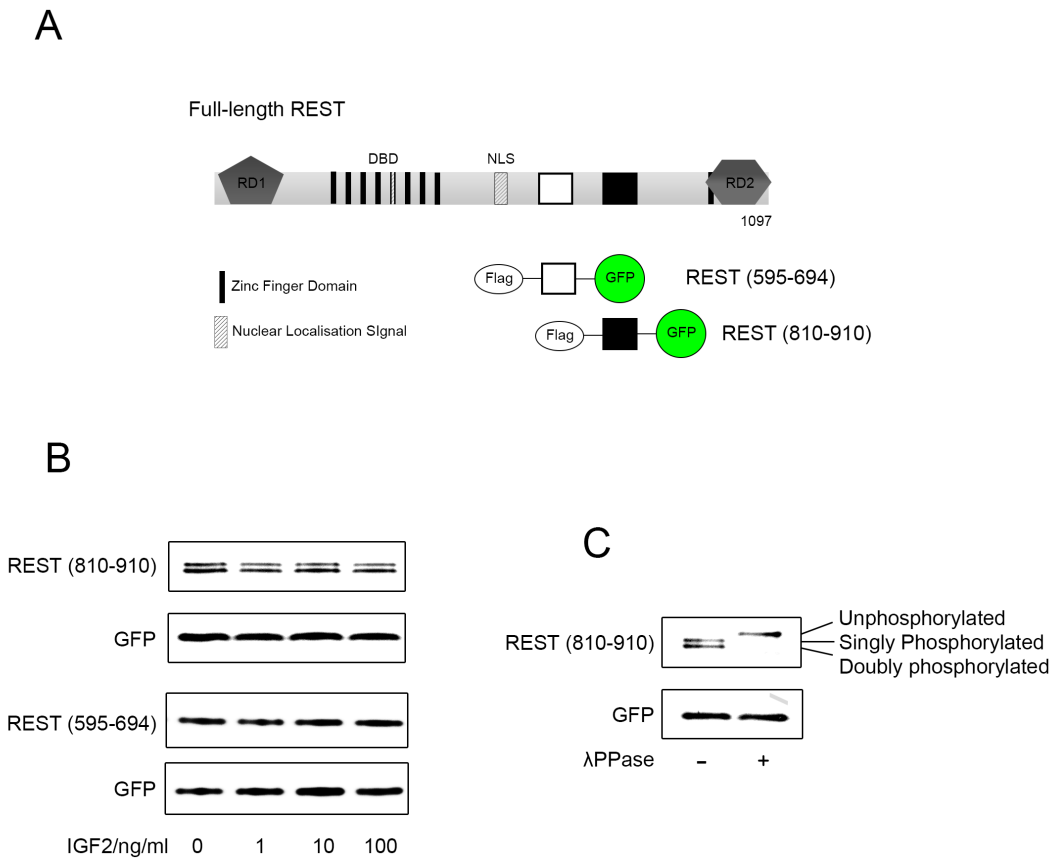


Figure 5.7 Validation of Flag-REST-GFP peptide reporter constructs.
A) A schematic of full-length REST and the REST-GFP reporter constructs. B) HEK293T cells were cotransfected with either reporter construct with GFP, serum-starved and treated with different doses of IGF2. Whole cell lysates were resolved on 15% acrylamide native gels and in-gel GFP fluorescence was detected using an imaging system. C) Whole cell lysate expressing REST (810-910) construct was treated with λ-phosphatase, resolved on a native gel and in-gel GFP fluorescence was detected.

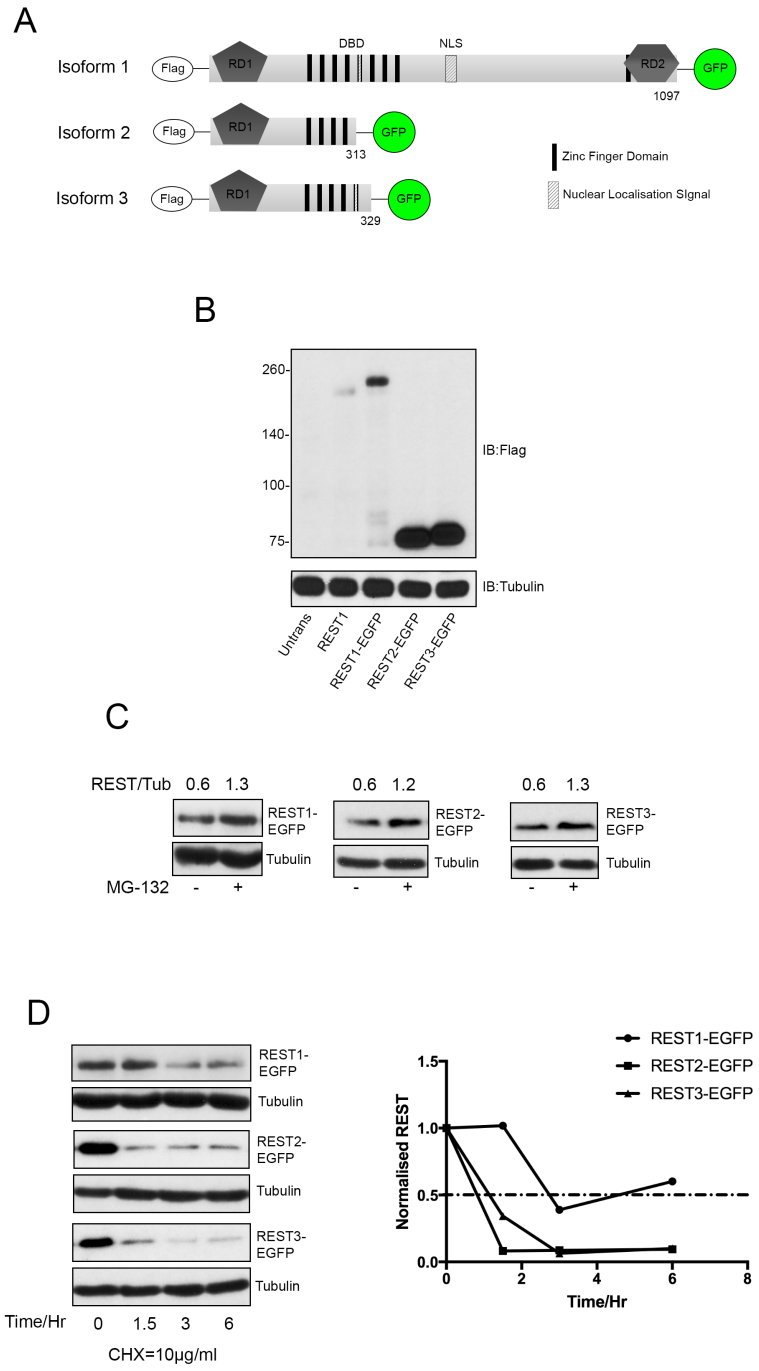


Figure 5.8 Generation and validation of Flag-GFP tagged-REST isoforms.
A) A schematic of N-Flag-C-GFP tagged-REST isoforms. B) N-Flag-C-GFP tagged-REST isoforms were expressed in HEK293T cells and immunoblotted using Flag antibody. C) HEK293T cells were transfected with N-Flag-C-GFP tagged-REST isoforms, and treated with MG-132 (0.75 µM) for 6 hours, followed by immunoblotting using Flag antibody D) HEK293T cells were transfected with N-Flag-C-GFP tagged-REST isoforms, and treated with cycloheximide (10 µg/ml) for the indicated times, followed by immunoblotting using Flag antibody. REST levels were quantified using Image J, normalised to α-tubulin levels and plotted against time. All Western blots shown are representative of two independent experiments.

5.4 Discussion

The in silico approach led to the identification of multiple sites which were both predicted to be phosphorylated and conserved. In addition, there were numerous sites already shown to be phosphorylated by Mass Spectrometry. This in silico strategy failed to narrow down sites to a selected group, which could be further characterised by mutation analysis. Hence, a mass spectrometry approach was taken to identify phosphosites on REST in my system.

Mass Spectrometry identified 17 sites spread over the entire REST sequence. Of these, eight were found to be previously undiscovered sites. All but two of these sites were identified without phosphoenrichment indicating they are present in high stoichiometry in cells.

A lentiviral construct with low REST expression was used for protein production to identify physiologically relevant sites. Flag-REST was found to be predominantly localised in the nucleus like endogenous REST and also recapitulated the effect of IGF2 and MG-132 on endogenous REST. That increased my confidence in the system being used for MS evaluation.

A direct corollary of using a low expression REST plasmid was that despite using a large amount of material and two successive enzymatic digestions, the best coverage obtained was still 64.8%. There are multiple uncovered peptide fragments, which might contain the IGF2-dependent phosphosite. The C-terminal region of REST, which has been shown to be a key regulatory region with two phosphodegrons is not covered completely well. Westbrook et al. use an N1022R mutation to enable identification of the tryptic peptide in the C-terminus (Westbrook et al., 2008). Similar mutations after sequence analysis might be necessary to achieve coverage of the missing regions. Alternatively, a third enzyme might be added for example LysC

which has a high specificity for lysine residues and can cleave lysine followed by proline residues, hence is often used in tandem with trypsin to decrease missed cleavages. Furthermore, the REST phosphoband is relatively less abundant than the unphosphorylated REST band. Hence, due to the low stoichiometry of phosphorylation, identification of the remaining phosphosites might prove to be challenging, and additional steps like IMAC might have to be added. Since nuclear REST was found to be predominantly unphosphorylated, another viable possibility would be to use cytoplasmic and nuclear extracts separately for phosphosite identification as opposed to whole cell extract. This would increase the stoichiometry of phosphorylated peptides in the sample for MS.

Treatment of REST with IGF2 failed to identify a differential phosphosite most likely due to the lack of full coverage of the protein. Multiple MS runs with different conditions are therefore needed to identify the site. Also, fewer phosphosites were detected in REST upon serum-starving cells, as compared to previous runs, and this might be a consequence of serum withdrawal. Another point of concern in the experiment conducted was the presence of MG-132, along with IGF2, which might have stabilised multiple IGF2 sites. A true representation of IGF2 enrichment would be comparison of an IGF2 only treated sample with an untreated sample and hence, a possible reason for failure of the experiment.

Another strategy for phosphosite identification is a top-down approach by using in vitro assays using kinases and phophatases involved in the IGF2 signalling pathway. This would allow identification of the kinases and phosphatases involved in phosphorylating and dephosphorylating REST. This information can then be used to identify kinase or phosphatase substrate motifs, followed by confirmation by mutational analysis.

A third strategy is generating truncated versions of REST sequence followed by analysis of the effect of IGF2 to narrow down regions followed by mutation analysis of phosphosites in said region. A similar strategy was employed by Nesti et al. to identify the S861 and S864 phosphodegron (Nesti et al., 2014). I used the reporter constructs provided by Mandel and validated the system in our cell line model. I also generated full-length and C-terminal truncated REST isoforms expressing Flag and eGFP tags. Both truncated isoforms were stabilised by the proteasome and were turned over rapidly in cells, suggesting the presence of a C-terminal-independent mechanism of destabilisation, which might be proteasome-dependent. Furthermore, the full-length REST construct is double the predicted size, like endogenous REST, whereas the truncated isoforms only show a minor difference in size from the predicted size. If the model I present in Chapter 4 is true, it will imply the presence of the dimerisation domain in the C-terminus of REST.

The next step would be the functional validation of the tagged isoforms using a DNA binding assay and investigation of their response to IGF2 treatment. Once it has been established whether the IGF2 response is dependent on the C-terminus or N-terminus of REST protein, the resultant region will be truncated further to generate small approximately 100 amino acid peptides. Domain boundaries, sequence conservation of REST and mass spectrometry data would be used as guidance here. This would enable honing down to the IGF2 phosphosite, assuming there is one. Since REST might have multiple phosphorylation sites with opposing effects, such a strategy of simplifying the system to verify phosphosites independently is essential. Compartmentalisation of phosphorylation can also be studied by fractionating REST peptide transfected cell extracts.

It might turn out that the previously identified phosphosites, including S861 and S864 which are phosphorylated by MAPK, and S1013, 1024, and 1027, which constitute two phosphodegrons mediating REST degradation, are responsive to IGF2. IGF2 signalling might prove to control REST stability via one of these sites. Hence, it would be prudent to mutate these sites and check if they abrogate REST response to IGF2.

Of the eight novel sites identified by MS, three were completely conserved amongst vertebrates. That and their high stoichiometry in cells indicate important function. These sites will form the basis of future for future work to fully characterise PTM of REST and their functional relevance in vivo.

In the end, the prospect remains that IGF2 might not directly phosphorylate REST, but mediate its effect indirectly through interacting partners of REST. For example, it might be acting to stabilise HAUSP potentially by phosphorylation, the deubiquitinase that stabilises REST, or conversely directly destabilising βTrCP, which targets REST for proteasomal degradation. Hence, an analysis of the known REST regulators and their response to IGF2 would be performed in the first instance, followed by MS identification of phosphorylation site of the resulting candidate if any.

6 TARGETED DELETION OF REST GENE IN MAMMALIAN CELLS BY CRISPR/CAS9 SYSTEM

6.1 Introduction

Gene knock down or knock out is widely used for genetic functional analysis. The most popular strategy for reducing a gene product is the process of RNA interference, which utilises short hairpin RNA (shRNA) or short interfering RNA (siRNA). These are processed into single-stranded RNA by the enzyme Dicer. Single-stranded RNA causes gene silencing by targeting the RNA-induced silencing complex (RISC) to complementary mRNA molecules, leading to cleavage by Argonaute, the active part of RISC, causing their degradation (Ketting, 2011). This results in post-transcriptional gene silencing without altering the genomic sequence. Gene silencing is not 100% as some functional RNA might remain and be translated, hence the term knockdown. This leads to a reduction in gene function but not complete elimination. Contrary to gene knock down, gene knock out results in complete elimination of protein levels.

This is achieved by directly targeting the genomic sequence and is either called genome engineering or genome editing.

6.1.1 Genome Editing

Genome editing is described as the ability to alter and manipulate DNA sequences in living cells precisely. Targeted genome editing is a powerful tool for understanding gene function, both in *vitro* and in *vivo* by the creation of animal models, gene therapy for genetic diseases and animal breeding for desired traits.

Several targeted genome-editing strategies have been created in recent years. Zinc-finger nucleases (ZFNs) were the first generation targeted endonucleases to be used extensively in plants and animals (Remy et al., 2010). Another type of engineered endonuclease is transcription activator-like effector nucleases (TALENs), which have been employed widely for targeting genome editing (Bogdanove and Voytas, 2011). Both tools use a strategy of tethering endonucleases with modular DNA-binding domains, the DNA-binding domains guiding the endonuclease to induce DNA double-stranded breaks (DSBs) at the targeted genomic locus. These are then repaired by the error-prone non-homologous end joining pathway (NHEJ) which results in insertions or deletions (indels) at the repair site, 66% of these would be expected to lead to a frameshift in the case of the target being an exonic sequence (Segal and Meckler, 2013). Both these technologies require engineering of proteins with reprogrammed cleavage sites for each application, making the process quite tedious. Furthermore, accessibility of the chromatin site adds another limitation. The off-target effects of both technologies are quite high, with ZFNs having higher chances due to smaller recognition sequences.

The latest addition to the genome engineering toolbox is the clustered regulatory interspaced short palindromic repeat/Cas (CRISPR/Cas) system. This system is based

on the bacterial and archaeal system of adaptive immunity against viruses and plasmids (Jinek et al., 2012; Horvath and Barrangou, 2010). The type II CRISPR/Cas9 of Streptococcus pyogenesis involves a single effector nuclease, which cleaves dsDNA. The specificity of DSB is mediated by two RNAs, CRISPR RNA (crRNA) and transactivating crRNA (tracrRNA). The crRNA sequences are transcribed from DNA sequences called protospacers, which are short sequences of foreign DNA interspersed with short palindromic repeats and clustered in the bacterial genome as a CRISPR array. The processing of transcribed pre-crRNA into crRNA is dependent on tracrRNA, which has a complementary sequence to the palindromic repeats. Both crRNA and tracrRNA bind Cas9 to activate guide it to cut specific non-host sequences. Due to the simplicity of the system, it has been developed into a genome engineering tool (Jinek et al., 2012).

TracrRNA and crRMA are combined to form a chimeric guide RNA (gRNA) and expressed simultaneously with Cas9 protein (Walsh and Hochedlinger, 2013). The gRNA binds to the target sequence via Watson Crick base pairing and localises Cas9 to the target sequence which is followed by a 5'-NGG PAM motif (Ran et al., 2013b). The resulting DSB is repaired by either NHEJ or homologous recombination (HR). HR requires a donor template in the form of either the wild-type allele or an exogenous template and can be used to insert sequences in the genome. The error-prone NHEJ creates insertions, deletions or frame-shift mutations in cells.

CRISPR/Cas9 system has numerous advantages over the other two techniques. These include the efficiency of targeting, the ability to target more than one sequence at once, low cost and simplicity due to the need of only a small guide RNA to be designed and the ability to target Cas9 to multiple sites by simply altering the gRNA sequence.

Since the guide sequence can tolerate a number of mismatches with the target DNA, this leads to undesired off-target effects. One strategy to improve specificity is by using a Cas9 D10A mutant, which can only induce single-strand nicks in DNA. Mutant Cas9 introduces a nick at the target site, and two appropriately placed target sites can lead to a double-strand break, which can be efficiently repaired by the cellular DNA repair machinery. This improves Cas9 targeting specificity and reduces off-target effects by 50 to 1500-fold in cell lines since the number of specifically recognised bases is increased (Ran et al., 2013a).

CRISPR/Cas9 technology is a rapidly growing field, and subsequent to these experiments, it has been modified into a sequence-specific tool to regulate gene transcription. Named CRISPR interference (CRISPRi) and CRISPR activation (CRISPRa), these technologies can be used for genome-wide transcriptional repression and activation, respectively (Dominguez, Lim and Qi, 2016).

6.1.2 REST knock down and knock out studies

Animal models have been used to show the essential role of REST. Loss of function of REST leads to embryonic lethality as demonstrated by homozygous knockout mice (Jones and Meech, 1999). Disruption of REST gene expression in zebrafish using zinc finger nucleases, however, reveals normal neurogenesis for the most part in mutants (Kok et al., 2012). A second study, which employs REST conditional knockout mice, demonstrates that mice lacking REST in the brain grow into normal adults and concludes that REST is expendable for embryonic neurogenesis (Aoki et al., 2012). Nevertheless, it plays a vital role is suppressing neuronal gene expression in cultured murine neuronal cells as well as in non-neuronal cells in *vitro*.

Transient knock down of REST has also provided valuable insight into REST function. Decreased REST expression is accompanied by transformation of human

mammary epithelial cells (HMECs) (Coulson et al., 2000; Wagoner et al., 2010; Westbrook et al., 2005) as well as oncogenic transformation of immortalised primary human cells (Mulligan et al., 2008). In glioblastoma cells, REST knockdown leads to reduction of the self-renewal potential of cells and triggers neuronal differentiation and cell death (Conti et al., 2012). Similarly, REST knockdown in mouse embryonic stem cells (mESCs) was shown to reduce self-renewal potential and upregulation of differentiation markers (Singh et al., 2008). In human embryonic stem cells (hESCs), however, REST was found to be unessential for maintenance of self-renewal but was potentially related to regulation of lineage differentiation (Thakore-Shah et al., 2015).

Prior to this study, no stable human REST KO cell has been generated. Stable cell lines have numerous benefits over transient methods. The efficiency of gene knock out can be 100%. Hence the effect of complete elimination of a protein can be studied. Furthermore, since stable KO is permanent, it allows the study of long-term effects of gene elimination. To analyse the functional effect of REST complementation mutants in cells, it is essential to remove endogenous REST. Since RNAi REST knockdown, using either siRNA or shRNA, is transient and not 100% efficient, remaining endogenous REST would interfere with proper functional analysis. Thus, it is a poor strategy for eliminating endogenous REST.

