

**WestminsterResearch**

<http://www.westminster.ac.uk/research/westminsterresearch>

**Novel cell-penetrating peptide-based vectors for siRNA delivery**

**Myrsini Tsimon**

Faculty of Science and Technology

This is an electronic version of a PhD thesis awarded by the University of Westminster. © The Author, 2014.

This is an exact reproduction of the paper copy held by the University of Westminster library.

---

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners.

Users are permitted to download and/or print one copy for non-commercial private study or research. Further distribution and any use of material from within this archive for profit-making enterprises or for commercial gain is strictly forbidden.

---

Whilst further distribution of specific materials from within this archive is forbidden, you may freely distribute the URL of WestminsterResearch: (<http://westminsterresearch.wmin.ac.uk/>).

In case of abuse or copyright appearing without permission e-mail [repository@westminster.ac.uk](mailto:repository@westminster.ac.uk)

Novel cell-penetrating  
peptide-based vectors for  
siRNA delivery

**MYRSINI TSIMON**

A THESIS SUBMITTED IN PARTIAL  
FULFILLMENT OF THE REQUIREMENTS OF  
THE UNIVERSITY OF WESTMINSTER FOR THE  
DEGREE OF DOCTOR OF PHILOSOPHY

June 2014

# Abstract

Harnessing the RNAi pathway with synthetic siRNA as a potent and selective mode of post-transcriptional gene silencing and has therapeutic potential in personalised medicine; however, the large size and negative charge of siRNA creates a hurdle for intracellular delivery that has thus far limited its development as a therapeutic. Cell penetrating peptides (CPPs), such as the Antennapedia homeodomain (AntpHD) and its third helix, Penetratin, are well characterised cationic motifs used previously to deliver covalently linked nucleic acid cargo, such as siRNA *in vitro* and *in vivo*, may offer a strategy to address this challenge.

This thesis aimed to design, purify and characterize novel recombinant fusion proteins, that would have broad applicability as carrier molecules for non-covalent siRNA delivery. The fusion proteins were comprised of a cell penetrating peptide sequence (either AntpHD, Penetratin, HIV-Tat or EB1) fused to either the first dsRNA binding motif 1 (DRBM) or the tandem motifs (DRBMx2) from human PKR, which bind dsRNA with high avidity in a sequence independent manner. A panel of constructs were cloned, expressed in a bacterial cell system, and purified by affinity chromatography under both native and denaturing conditions. Several of the constructs were either poorly expressed, insoluble or prone to precipitation during purification or dialysis; however, construct C5.1 was successfully purified and its identity confirmed by mass spectrometry. Construct C5.1 bound siRNA only at a high ratio of protein to siRNA due to the presence of co-purifying nucleic acid, whereas constructs C12.2 and C13.2 bound siRNA at low molar ratios. Both C12.2 and C5.1 were efficiently internalized in either live or fixed HEK293 and HepG2 cells; however, the proteins appeared to be sequestered in endosomes whether in the presence or absence of cargo. Cytotoxicity of the fusion proteins in HEK293 cells increased in the order of C12.2<C5.1<C13.2<C11.2 whereas in HepG2 cells, C11.2 was significantly more cytotoxic than C12.2; suggesting that the proteins exhibit cell type-specific cytotoxicity. Moreover, C11.2 and C12.2 altered HepG2 cell morphology in the presence of siRNA, compared to C5.1 and C13.2. In a HEK293-dEGFP reporter cell line, a complex of C12.2 and siRNA induced a significant decrease in dEGFP expression, which was not observed with C5.1-siRNA; however, unexpectedly similar effects were observed with C12.2-scrambled siRNA suggesting non-specific siRNA effects. Effects on PTP1B expression were also examined with all purified proteins complexed with the optimum molar concentration of PTPN1 siRNA; none were able to exhibit PTP1B expression knockdown at the protein level.

In this thesis, a comprehensive strategy for the design, purification and testing of novel siRNA carriers has been developed. A number of recombinant siRNA carriers have been successfully produced and characterised. The results highlight common issues encountered with the development CPP-based siRNA delivery vectors; nonetheless, the ability of C12.2 to mediate RNAi-mediated knockdown demonstrates the potential of the development of CPPs as non-covalent siRNA delivery vectors.

# Acknowledgements

I would like to extend my sincere gratitude for the uninterrupted guidance, patience, and mentoring by Dr Alastair Barr, my Director of Studies. He has been a resounding voice of reason and a compass throughout the past three years. I would like to thank him for giving me the opportunity to learn from him whilst allowing me to work independently.

I would also like to thank Dr John Murphy for his continuous encouragement, support and advice, as well as Dr Anatoliy Markiv his invaluable input throughout the past four years and for not allowing me to rest on my laurels...ever. He has helped me more than he knows.

I'd like to extend my gratitude to Professor Taj Keshavarz for his kind words of wisdom. I could not have finished this project without his support. I am indebted to the Dean of the School of Life Sciences, Professor Jane Lewis; the university of Westminster for their financial support, and to Diabetes UK for their generous grant. I would also like to thank the Leventis Foundation for their contribution towards my studies.

Last but not least; a like to thank Andy Jenks for his help with the flow cytometer, Brad Elliott for his help with statistics, as well as Rana, Faye, Katie, Artun and the rest of my colleagues for their optimism. I owe my (partial) sanity to them. Moreover, I would like to acknowledge Vanita, Paulet, Jonathan and the rest of the technicians for all their assistance throughout the past four years.