6.2 Aims

The purpose of this chapter is to evaluate knock out of *REST* gene, both to define functional consequences and to provide a system for genetic complementation of mutant forms of REST.

6.3 Results

6.3.1 Design and construction of CRISPR/Cas9 system for REST editing

REST is located on human chromosome 4 and is composed of three alternative non-coding exons, each with a different promoter, three coding exons, and an alternative exon, associated with disease-relevant and neural splice variants of REST (Figure 6.1A). To achieve a complete knockout of REST, the first coding exon was selected for CRISPR targeting. The Zhang tool at http://crispr.mit.edu was used for sgRNA design. We decided to use a double nickase strategy, which employs a D10A nickase mutant of Cas9. Mutant Cas9 introduces a nick at the target site, and two appropriately placed target sites can lead to a double-strand break, which can be efficiently repaired by the cellular DNA repair machinery. This improves Cas9 targeting specificity and reduces off-target effects since the number of specifically recognised bases is increased. Two pairs of sgRNA were designed using the Zhang tool, both targeting the first coding exon near the start site (Figure 6.1B). These were cloned separately into pSpCas9n(BB)-2A-GFP (PX461), which simultaneously expresses the Cas9 nickase mutant, sgRNA, and GFP.

To test the cleavage efficiency of the sgRNA sequences chosen, the paired expression vectors were co-transfected in HEK293T cells. Genomic DNA was extracted from cells 48 hours post-transfection for analysis using high resolution melt analysis (HRMA). HRMA involves a Real-time PCR amplification step in the presence of saturation dyes that fluoresce only in the presence of double-stranded DNA. Following DNA amplification, the sample is heated and DNA fluorescence measured as a property of temperature, generating a melt curve. Altered genetic sequences can result in different melting curves, allowing detection of mutations (Figure 6.2A).

HRMA was chosen due to its ability to detect even small base pair changes and simplicity in assaying a large number of samples.

Both sgRNA pairs were found to shift the melt curves, with pair 2 being more efficient with a confidence interval of 25-28% whereas Pair 1 was shown to cause a difference with a confidence interval of 76-82% (Figure 6.2B).

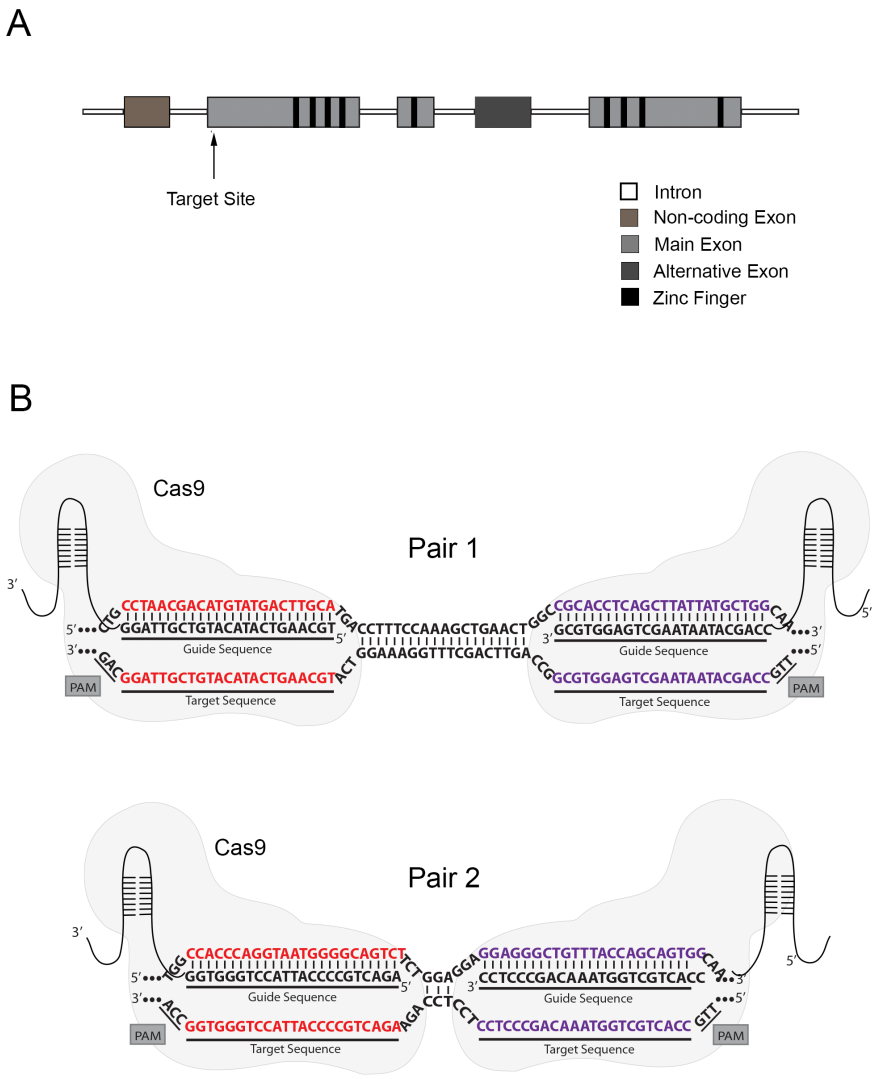
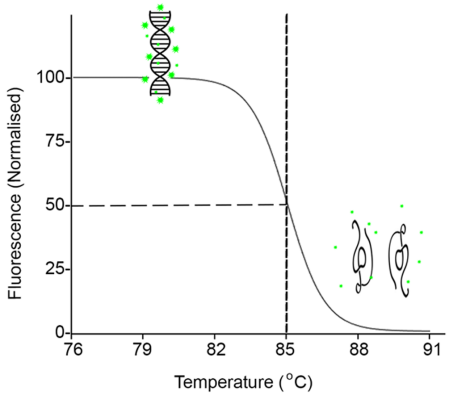
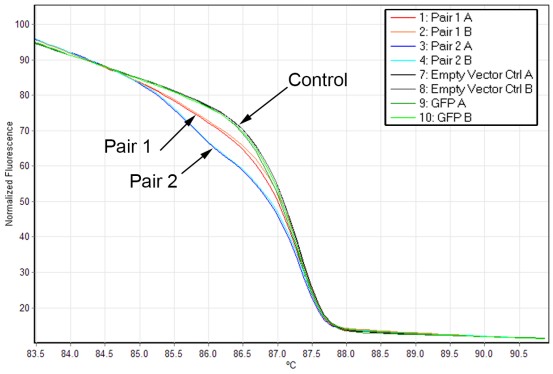


Figure 6.1 Targeting Strategy for CRISPR/Cas9 mediated REST knock out.
A) A schematic showing REST gene with three main exons, an alternative exon and one of three alternative non-coding exons. The target site for Cas9/sgRNA binding is shown. B) Two sgRNA pairs chosen are depicted here along with their binding sites and PAM sequence.

A



B



No.	Color	Name	Confidence %
1		Pair 1 A	76.5
2		Pair 1 B	82.18
3		Pair 2 A	25.22
4		Pair 2 B	28.68
7		Empty Vector Control A	95.08
8		Empty Vector Control B	96.00
9		GFP A	100.00
10		GFP B	98.26

Figure 6.2 Validation of sgRNA for CRISPR.
A) Schematic illustrating the principle behind HRMA. HRMA measures loss in DNA fluorescence as double-stranded DNA is converted to single-stranded DNA, generating a melt curve. Altered genetic sequences can result in different melting curves and melting temperatures allowing detection of mutations. B) HMRA analysis of two sgRNA pairs. Numbers indicate pairs and letters indicate repeats. Curves show that Pair 2 is more potent at inducing mutations.

6.3.2 Generation of Generation of Heterozygous REST KO mutants

6.3.2.1 Methodology

Three cell lines were chosen for editing of REST. SK-N-MC cells were selected considering their IGF2 dependency of REST and the EWS-FLI1 dependency in Ewing Sarcoma (Sankar et al., 2013). MCF7 cells were chosen to explore the dependency of breast carcinoma on REST, and HEK293T cells were used as a positive control due to their relatively easier genetic manipulation.

Cells were co-transfected with either of the Cas9-sgRNA expressing vector pairs. At one day post-transfection, cells were FACS sorted for GFP fluorescence. GFP-positive cells were grown in bulk culture, allowed to recover for an additional week and subsequently single-cell cloned into 96 well plates. Clones were expanded to 12-well plates over the next 1-3 weeks. Plates were then replicated, cells from one plate frozen down and cells from the other used for isolation of genomic DNA. The clones were assessed for mutation by using HRMA. Clones that were identified as mutant were verified by PCR analysis using primers spanning the target sites. WT clones showed one band on an agarose gel whereas mutant clones showed a wild-type band and a second band resulting from an insertion or a deletion. Individual alleles of mutant clones were separated by running out on an agorase gel, gel extracted, purified and subjected to Sanger sequencing. Clones resulting in mixed sequences were TA cloned into a pGEX vector and five colonies from each clone were sequenced. Clones showing one band on PCR were also sequenced to avoid missing mutations too small to be detected on a gel. A detailed workflow of the method is shown in Figure. 6.3.

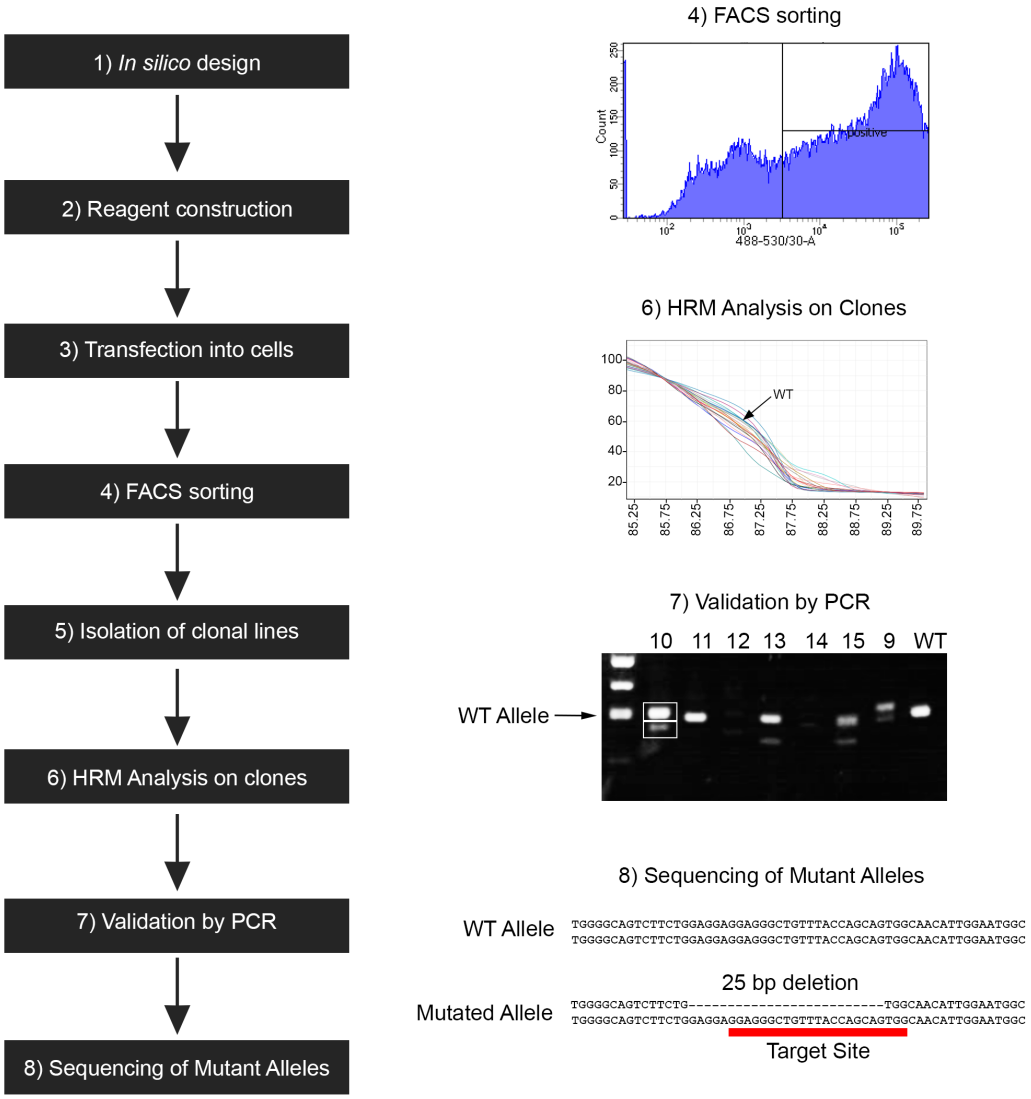


Figure 6.3 Methodology for creating REST null cell lines
Left: a workflow of the strategy used for generation of REST null cell lines. Right: representative examples of some steps. 4 shows FACS sorting of HEK293T cells, 6 shows HRMA analysis on HEK293T clones, 7 shows validation of promising clones by PCR analysis, and 8 shows sequences of two alleles of clone 10.

6.3.2.2 Highly efficient large fragment deletion in REST using CRISPR/Cas9

I analysed 29 HEK293T clones, 39 SK-N-MC clones and 33 MCF7 clones for mutations in REST sequence. Of the 29 HEK293T clones analysed, 11 were found to be positive for indels, an efficiency of 38%. Similarly, 21 out of 39 SK-N-MC clones screened were identified as mutated, a targeting efficiency of 54%. Contrary to this, no mutation was detected in all of 33 MCF7 clones tested.

Analysis of the indels showed that they were present in only one allele and that deletions were more common than insertions across HEK293T and SK-N-MC and ranged from anywhere between 3–90 bp. These indels were found to lie around the guide RNA binding sites. The sequences of 4 SK-N-MC and 4 HEK293T clones with indels leading to frameshift mutations and stop codons are showed in Figure 6.4. HEK293T clone 15 and 20 and SK-N-MC clone A9D were chosen for further analysis.

A

Cell line	Clones screened	WT	Heterozygous Mutations		
			Insertion	Deletion	Total
HEK293T	29	18	3	8	11
SKNMC	39	18	6	15	21
MCF7	33	33	-	-	0

B

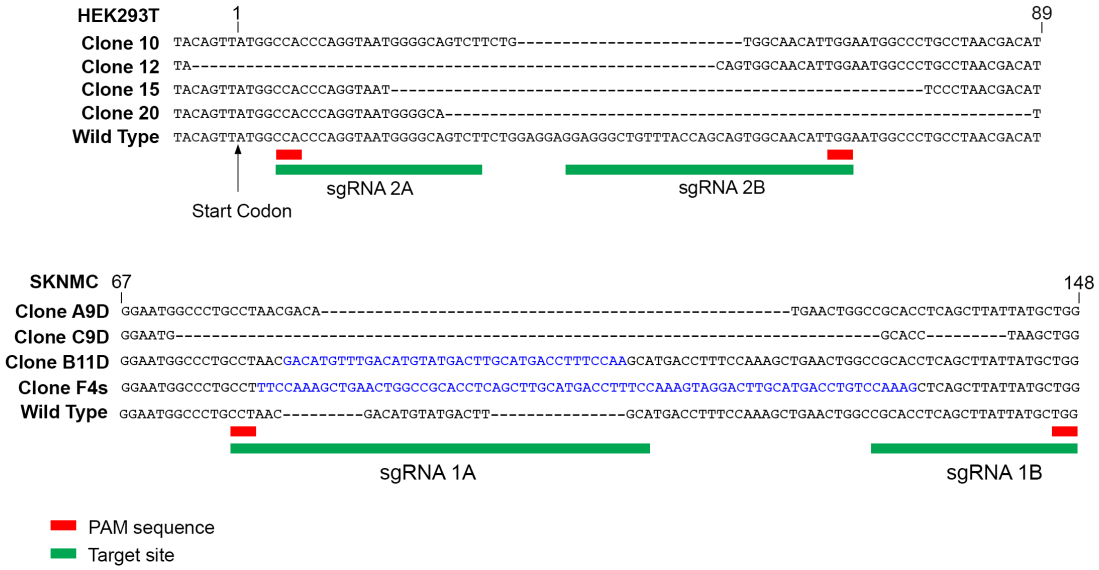


Figure 6.4 Generation of Heterozygous REST mutants.
A) Table showing the number of clones screened. B) Sequences of mutant allele for 4 HEK293T and SK-N-MC heterozygous REST clones. Binding sites for sgRNA are shown in green and PAM sequence in red.

6.3.3 Validation of Heterozygous REST KO mutants

Immunoblotting of the clones showed no difference in REST levels suggesting that both HEK293T and SK-N-MC are not haploinsufficient for REST (Figure 6.5A). The clones were further verified by a reporter assay as well as repression of a REST target gene.

6.3.3.1 RE1 Reporter Assay

A reporter assay was employed for analysing REST activity in the heterozygote clones. This assay was developed by the Perrier lab (Charbord et al., 2013) and consists of an RE1-plasmid containing 12 RE1 sites followed by the PGK promoter controlling the expression of *Renilla* Luciferase (Figure 6.5B). REST binding to RE1 vectors represses the expression of *Renilla* Luciferase. A control vector (RE1m/RE1mutated) has six residues of the consensus RE1 sequence, scoring the highest position-scoring-matrix (PSM), mutated to abrogate REST binding, resulting in de-repression of *Renilla* Luciferase.

Cells are transfected with either RE1 or RE1m plasmids along with pGL4.73 vector that expresses *Firefly* Luciferase. Luciferase signal was measured 48 hours post transfection and normalised to firefly luciferase to control for transfection efficiency and cell viability. Wild type HEK293T and SK-N-MC both showed lower luciferase signal in cells transfected with RE1 plasmids, suggesting that REST mediates repression of the luciferase expression cassette through the RE1 sites. This suggested the RE1-mediated repression of the luciferase expression cassette by endogenous REST in transfected cells. A de-repression value resulting from RE1m/RE1 signal can be used as a measure of REST levels in cells. Higher REST levels would result in a lower RE1 signal, and a higher RE1m/RE1 value and conversely, lower REST levels would result in lower RE1m/RE1 value.

Next, the reporter assay was performed in the heterozygote clones and results are shown in Figure 6.5C. Of the two HEK293T heterozygote clones tested, clone 15 was found to have the same de-repression, but there was a slight reduction in the de-repression signal in clone 20 ($p<0.05$), indicating lower REST function. SK-N-MC heterozygote clone tested showed no difference in de-repression.

6.3.3.2 L1Cam De-repression

L1Cam, neuronal cellular adhesion molecule, was identified to be de-repressed by REST knockdown in NIH3T3 cells by a qPCR screen for known REST target genes (See Appendix 7). De-repression of REST resulted in a 15-fold increased in L1Cam mRNA. This and other studies have shown L1Cam to be a significant REST target (Mikulak et al., 2012). Decrease in REST levels would be mirrored by increase in L1Cam levels. L1Cam is a cell surface receptor and can be stained relatively quickly using FACS. Hence, L1Cam was used as an output for measuring REST levels in my heterozygote clones.

Staining of HEK293T heterozygote clones show an increase in L1Cam levels in both clones, increase being more pronounced in clone 15 (Figure 6.5D). SK-N-MC clone A9D was found to have the greatest effect on the increase in L1Cam levels (Figure 6.5D). These data suggest that although REST protein levels might not be different in these clones by Western blotting, REST activity appears altered at least for L1Cam.