Finally, I have been very lucky in that my doctoral studies have given me great satisfaction. This project would have been impossible without the unconditional and unwavering emotional, financial, and mental support from my family –Anna, Kyras, Melissa and Mario. Last but not least, I would also like to thank Demis, Judy, and yiayia Dina, for their faith in me. I hope I have made them proud. This work is dedicated to them.

# Table of Contents

Abstract .....	1
Acknowledgements .....	2
Table of Contents .....	3
List of Figures .....	8
List of Tables .....	12
List of Abbreviations.....	14
List of Formulae.....	18
Chapter 1. General Introduction .....	19
1.1 The RNAi pathway .....	20
1.2 siRNA as a therapeutic drug: challenges and limitations.....	27
1.3 Engineering stabilizing modifications for siRNA.....	32
1.4 siRNA Delivery Strategies.....	35
1.5 Cell Penetrating Peptides .....	37
1.6 HIV-Tat - a (brief) timeline.....	41
1.7 The Antennapedia homeodomain and Penetratin .....	43
1.8 Overcoming endosomal entrapment .....	48
1.9 Double stranded RNA binding domains .....	50
1.10 Human Protein Kinase R.....	53
1.11 PTP1B in Type 2 Diabetes: a potential target for siRNA-based intervention .....	56
1.12 Aims and Objectives .....	61

Chapter 2.	Characterisation of recombinant proteins comprised of the AntpHD or Penetratin and DRBM1 from human PKR .....	63
2.1	Introduction .....	64
2.2	Results .....	69
2.2.1	Subcloning the recombinant constructs into expression vectors.....	69
2.2.2	Recombinant protein expression on a 50ml scale and purification by affinity chromatography.....	71
2.2.3	Poly(His)- tagged protein purifications from 500 ml <i>E.coli</i> cultures under native conditions .....	77
2.2.4	GST-tagged protein purifications from 500 ml <i>E.coli</i> cultures.....	79
2.2.5	Large scale purifications of GST-tagged C5.1, C6.1-myc, C6.1 and C6.1-myc . .....	86
2.2.6	C5.1 and C5.1-myc stability assessment and detection by western blotting ...	90
2.2.7	Mass and purity determination for C5.1 and C5.1-myc by Mass Spectroscopy.. .....	92
2.2.8	C5.1-myc shows a dose-dependent intracellular localisation in HEK293 and HepG2 cells.....	95
2.2.9	Assessing the potential of C5.1to deliver siRNA in live HepG2 cells with a lectin counterstain .....	99
2.2.10	C5.1 shows diminished siRNA binding through the DRBM1 .....	104
2.3	Discussion .....	106
Chapter 3.	Characterisation of endosomolytic, cell penetrating, dsRNA-binding fusion proteins .....	118
3.1	Introduction .....	119
3.2	Results .....	122
3.2.1	Subcloning EB1-DRBMx2 constructs into plasmid vectors.....	122
3.2.2	Small-scale expression optimization of His-tagged C7.2 and C8.2.....	125
3.2.3	Small-scale expression optimization of GST-tagged C9.2 and C10.2.....	127
3.2.4	Large scale purifications of C7.2 and C8.2 under denaturing conditions.....	131

3.3	Discussion .....	133
Chapter 4.	Characterisation of fusion proteins containing both dsRNA binding domains from PKR. ....	139
4.1	Introduction .....	140
4.2	Results .....	144
4.2.1	Purification and Refolding of C11.2, C12.2 and C13.2.....	144
4.2.2	Functional analysis of purified C11.2, C12.2 and C13.2.....	146
4.3	Discussion .....	167
Chapter 5.	Conclusions and future directions.....	175
Chapter 6.	Materials and Methods.....	182
6.1	Materials.....	183
6.1.1	Bacterial Strains .....	183
6.1.2	Plasmids .....	183
6.1.3	Cloning primers.....	184
6.1.4	Sequencing primers.....	185
6.1.5	Protein purification buffers, media, solutions .....	186
6.2	Methods.....	189
6.2.1	Plasmids .....	189
6.2.2	Competent <i>E. coli</i> strain preparation.....	189
6.2.3	cDNA Cloning by PCR.....	189
6.2.4	Cloning into plasmid vectors .....	191
6.2.5	Gel extraction and ligation into plasmid vectors.....	191
6.2.6	Transformation of competent cells.....	192
6.2.7	Plasmid propagation.....	192

6.2.8	Agarose gel electrophoresis .....	193
6.2.9	DNA purifications.....	194
6.2.10	Validation of target sequence amplification by colony PCR.....	194
6.2.11	Small scale purifications of C1.1-C4.1 by immobilized metal ion chromatography (IMAC) under native conditions.....	195
6.2.12	Small scale purifications of GST-tagged proteins by affinity chromatography .. .....	196
6.2.13	Large scale purifications of GST-tagged C5.1, C6.1, C5.1-myc and C6.1-myc . .....	197
6.2.14	Denaturing cobalt IMAC purification of His-tagged proteins.....	198
6.2.15	Trichloroacetic Acid – Ethanol (TCA-EtOH) Protein Purification .....	199
6.2.16	Polyacrylamide Gel Electrophoresis, protein transfer and Western Blotting .....	200
6.2.17	Electromobility shift assay (EMSA) with C5.1, C12.2 and C13.2.....	201
6.2.18	Cell culture.....	202
6.2.19	Trypan Blue exclusion assay and hemocytometry.....	203
6.2.20	Generation of a polyclonal HEK293 Reporter cell line.....	204
6.2.21	Cell viability assay.....	206
6.2.22	Protein-siRNA complex formation .....	207
6.2.23	Transient knockdown assays in a HEK293 reporter cell line .....	207
6.2.24	Analysis of dEGFP knockdown by flow cytometry .....	208
6.2.25	PTP1B knockdown assays in HepG2 cells. ....	208
6.2.26	Whole cell lysate preparation.....	210
6.2.27	Localisation studies in fixed HepG2 and HEK293 cells.....	210
6.2.28	Localisation studies in live HepG2 cells.....	212
6.2.29	Confocal microscopy .....	212