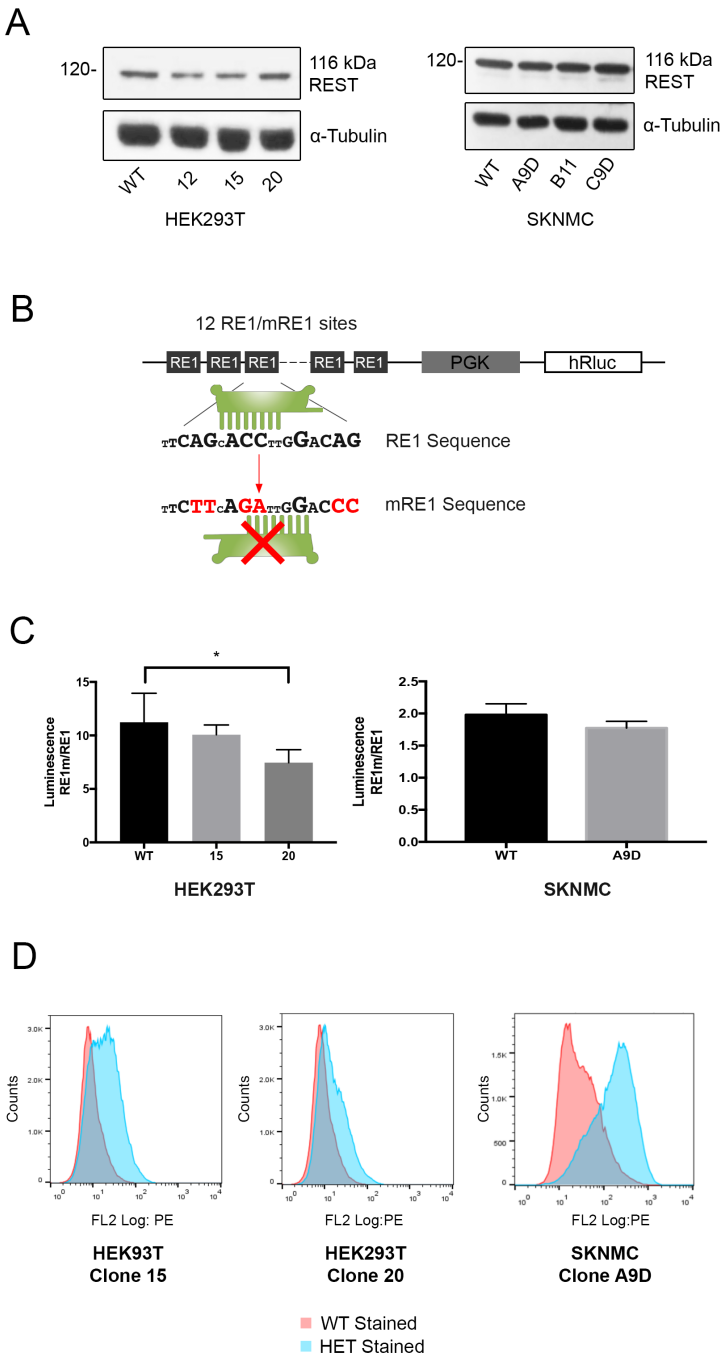


Figure 6.5 Validation of Heterozygous REST mutants.
A) Immunoblotting of heterozygous clones using anti-REST (17-641) and anti-α-tubulin antibodies. B) A schematic showing REST reporter with RE1 and mutated RE1 sequences. C) REST activity in wild-type and heterozygous clones represented as de-repression of luciferase activity. Error bars represent mean ± SEM ($n = 4$). *: $p < 0.05$ by unpaired two-tailed Student's *t*-test. D) L1Cam labelling in heterozygous clones compared to wildtype and analysed by one-dimensional flow cytometry.

6.3.4 Usage of X5050, a putative REST inhibitor, for achieving partial REST knock down

One way of reducing REST levels in cells is by using a REST inhibitor. One such inhibitor, X5050, was identified by high-throughput screening of a library of compounds (Charbord et al., 2013). X5050 was found to decrease REST levels by targeting its degradation, the mechanism of action is however not known. Neural cells treated with X5050 were shown to upregulate neuronal genes by differential transcriptomic analysis (Charbord et al., 2013).

Heterozygote REST clones have one functional copy of REST. However, levels of REST protein are unchanged suggesting a possible mechanism of autoregulation. Using X5050 in these heterozygous REST cells might lead to a greater decrease in REST levels than wild-type cells owing to the presence of one functional REST copy.

Wild-type HEK293T and heterozygote clones 15 and 20 were treated with varying doses of X5050 for 24 hours. Whole cell lysates were immunoblotted for REST (Figure 6.6A). Immunoblotting shows that REST levels drop in wild-type cells to almost 50%. However, REST levels in heterozygote clones remain unchanged, suggesting that autoregulation bypasses the degradation mechanism employed by X5050.

To investigate if X5050 treatment leads to a difference in cell proliferation or differentiation, cell viability and differentiation were both measured 48 hours after X5050 treatment, using an MTA assay and L1Cam staining respectively (Figure 6.6B and C). A difference in proliferation was observed in both wild type and heterozygotes, but this difference was not correlated with REST levels, suggesting a non-specific effect of the drug. No change in differentiation was seen in either wild type or heterozygote cells. The dosage required to detect changes in cells appears greater that 100 μM, suggesting toxicity and potential indirect effects independent of

REST. Paradoxically, a REST knock out model could test X5050 REST-independent effects.

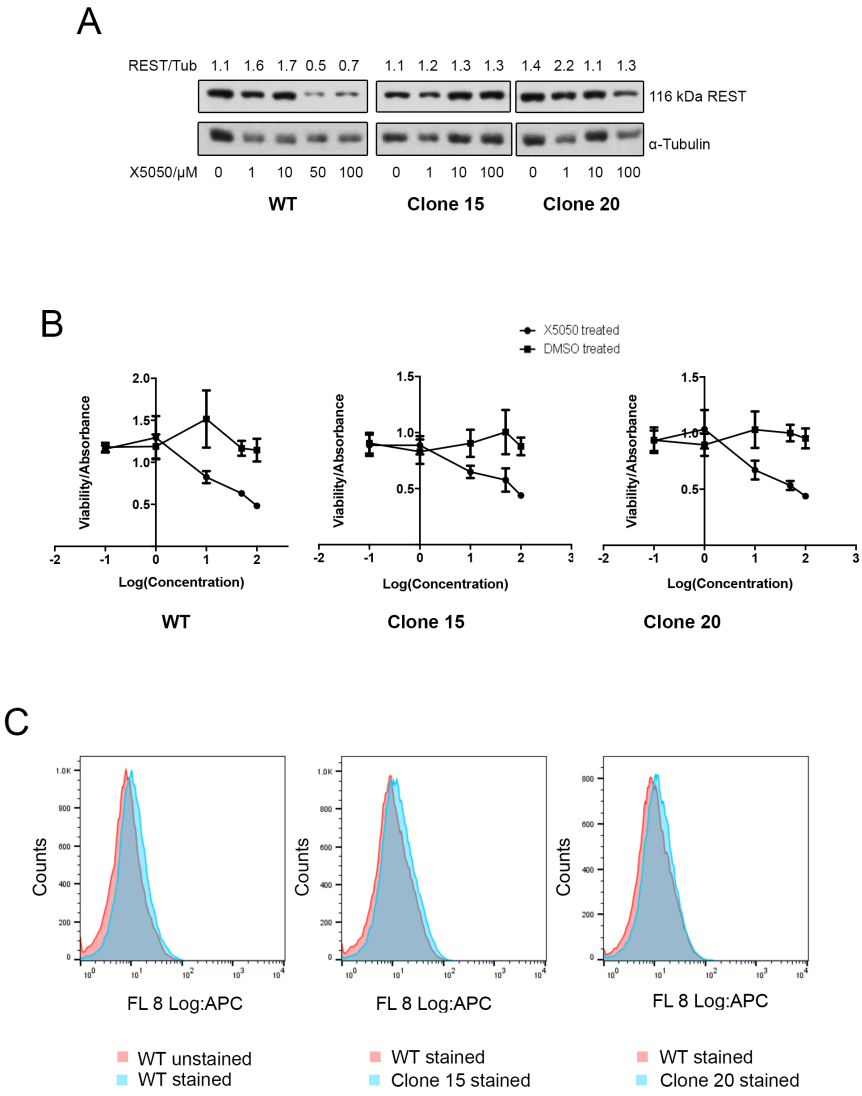


Figure 6.6 Using X5050 to achieve transient REST knockdown in HEK293T.
A) REST levels in wild type and heterozygous clones after treatment with different doses of X5050 are shown. REST levels were quantified by Image J, normalised to α-tubulin levels and indicated above the blots. B) Decrease in cell proliferation in X5050 versus DMSO-treated cells. Error bars represent mean ± SEM ($n = 3$). $p < 0.05$ C) No change in L1Cam expression in X5050 treated wild type or heterozygous cells.

6.3.5 Generation of Homozygous REST mutants

Autoregulation of REST levels in heterozygote clones and failure to achieve a knockdown using X5050, led to the need to generate homozygote REST mutants. This was achieved by making use of the heterozygote clones, and repeating the process of transfecting Cas9/sgRNA and selection of mutant clones (Figure 6.7A). HEK293T heterozygote clone 15 and 20 were chosen because they lacked the binding site for one of the guide RNA, sgRNA 2A. Hence, the mutant allele, would not be able to bind both Cas9, resulting in only one nick and preventing a double-strand break. This would leave only the wild-type allele available for editing.

An observation from previous experiments was the essential requirement of a selection strategy to enrich for genetically modified cells. Hence, a selection strategy for REST knock out cells was also developed based on REST function.

6.3.5.1 Killer Strategy- Generation of a functional selection strategy

The 'killer' strategy was based on a selection plasmid, which has the sequence of puromycin resistance gene, puromycin-N-acetyl-transferase (PAC), under the control of 12 RE1 sites followed by PGK promoter, named the Killer Vector (Figure 6.7B). Lack of REST activity in REST KO cells would lead to de-repression of the PAC gene, making cells resistant to puromycin whereas REST producing cells would be vulnerable. This would allow for selection of REST KO cells only.

This construct was tested by mutating the RE1 sites to check if that leads to de-repression of PAC and makes cells resistant to puromycin (mutKiller Vector). HEK293T cells were transfected with either Killer or mutKiller vector, treated with different puromycin doses after 48 hours and cell viability measured after another 24 hours (Figure 6.7C). The mutKiller-transfected cells were found to be more resistant to puromycin, validating the strategy. The optimal dosage of puromycin was found to

be 1 μg/ml. This dosage would allow for maximal survival of KO cells and maximal killing of wild type cells

6.3.5.2 Inability to detect homozygous REST mutants

HEK293T heterozygous REST clone 15 and 20 were re-transfected with plasmids expressing Cas9 and sgRNA 2A/2B and FACS sorted for GFP 24 hours post-transfection. Cells were allowed to recover for a week and then transfected with Killer vector. 24 hours post-transfection, cells were exposed to puromycin selection for another week. Surviving cells were grown and subsequently single-cell cloned into 96 well plates. Expanded clones were then tested for mutation in REST using PCR analysis.

45 clones from each heterozygous cell line were analysed for REST knock out using PCR analysis and Western blotting. I failed to identify a knockout. This suggests either of three possibilities, 1) the efficiency of homozygous knockout is very low, 2) REST knock out is lethal to cells, or 3) REST knockout leads to alteration of cell phenotype, possibly growth or differentiation. All three possibilities make it difficult to select REST null cells in growth conditions.

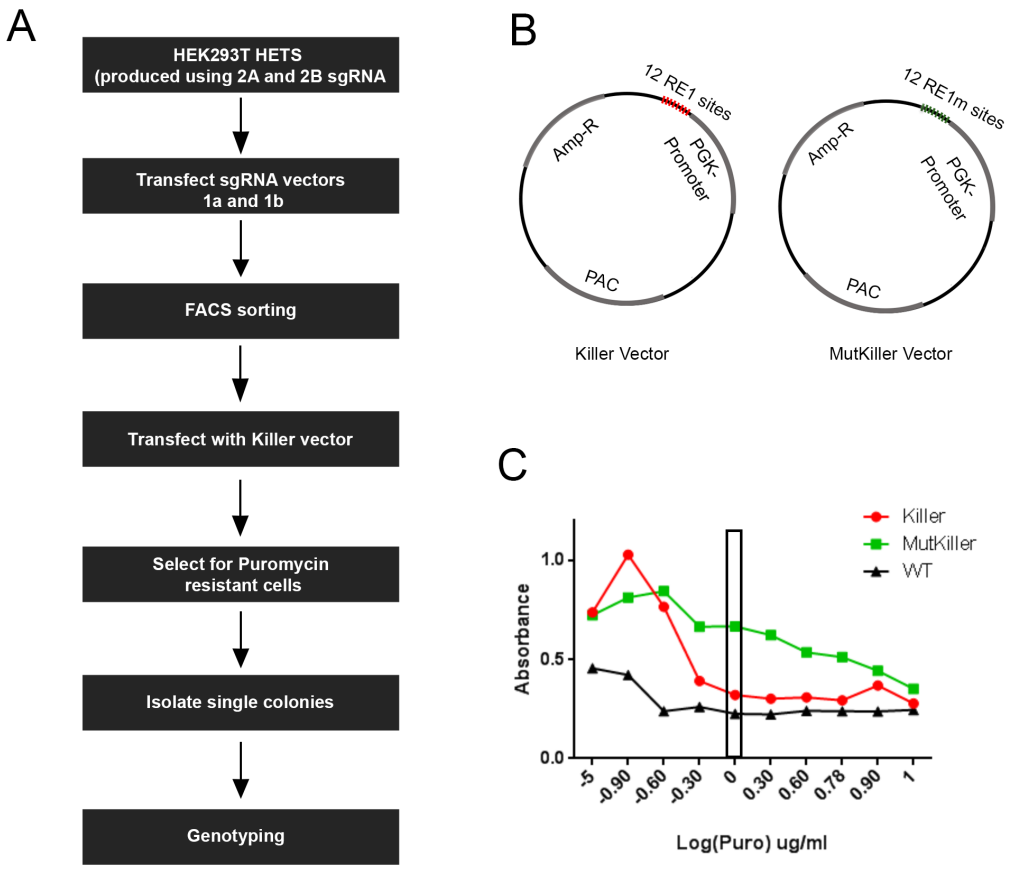


Figure 6.7 Killer Strategy- development of a functional selection strategy.
A) A workflow of the strategy used for generation of REST null cell lines. B) A depiction of the scheme of Killer and mutKiller plasmids used for selection of REST null cells. C) Validation of the selection strategy. HEK293T transfected with either Killer or mutKiller vectors were treated with puromycin for 24 hours and cell viability measured using MTA assay. Optimal dosage for selection with the maximal difference in cell viability is indicated with a box.

6.4 Discussion

This study describes an attempt to use the CRISPR/Cas9 technology to achieve a REST knock out cell line. The findings in this chapter indicate that gRNAs targeted to the first coding exon of REST gene cause sequence-specific NHEJ, mediated by Cas-9, resulting in the generation of indels with an efficiency ranging from 38-54%. These indels can cause frameshift mutations and lead to the generation of heterozygous REST knockout cells.

The heterozygous KO cells in both cell lines have similar levels of REST to wild-type cells and are functionally not haploinsufficient. This suggests a mechanism of autoregulation returning REST levels to wild-type levels and implies strict control of REST levels. To our knowledge, this study presents the first demonstration of this phenomenon. This autoregulation might be either transcriptional or translational, and further characterisation of heterozygous KO cells is needed to delineate the autoregulatory pathway. These results are in contrast to studies in mESC cells, which report a decrease in REST levels in REST +/- ESC cells (Singh et al., 2008; Buckley et al., 2009). The difference might be explained by the use of immortalised cell lines instead of ESC.

The failure to generate any mutations in MCF7 cells was surprising. To exclude the possibility of SNP/SNPs in the gRNA binding regions, which interfere with binding, the first coding exon of MCF7 was sequenced and no SNP was identified. It might be likely that mutations in genes essential for the pathway, especially those forming the NHEJ machinery, makes the cell line unamenable to gene editing via CRISPR. Furthermore, no precedence of CRISPR application in MCF7 was found in the literature.

One of the biggest challenges in CRISPR application is the selection of genetically modified cells (Ramakrishna et al., 2014). The efficiency of CRISPR can be quite low, depending on the sequence of the gRNA used. This results in the need to screen multiple clones to identify a positively edited clone. Hence, methods to enrich mutated cells can simplify the process significantly. To this effect, a selection strategy was created that relied on the sequence-specific binding and transcriptional repression activity of REST. Despite this, it was not possible to identify a REST null clone.

The targeting efficiency of one allele was found to vary from 38-54%. Keeping this in mind, the chances of both alleles to be mutated at random are 14-29%. Furthermore, a comparison of deletion frequencies across 14 gRNAs targeting non-essential genes shows that the observed biallelic deletion frequency exceeds probabilistic expectation (Canver et al., 2014). However, this probabilistic estimation can only be true if other biases in the system are discounted.

Failure to select REST null can be due to numerous possibilities. It might be possible that REST knock out is lethal to cells or induces apoptosis. This is supported by the importance of REST in embryonic development, REST knock out mice being embryonically lethal. Furthermore, expression of VP16-REST, a recombinant form of REST, which has repressor domains replaced by activator domains of a viral activator and activates REST target genes, has been shown to induce apoptosis in medulloblastoma (Fuller et al., 2005). Contrary to this, some studies report generation of REST -/- mESC (Yamada et al., 2010; Jorgensen et al., 2009). A second possibility is a decrease in proliferation and growth. Thirdly, there is a potential for REST KO inducing a neuronal phenotype, resulting from de-repression of genes. This is supported by REST null ESC found to de-repress REST target genes in several studies as well as REST knockdown leading to up regulation of neuronal genes

(Yamada et al., 2010; Jorgensen et al., 2009; Conti et al., 2012; Singh et al., 2008). This is especially true for the Ewing cell line SK-N-MC, which has a neuronal origin. Both of the above possibilities would explain the difficulty in obtaining REST null cells by screening. Indeed, during the screening of clones, I noticed different growth curves of clones; a few clones were extremely slow to grow and died before they could be transferred to larger wells.

These possibilities mentioned above are all equally probable. Since REST effects and function varies drastically from cell line to cell line, different mechanisms might be responsible for different cell lines. To identify which of these possibilities is true, one strategy would be to use RNAi knockdown of REST. REST knockdown in the cell lines followed by an investigation of cell phenotype including proliferation and neuronal differentiation would show whether depletion of REST leads to slowing down of growth or a neuronal differentiation phenotype. A conditional knockdown like the one used by Thakore-Shah et al. is essential to avoid the lethality phenotype and to avoid cells compensating for the mechanism (Thakore-Shah et al., 2015). Lethality induced by REST knock out can be investigated by generating a stable cell line with inducible Cre-LoxP REST for CRISPR. Cell death following floxing out of REST would confirm that REST knockout is lethal to cells. Another options for functional analysis of REST knock out would be to use CRISPRi to achieve genome-wide transcriptional repression of REST target genes.

The use of CRISPR for genome editing raises an important concern for off-target effects (Ran et al., 2013a). In the present study, although attempts were made to decrease these off-target effects by using a paired nickase strategy to increase specificity, some off-target effects might still be possible. These might be a reason for phenotypic differences noticed in some clones. There is a need for whole-genome

sequencing of resulting clones to determine the non-specific activity of gRNAs as well as to identify off-target indels.

Finally, changes in gene dosages are often accompanied by compensatory mechanisms, leading to profound phenotypic differences between genetic mutations and gene knockdowns, pointing towards the importance of conditional knockouts/knockdowns over shorter periods (Rossi et al., 2015). Furthermore, the consequence of knockouts can change over passages significantly (Roy et al., 2015). These observations bring home the need for caution when interpreting gene knockout studies.