Appendix I	.....	214
I.1	Bioinformatic analysis on constructs C1.1-C6.1.....	215
I.2	Bioinformatic analysis on myc-tagged constructs .....	217
I.3	Bioinformatic analysis on constructs C11.2- C13.....	220
I.4	Figure 2.19, enlarged.....	221
Appendix II	.....	222
II.1	Plasmid Vector Maps .....	223
II.2	Statistical Analysis of MTT assays in HEK293 and HepG2 cells by 2-way Anova (SPSS). .....	227
II.2.1	MTT assay in HEK293 cells (n=3). P-value limits were set at $\leq 0.05$ .....	227
II.2.2	MTT assay in HepG2 cells (n=9).....	231
Bibliography	.....	235

# List of Figures

Figure 1.1	Mechanism of RNA interference in mammalian cells.....	26
Figure 1.2	Common chemical siRNA modifications.. ..	34
Figure 1.3	Secondary structure of the Antennapedia homeodomain protein (AntpHD)... .....	44
Figure 1.4	Structure of the two dsRNA binding motifs in human PKR that compose its dsRNA binding domain (DRBD).....	54
Figure 1.5	Structural domains and PTP1B regulation.....	60
Figure 2.1	Schematic of the fusion protein constructs in a modified pET32-a vector (C1.1-C4.1) or a pGEX-6P-2 vector (C5.1-C6.1).....	67
Figure 2.2	Schematic of the myc-tagged recombinant proteins with an <i>N</i> -terminus Glutathione <i>S</i> -transferase (GST) tag in a pGEX-6P-2 vector.....	68
Figure 2.3	Insert screening by colony PCR following ligation of insert in pET32-a (C1.1-C4.1) or pGEX-6P-2 (C5.1-C6.1).....	70
Figure 2.4	PCR amplification of GST - tagged C5.1 and C6.1 with a c-myc epitope tag .....	70
Figure 2.5	Small-scale expression analysis of His <sub>10</sub> constructs C1.1 and C2.1 by SDS- PAGE.. ..	73
Figure 2.6	Small-scale expression analysis of His <sub>6</sub> constructs C3.1 and C4.1 by SDS- PAGE.. ..	74
Figure 2.7	Small-scale expression analysis of GST-fusion constructs C5.1 and C6.1 by SDS-PAGE.....	75
Figure 2.8	Small-scale expression analysis of GST-fusion constructs C5.1-myc and C6.1-myc by SDS-PAGE.....	76
Figure 2.9	Expression analysis of C1.1, C2.1 and C4.1 from 0.5L cultures by SDS- PAGE.. ..	78
Figure 2.10	Expression analysis of GST-tagged constructs C5.1 and C5.1-myc from 0.5L cultures by SDS-PAGE.....	81

Figure 2.11	Expression analysis of GST-tagged constructs C6.1 and C6.1-myc from 0.5L cultures by SDS-PAGE.....	82
Figure 2.12	0.5L expression analysis of C5.1 and C5.1-myc in 1xPBS following GST-tag cleavage, by SDS-PAGE.....	83
Figure 2.13	0.5L expression analysis of C6.1 and C6.1-myc in 1xPBS following GST-tag cleavage, by SDS-PAGE.....	84
Figure 2.14	4L expression analysis of C5.1 and C5.1-myc following GST-tag cleavage, by SDS-PAGE.....	87
Figure 2.15	3L expression analysis of C6.1 and C6.1-myc following GST-tag cleavage, by SDS-PAGE.....	88
Figure 2.16	Stability analysis by centrifugation and immunoblot of purified C5.1, C5.1-myc, C6.1 and C6.1-myc in 1xPBS .....	91
Figure 2.17	Intact C5.1 molecular weight assessed by MALDI-TOF.....	93
Figure 2.18	Deconvoluted ESI-QTOF spectra for C5.1 (A) and C5.1-myc (B) .....	94
Figure 2.19	Intracellular localisation of C5.1-myc in fixed HEK293 cells is time-dependent and dose-dependent. ....	97
Figure 2.20	Intracellular localization of C5.1-myc in fixed HepG2 cells is time-dependent and dose-dependent. ....	98
Figure 2.21	C5.1-siRNA localization in live HepG2 cells in the absence of a membrane counterstain .....	101
Figure 2.22	Intracellular localization of C5.1-siRNA in live HepG2 cells counterstained with 1 mg/ml wheat germ agglutinin (WGA)-dsRed for 30 minutes .....	102
Figure 2.23	Intracellular localization of C5.1-siRNA-FITC in live HepG2 counterstained with 2.5 mg/ml wheat germ agglutinin (WGA)-dsRed for 10 minutes .....	103
Figure 2.24	Electrophoretic mobility shift assay with C5.1 and siRNA at various molar ratios.....	105
Figure 3.1	Schematic of the EB1-based fusion proteins cloned into a modified pET32-a or pGEX-6P-2 vector. ....	121
Figure 3.2	Insert screening by colony PCR following ligation in pGEX-6P-2.....	124
Figure 3.3	Insert screening by colony PCR following ligation in pET32-a, by a vector-specific forward primer and a gene-specific reverse primer.....	124