7 DISCUSSION

7.1 Summary and Significance

IGF2 is a mitogenic peptide that is crucial for normal fetal growth (Foulstone et al., 2005). IGF2 dosage supply is different in the somites and the choroid plexus, and this is achieved by genomic imprinting, a mechanism established in the germ line (DeChiara, Robertson and Efstratiadis, 1991). Very little is known about the partition of IGF2 expression between the neuronal and non-neuronal expression, and its implication.

Lately, the hypothesis that the IGF system has a role in neurogenesis has gathered a lot of support (Ferron et al., 2015). Dosage control of IGF2 has been shown to be essential in regulating the different neurological niches of the adult brain. IGF2 imprinting has been demonstrated to be relevant for its function as bi-allelic expression is essential for SVZ neurogenesis whereas the hippocampus makes use of lower levels of IGF2 for autocrine signalling (Ferron et al., 2015). IGF2 dose supply has also been implicated in growth and development. Consequently, deregulation of IGF2 dose supply results in growth disorders, such as Russell-Silver and Beckwith-

Wiedemann syndromes, or neurological syndromes such as medulloblastoma (Morison and Reeve, 1998; Corcoran et al., 2008; Hur et al., 2016).

The present work was initiated to identify genes expressed downstream of IGF2 which are differentially regulated by IGF2 genetic dose supply, and thus might control IGF2 signalling outputs. RNA expression analysis from embryos with altered allelic dosage of *IGF2* leads to changes in IGF2 peptide dose supply. Moreover, a cell line model, exhibiting autocrine IGF2 signalling, resulted in the identification of differentially expressed genes which are also IGF2 dosage dependent. REST was identified as a key hub transcription factor controlling the expression of these differentially expressed genes, and exerting changes in expression at least 24-48 hours after IGF2 exposure. Further molecular analysis was used to verify this finding. REST protein was found to be stabilised within 6 hours of exogenous exposure to IGF2, in a proteasome and PI3K-dependent manner. The response was also found to be dependent on the cell lines used, the timing of the treatments and feedback signaling, which appears to be a common theme in signaling pathways. An elegant mechanism of REST control was proposed involving REST de-phosphorylation causing increased stability of REST, thus increasing global REST levels leading to an increase in the nuclear and potentially functional form of REST.

REST is an important master regulator of neurogenesis, with an emerging role in tumourigenesis. The growth factor IGF2 is proposed to stabilise REST by de-phosphorylation, preventing its ubiquitination and degradation by the proteasome. Increased REST levels, in turn, lead to cell proliferation and repress neuronal differentiation. Blocking IGF2 and/or REST blocks cell proliferation, inducing the cell to undergo neuronal differentiation. This hypothesis partly explains why it is important not to have IGF2 expressed widely in the adult brain, as stabilised REST

would result in the inhibition of differentiation and stimulation of proliferation. Importantly, stabilised REST is widely observed in brain tumours. Hence, I have identified a potential key regulatory switch, which might regulate a balance between growth and neuronal differentiation (See Figure 7.1). Further evaluation of this model could be carried out by ectopic expression of IGF2 in the brain and evaluation of the resulting phenotypes, followed by rescue experiments with knock out of REST to prove the REST-dependency of the mechanism.

The proposed mechanism also has implications for cancer as REST has been shown to have an emerging role in tumorigenesis, with both oncogenic and tumour repressor functions (Westbrook et al., 2005; Su et al., 2006; Lawinger et al., 2000). The data in this thesis also points towards an oncogenic function. Cancers with deregulated IGF2 levels, frequently by LOI, might potentiate its effects by increased REST levels causing proliferation. This has a therapeutic implication for tumour cells. A way to prevent tumour growth would be to block REST to inhibit cell proliferation and induce neuronal differentiation. By inference, the inability to select REST null clones might be related to this switch to differentiation.

Recent studies have provided further evidence to support my findings. Wilms' tumour, the most frequent childhood renal cancer, arises from embryonal cells with dysregulated growth and differentiation during development (Rainier et al., 1993). It often occurs in patients with Beckwith-Wiedemann syndrome at a very high rate and IGF2 overexpression in combination with other mutations like Wilms' tumour 1 (Wt1) have been shown to have a driver role in tumour pathogenesis (Hu et al., 2011). Whole exome sequencing of Wilms' tumour cases identified 11 mutations in *REST* gene in four familial and nine non-familial cases, accounting for ~2% of Wilms' tumour cases (Mahamdallie et al., 2015). Ten of these mutations were clustered in the

DNA-binding domain of REST and resulted in a partial or complete loss of function of REST. It was proposed that LOF of REST might either de-repress neural or neuro-ectodermal transcriptional genes in developing kidney or primordial tissues, predisposing to Wilms' tumour. This dysfunction might be directly or indirectly linked to overexpression of IGF2, providing further impetus for my research.

To my knowledge, this is the first study that systematically characterises the HMW form of REST. My data points towards a dimerisation mechanism mediated by a covalent linkage. The dimer form of REST appears to be the functional form of REST present in the nucleus. This has important implications as most previously published studies fail to mention the size of the immune-reactive REST band being studied. The band size being studied might turn out to be a variability explaining some paradoxical observations made in REST functional studies, emphasising the need to employ caution when interpreting previous data.

This study also describes the first ever attempt that I am aware of for the use of CRISPR/Cas9 technology to achieve a REST knock out. Using sgRNA targeted to the first coding exon of REST, I report an efficiency of 38-54% in indel generation, resulting in frame-shift mutations and heterozygous REST knock out cells in both HEK293T and SK-N-MC cells. These heterozygous cells appear auto-regulated, suggesting the presence of feedback mechanisms of REST control, such as REST control of PI3K/AKT signaling identified by Westbrook et al. (Westbrook et al., 2005). The failure to screen REST null homozygotes indicates either knockout-related cell lethality or a differentiation phenotype, which needs further validation.

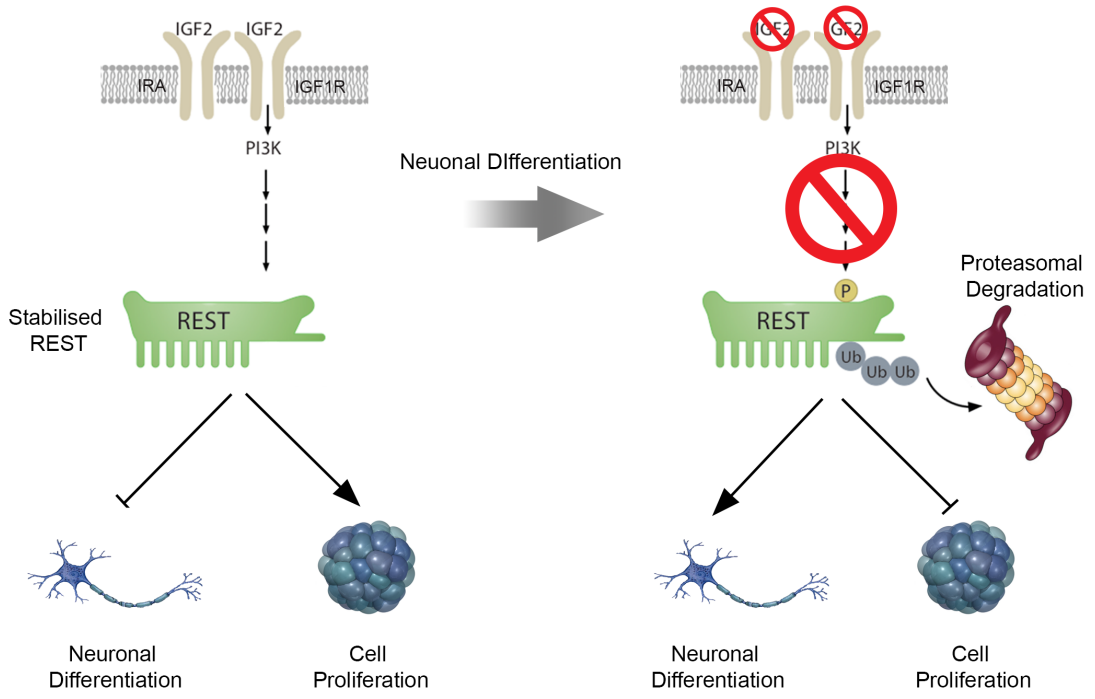


Figure 7.1 Proposed Control of Cell Proliferation and Differentiation by IGF2.
The growth factor IGF2 stabilises REST by modifying de-phosphorylation it, preventing its ubiquitination and degradation by the proteasome. Increased REST levels lead to cell proliferation and represses neuronal differentiation. Blocking IGF2 results in loss of destabilising REST via the proteasome, blocking cell proliferation and inducing the cell to undergo neuronal differentiation.

7.2 Future Directions

7.2.1 REST Truncation Mutants for Phosphorylation Studies

One observation from my studies was the inherent complications in the REST system, varying from cell-type effects to temporal, dose-dependent and isoform-dependent outcomes. This is a direct corollary of REST being a master regulator gene and hence the need to integrate various signals from multiple signalling pathways, which might be context- and cell type-dependent. Also, REST has a relatively short half-life, and its turnover is strictly regulated. A further complication arises from its role in the cell cycle and regulation of its phase-dependent protein levels.

Therefore, there are probably multiple mechanisms that control REST stability, in addition to the ones already identified. There are multiple phosphorylation sites identified by my mass spectrometry analysis in REST which might have complementary or opposing regulatory effects combined with potential compartmentalisation of the phosphorylated and un-phosphorylated states of REST. Consequently, my Mass Spectrometry approach failed to identify the REST phosphorylation site responsive to IGF2; the high coverage required being a major factor. Hence, it was found imperative to simplify the strategy for identification, and one option included dividing REST protein into smaller regions. There is precedence in literature for the success of this strategy (Nesti et al., 2014). The starting point of this was investigating whether the IGF2 effect is localised in the N or C-terminus. I have constructed constructs with three different isoforms of REST tagged with N-terminal Flag and C-terminal GFP tags, which will be used for this purpose. This will be followed by further analysis of the region identified, by generation of smaller peptide fragments and their investigation.

If one or a few PTMs that respond to IGF2 signalling are identified, they will be confirmed by mutagenesis and abrogation of REST response to IGF2. Subsequent identification of proteins involved in marking PTM will be carried out, first by attempting to pull down interacting proteins in a phosphorylation and IGF2-dependent manner. Here, the aim will be to pull out Flag-REST-protein complexes using normal and mutant protein respectively and characterising complexes using MS. The two groups of proteins can then be compared to identify the stabilising partner of REST. Candidates may be confirmed by knockdown studies to observe differential effects on REST stability. Secondly, consensus phosphorylation sites may give us information about the involvement of kinases or phosphatases in the process. The specific either kinase or phosphatase involved may be identified by a siRNA/CRISPRi screen of known kinases and phosphatases, using a phosphorylation-dependent antibody, or a functional screen based on luciferase expression. If, however, no stabilising PTM is found on REST, then other candidates will be sought. These might include de-ubiquitinases like HAUSP and USP15, which are involved in stabilisation of REST by preventing ubiquitination (Faronato et al., 2013; Huang et al., 2011). It is not yet known what controls either deubiquitinase, and it might be that one of them is either a regulatory target of IGF2 or of a component of the IGF2 signalling pathway. Experiments would include co-immunoprecipitation with REST, identification of putative regulatory phosphosites and screens for kinases/phosphatases capable of interacting with de-ubiquitinases.

7.2.2 CRISPR/Cas9 Validation

An exit point from our Crispr/Cas9 validation was the inability to select REST null homozygotes. Delineating the reason behind this observation might provide important clues about REST function. For this, conditional REST shRNA or CRISPRi

constructs will be used to knockdown REST and investigate the resultant phenotype of REST deficient cells concerning growth, differentiation, and apoptosis in our cell line models. If cells demonstrate a lethality phenotype, a conditional CRISPR/Cas9 system will be used to insert loxP sites in the *REST* gene, followed by recombination using Cre recombinase, and confirmation of cell death. If REST knockdown cells are observed to be differentiating, a selection strategy using multiple neuronal differentiation markers will be used to isolate differentiating cells, followed by using single cell genomics to perform transcriptome analysis on the isolated cells.

Alternatively, optogenetic induced inhibition of REST might also be used to visualise REST suppression and its consequence over time.

7.2.3 Structure of REST

An important issue that is much needed in the field is a molecular structure of REST. This would lead to deeper insights into REST protein function and mechanism, including the identity of the HMW REST form, ultimately aiding in the design of therapeutics.

Single-particle cryo-electron microscopy (EM) is a technique that is gaining rapid focus in the field of structural biology and can be used to create high-resolution models of molecules previously found to be difficult with conventional approaches in the field including X-ray crystallography (Doerr, 2016). Furthermore, samples with significant flexibility can be studied, by separation of particle images into homogenous subsets after image acquisition (Villarreal and Stewart, 2014). Recent improvements in the field have made it possibly to study the structural details of large protein complexes including DNA-PKcs/Ku/DNA (Villarreal and Stewart, 2014), ribosome-SecYE (Frauenfeld et al., 2011), spliceosomal complexes (Nguyen et al.,

2016) and the AMPA receptor–stargazing complex (Twomey et al., 2016), amongst others.

Complexes containing GFP-tagged REST will be isolated using biotinylated oligos with restriction enzyme sites followed by multiple RE1 sites. DNA-bound complexes will be released from beads using restriction enzyme digestion. This strategy was used successfully to isolate Estrogen receptor α (ERα)-DNA complexes (Schultz-Norton et al., 2008). These complexes will be utilised for structure determination using single-particle cryo-EM.

7.2.4 Ewing sarcoma- REST Dependency

A specific genetic translocation, EWS-FLI1, present in ~85-90% of Ewing sarcoma patients, is thought to be a driver switch for Ewing sarcoma. This translocation is specific for Ewing sarcoma and combines the RNA targeting function of EWS with the DNA binding function of FLI1, resulting in loss of function or haploinsufficiency of Ewing sarcoma and differential target gene expression (Sankar et al., 2013a). This study defines REST to be a specific epigenetic target in Ewing sarcoma. REST in Ewing sarcoma can be stabilised by IGF2 signalling and blockade of the proteasome. Blockade of IGF signalling resulted in destabilisation of REST (results not shown), and elimination of REST gene led to auto-regulation, and potentially loss of survival or neuronal differentiation in Ewing sarcoma cells. These data are the potential basis for further exploration and to accelerate an exciting evaluation of whether REST is a either molecular cornerstone or Achilles heel for Ewing sarcoma.

Previous experiments support this Ewing sarcoma-REST link. REST expression was identified in Ewing sarcoma, with EWS-FLI1 inducing expression of REST (Zhou, Yu and Kleinerman, 2014). Furthermore, the N-terminal region of EWS was found to interact with REST, resulting in inhibition of neuronal phenotype (Sankar et al.,

2013b). REST complex with Co-REST occurs with LSD1 (KDM1a), which is expressed at high levels in Ewing Sarcoma, and inhibition of which attenuates Ewing sarcoma growth in xenografts (Theisen et al., 2016). These all implicate direct mechanisms of interaction between REST and neuronal differentiation in Ewing sarcoma.

Several cutting-edge technologies will be adopted to understand REST function in Ewing sarcoma. The first is using CRISPR/Cas9 technology to disrupt REST function, to validate whether this always results in loss of Ewing sarcoma viability. To do this, a panel of Ewing sarcoma cell lines donated by patients will be used and a modified version of CRISPR called CRISPRi will be used, to circumvent the possibility of knock out-related lethality. This technology combines the CRISPR technology with RNAi to achieve inhibition of target genes using an inactive Cas9 and a single sgRNA (Dominguez, Lim and Qi, 2016).

The second technology is the use of light to monitor and modify how Ewing sarcoma cells work. The technological advances in super-resolution light imaging have allowed us to visualise single molecules in live cells in real time. These molecules can be activated with sophisticated chemical modification that reacts to light, so that we can choose a cell, activate a molecule with a specific wavelength of light, and video the consequences. This technology will be used to block the effects of REST using 'optogenetics', where laser light can flip a very special designer molecule to block REST, and then to monitor the effects in that cell over 24 hours. In this way, the immediate effects of REST damage on the Ewing sarcoma cell can be determined, in effect this is the 'seeing is believing' experiment.

This strategy has been used successfully to inhibit REST and activate transcription of REST target genes (Paonessa et al., 2016). Paonessa et al. used two REST inhibitor

domains, based on the minimal REST binding sequence of mSin3a and the minimal REST binding domain of REST-interacting LIM domain protein (RILP). These inhibitory domains were fused to light-sensitive Avena sativa light-oxygen-voltage-sensing (LOV) 2-phototrophin 1 (AsLOV2). Expression of these chimeric AsLOV2 proteins in cells, followed by blue light activation, can either inhibit REST recruitment of mSin3a or REST-DNA binding, respectively.

Finally, we will test whether damaging REST in the context of a growing Ewing sarcoma cures the tumour. To do this, mouse tumour xenograft models, originating from Ewing sarcoma cells, will be utilised, and before the tumour grows to a size that can harm the mouse, REST function will be blocked in all the cells of the tumour and the consequences monitored in real life. The information from the mouse studies is essential to provide proof-of-principle for further work in designing smart drugs that block REST in Ewing sarcoma selectively, and to use these in humans.

8 REFERENCES

Abramovitz, L., Shapira, T., Ben-Dror, I., Dror, V., Granot, L., Rousso, T., Landoy, E., Blau, L., Thiel, G. and Vardimon, L. (2008) Dual role of NRSF/REST in activation and repression of the glucocorticoid response. *The Journal of biological chemistry,* 283 (1), pp. 110-119.

Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature,* 422 (6928), pp. 198-207.

Andres, M.E., Burger, C., Peral-Rubio, M.J., Battaglioli, E., Anderson, M.E., Grimes, J., Dallman, J., Ballas, N. and Mandel, G. (1999) CoREST: a functional corepressor required for regulation of neural-specific gene expression. *Proceedings of the National Academy of Sciences of the United States of America,* 96 (17), pp. 9873-9878.

Aoki, H., Hara, A., Era, T., Kunisada, T. and Yamada, Y. (2012) Genetic ablation of Rest leads to in vitro-specific derepression of neuronal genes during neurogenesis. *Development (Cambridge, England),* 139 (4), pp. 667-677.

Apweiler, R., Hermjakob, H. and Sharon, N. (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica et biophysica acta,* 1473 (1), pp. 4-8.

Ayed, A., Mulder, F.A., Yi, G.S., Lu, Y., Kay, L.E. and Arrowsmith, C.H. (2001) Latent and active p53 are identical in conformation. *Nature structural biology,* 8 (9), pp. 756-760.

Bahn, S., Mimmack, M., Ryan, M., Caldwell, M.A., Jauniaux, E., Starkey, M., Svendsen, C.N. and Emson, P. (2002) Neuronal target genes of the neuron-restrictive silencer factor in neurospheres derived from fetuses with Down's syndrome: a gene expression study. *Lancet (London, England),* 359 (9303), pp. 310-315.

Baiula, M., Carbonari, G., Dattoli, S.D., Calienni, M., Bedini, A. and Spampinato, S. (2012) REST is up-regulated by epidermal growth factor in HeLa cells and inhibits apoptosis by influencing histone H3 acetylation. *Biochimica et biophysica acta,* 1823 (8), pp. 1252-1263.

Ballas, N., Battaglioli, E., Atouf, F., Andres, M.E., Chenoweth, J., Anderson, M.E., Burger, C., Moniwa, M., Davie, J.R., Bowers, W.J., Federoff, H.J., Rose, D.W., Rosenfeld, M.G., Brehm, P. and Mandel, G. (2001) Regulation of neuronal traits by a novel transcriptional complex. *Neuron,* 31 (3), pp. 353-365.