Figure 3.4	Optimization of His <sub>6</sub> -tagged C7.2 and C8.2 protein expression following incubation at 18°C and 37°C.....	126
Figure 3.5	Optimization of GST-tagged C9.2 and C10.2 protein expression following induction with 0.5 mM and 1.0 mM IPTG. ....	129
Figure 3.6	Optimization of GST-tagged C9.2 and C10.2 protein expression following induction incubation at 18°C or 37°C.....	130
Figure 3.7	Analysis of purified C7.2 and C8.2 by SDS-PAGE at various stages during refolding and buffer exchange.. ....	132
Figure 4.1	Schematic of fusion proteins comprised of the HIV Tat; Penetratin; DRBM1 and DRBM2; His <sub>6</sub> affinity tag.....	143
Figure 4.2	Purification of C11.2, C12.2 and C13.2 from 0.5L bacterial cultures under denaturing conditions.....	145
Figure 4.3	Electrophoretic mobility shift assays with C12.2 (A) or C5.1 and C13.2 (B) complexed with siRNA at various molar ratios. ....	147
Figure 4.4	Schematic of the chemical reaction of MTT reduction to formazan, catalyzed by mitochondrial NADH oxidoreductase. ....	149
Figure 4.5	Cell viability in HEK293 and HepG2 cells following a 24 hour incubation with increasing concentrations of C5.1, C11.2, C12.2 and C13.2 is cell-line dependent.. ....	151
Figure 4.6	Intracellular localisation of C5.1 and C12.2 complexed with 100 nM siRNA-fluorescein in live HepG2 cells. ....	153
Figure 4.7	Cell viability of HEK293-dEGFP cells treated for 10 days with increasing concentrations of G418 following transformation with pdEGFP-N1.....	155
Figure 4.8	FACS analysis of stable HEK293-dEGFP clones.....	155
Figure 4.9	FACS histogram analysis of stably transfected HEK293-dEGFP cell populations selected with 1.0 mg/ml G418.....	156
Figure 4.10	FACS histogram analysis of stably transfected HEK293-dEGFP cell populations maintained with 0.5 mg/ml G418 shows a shift to a lower dEGFP phenotype. ....	156
Figure 4.11	dEGFP expression analysis following treatment with 10 nM siRNA complexed with C5.1 or C12.2 for 48 hours.....	158
Figure 4.12	FACS analysis of a HEK293-dEGFP cell line following treatment with protein-siRNA complexes for 24 hours .....	160

Figure 4.13	dEGFP expression analysis following treatment with 100 nM siRNA complexed with C5.1 or C12.2 for 24 hours.....	161
Figure 4.14	Phase-contrast microscopy of HepG2 cells transduced with 25 nM PTPN1 siRNA complexed with proteins at a 50:1 molar ratio for 72 hours..	164
Figure 4.15	HepG2 cell lysate immunoblot analysis with an $\alpha$ -PTP1B antibody.....	166
Figure 4.16	Immunoblot analysis on HepG2 cell lysates following PTP1B knockdown for 72 hours. ....	166
Figure 5.1	Modes of action of the attenuation of protein activity.....	176

# List of Tables

Table 1.1	SiRNA-based therapeutics currently in clinical trials (2013). .....	31
Table 1.2	Examples of cell penetrating peptides that have shown efficient cargo transduction <i>in vitro</i> .....	40
Table 1.3	Notable examples of cargo successfully transduced by Penetratin.....	47
Table 1.4	Examples and cellular functions of the members of the dsRNA binding protein family (DRBPs).....	52
Table 2.1	Yield analysis of GST- tagged fusion proteins obtained during purification from 0.5L <i>E. coli</i> cultures .....	85
Table 2.2	Large-scale protein purification yields for C5.1, C5.1-myc, C6.1 and C6.1-myc .....	89
Table 2.3	Final concentrations and yields of cleaved proteins .....	91
Table 4.1	Final yields for purified proteins C11.2-C13.2 from 0.5 L cultures .....	144
Table 6.1	<i>E.coli</i> strains and associated genotypes .....	183
Table 6.2	Plasmid vectors and inserts .....	183
Table 6.3	Primer list for PCR amplification of constructs C1.1-C6.1 .....	184
Table 6.4	Primer list for PCR amplification of constructs C7.2-C10.2 .....	184
Table 6.5	Sequencing primer list.....	185
Table 6.6	Antibiotics used in bacterial cell cultures or in cell culture .....	186
Table 6.7	General buffers and solutions .....	186
Table 6.8	Reagents used in Polyacrylamide gel electrophoresis (SDS-PAGE).....	187
Table 6.9	Protein purification buffers used in IMAC under native conditions.....	187

Table 6.10	Protein purification buffers used in GST affinity chromatography .....	187
Table 6.11	Protein purification buffers used in IMAC under denaturing conditions ..	188
Table I.1	Amino acid sequences of constructs C1.1-C6.1.....	215
Table I.2	Amino acid sequences for constructs C5.1-myc and C6.1-myc .....	217
Table I.3	Amino acid sequences of GST-tagged constructs following purification and tag cleavage by Prescission protease.....	218
Table I.4	Amino acid sequences of constructs C7-C10.2.....	219
Table I.5	Amino acid sequences for constructs C11.2-C13.2 .....	220
Table II.2.1.1	Tests of Between-Subjects Effects in HEK293 cells. ....	227
Table II.2.1.2	Pairwise comparisons between proteins by 2-way Anova.....	228
Table II.2.1.3	Dunett's post-hoc test based on observed means following MTT assays in HEK293 cells .....	231
Table II.2.1.4	Tests of Between-Subjects Effects in HepG2 cells.....	231
Table II.2.1.5	Pairwise comparisons between proteins by 2-way Anova.....	232
Table II.2.1.6	Dunett's post-hoc analysis for each protein against C11.2 to indicate protein-specific effects in HepG2 cells.....	234
Table II.2.1.7	Dunett's post-hoc test based for each concentration against the highest concentration used to indicate dose-dependent effects in HepG2 cells .....	234

# List of Abbreviations

$\Delta T_m$ : Change in melting temperature ( $^{\circ}\text{C}$ )

2'-F: 2'-Fluorine substitution

2'-OMe: 2'-O methyl substitution

ADAR1: adenosine deaminase acting on RNA 1

ADP: Adenosine diphosphate

AGO1-4: Argonaute protein 1-4

AMD: age-related macular degeneration

ATP: Adenosine triphosphate

BCA: Bicinchoninic acid

CPP: Cell penetrating peptide

DDP-4: dipeptidyl peptidase 4 (DPP-4)

DGCR8: Di-George syndrome critical region 8

DME: diabetic macular edema

DNA: deoxyribonucleic acid

DRBD: dsRNA binding domain ( same as DRBM, dsRNA binding motif)

DRBP: dsRNA binding protein

dsRNA: double stranded RNA

ES cells: Embryonic Stem cells

F: phenylalanine

GLP-1: glucagon-like peptide-1

Grb-SOS: Grb-Son of Sevenless

HCV: Hepatic Hepatitis C virus

HSP90: Heat-shock protein 90

ICH: Institute of Child Health

IR: insulin receptor

IRS 1 and 2: insulin receptor substrates 1 and 2

JAK2: Janus kinase 2

LNA: Locked nucleic acid

LTR: long terminal repeat

MAPK: Map kinase

MID: Middle domain of Argonaute proteins

miR-122: microRNA-122

miRNA: microRNA

mRNA: messenger RNA

N: Amino domain of Argonaute proteins

NFAR 1 and 2: Nuclear factors associated with dsRNA 1 and 2

NMR: nuclear magnetic resonance

Nt: nucleotide(s)

PACT: protein activator of protein kinase PKR

PAZ: Piwi-Argonaute-Zwille domain

PDK1 and 2: 3-phosphoinositide-dependent protein kinase 1 and 2

PEG: polyethylene glycol

Pi: Phosphate group

PI3K: phosphatidyl inositol 3 kinase

PIP3: phosphatidyl-inositol 3,4,5-triphosphate

PKB: protein kinase B

PKN3: protein kinase N3

PKR: Protein kinase R

PLK1: polio-like kinase 1

PNA: peptide nucleic acid

pre-miRNA: precursor microRNA precursor

pri-miRNA: primary microRNA

PRR: Pattern Recognition receptor

PTD: Protein transduction domain

PTGS: post-transcriptional gene silencing

PTP1B: protein tyrosine phosphatase 1B

PTPN1: tyrosine-protein phosphatase non-receptor type 1 gene

PTPN2: tyrosine-protein phosphatase non-receptor type 2 gene

rasGAP: ras-GTPase-activating protein

RHA: RNA helicase A

RISC: RNA-induced silencing complex

RNA: ribonucleic acid

RNAi: RNA interference

rRNA: ribosomal RNA

RSV: respiratory syncytial virus

siRNA: short interfering RNA

SNALP: stable nucleic acid lipid particle

SPNR: spermatid perinuclear RNA binding protein

ssRNA: single stranded RNA

STAT3: Signal transducer and activator of transcription 3

TAR: trans-acting response element

TCPTP: T-cell Protein Tyrosine Phosphatase

T<sub>m</sub>: melting temperature (°C)

TRBP: HIV-TAR RNA-binding protein from HIV

TRBP: TAR RNA binding protein

tRNA: transfer RNA

TTR: transthyretin;

UCL: University College London

UTR: untranslated region

VEGF: vascular endothelial growth factor

W: tryptophan

εNHα : four trifluoromethylquinoline-based derivatives

# List of Formulae

1. Live cell count formula by the Trypan Blue method followed by hemocytometry:

*Average cell number in 4 squares  $\times n \times 10^4 = \text{number of cells/ml}$*

*Where  $n = \text{dilution factor (=5)}$  and  
 $10^4 = \text{number of } \mu\text{l in 1 ml.}$*

2. Molar concentration formula:

*Molar concentration of protein (mmol/L)  
= Concentration at  $A_{280}$  (mg/ml)  $\times$  MW (mMol/mg)  $\times$  Volume (1,000ml/L)  
 $\therefore [(\text{mg/ml}) / \text{MW (Da)}] \times 1,000$*