References

Ballas, N., Grunseich, C., Lu, D.D., Speh, J.C. and Mandel, G. (2005) REST and its corepressors mediate plasticity of neuronal gene chromatin throughout neurogenesis. *Cell,* 121 (4), pp. 645-657.

Battaglioli, E., Andres, M.E., Rose, D.W., Chenoweth, J.G., Rosenfeld, M.G., Anderson, M.E. and Mandel, G. (2002) REST repression of neuronal genes requires components of the hSWI.SNF complex. *The Journal of biological chemistry,* 277 (43), pp. 41038-41045.

Bedini, A., Baiula, M. and Spampinato, S. (2008) Transcriptional activation of human mu-opioid receptor gene by insulin-like growth factor-I in neuronal cells is modulated by the transcription factor REST. *Journal of neurochemistry,* 105 (6), pp. 2166-2178.

Belfiore, A. (2007) The role of insulin receptor isoforms and hybrid insulin/IGF-I receptors in human cancer. *Current pharmaceutical design,* 13 (7), pp. 671-686.

Bellon, S.F., Rodgers, K.K., Schatz, D.G., Coleman, J.E. and Steitz, T.A. (1997) Crystal structure of the RAG1 dimerization domain reveals multiple zinc-binding motifs including a novel zinc binuclear cluster. *Nature structural biology,* 4 (7), pp. 586-591.

Benarroch, E.E. (2012) Insulin-like growth factors in the brain and their potential clinical implications. *Neurology,* 79 (21), pp. 2148-2153.

Bingham, A.J., Ooi, L., Kozera, L., White, E. and Wood, I.C. (2007) The repressor element 1-silencing transcription factor regulates heart-specific gene expression using multiple chromatin-modifying complexes. *Molecular and cellular biology,* 27 (11), pp. 4082-4092.

Blom, N., Gammeltoft, S. and Brunak, S. (1999) Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *Journal of Molecular Biology,* 294 (5), pp. 1351-1362.

Bogdanove, A.J. and Voytas, D.F. (2011) TAL effectors: customizable proteins for DNA targeting. *Science (New York, N.Y.),* 333 (6051), pp. 1843-1846.

Bracko, O., Singer, T., Aigner, S., Knobloch, M., Winner, B., Ray, J., Clemenson, G.D.,Jr, Suh, H., Couillard-Despres, S., Aigner, L., Gage, F.H. and Jessberger, S. (2012) Gene expression profiling of neural stem cells and their neuronal progeny reveals IGF2 as a regulator of adult hippocampal neurogenesis. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* 32 (10), pp. 3376-3387.

Breitkopf, S.B. and Asara, J.M. (2012) Determining in vivo phosphorylation sites using mass spectrometry. *Current protocols in molecular biology / edited by Frederick M.Ausubel ...[et al.],* Chapter 18, pp. Unit18.19.1-27.

Bruce, A.W., Donaldson, I.J., Wood, I.C., Yerbury, S.A., Sadowski, M.I., Chapman, M., Gottgens, B. and Buckley, N.J. (2004) Genome-wide analysis of repressor element 1 silencing transcription factor/neuron-restrictive silencing factor (REST/NRSF) target genes. *Proceedings of the National Academy of Sciences of the United States of America,* 101 (28), pp. 10458-10463.

Bruce, A.W., Lopez-Contreras, A.J., Flicek, P., Down, T.A., Dhami, P., Dillon, S.C., Koch, C.M., Langford, C.F., Dunham, I., Andrews, R.M. and Vetrie, D. (2009) Functional diversity for REST (NRSF) is defined by in vivo binding affinity hierarchies at the DNA sequence level. *Genome research,* 19 (6), pp. 994-1005.

Bruel-Jungerman, E., Rampon, C. and Laroche, S. (2007) Adult hippocampal neurogenesis, synaptic plasticity and memory: facts and hypotheses. *Reviews in the neurosciences,* 18 (2), pp. 93-114.

Buckley, N.J., Johnson, R., Sun, Y.M. and Stanton, L.W. (2009) Is REST a regulator of pluripotency? *Nature,* 457 (7233), pp. E5-6; discussion E7.

Burns, J.L. and Hassan, A.B. (2001) Cell survival and proliferation are modified by insulin-like growth factor 2 between days 9 and 10 of mouse gestation. *Development (Cambridge, England),* 128 (19), pp. 3819-3830.

Canver, M.C., Bauer, D.E., Dass, A., Yien, Y.Y., Chung, J., Masuda, T., Maeda, T., Paw, B.H. and Orkin, S.H. (2014) Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. *The Journal of biological chemistry,* 289 (31), pp. 21312-21324.

Canzonetta, C., Mulligan, C., Deutsch, S., Ruf, S., O'Doherty, A., Lyle, R., Borel, C., Lin-Marq, N., Delom, F., Groet, J., Schnappauf, F., De Vita, S., Averill, S., Priestley, J.V., Martin, J.E., Shipley, J., Denyer, G., Epstein, C.J., Fillat, C., Estivill, X., Tybulewicz, V.L., Fisher, E.M., Antonarakis, S.E. and Nizetic, D. (2008) DYRK1A-dosage imbalance perturbs NRSF/REST levels, deregulating pluripotency and embryonic stem cell fate in Down syndrome. *American Journal of Human Genetics,* 83 (3), pp. 388-400.

Chao, W. and D'Amore, P.A. (2008) IGF2: epigenetic regulation and role in development and disease. *Cytokine & growth factor reviews,* 19 (2), pp. 111-120.

Charalambous, M., Menheniott, T.R., Bennett, W.R., Kelly, S.M., Dell, G., Dandolo, L. and Ward, A. (2004) An enhancer element at the Igf2/H19 locus drives gene expression in both imprinted and non-imprinted tissues. *Developmental biology,* 271 (2), pp. 488-497.

Charbord, J., Poydenot, P., Bonnefond, C., Feyeux, M., Casagrande, F., Brinon, B., Francelle, L., Auregan, G., Guillermier, M., Cailleret, M., Viegas, P., Nicoleau, C., Martinat, C., Brouillet, E., Cattaneo, E., Peschanski, M., Lechuga, M. and Perrier, A.L. (2013) High throughput screening for inhibitors of REST in neural derivatives of human embryonic stem cells reveals a chemical compound that promotes expression of neuronal genes. *Stem cells (Dayton, Ohio),* 31 (9), pp. 1816-1828.

Chen, Z.F., Paquette, A.J. and Anderson, D.J. (1998) NRSF/REST is required in vivo for repression of multiple neuronal target genes during embryogenesis. *Nature genetics,* 20 (2), pp. 136-142.

Chong, J.A., Tapia-Ramirez, J., Kim, S., Toledo-Aral, J.J., Zheng, Y., Boutros, M.C., Altshuller, Y.M., Frohman, M.A., Kraner, S.D. and Mandel, G. (1995) REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons. *Cell,* 80 (6), pp. 949-957.

Christofori, G., Naik, P. and Hanahan, D. (1994) A second signal supplied by insulin-like growth factor II in oncogene-induced tumorigenesis. *Nature,* 369 (6479), pp. 414-418.

Church, D.N., Phillips, B.R., Stuckey, D.J., Barnes, D.J., Buffa, F.M., Manek, S., Clarke, K., Harris, A.L., Carter, E.J. and Hassan, A.B. (2012) Igf2 ligand dependency of Pten(+/-) developmental and tumour phenotypes in the mouse. *Oncogene,* 31 (31), pp. 3635-3646.

Clemmons, D.R. (1997) Insulin-like growth factor binding proteins and their role in controlling IGF actions. *Cytokine & growth factor reviews,* 8 (1), pp. 45-62.

Collins, M.O., Yu, L., Campuzano, I., Grant, S.G. and Choudhary, J.S. (2008) Phosphoproteomic analysis of the mouse brain cytosol reveals a predominance of protein phosphorylation in regions of intrinsic sequence disorder. *Molecular & cellular proteomics : MCP,* 7 (7), pp. 1331-1348.

Conaco, C., Otto, S., Han, J.J. and Mandel, G. (2006) Reciprocal actions of REST and a microRNA promote neuronal identity. *Proceedings of the National Academy of Sciences of the United States of America,* 103 (7), pp. 2422-2427.

Conti, L., Crisafulli, L., Caldera, V., Tortoreto, M., Brilli, E., Conforti, P., Zunino, F., Magrassi, L., Schiffer, D. and Cattaneo, E. (2012) REST controls self-renewal and tumorigenic competence of human glioblastoma cells. *PloS one,* 7 (6), pp. e38486.

Corcoran, R.B., Bachar Raveh, T., Barakat, M.T., Lee, E.Y. and Scott, M.P. (2008) Insulin-like growth factor 2 is required for progression to advanced medulloblastoma in patched1 heterozygous mice. *Cancer research,* 68 (21), pp. 8788-8795.

Cory, A.H., Owen, T.C., Barltrop, J.A. and Cory, J.G. (1991) Use of an aqueous soluble tetrazolium/formazan assay for cell growth assays in culture. *Cancer communications,* 3 (7), pp. 207-212.

Coulson, J.M. (2005) Transcriptional regulation: cancer, neurons and the REST. *Current biology : CB,* 15 (17), pp. R665-8.

Coulson, J.M., Edgson, J.L., Woll, P.J. and Quinn, J.P. (2000a) A splice variant of the neuron-restrictive silencer factor repressor is expressed in small cell lung cancer: a potential role in derepression of neuroendocrine genes and a useful clinical marker. *Cancer research,* 60 (7), pp. 1840-1844.

Coulson, J.M., Edgson, J.L., Woll, P.J. and Quinn, J.P. (2000b) A splice variant of the neuron-restrictive silencer factor repressor is expressed in small cell lung cancer: a potential role in derepression of neuroendocrine genes and a useful clinical marker. *Cancer research,* 60 (7), pp. 1840-1844.

Coulson, J.M., Stanley, J. and Woll, P.J. (1999) Tumour-specific arginine vasopressin promoter activation in small-cell lung cancer. *British journal of cancer,* 80 (12), pp. 1935-1944.

Cui, H., Cruz-Correa, M., Giardiello, F.M., Hutcheon, D.F., Kafonek, D.R., Brandenburg, S., Wu, Y., He, X., Powe, N.R. and Feinberg, A.P. (2003) Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. *Science (New York, N.Y.),* 299 (5613), pp. 1753-1755.

de Pagter-Holthuizen, P., Jansen, M., van Schaik, F.M., van der Kammen, R., Oosterwijk, C., Van den Brande, J.L. and Sussenbach, J.S. (1987) The human insulin-like growth factor II gene contains two development-specific promoters. *FEBS letters,* 214 (2), pp. 259-264.

De Souza, A.T., Hankins, G.R., Washington, M.K., Fine, R.L., Orton, T.C. and Jirtle, R.L. (1995) Frequent loss of heterozygosity on 6q at the mannose 6-phosphate/insulin-like growth factor II receptor locus in human hepatocellular tumors. *Oncogene,* 10 (9), pp. 1725-1729.

DeChiara, T.M., Efstratiadis, A. and Robertson, E.J. (1990) A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. *Nature,* 345 (6270), pp. 78-80.

DeChiara, T.M., Robertson, E.J. and Efstratiadis, A. (1991a) Parental imprinting of the mouse insulin-like growth factor II gene. *Cell,* 64 (4), pp. 849-859.

DeChiara, T.M., Robertson, E.J. and Efstratiadis, A. (1991b) Parental imprinting of the mouse insulin-like growth factor II gene. *Cell,* 64 (4), pp. 849-859.

Dephoure, N., Gould, K.L., Gygi, S.P. and Kellogg, D.R. (2013) Mapping and analysis of phosphorylation sites: a quick guide for cell biologists. *Molecular biology of the cell,* 24 (5), pp. 535-542.

Di Toro, R., Baiula, M. and Spampinato, S. (2005) Expression of the repressor element-1 silencing transcription factor (REST) is influenced by insulin-like growth factor-I in differentiating human neuroblastoma cells. *The European journal of neuroscience,* 21 (1), pp. 46-58.

Doerr, A. (2016) Single-particle cryo-electron microscopy. *Nature methods,* 13 (1), pp. 23.

References

Doetsch, F., Caille, I., Lim, D.A., Garcia-Verdugo, J.M. and Alvarez-Buylla, A. (1999) Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. *Cell,* 97 (6), pp. 703-716.

Dominguez, A.A., Lim, W.A. and Qi, L.S. (2016) Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nature reviews.Molecular cell biology,* 17 (1), pp. 5-15.

Dunker, A.K., Lawson, J.D., Brown, C.J., Williams, R.M., Romero, P., Oh, J.S., Oldfield, C.J., Campen, A.M., Ratliff, C.M., Hipps, K.W., Ausio, J., Nissen, M.S., Reeves, R., Kang, C., Kissinger, C.R., Bailey, R.W., Griswold, M.D., Chiu, W., Garner, E.C. and Obradovic, Z. (2001) Intrinsically disordered protein. *Journal of molecular graphics & modelling,* 19 (1), pp. 26-59.

Engelman, J.A. (2009) Targeting PI3K signalling in cancer: opportunities, challenges and limitations. *Nature reviews.Cancer,* 9 (8), pp. 550-562.

Fabbro, D., Cowan-Jacob, S.W. and Moebitz, H. (2015) Ten things you should know about protein kinases: IUPHAR Review 14. *British journal of pharmacology,* 172 (11), pp. 2675-2700.

Faronato, M., Patel, V., Darling, S., Dearden, L., Clague, M.J., Urbe, S. and Coulson, J.M. (2013) The deubiquitylase USP15 stabilizes newly synthesized REST and rescues its expression at mitotic exit. *Cell cycle (Georgetown, Tex.),* 12 (12), pp. 1964-1977.

Ferron, S.R., Radford, E.J., Domingo-Muelas, A., Kleine, I., Ramme, A., Gray, D., Sandovici, I., Constancia, M., Ward, A., Menheniott, T.R. and Ferguson-Smith, A.C. (2015) Differential genomic imprinting regulates paracrine and autocrine roles of IGF2 in mouse adult neurogenesis. *Nature communications,* 6, pp. 8265.

Fischer, E.H. and Krebs, E.G. (1955) Conversion of phosphorylase b to phosphorylase a in muscle extracts. *The Journal of biological chemistry,* 216 (1), pp. 121-132.

Floyd, Z.E., Trausch-Azar, J.S., Reinstein, E., Ciechanover, A. and Schwartz, A.L. (2001) The nuclear ubiquitin-proteasome system degrades MyoD. *The Journal of biological chemistry,* 276 (25), pp. 22468-22475.

Foulstone, E., Prince, S., Zaccheo, O., Burns, J.L., Harper, J., Jacobs, C., Church, D. and Hassan, A.B. (2005) Insulin-like growth factor ligands, receptors, and binding proteins in cancer. *The Journal of pathology,* 205 (2), pp. 145-153.

Frauenfeld, J., Gumbart, J., Sluis, E.O., Funes, S., Gartmann, M., Beatrix, B., Mielke, T., Berninghausen, O., Becker, T., Schulten, K. and Beckmann, R. (2011) Cryo-EM structure of the ribosome-SecYE complex in the membrane environment. *Nature structural & molecular biology,* 18 (5), pp. 614-621.

Fu, V.X., Schwarze, S.R., Kenowski, M.L., Leblanc, S., Svaren, J. and Jarrard, D.F. (2004) A loss of insulin-like growth factor-2 imprinting is modulated by CCCTC-binding factor down-regulation at senescence in human epithelial cells. *The Journal of biological chemistry,* 279 (50), pp. 52218-52226.

Fuller, G.N., Su, X., Price, R.E., Cohen, Z.R., Lang, F.F., Sawaya, R. and Majumder, S. (2005) Many human medulloblastoma tumors overexpress repressor element-1 silencing transcription (REST)/neuron-restrictive silencer factor, which can be functionally countered by REST-VP16. *Molecular cancer therapeutics,* 4 (3), pp. 343-349.

Gage, F.H., Kempermann, G., Palmer, T.D., Peterson, D.A. and Ray, J. (1998) Multipotent progenitor cells in the adult dentate gyrus. *Journal of neurobiology,* 36 (2), pp. 249-266.

Gallagher, E.J. and LeRoith, D. (2010) The proliferating role of insulin and insulin-like growth factors in cancer. *Trends in endocrinology and metabolism: TEM,* 21 (10), pp. 610-618.

Green, E.D., Adelt, G., Baenziger, J.U., Wilson, S. and Van Halbeek, H. (1988) The asparagine-linked oligosaccharides on bovine fetuin. Structural analysis of N-glycanase-released oligosaccharides by 500-megahertz 1H NMR spectroscopy. *The Journal of biological chemistry,* 263 (34), pp. 18253-18268.

Grimes, J.A., Nielsen, S.J., Battaglioli, E., Miska, E.A., Speh, J.C., Berry, D.L., Atouf, F., Holdener, B.C., Mandel, G. and Kouzarides, T. (2000) The co-repressor mSin3A is a functional component of the REST-CoREST repressor complex. *The Journal of biological chemistry,* 275 (13), pp. 9461-9467.

Guardavaccaro, D., Frescas, D., Dorrello, N.V., Peschiaroli, A., Multani, A.S., Cardozo, T., Lasorella, A., Iavarone, A., Chang, S., Hernando, E. and Pagano, M. (2008a) Control of chromosome stability by the beta-TrCP-REST-Mad2 axis. *Nature,* 452 (7185), pp. 365-369.

Guardavaccaro, D., Frescas, D., Dorrello, N.V., Peschiaroli, A., Multani, A.S., Cardozo, T., Lasorella, A., Iavarone, A., Chang, S., Hernando, E. and Pagano, M. (2008b) Control of chromosome stability by the beta-TrCP-REST-Mad2 axis. *Nature,* 452 (7185), pp. 365-369.

Gunsalus, K.T., Wagoner, M.P., Meyer, K., Potter, W.B., Schoenike, B., Kim, S., Alexander, C.M., Friedl, A. and Roopra, A. (2012) Induction of the RNA regulator LIN28A is required for the growth and pathogenesis of RESTless breast tumors. *Cancer research,* 72 (13), pp. 3207-3216.

Gupta, R. and Brunak, S. (2002) Prediction of glycosylation across the human proteome and the correlation to protein function. *Pacific Symposium on Biocomputing.Pacific Symposium on Biocomputing,* , pp. 310-322.

Gurrola-Diaz, C., Lacroix, J., Dihlmann, S., Becker, C.M. and von Knebel Doeberitz, M. (2003) Reduced expression of the neuron restrictive silencer factor permits transcription of glycine receptor alpha1 subunit in small-cell lung cancer cells. *Oncogene,* 22 (36), pp. 5636-5645.

Hankins, G.R., De Souza, A.T., Bentley, R.C., Patel, M.R., Marks, J.R., Iglehart, J.D. and Jirtle, R.L. (1996) M6P/IGF2 receptor: a candidate breast tumor suppressor gene. *Oncogene,* 12 (9), pp. 2003-2009.

Harrington, L.S., Findlay, G.M., Gray, A., Tolkacheva, T., Wigfield, S., Rebholz, H., Barnett, J., Leslie, N.R., Cheng, S., Shepherd, P.R., Gout, I., Downes, C.P. and Lamb, R.F. (2004) The TSC1-2 tumor suppressor controls insulin-PI3K signaling via regulation of IRS proteins. *The Journal of cell biology,* 166 (2), pp. 213-223.

Hassan, A.B. and Howell, J.A. (2000) Insulin-like growth factor II supply modifies growth of intestinal adenoma in Apc(Min/+) mice. *Cancer research,* 60 (4), pp. 1070-1076.

Hatano, Y., Yamada, Y., Hata, K., Phutthaphadoong, S., Aoki, H. and Hara, A. (2011) Genetic ablation of a candidate tumor suppressor gene, Rest, does not promote mouse colon carcinogenesis. *Cancer science,* 102 (9), pp. 1659-1664.