*Where  $A_{280}$  is the sample absorbance at 280nm and MW = molecular weight*

3. Concentration formula:

$$C_i (\mu\text{M}) \times V_i (\mu\text{l}) = C_f (\mu\text{M}) \times V_f (\mu\text{l})$$

*Where  $C_i = \text{initial concentration}$ ,  $V_i = \text{initial volume}$ ,  $C_f = \text{Final concentration}$  and  
 $V_f = \text{final volume}$*

4. Beer's Law:

*Molar absorptivity is constant (and the absorbance is proportional to concentration) for a given substance dissolved in a given solvent and measured at a given wavelength.*

$$A = \epsilon \times b \times C,$$

*Where  $A = \text{absorbance value}$ ,  $\epsilon = \text{extinction or molar absorption coefficient (L mol}^{-1} \text{ cm}^{-1}\text{)}$ ,  $b = \text{path length (cm)}$ ,  $C = \text{analyte concentration in moles/litre or Molarity}$*

*$\epsilon$  is proportional to the number of tyrosine (Y), tryptophan (W) and cysteine (C) amino acid composition at 280nm:*

$$\epsilon = (nW \times 5500) + (nY \times 1490) + (nC \times 125)$$

# Chapter 1. General Introduction

## 1.1 The RNAi pathway

In 1970, following the discovery of the double stranded DNA helix, Francis Crick proposed a model describing how DNA governs protein synthesis through an RNA intermediate. This model went on to become the ‘Central Dogma of biology’ (Crick, 1970), describing the linear flow of genetic information from the DNA level, to RNA and finally protein, in a hypothesis that is generally described as ‘one gene-one ribonucleic acid chain-one protein’. Messenger RNA (mRNA) was subsequently discovered by Jacob and Monod (1961), as the intermediate following gene transcription; and Crick’s hypothesis that one gene encodes for one protein product was rejected when Leder and Nirenberg ‘cracked’ the genetic code in 1964 and assigned nucleotide triplets (‘codons’) to each of the 20 amino acids (Leder and Nirenberg, 1964). The RNAi pathway, was not elucidated until much later, following the work of plant biologists with, oddly enough, petunias.

In the early nineties, Napoli and Jorgensen were investigating how to increase the purple coloration in petunias by introducing chalcone synthase as a transgene, in an attempt to overexpress the protein in the rate-limiting step in flavonoid biosynthesis. The resulting petunias came out white (Napoli *et al.*, 1990). This observation of gene co-suppression had never been reported before, although the mechanism was not elucidated until Fire *et al.* (1998) used double-stranded RNA (dsRNA) instead of single stranded RNA (ssRNA) in the nematode *C. elegans* to modify gene expression.

In 2006, Craig Mello and Andrew Fire were subsequently awarded the Nobel Prize in Medicine for their seminal work with RNA interference (RNAi), which highlighted an endogenous, evolutionarily conserved mechanism, that involves sequence-specific post-transcriptional gene silencing (PTGS) with dsRNA as a trigger, and not single-stranded

RNA (ssRNA) as had been previously reported by Guo and Kemphues (1995). According to their proposed model, double-stranded RNA (dsRNA) could interfere with gene expression by selectively binding to cognate mRNA and targeting it for degradation, thereby attenuating protein translation from the target mRNA.

In plants and animals, RNAi via microRNA (miRNA) transcription play an important role in cellular processes such as gene regulation, proliferation, development, tumour formation and homeostasis (Ambros and Lee, 2004). RNAi is also an innate antiviral mechanism; viral dsRNAs are generally long and perfectly complementary, inducing RISC loading, cleavage into 21-25 bp fragments (siRNAs) and cleavage upon cellular infection, inhibiting viral replication (Umbach and Cullen, 2009). In nematodes and plants, the initial antiviral response is potentiated by RNA-dependent polymerases, increasing the overall number of siRNAs available for RISC loading (Aoki *et al.*, 2007, Diaz-Pendon *et al.*, 2007). In mammals, viral infection activates the interferon (IFN) response via recognition by pattern-recognition receptors (PRRs). A systemic antiviral response is mounted via dissemination of the intercellular IFN response, which activates a positive feedback loop with the Janus kinase/Signal transducer and activator of transcription (JAK/STAT) pathway and the activation of RNase-encoding genes and protein kinase R (PKR) (Cullen, 1986, Umbach and Cullen, 2009). In the context of drug discovery, the value of siRNA is based on the assertion that drug targets, such as overexpressed oncogenes in cancer, or indeed the aberrant expression of any gene, can be knocked-down before it is translated into a mutated, misfolded or overexpressed protein. It therefore shows great promise as a therapeutic approach to many diseases whose underlying causes are genetic aberrations. In addition to oncologic applications, an RNAi strategy has been investigated in infectious diseases, such as malaria, targeting *plasmodium falciparum* (Malhotra *et al.*, 2002); HIV,

against the p24, Gag, and CCR5 proteins (Surabhi and Gaynor, 2002), amongst others, with promising results *in vitro*.

RNAi in humans can be triggered by dsRNA sequences 21-25 nt in length, via two distinct pathways: the natural and the synthetic pathway. The latter occurs via the introduction of stabilized and often chemically modified short interfering (siRNA) molecules, whereas the canonical RNAi pathway is initiated by the 'trimming' of long primary-microRNA (pri-miRNA) transcripts into a ~70 bp stem-loop or hairpin miRNA precursors (pre-miRNA) in the nucleus, by a protein complex comprised of the RNaseIII enzyme Drosha and DiGeorge syndrome critical region 8 (DGCR8), which collectively make up the Microprocessor complex (Ha and Kim, 2014). Pre-miRNA is exported from the nucleus by exportin 5, a dsRNA binding protein, for further processing by Dicer in the cytoplasm.

Once in the cytoplasm, pre-miRNA is cleaved into an miRNA duplex by Dicer, an RNaseIII endonuclease (Zhang *et al.*, 2004) which works in partnership with HIV-TAR RNA-binding protein (TRBP), Drosha and the protein activator of protein kinase PKR (PACT) (Fig. 1.1). Dicer proceeds to remove any loop structures and cleave pre-miRNA into shorter, 19-22 bp dsRNA duplexes with 2 nucleotide (nt) 3'-hydroxyl overhangs.

Similarly, Dicer processes endogenous, long dsRNA precursors into siRNA duplexes, 20-25 bp in length. Endogenous siRNA precursors may arise from *trans* interactions between gene and pseudogene transcripts, or from hairpin transcripts (Czech and Hannon, 2011). In mammals, these precursors have so far been found only in embryonic stem (ES) cells and mouse oocytes (Babiarz and Blelloch, 2008, Babiarz *et al.*, 2008, Tam *et al.*, 2008, Watanabe *et al.*, 2008). Exogenous siRNA, procured as synthetic 20-25 nt oligos are not

subject to Dicer's RNase III action and are this directly loaded onto the RNA-induced silencing complex (RISC).

Once cleaved, small dsRNAs are sorted by Dicer based on the precursor molecule, the structure of miRNA and siRNA duplexes, terminal nucleotides, thermodynamic properties and the Argonaute (AGO) protein they are destined for.

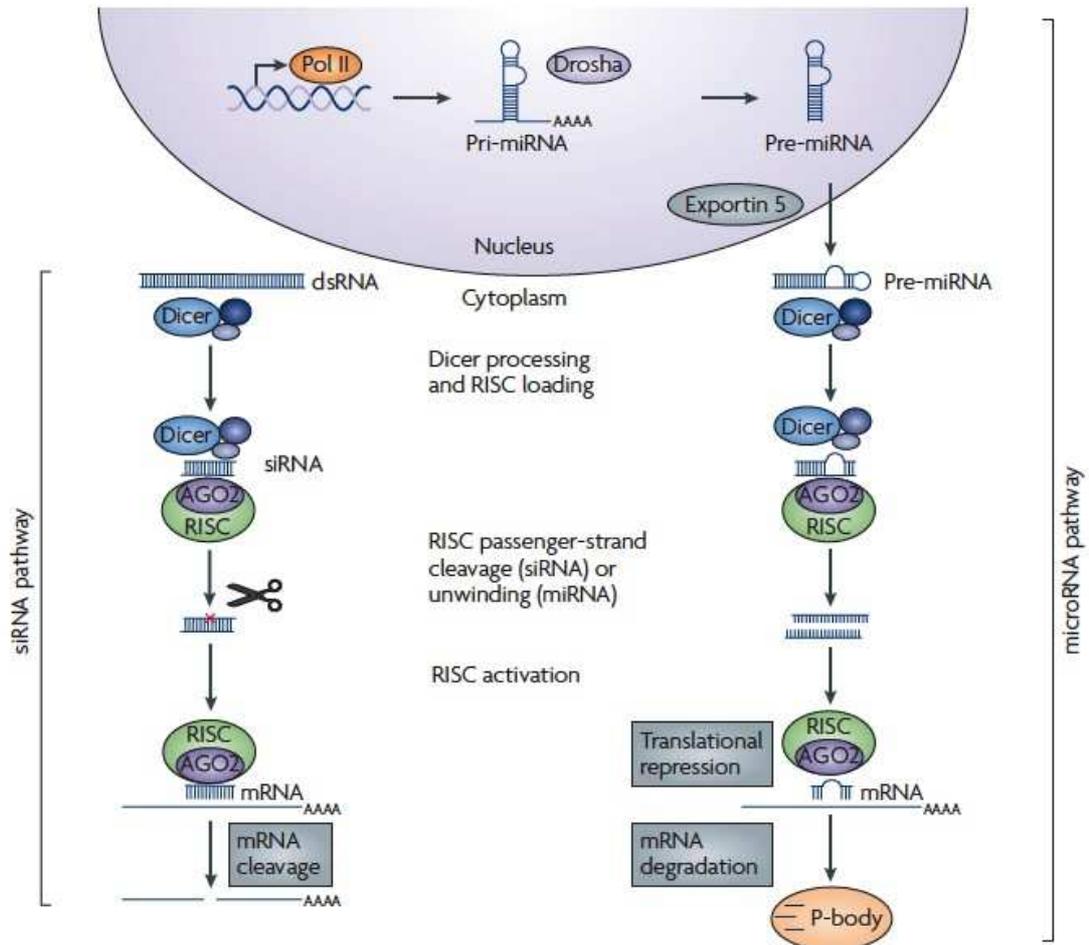
The Argonaute clade of mammals consists of eight members in the Ago (AGO1-AGO4), and Piwi (PIWIL) proteins (Ross and Kassir, 2014), although only AGO1-AGO4 have been shown to bind siRNA and miRNA indiscriminately (Burroughs *et al.*, 2011, Czech and Hannon, 2011, Liu *et al.*, 2004, Meister *et al.*, 2004). These are expressed ubiquitously in humans, although in different, tissue-specific, proportions (Reviewed by Meister, 2013) Post-translational gene silencing is mediated by AGO2, the only member of the AGO clade that possesses mRNA cleavage activity in mammals. Studies have shown that the silencing action of siRNAs targeting the coding sequence of several endogenous genes is inhibited in the absence of AGO2 in mice livers (Ruda *et al.*, 2014). Furthermore, studies by Gregory *et al.* (2005) and Maniataki and Mourelatos (2005) have shown that small dsRNA sorting by Dicer and AGO loading is a concerted effort, at least in animals, whereas Jinek and Doudna (2009) proposed that the structure of AGO proteins mediates dsRNA loading, unwinding and the formation of RNA induced silencing complex (RISC). AGO proteins are comprised of an amino-terminal (N), a PIWI-Argonaute-Zwille (PAZ), a middle (MID) and a PIWI domain. The N and PAZ domains collectively make up lobe 1, whereas the MID and PIWI domains make up lobe 2. Lobes 1 and 2 are separated by a flexible hinge region that mediates the structural rearrangement of AGO proteins upon dsRNA binding. According to their model, the PAZ domain anchors the 3' end of small dsRNA by bending it into a binding pocket, whereas the MID domain anchors the 5' end of dsRNA to a