Hermanson, O. (2008) Stem cells have different needs for REST. *PLoS biology,* 6 (10), pp. e271.

Hernandez, L., Kozlov, S., Piras, G. and Stewart, C.L. (2003) Paternal and maternal genomes confer opposite effects on proliferation, cell-cycle length, senescence, and tumor formation. *Proceedings of the National Academy of Sciences of the United States of America,* 100 (23), pp. 13344-13349.

Hietakangas, V., Ahlskog, J.K., Jakobsson, A.M., Hellesuo, M., Sahlberg, N.M., Holmberg, C.I., Mikhailov, A., Palvimo, J.J., Pirkkala, L. and Sistonen, L. (2003) Phosphorylation of serine 303 is a prerequisite for the stress-inducible SUMO modification of heat shock factor 1. *Molecular and cellular biology,* 23 (8), pp. 2953-2968.

Holm, T.M., Jackson-Grusby, L., Brambrink, T., Yamada, Y., Rideout, W.M.,3rd and Jaenisch, R. (2005) Global loss of imprinting leads to widespread tumorigenesis in adult mice. *Cancer cell,* 8 (4), pp. 275-285.

Hornbeck, P.V., Zhang, B., Murray, B., Kornhauser, J.M., Latham, V. and Skrzypek, E. (2015) PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic acids research,* 43 (Database issue), pp. D512-20.

Horvath, P. and Barrangou, R. (2010) CRISPR/Cas, the immune system of bacteria and archaea. *Science (New York, N.Y.),* 327 (5962), pp. 167-170.

Housley, M.P., Rodgers, J.T., Udeshi, N.D., Kelly, T.J., Shabanowitz, J., Hunt, D.F., Puigserver, P. and Hart, G.W. (2008) O-GlcNAc regulates FoxO activation in response to glucose. *The Journal of biological chemistry,* 283 (24), pp. 16283-16292.

Hu, Q., Gao, F., Tian, W., Ruteshouser, E.C., Wang, Y., Lazar, A., Stewart, J., Strong, L.C., Behringer, R.R. and Huff, V. (2011) Wt1 ablation and Igf2 upregulation in mice result in Wilms tumors with elevated ERK1/2 phosphorylation. *The Journal of clinical investigation,* 121 (1), pp. 174-183.

Huang, T.T., Wuerzberger-Davis, S.M., Wu, Z.H. and Miyamoto, S. (2003) Sequential modification of NEMO/IKKgamma by SUMO-1 and ubiquitin mediates NF-kappaB activation by genotoxic stress. *Cell,* 115 (5), pp. 565-576.

Huang, Y., Myers, S.J. and Dingledine, R. (1999) Transcriptional repression by REST: recruitment of Sin3A and histone deacetylase to neuronal genes. *Nature neuroscience,* 2 (10), pp. 867-872.

Huang, Z., Wu, Q., Guryanova, O.A., Cheng, L., Shou, W., Rich, J.N. and Bao, S. (2011) Deubiquitylase HAUSP stabilizes REST and promotes maintenance of neural progenitor cells. *Nature cell biology,* 13 (2), pp. 142-152.

Humphrey, S.J., James, D.E. and Mann, M. (2015) Protein Phosphorylation: A Major Switch Mechanism for Metabolic Regulation. *Trends in endocrinology and metabolism: TEM,* 26 (12), pp. 676-687.

Hunter, A.P. and Games, D.E. (1994) Chromatographic and mass spectrometric methods for the identification of phosphorylation sites in phosphoproteins. *Rapid communications in mass spectrometry : RCM,* 8 (7), pp. 559-570.

Hur, S.K., Freschi, A., Ideraabdullah, F., Thorvaldsen, J.L., Luense, L.J., Weller, A.H., Berger, S.L., Cerrato, F., Riccio, A. and Bartolomei, M.S. (2016) Humanized H19/Igf2 locus reveals diverged imprinting mechanism between mouse and human and reflects Silver-Russell syndrome phenotypes. *Proceedings of the National Academy of Sciences of the United States of America,* .

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.),* 337 (6096), pp. 816-821.

Johnson, D.S., Mortazavi, A., Myers, R.M. and Wold, B. (2007) Genome-wide mapping of in vivo protein-DNA interactions. *Science (New York, N.Y.),* 316 (5830), pp. 1497-1502.

Johnson, L.N. (2009) The regulation of protein phosphorylation. *Biochemical Society transactions,* 37 (Pt 4), pp. 627-641.

Johnson, L.N. and Lewis, R.J. (2001) Structural basis for control by phosphorylation. *Chemical reviews,* 101 (8), pp. 2209-2242.

Johnson, R. and Buckley, N.J. (2009) Gene dysregulation in Huntington's disease: REST, microRNAs and beyond. *Neuromolecular medicine,* 11 (3), pp. 183-199.

Johnson, R., Teh, C.H., Kunarso, G., Wong, K.Y., Srinivasan, G., Cooper, M.L., Volta, M., Chan, S.S., Lipovich, L., Pollard, S.M., Karuturi, R.K., Wei, C.L., Buckley, N.J. and Stanton, L.W. (2008) REST regulates distinct transcriptional networks in embryonic and neural stem cells. *PLoS biology,* 6 (10), pp. e256.

Jones, F.S. and Meech, R. (1999) Knockout of REST/NRSF shows that the protein is a potent repressor of neuronally expressed genes in non-neural tissues. *BioEssays : news and reviews in molecular, cellular and developmental biology,* 21 (5), pp. 372-376.

Jorgensen, H.F., Chen, Z.F., Merkenschlager, M. and Fisher, A.G. (2009a) Is REST required for ESC pluripotency? *Nature,* 457 (7233), pp. E4-5; discussion E7.

Jorgensen, H.F., Chen, Z.F., Merkenschlager, M. and Fisher, A.G. (2009b) Is REST required for ESC pluripotency? *Nature,* 457 (7233), pp. E4-5; discussion E7.

Jothi, R., Cuddapah, S., Barski, A., Cui, K. and Zhao, K. (2008) Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data. *Nucleic acids research,* 36 (16), pp. 5221-5231.

Juul, A. (2003) Serum levels of insulin-like growth factor I and its binding proteins in health and disease. Growth hormone & IGF research : official journal of the Growth Hormone Research Society and the International IGF Research Society, 13 (4), pp. 113-170.

Kaneko, N., Hwang, J.Y., Gertner, M., Pontarelli, F. and Zukin, R.S. (2014) Casein kinase 1 suppresses activation of REST in insulted hippocampal neurons and halts ischemia-induced neuronal death. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* 34 (17), pp. 6030-6039.

Kennedy, S.G., Wagner, A.J., Conzen, S.D., Jordan, J., Bellacosa, A., Tsichlis, P.N. and Hay, N. (1997) The PI 3-kinase/Akt signaling pathway delivers an anti-apoptotic signal. *Genes & development,* 11 (6), pp. 701-713.

Ketting, R.F. (2011) The many faces of RNAi. *Developmental cell,* 20 (2), pp. 148-161.

Klaus, A. and Birchmeier, W. (2008) Wnt signalling and its impact on development and cancer. *Nature reviews.Cancer,* 8 (5), pp. 387-398.

Klemm, J.D., Schreiber, S.L. and Crabtree, G.R. (1998) Dimerization as a regulatory mechanism in signal transduction. *Annual Review of Immunology,* 16, pp. 569-592.

Kohyama, J., Sanosaka, T., Tokunaga, A., Takatsuka, E., Tsujimura, K., Okano, H. and Nakashima, K. (2010) BMP-induced REST regulates the establishment and maintenance of astrocytic identity. *The Journal of cell biology,* 189 (1), pp. 159-170.

Kok, F.O., Taibi, A., Wanner, S.J., Xie, X., Moravec, C.E., Love, C.E., Prince, V.E., Mumm, J.S. and Sirotkin, H.I. (2012) Zebrafish rest regulates developmental gene expression but not neurogenesis. *Development (Cambridge, England),* 139 (20), pp. 3838-3848.

Krebs, E.G. and Fischer, E.H. (1955) Phosphorylase activity of skeletal muscle extracts. *The Journal of biological chemistry,* 216 (1), pp. 113-120.

Kreisler, A., Strissel, P.L., Strick, R., Neumann, S.B., Schumacher, U. and Becker, C.M. (2010) Regulation of the NRSF/REST gene by methylation and CREB affects the cellular phenotype of small-cell lung cancer. *Oncogene,* 29 (43), pp. 5828-5838.

Kuwabara, T., Hsieh, J., Nakashima, K., Taira, K. and Gage, F.H. (2004) A small modulatory dsRNA specifies the fate of adult neural stem cells. *Cell,* 116 (6), pp. 779-793.

Lawinger, P., Venugopal, R., Guo, Z.S., Immaneni, A., Sengupta, D., Lu, W., Rastelli, L., Marin Dias Carneiro, A., Levin, V., Fuller, G.N., Echelard, Y. and Majumder, S. (2000) The neuronal repressor REST/NRSF is an essential regulator in medulloblastoma cells. *Nature medicine,* 6 (7), pp. 826-831.

Lee, J.H., Shimojo, M., Chai, Y.G. and Hersh, L.B. (2000) Studies on the interaction of REST4 with the cholinergic repressor element-1/neuron restrictive silencer element. *Brain research.Molecular brain research,* 80 (1), pp. 88-98.

Lehtinen, M.K., Zappaterra, M.W., Chen, X., Yang, Y.J., Hill, A.D., Lun, M., Maynard, T., Gonzalez, D., Kim, S., Ye, P., D'Ercole, A.J., Wong, E.T., LaMantia, A.S. and Walsh, C.A. (2011) The cerebrospinal fluid provides a proliferative niche for neural progenitor cells. *Neuron,* 69 (5), pp. 893-905.

Leighton, P.A., Ingram, R.S., Eggenschwiler, J., Efstratiadis, A. and Tilghman, S.M. (1995) Disruption of imprinting caused by deletion of the H19 gene region in mice. *Nature,* 375 (6526), pp. 34-39.

Lenk, U. and Sommer, T. (2000) Ubiquitin-mediated proteolysis of a short-lived regulatory protein depends on its cellular localization. *The Journal of biological chemistry,* 275 (50), pp. 39403-39410.

Lepagnol-Bestel, A.M., Zvara, A., Maussion, G., Quignon, F., Ngimbous, B., Ramoz, N., Imbeaud, S., Loe-Mie, Y., Benihoud, K., Agier, N., Salin, P.A., Cardona, A., Khung-Savatovsky, S., Kallunki, P., Delabar, J.M., Puskas, L.G., Delacroix, H., Aggerbeck, L., Delezoide, A.L., Delattre, O., Gorwood, P., Moalic, J.M. and Simonneau, M. (2009) DYRK1A interacts with the REST/NRSF-SWI/SNF chromatin remodelling complex to deregulate gene clusters involved in the neuronal phenotypic traits of Down syndrome. *Human molecular genetics,* 18 (8), pp. 1405-1414.

Li, S., Mattar, P., Dixit, R., Lawn, S.O., Wilkinson, G., Kinch, C., Eisenstat, D., Kurrasch, D.M., Chan, J.A. and Schuurmans, C. (2014) RAS/ERK signaling controls proneural genetic programs in cortical development and gliomagenesis. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* 34 (6), pp. 2169-2190.

Lois, C. and Alvarez-Buylla, A. (1994) Long-distance neuronal migration in the adult mammalian brain. *Science (New York, N.Y.),* 264 (5162), pp. 1145-1148.

Lu, T., Aron, L., Zullo, J., Pan, Y., Kim, H., Chen, Y., Yang, T.H., Kim, H.M., Drake, D., Liu, X.S., Bennett, D.A., Colaiacovo, M.P. and Yankner, B.A. (2014) REST and stress resistance in ageing and Alzheimer's disease. *Nature,* 507 (7493), pp. 448-454.

Lv, H., Pan, G., Zheng, G., Wu, X., Ren, H., Liu, Y. and Wen, J. (2010) Expression and functions of the repressor element 1 (RE-1)-silencing transcription factor (REST) in breast cancer. *Journal of cellular biochemistry,* 110 (4), pp. 968-974.

Mahamdallie, S.S., Hanks, S., Karlin, K.L., Zachariou, A., Perdeaux, E.R., Ruark, E., Shaw, C.A., Renwick, A., Ramsay, E., Yost, S., Elliott, A., Birch, J., Capra, M., Gray, J., Hale, J., Kingston, J., Levitt, G., McLean, T., Sheridan, E., Renwick, A., Seal, S., Stiller, C., Sebire, N., Westbrook, T.F. and Rahman, N. (2015) Mutations in the transcriptional repressor REST predispose to Wilms tumor. *Nature genetics,* 47 (12), pp. 1471-1474.

Manning, G., Whyte, D.B., Martinez, R., Hunter, T. and Sudarsanam, S. (2002) The protein kinase complement of the human genome. *Science (New York, N.Y.),* 298 (5600), pp. 1912-1934.

McCann, A.H., Miller, N., O'Meara, A., Pedersen, I., Keogh, K., Gorey, T. and Dervan, P.A. (1996) Biallelic expression of the IGF2 gene in human breast disease. *Human molecular genetics,* 5 (8), pp. 1123-1127.

Medinas, D.B., Gozzo, F.C., Santos, L.F., Iglesias, A.H. and Augusto, O. (2010) A ditryptophan cross-link is responsible for the covalent dimerization of human superoxide dismutase 1 during its bicarbonate-dependent peroxidase activity. *Free radical biology & medicine,* 49 (6), pp. 1046-1053.

Mikulak, J., Negrini, S., Klajn, A., D'Alessandro, R., Mavilio, D. and Meldolesi, J. (2012) Dual REST-dependence of L1CAM: from gene expression to alternative splicing governed by Nova2 in neural cells. *Journal of neurochemistry,* 120 (5), pp. 699-709.

Minuto, F., Palermo, C., Arvigo, M. and Barreca, A.M. (2005) The IGF system and bone. *Journal of endocrinological investigation,* 28 (8 Suppl), pp. 8-10.

Moorehead, R.A., Hojilla, C.V., De Belle, I., Wood, G.A., Fata, J.E., Adamson, E.D., Watson, K.L., Edwards, D.R. and Khokha, R. (2003a) Insulin-like growth factor-II regulates PTEN expression in the mammary gland. *The Journal of biological chemistry,* 278 (50), pp. 50422-50427.

Moorehead, R.A., Hojilla, C.V., De Belle, I., Wood, G.A., Fata, J.E., Adamson, E.D., Watson, K.L., Edwards, D.R. and Khokha, R. (2003b) Insulin-like growth factor-II regulates PTEN expression in the mammary gland. *The Journal of biological chemistry,* 278 (50), pp. 50422-50427.

Moorhead, G.B., De Wever, V., Templeton, G. and Kerk, D. (2009) Evolution of protein phosphatases in plants and animals. *The Biochemical journal,* 417 (2), pp. 401-409.

Moremen, K.W., Tiemeyer, M. and Nairn, A.V. (2012) Vertebrate protein glycosylation: diversity, synthesis and function. *Nature reviews.Molecular cell biology,* 13 (7), pp. 448-462.

Morison, I.M. and Reeve, A.E. (1998) Insulin-like growth factor 2 and overgrowth: molecular biology and clinical implications. *Molecular medicine today,* 4 (3), pp. 110-115.

Muller, S., Berger, M., Lehembre, F., Seeler, J.S., Haupt, Y. and Dejean, A. (2000) c-Jun and p53 activity is modulated by SUMO-1 modification. *The Journal of biological chemistry,* 275 (18), pp. 13321-13329.

Mulligan, P., Westbrook, T.F., Ottinger, M., Pavlova, N., Chang, B., Macia, E., Shi, Y.J., Barretina, J., Liu, J., Howley, P.M., Elledge, S.J. and Shi, Y. (2008) CDYL bridges REST and histone methyltransferases for gene repression and suppression of cellular transformation. *Molecular cell,* 32 (5), pp. 718-726.

Murai, K., Naruse, Y., Shaul, Y., Agata, Y. and Mori, N. (2004) Direct interaction of NRSF with TBP: chromatin reorganization and core promoter repression for neuron-specific gene transcription. *Nucleic acids research,* 32 (10), pp. 3180-3189.

Nechiporuk, T., McGann, J., Mullendorff, K., Hsieh, J., Wurst, W., Floss, T. and Mandel, G. (2016) The REST remodeling complex protects genomic integrity during embryonic neurogenesis. *eLife,* 5, pp. e09584.

Nesti, E., Corson, G.M., McCleskey, M., Oyer, J.A. and Mandel, G. (2014) C-terminal domain small phosphatase 1 and MAP kinase reciprocally control REST stability and neuronal differentiation. *Proceedings of the National Academy of Sciences of the United States of America,* 111 (37), pp. E3929-36.

Nguyen, T.H., Galej, W.P., Fica, S.M., Lin, P.C., Newman, A.J. and Nagai, K. (2016) CryoEM structures of two spliceosomal complexes: starter and dessert at the spliceosome feast. *Current opinion in structural biology,* 36, pp. 48-57.

Nishihara, S., Tsuda, L. and Ogura, T. (2003) The canonical Wnt pathway directly regulates NRSF/REST expression in chick spinal cord. *Biochemical and biophysical research communications,* 311 (1), pp. 55-63.

Ohlsson, R., Hedborg, F., Holmgren, L., Walsh, C. and Ekstrom, T.J. (1994) Overlapping patterns of IGF2 and H19 expression during human development: biallelic IGF2 expression correlates with a lack of H19 expression. *Development (Cambridge, England),* 120 (2), pp. 361-368.

Ojala, P.M., Yamamoto, K., Castanos-Velez, E., Biberfeld, P., Korsmeyer, S.J. and Makela, T.P. (2000) The apoptotic v-cyclin-CDK6 complex phosphorylates and inactivates Bcl-2. *Nature cell biology,* 2 (11), pp. 819-825.

Olsen, J.V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P. and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell,* 127 (3), pp. 635-648.

Olsen, J.V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M.L., Jensen, L.J., Gnad, F., Cox, J., Jensen, T.S., Nigg, E.A., Brunak, S. and Mann, M. (2010) Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Science signaling,* 3 (104), pp. ra3.

Ooi, L., Belyaev, N.D., Miyake, K., Wood, I.C. and Buckley, N.J. (2006) BRG1 chromatin remodeling activity is required for efficient chromatin binding by repressor element 1-silencing transcription factor (REST) and facilitates REST-mediated repression. *The Journal of biological chemistry,* 281 (51), pp. 38974-38980.

Otto, S.J., McCorkle, S.R., Hover, J., Conaco, C., Han, J.J., Impey, S., Yochum, G.S., Dunn, J.J., Goodman, R.H. and Mandel, G. (2007) A new binding motif for the transcriptional repressor REST uncovers large gene networks devoted to neuronal functions. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* 27 (25), pp. 6729-6739.

Palm, K., Metsis, M. and Timmusk, T. (1999) Neuron-specific splicing of zinc finger transcription factor REST/NRSF/XBR is frequent in neuroblastomas and conserved in human, mouse and rat. *Brain research.Molecular brain research,* 72 (1), pp. 30-39.

Paonessa, F., Criscuolo, S., Sacchetti, S., Amoroso, D., Scarongella, H., Pecoraro Bisogni, F., Carminati, E., Pruzzo, G., Maragliano, L., Cesca, F. and Benfenati, F. (2016) Regulation of neural gene transcription by optogenetic inhibition of the RE1-silencing transcription factor. *Proceedings of the National Academy of Sciences of the United States of America,* 113 (1), pp. E91-100.

Pichler, A. and Melchior, F. (2002) Ubiquitin-related modifier SUMO1 and nucleocytoplasmic transport. *Traffic (Copenhagen, Denmark),* 3 (6), pp. 381-387.

Pinkse, M.W., Uitto, P.M., Hilhorst, M.J., Ooms, B. and Heck, A.J. (2004) Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. *Analytical Chemistry,* 76 (14), pp. 3935-3943.

Posewitz, M.C. and Tempst, P. (1999) Immobilized gallium(III) affinity chromatography of phosphopeptides. *Analytical Chemistry,* 71 (14), pp. 2883-2892.

Pouyssegur, J., Volmat, V. and Lenormand, P. (2002a) Fidelity and spatio-temporal control in MAP kinase (ERKs) signalling. *Biochemical pharmacology,* 64 (5-6), pp. 755-763.

Pouyssegur, J., Volmat, V. and Lenormand, P. (2002b) Fidelity and spatio-temporal control in MAP kinase (ERKs) signalling. *Biochemical pharmacology,* 64 (5-6), pp. 755-763.

Rainier, S., Johnson, L.A., Dobry, C.J., Ping, A.J., Grundy, P.E. and Feinberg, A.P. (1993) Relaxation of imprinted genes in human cancer. *Nature,* 362 (6422), pp. 747-749.

Ramakrishna, S., Cho, S.W., Kim, S., Song, M., Gopalappa, R., Kim, J.S. and Kim, H. (2014) Surrogate reporter-based enrichment of cells containing RNA-guided Cas9 nuclease-induced mutations. *Nature communications,* 5, pp. 3378.

Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y. and Zhang, F. (2013a) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell,* 154 (6), pp. 1380-1389.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A. and Zhang, F. (2013b) Genome engineering using the CRISPR-Cas9 system. *Nature protocols,* 8 (11), pp. 2281-2308.

Ravenel, J.D., Broman, K.W., Perlman, E.J., Niemitz, E.L., Jayawardena, T.M., Bell, D.W., Haber, D.A., Uejima, H. and Feinberg, A.P. (2001) Loss of imprinting of insulin-like growth factor-II (IGF2) gene in distinguishing specific biologic subtypes of Wilms tumor. *Journal of the National Cancer Institute,* 93 (22), pp. 1698-1703.

Reik, W., Constancia, M., Dean, W., Davies, K., Bowden, L., Murrell, A., Feil, R., Walter, J. and Kelsey, G. (2000) Igf2 imprinting in development and disease. *The International journal of developmental biology,* 44 (1), pp. 145-150.

Remy, S., Tesson, L., Menoret, S., Usal, C., Scharenberg, A.M. and Anegon, I. (2010) Zinc-finger nucleases: a powerful tool for genetic engineering of animals. *Transgenic research,* 19 (3), pp. 363-371.

Roopra, A., Qazi, R., Schoenike, B., Daley, T.J. and Morrison, J.F. (2004) Localized domains of G9a-mediated histone methylation are required for silencing of neuronal genes. *Molecular cell,* 14 (6), pp. 727-738.

Ross, S., Best, J.L., Zon, L.I. and Gill, G. (2002) SUMO-1 modification represses Sp3 transcriptional activation and modulates its subnuclear localization. *Molecular cell,* 10 (4), pp. 831-842.

Rossbach, M. (2011) Non-Coding RNAs in Neural Networks, REST-Assured. *Frontiers in genetics,* 2, pp. 8.

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Holper, S., Kruger, M. and Stainier, D.Y. (2015) Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature,* 524 (7564), pp. 230-233.

Roy, A., Goodman, J.H., Begum, G., Donnelly, B.F., Pittman, G., Weinman, E.J., Sun, D. and Subramanya, A.R. (2015) Generation of WNK1 knockout cell lines by CRISPR/Cas-mediated genome editing. *American journal of physiology.Renal physiology,* 308 (4), pp. F366-76.

Ruiz i Altaba, A., Sanchez, P. and Dahmane, N. (2002) Gli and hedgehog in cancer: tumours, embryos and stem cells. *Nature reviews.Cancer,* 2 (5), pp. 361-372.

Sankar, S., Bell, R., Stephens, B., Zhuo, R., Sharma, S., Bearss, D.J. and Lessnick, S.L. (2013a) Mechanism and relevance of EWS/FLI-mediated transcriptional repression in Ewing sarcoma. *Oncogene,* 32 (42), pp. 5089-5100.

Sankar, S., Gomez, N.C., Bell, R., Patel, M., Davis, I.J., Lessnick, S.L. and Luo, W. (2013b) EWS and RE1-Silencing Transcription Factor Inhibit Neuronal Phenotype Development and Oncogenic Transformation in Ewing Sarcoma. *Genes & cancer,* 4 (5-6), pp. 213-223.

Schneider-Poetsch, T., Ju, J., Eyler, D.E., Dang, Y., Bhat, S., Merrick, W.C., Green, R., Shen, B. and Liu, J.O. (2010) Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nature chemical biology,* 6 (3), pp. 209-217.

Schoenherr, C.J. and Anderson, D.J. (1995) The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. *Science (New York, N.Y.),* 267 (5202), pp. 1360-1363.

Schultz-Norton, J.R., Ziegler, Y.S., Likhite, V.S., Yates, J.R. and Nardulli, A.M. (2008) Isolation of novel coregulatory protein networks associated with DNA-bound estrogen receptor alpha. *BMC molecular biology,* 9, pp. 97-2199-9-97.

Scott, C.D. and Firth, S.M. (2004) The role of the M6P/IGF-II receptor in cancer: tumor suppression or garbage disposal? *Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme,* 36 (5), pp. 261-271.

Segal, D.J. and Meckler, J.F. (2013) Genome engineering at the dawn of the golden age. *Annual review of genomics and human genetics,* 14, pp. 135-158.

Shapiro, L.R., Duncan, P.A., Davidian, M.M. and Singer, N. (1982) The placenta in familial beckwith-wiedemann syndrome. *Birth defects original article series,* 18 (3B), pp. 203-206.

Sharma, K., D'Souza, R.C., Tyanova, S., Schaab, C., Wisniewski, J.R., Cox, J. and Mann, M. (2014) Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. *Cell reports,* 8 (5), pp. 1583-1594.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A. and Shi, Y. (2004) Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell,* 119 (7), pp. 941-953.

Shimojo, M. (2008) Huntingtin regulates RE1-silencing transcription factor/neuron-restrictive silencer factor (REST/NRSF) nuclear trafficking indirectly through a complex with REST/NRSF-interacting LIM domain protein (RILP) and dynactin p150 Glued. *The Journal of biological chemistry,* 283 (50), pp. 34880-34886.

Shimojo, M. and Hersh, L.B. (2003) REST/NRSF-interacting LIM domain protein, a putative nuclear translocation receptor. *Molecular and cellular biology,* 23 (24), pp. 9025-9031.

Shimojo, M., Lee, J.H. and Hersh, L.B. (2001) Role of zinc finger domains of the transcription factor neuron-restrictive silencer factor/repressor element-1 silencing transcription factor in DNA binding and nuclear localization. *The Journal of biological chemistry,* 276 (16), pp. 13121-13126.

Shimojo, M., Paquette, A.J., Anderson, D.J. and Hersh, L.B. (1999) Protein kinase A regulates cholinergic gene expression in PC12 cells: REST4 silences the silencing activity of neuron-restrictive silencer factor/REST. *Molecular and cellular biology,* 19 (10), pp. 6788-6795.

Shirangi, T.R., Zaika, A. and Moll, U.M. (2002) Nuclear degradation of p53 occurs during down-regulation of the p53 response after DNA damage. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology,* 16 (3), pp. 420-422.

Shuai, K., Horvath, C.M., Huang, L.H., Qureshi, S.A., Cowburn, D. and Darnell, J.E.,Jr (1994) Interferon activation of the transcription factor Stat91 involves dimerization through SH2-phosphotyrosyl peptide interactions. *Cell,* 76 (5), pp. 821-828.

Singh, S.K., Kagalwala, M.N., Parker-Thornburg, J., Adams, H. and Majumder, S. (2008) REST maintains self-renewal and pluripotency of embryonic stem cells. *Nature,* 453 (7192), pp. 223-227.

Soroceanu, L., Kharbanda, S., Chen, R., Soriano, R.H., Aldape, K., Misra, A., Zha, J., Forrest, W.F., Nigro, J.M., Modrusan, Z., Feuerstein, B.G. and Phillips, H.S. (2007) Identification of IGF2 signaling through phosphoinositide-3-kinase regulatory subunit 3 as a growth-promoting axis in glioblastoma. *Proceedings of the National Academy of Sciences of the United States of America,* 104 (9), pp. 3466-3471.

Su, X., Gopalakrishnan, V., Stearns, D., Aldape, K., Lang, F.F., Fuller, G., Snyder, E., Eberhart, C.G. and Majumder, S. (2006) Abnormal expression of REST/NRSF and Myc in neural stem/progenitor

cells causes cerebellar tumors by blocking neuronal differentiation. *Molecular and cellular biology,* 26 (5), pp. 1666-1678.

Sun, F.L., Dean, W.L., Kelsey, G., Allen, N.D. and Reik, W. (1997) Transactivation of Igf2 in a mouse model of Beckwith-Wiedemann syndrome. *Nature,* 389 (6653), pp. 809-815.

Sun, L., Liu, A. and Georgopoulos, K. (1996) Zinc finger-mediated protein interactions modulate Ikaros activity, a molecular control of lymphocyte development. *The EMBO journal,* 15 (19), pp. 5358-5369.

Tahiliani, M., Mei, P., Fang, R., Leonor, T., Rutenberg, M., Shimizu, F., Li, J., Rao, A. and Shi, Y. (2007) The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation. *Nature,* 447 (7144), pp. 601-605.

Tapia-Ramirez, J., Eggen, B.J., Peral-Rubio, M.J., Toledo-Aral, J.J. and Mandel, G. (1997) A single zinc finger motif in the silencing factor REST represses the neural-specific type II sodium channel promoter. *Proceedings of the National Academy of Sciences of the United States of America,* 94 (4), pp. 1177-1182.

Thakore-Shah, K., Koleilat, T., Jan, M., John, A. and Pyle, A.D. (2015) REST/NRSF Knockdown Alters Survival, Lineage Differentiation and Signaling in Human Embryonic Stem Cells. *PloS one,* 10 (12), pp. e0145280.

Theisen, E.R., Pishas, K.I., Saund, R.S. and Lessnick, S.L. (2016) Therapeutic opportunities in Ewing sarcoma: EWS-FLI inhibition via LSD1 targeting. *Oncotarget,* 7 (14), pp. 17616-17630.

Thiel, G., Lietz, M. and Cramer, M. (1998) Biological activity and modular structure of RE-1-silencing transcription factor (REST), a repressor of neuronal genes. *The Journal of biological chemistry,* 273 (41), pp. 26891-26899.

Thorvaldsen, J.L., Duran, K.L. and Bartolomei, M.S. (1998) Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. *Genes & development,* 12 (23), pp. 3693-3702.

Tsai, M.J. and O'Malley, B.W. (1994) Molecular mechanisms of action of steroid/thyroid receptor superfamily members. *Annual Review of Biochemistry,* 63, pp. 451-486.

Twomey, E.C., Yelshanskaya, M.V., Grassucci, R.A., Frank, J. and Sobolevsky, A.I. (2016) Elucidation of AMPA receptor-stargazin complexes by cryo-electron microscopy. *Science (New York, N.Y.),* 353 (6294), pp. 83-86.

Ulloa-Aguirre, A., Timossi, C., Damian-Matsumura, P. and Dias, J.A. (1999) Role of glycosylation in function of follicle-stimulating hormone. *Endocrine,* 11 (3), pp. 205-215.

Vaidyanathan, K., Durning, S. and Wells, L. (2014) Functional O-GlcNAc modifications: implications in molecular regulation and pathophysiology. *Critical reviews in biochemistry and molecular biology,* 49 (2), pp. 140-163.

Villarreal, S.A. and Stewart, P.L. (2014) CryoEM and image sorting for flexible protein/DNA complexes. *Journal of structural biology,* 187 (1), pp. 76-83.

Vincent, A.M. and Feldman, E.L. (2002) Control of cell survival by IGF signaling pathways. Growth hormone & IGF research : official journal of the Growth Hormone Research Society and the International IGF Research Society, 12 (4), pp. 193-197.

Volvert, M.L., Prevot, P.P., Close, P., Laguesse, S., Pirotte, S., Hemphill, J., Rogister, F., Kruzy, N., Sacheli, R., Moonen, G., Deiters, A., Merkenschlager, M., Chariot, A., Malgrange, B., Godin, J.D. and

Nguyen, L. (2014) MicroRNA targeting of CoREST controls polarization of migrating cortical neurons. *Cell reports,* 7 (4), pp. 1168-1183.

Vu, A. (2009) Nuclear and Cytoplasmic Roles of the Transcription factor REST.

Vu, T.H. and Hoffman, A.R. (1994) Promoter-specific imprinting of the human insulin-like growth factor-II gene. *Nature,* 371 (6499), pp. 714-717.

Wagoner, M.P., Gunsalus, K.T., Schoenike, B., Richardson, A.L., Friedl, A. and Roopra, A. (2010) The transcription factor REST is lost in aggressive breast cancer. *PLoS genetics,* 6 (6), pp. e1000979.

Walsh, D.A., Perkins, J.P. and Krebs, E.G. (1968) An adenosine 3',5'-monophosphate-dependant protein kinase from rabbit skeletal muscle. *The Journal of biological chemistry,* 243 (13), pp. 3763-3765.

Walsh, R.M. and Hochedlinger, K. (2013) A variant CRISPR-Cas9 system adds versatility to genome engineering. *Proceedings of the National Academy of Sciences of the United States of America,* 110 (39), pp. 15514-15515.

Walther, T.C., Fornerod, M., Pickersgill, H., Goldberg, M., Allen, T.D. and Mattaj, I.W. (2001) The nucleoporin Nup153 is required for nuclear pore basket formation, nuclear pore complex anchoring and import of a subset of nuclear proteins. *The EMBO journal,* 20 (20), pp. 5703-5714.

Wan, X., Harkavy, B., Shen, N., Grohar, P. and Helman, L.J. (2007) Rapamycin induces feedback activation of Akt signaling through an IGF-1R-dependent mechanism. *Oncogene,* 26 (13), pp. 1932-1940.

Ward, A. (1997) Beck-Wiedemann syndrome and Wilms' tumour. *Molecular human reproduction,* 3 (2), pp. 157-168.

Weissman, A.M. (2008) How much REST is enough? *Cancer cell,* 13 (5), pp. 381-383.

Westbrook, T.F., Hu, G., Ang, X.L., Mulligan, P., Pavlova, N.N., Liang, A., Leng, Y., Maehr, R., Shi, Y., Harper, J.W. and Elledge, S.J. (2008) SCFbeta-TRCP controls oncogenic transformation and neural differentiation through REST degradation. *Nature,* 452 (7185), pp. 370-374.

Westbrook, T.F., Martin, E.S., Schlabach, M.R., Leng, Y., Liang, A.C., Feng, B., Zhao, J.J., Roberts, T.M., Mandel, G., Hannon, G.J., Depinho, R.A., Chin, L. and Elledge, S.J. (2005) A genetic screen for candidate tumor suppressors identifies REST. *Cell,* 121 (6), pp. 837-848.

White, M.F. (2006) Regulating insulin signaling and beta-cell function through IRS proteins. *Canadian journal of physiology and pharmacology,* 84 (7), pp. 725-737.

Willert, J., Epping, M., Pollack, J.R., Brown, P.O. and Nusse, R. (2002) A transcriptional response to Wnt protein in human embryonic carcinoma cells. *BMC developmental biology,* 2, pp. 8.

Wilson, E.M. and Rotwein, P. (2006) Control of MyoD function during initiation of muscle differentiation by an autocrine signaling pathway activated by insulin-like growth factor-II. *The Journal of biological chemistry,* 281 (40), pp. 29962-29971.

Xu, W., Doshi, A., Lei, M., Eck, M.J. and Harrison, S.C. (1999) Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Molecular cell,* 3 (5), pp. 629-638.

Yamada, Y., Aoki, H., Kunisada, T. and Hara, A. (2010) Rest promotes the early differentiation of mouse ESCs but is not required for their maintenance. *Cell stem cell,* 6 (1), pp. 10-15.

Yang, S.H., Jaffray, E., Hay, R.T. and Sharrocks, A.D. (2003) Dynamic interplay of the SUMO and ERK pathways in regulating Elk-1 transcriptional activity. *Molecular cell,* 12 (1), pp. 63-74.

Yang, S.H. and Sharrocks, A.D. (2004) SUMO promotes HDAC-mediated transcriptional repression. *Molecular cell,* 13 (4), pp. 611-617.

Yang, W.H., Kim, J.E., Nam, H.W., Ju, J.W., Kim, H.S., Kim, Y.S. and Cho, J.W. (2006) Modification of p53 with O-linked N-acetylglucosamine regulates p53 activity and stability. *Nature cell biology,* 8 (10), pp. 1074-1083.

Yates, J.R., Ruse, C.I. and Nakorchevsky, A. (2009) Proteomics by mass spectrometry: approaches, advances, and applications. *Annual Review of Biomedical Engineering,* 11, pp. 49-79.

Yeo, M., Lee, S.K., Lee, B., Ruiz, E.C., Pfaff, S.L. and Gill, G.N. (2005) Small CTD phosphatases function in silencing neuronal gene expression. *Science (New York, N.Y.),* 307 (5709), pp. 596-600.

You, A., Tong, J.K., Grozinger, C.M. and Schreiber, S.L. (2001) CoREST is an integral component of the CoREST- human histone deacetylase complex. *Proceedings of the National Academy of Sciences of the United States of America,* 98 (4), pp. 1454-1458.

Zha, J. and Lackner, M.R. (2010) Targeting the insulin-like growth factor receptor-1R pathway for cancer therapy. *Clinical cancer research : an official journal of the American Association for Cancer Research,* 16 (9), pp. 2512-2517.

Zhang, P., Pazin, M.J., Schwartz, C.M., Becker, K.G., Wersto, R.P., Dilley, C.M. and Mattson, M.P. (2008) Nontelomeric TRF2-REST interaction modulates neuronal gene silencing and fate of tumor and stem cells. *Current biology : CB,* 18 (19), pp. 1489-1494.

Zhao, N. and Enns, C.A. (2013) N-linked glycosylation is required for transferrin-induced stabilization of transferrin receptor 2, but not for transferrin binding or trafficking to the cell surface. *Biochemistry,* 52 (19), pp. 3310-3319.

Zhao, Q., Xie, Y., Zheng, Y., Jiang, S., Liu, W., Mu, W., Liu, Z., Zhao, Y., Xue, Y. and Ren, J. (2014) GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs. *Nucleic acids research,* 42 (Web Server issue), pp. W325-30.

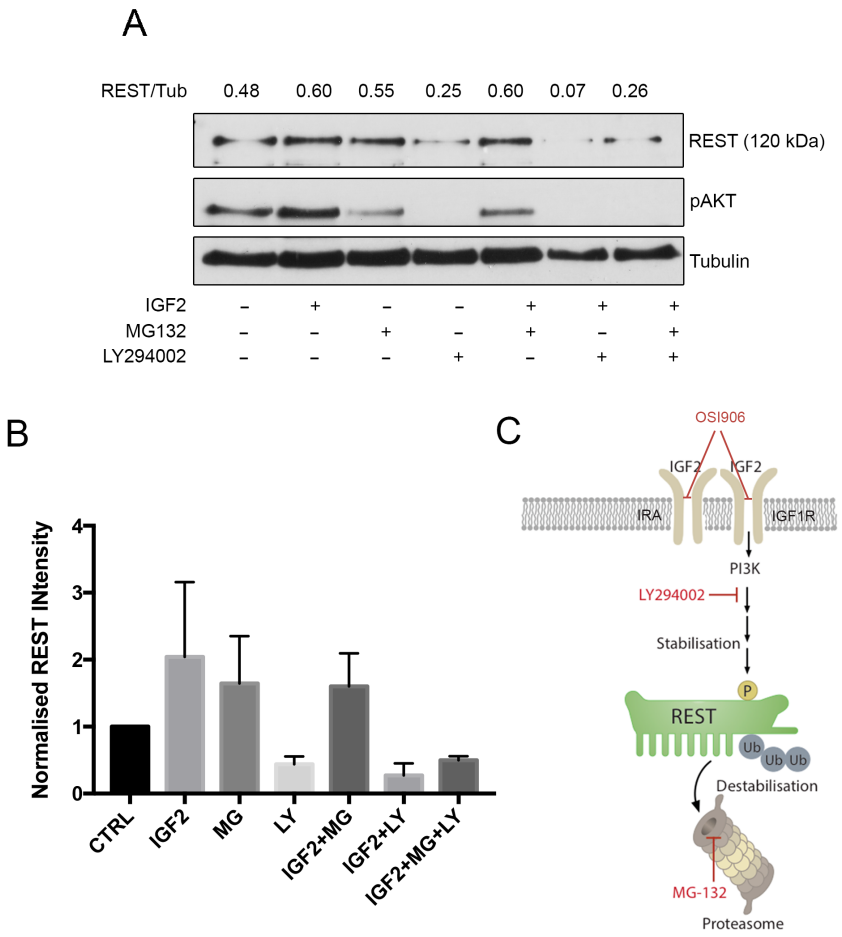
Zhou, H., Ye, M., Dong, J., Corradini, E., Cristobal, A., Heck, A.J., Zou, H. and Mohammed, S. (2013) Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography. *Nature protocols,* 8 (3), pp. 461-480.

Zhou, Z., Yu, L. and Kleinerman, E.S. (2014) EWS-FLI-1 regulates the neuronal repressor gene REST, which controls Ewing sarcoma growth and vascular morphology. *Cancer,* 120 (4), pp. 579-588.

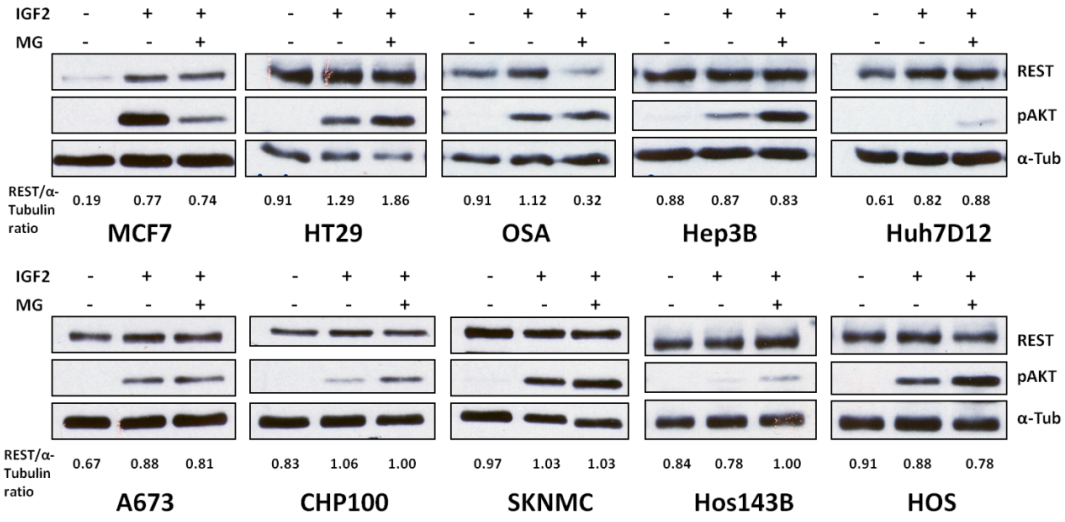
Ziegler, A.N., Schneider, J.S., Qin, M., Tyler, W.A., Pintar, J.E., Fraidenraich, D., Wood, T.L. and Levison, S.W. (2012) IGF-II promotes stemness of neural restricted precursors. *Stem cells (Dayton, Ohio),* 30 (6), pp. 1265-1276.

Zuccato, C., Tartari, M., Crotti, A., Goffredo, D., Valenza, M., Conti, L., Cataudella, T., Leavitt, B.R., Hayden, M.R., Timmusk, T., Rigamonti, D. and Cattaneo, E. (2003) Huntingtin interacts with REST/NRSF to modulate the transcription of NRSE-controlled neuronal genes. *Nature genetics,* 35 (1), pp. 76-83.

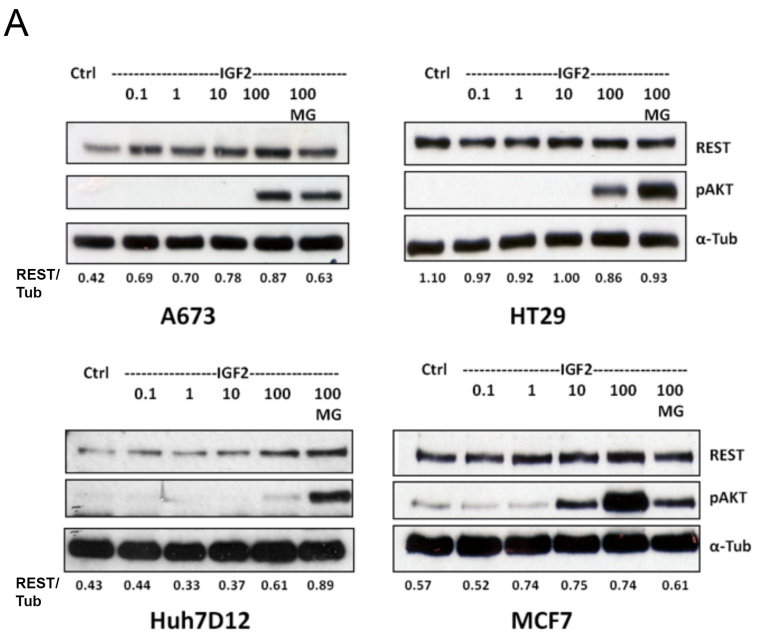
9 APPENDICES



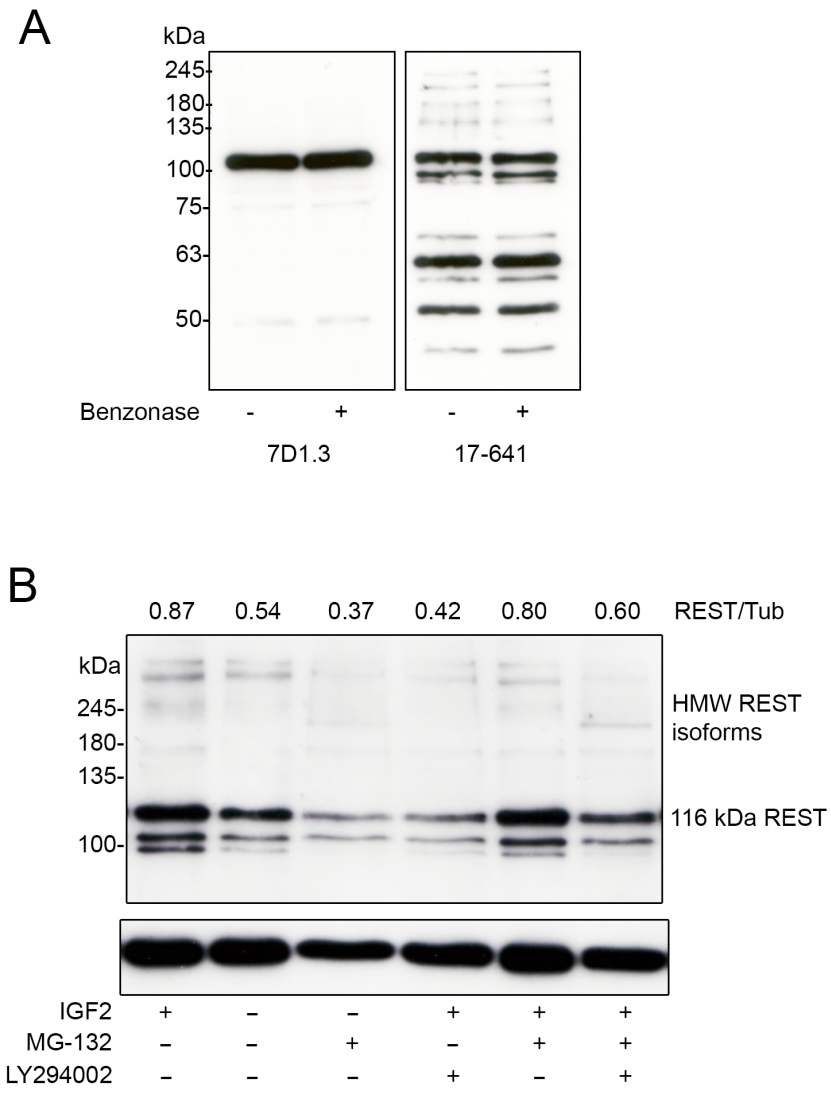
Appendix 1 IGF2 stabilises REST protein in a mouse fibroblast NIH3T3 cell line.
IGF2 stabilises REST protein in the mouse fibroblast NIH3T3 cell line. A) NIH3T3-ProIGF2 cell line was generated from retroviral transduction of mouse fibroblast NIH3T3 cells with a tetracycline-inducible form of the human *IGF2* gene. Withdrawal of doxycycline causes accumulation of IGF2 in the culture media, hence modelling the *in vivo* paracrine action of this growth factor. NIH3T3-ProIGF2 cells were serum-starved for 24 hours, washed and treated with IGF2 (100ng/ml), MG132 (0.75μM) and LY294002 (50μM) in different combinations. Protein from cell lysates was then detected for REST (7D1.3 antibody), pAKT and α-tubulin using antibodies. A representative blot from three biological replicates is shown. REST levels were quantified using Image J and normalised to α-tubulin levels. C) REST levels from three independent experiments were quantified, normalised to α-tubulin and control REST intensity and plotted. C) Schematic representation of the proposed hypothesis of regulation of REST via IGF2 signalling. Phosphorylated serine residues in the C-terminus of REST act as a binding site for the E3 ubiquitin ligase, β-TrCP, which polyubiquitinates (Ub) REST, leading to its degradation by the proteasome. IGF2 signalling via the PI3K pathway leads to PTM, shown as putative phosphorylation (P), which stabilises REST.



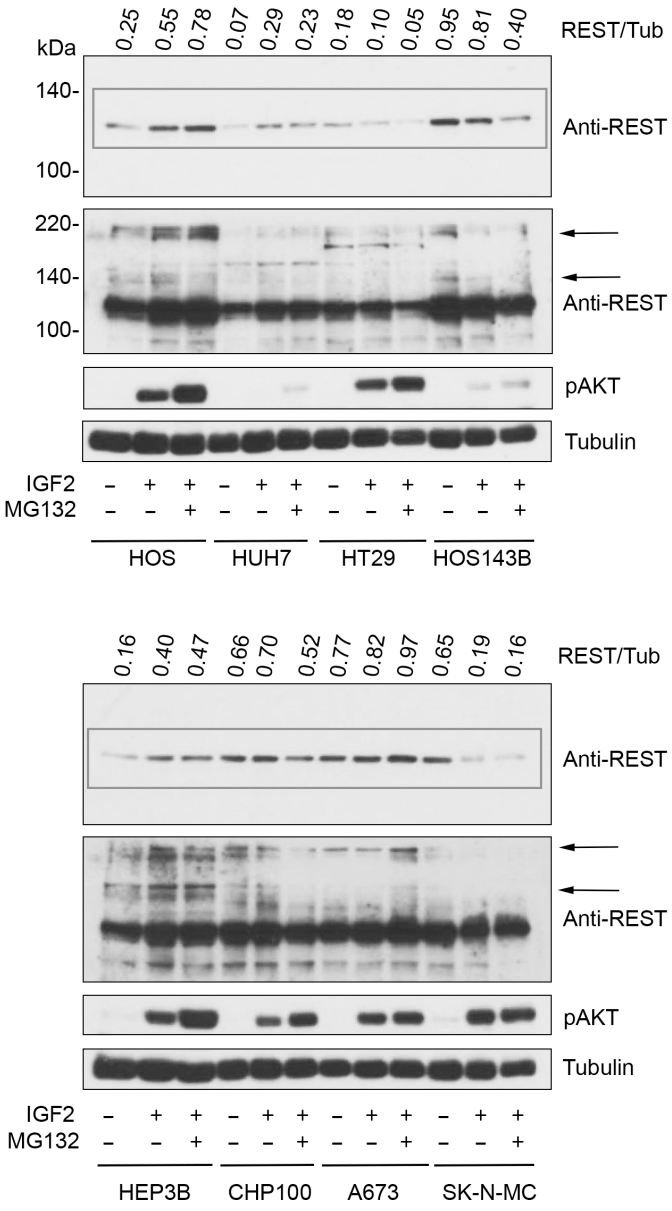
Appendix 2 Human cell line screen for REST response to IGF2 and MG132.
MCF7, Hos, Huh7, Hep3B, HT29, HOS143B, SKNMC, A673 cells were serum-starved for 24 hours, washed and treated with IGF2 (100ng/ml) and MG132 (0.75 μM) for 6 hours. Whole cell lysates were immunoblotted for REST (7D1.3 antibody), pAKT and α-tubulin using antibodies. REST levels were quantified using Image J and normalised to α-tubulin levels.



Appendix 3 Titration of cell lines with varying IGF2 doses.
Cells were serum-starved for 24 hours, washed and treated with different IGF2 concentrations ranging from 0.1 ng/ml to 100 ng/ml for 6 hours, with or without 0.75 μM MG-132 (MG). Protein from cell lysates was then detected for REST, pAKT and α-tubulin using antibodies. Representative Western blot from two independent experiments is shown.



Appendix 4 REST exists as higher MW isoforms.
A) NIH3T3 cell lysates were treated with Benzonase nuclease and immunoblotted for REST using the monoclonal anti-REST 7D1.3 antibody and the polyclonal anti-REST 17-641 antibody. B) Lysates from Figure 9.1 were immunoblotted using anti-REST (17-641). REST levels were quantified using Image J and normalised to α-tubulin levels.



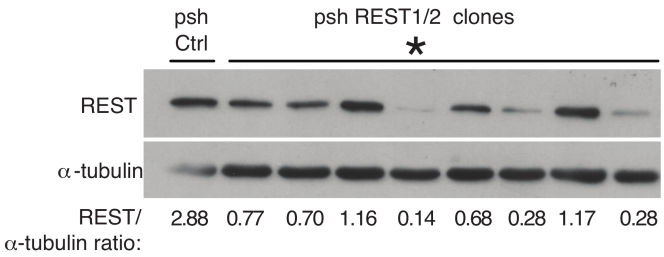
Appendix 5 Higher MW bands show a concomitant change in levels along with 116 kDa REST form.

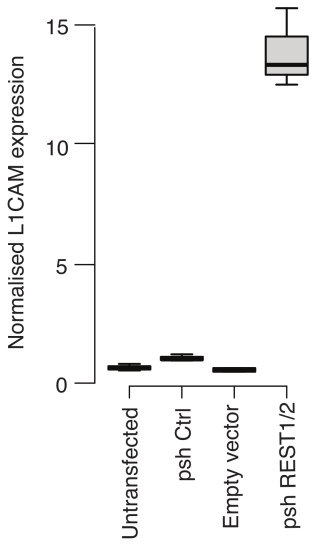
Cell lysates used in Figure. 9.2 were immunoblotted for REST using the Polyclonal anti-REST 17-641 antibody, pAKT and α-tubulin. Arrows indicate higher molecular weight isoforms. Upper REST panel: low exposure, lower REST panel: high exposure. REST levels (outlined by a rectangular box) were quantified by Image J, normalised to α-tubulin and are shown on the blot.

Gene	Significance	MCF7	HOS	HUH7	HEP3B	HT29	HOS 143B	SKNMC	A673
PIK3CA	PIK3K Signalling	E545K				P449T			
NRAS	MAPK Signalling								
KRAS	MAPK Signalling						A59T		
BRAF	MAPK Signalling					V600E(+T119S)			V600E
CDKN2A	Cell cycle control				G122V				
TP53	Genome stability		R156P	Y220C		R273H	R156P		P72R(X)(+A119fs*5)
ASXL1	Transcriptional regulation				L1490F(X)				N986S(X)
FLT3	Cytokine signalling	C368Y/S							
JAK3	Cytokine signalling				P132T(X)				
KMT2D	histone methyl transferase, transcriptional regulation,		P813L(X)	P2668L/D1767E			P813L		
KMT2C	histone methyl transferase, transcriptional regulation	C988F(X)	C988F(X)	L3234F	C988F	C988F(+P2415S)	C988F		
ABL1	Tyrosine kinase, cell division and differentiation					P829L(X)			
APC	Wnt Signalling					E853*(+T1556fs*3)			
CYP2D6	Metabolism	R365H(X)				R356H(X)			
ERBB2	EGFR signalling								
KIT	Cytokine receptor					M541L(X)		M541L	
SMAD4	TGFβ signalling					Q311*			
ATM	Genome stability								
NOTCH1	Notch signalling				G2152R(X)				
RB1	G1 checkpoint protein				S576*				
NF1	EGFR signalling		V689M				V689M		
Basal REST levels		L	L	L	L	H	H	H	H

Appendix 6 Common cancer-associated mutations do not correlate with REST stabilisation.

DNA was extracted from MCF7, Hos, Huh7, Hep3B, HT29, HOS143B, SKNMC, A673 cell lines. Extracted DNA was run on Illumina MiSEq, by Oxford Gene Technology, using TruSeq version 3 chemistry (generating 2x150 base reads). Agilent SureSelect whole exome baits were designed to capture 60 genes commonly mutated in cancer (SureSeq Solid Tumour Panel). The analysis was performed by mapping reads to the most recent build of human genome hg19/b37 using the Burrows-Wheeler Aligner (BWA). Variant analysis was done using the Genome Analysis Tool Tit (GATK) version 1.6. Variants were annotated with gene and gene function data from Ensemble. Known variants from dbSNP 137 were annotated within the dataset to identify novel variants. Mismatches to Cosmic annotation are shown in brackets with (x) indicating absence in Cosmic.





Appendix 7 REST knock down causes de-repression of L1Cam.
Top: Knock-down of REST using shRNA vectors. NIH3T3-ProIGF2 cells were transfected with a pair of pSuper.neo/GFP vectors containing shRNA targeting REST (psh REST1/2, a kind gift from Helle Jørgensen) (Jorgensen et al., 2009). Cells positive for GFP were FACS-sorted into a 96-well plate, expanded in culture and immunoblotted with antibodies against REST and α-tubulin. Lanes 2–9 are cultures derived from individual wells of the 96-well plate and lane 1 is a lysate of NIH3T3-ProIGF2 cells transfected with the psh Ctrl (control) vector containing a scrambled sequence. Knock-down was most efficient in the lane 5 culture marked with an asterisk and cDNA was prepared from this clone for subsequent qPCR experiments. Bottom: Knock down of REST leads to de-repression of L1CAM expression as determined by qPCR. Boxplots show the interquartile range (grey shaded box), the mean (thick horizontal bar) and standard deviation (whiskers) for data from 3 independent experiments. ΔC_t values for L1CAM expression have been normalised to the geometric mean of ΔC_t values for three housekeeping genes (Actin, GAPDH and Hprt). qPCR assays were done using a Rotor-Gene Q instrument (Qiagen). Figure courtesy of David Barnes

Appendices