conserved tyrosine residue by stacking interactions and hydrogen bonds. The N domain is required for loading and unwinding of the duplex during RISC assembly (Kwak and Tomari, 2012), whereas the PIWI domain in AGO proteins functions as an endonuclease to cleave target mRNA, following small dsRNA binding (Jinek and Doudna, 2009).

Cleavage of the passenger strand and subsequent unwinding of the dsRNA duplex is governed by the 'Asymmetry rule' whereby Argonaute2 (AGO2) action is guided by the strength of base-pair binding on the 5' end of the duplex - its thermodynamic stability (Schwarz *et al.*, 2003) - with the less stably paired 5' end strand preferentially loaded onto AGO. TRBP acts as an asymmetry sensor, and allows the guide strand to be loaded onto RISC following dicing (Noland *et al.*, 2011). In order for the mature RISC to form, siRNA or miRNA transfer to AGO is mediated by heat shock protein 90 (HSP90). HSP90 allows AGO to remain in an open state to accommodate dsRNA by hydrolyzing Adenosine triphosphate (ATP) to Adenosine diphosphate (ADP) and phosphate (Pi) (Iki *et al.*, 2010, Johnston *et al.*, 2010). RISC is a multimeric protein complex, comprised of Dicer, TRBP and the 'slicing' enzyme AGO2 (Meister *et al.*, 2004). Once RISC is assembled, the N-domain of the Ago proteins unwind the complex, although reports by Meister *et al.* (2005) and Gu *et al.* (2011) support the action of RISC-associated helicases, such as RNA helicase A (RHA) when unwinding the dsRNA helix.

In the endogenous miRNA pathway, the 'seed' sequence, comprised of nt 2-8 within a longer miRNA, is perfectly complementary to the cognate mRNA. Complementarity between at least 7 bp of siRNA-mRNA is required for efficient silencing, thus providing some specificity to this ubiquitous mechanism (Elbashir *et al.*, 2001b). Where partial sequence complementarity occurs, *i.e.* where the RNA duplex contains internal loops or bulges, translational repression occurs via RNAi. This occurs when miRNAs are loaded

onto RISC, rather than siRNAs; RISC in this case is guided to partially complementary sequences typically found in the 3' untranslated region (3' UTR) of target mRNA, leading to translational silencing or exonuclease degradation (Huntzinger and Izaurralde, 2011). In the case of siRNAs (either endogenously generated or exogenously introduced), where perfect Watson-Crick base pair complementarity exists, RISC is directed to induce the degradation of primary mRNA transcripts by AGO2-mediated cleavage (Kim *et al.*, 2009b, Umbach and Cullen, 2009). Moreover, the greater the number of RISCs bound to a target transcript, the greater the inhibitory effect of the RNAi response (Doench *et al.*, 2003).



**Figure 1.1 Mechanism of RNA interference in mammalian cells.** RNA interference (RNAi) pathways are guided by small RNAs that include small interfering RNA (siRNA) and microRNAs (miRNAs). The siRNA pathway begins with cleavage of long double-stranded RNA (dsRNA) by the Dicer enzyme complex into siRNA, which are then incorporated into Argonaute 2 (AGO2) and the RNAi-induced silencing complex (RISC). If the RNA duplex loaded onto RISC has perfect sequence complementarity, AGO2 cleaves the passenger (sense) strand so that active RISC containing the guide (antisense) strand is produced. The siRNA guide strand recognizes target sites to direct mRNA cleavage. RNAi therapeutics typically involve the delivery of synthetic siRNA into the cell cytoplasm. The microRNA pathway begins with endogenously encoded primary microRNA transcripts (pri-miRNAs) that are transcribed by RNA polymerase II (Pol II) and are processed by the Drosha enzyme complex to yield precursor miRNAs (pre-miRNAs). These precursors are then exported to the cytoplasm by exportin 5 and subsequently bind to the Dicer enzyme complex, which processes the pre-miRNA for loading onto the AGO2–RISC complex. When the RNA duplex loaded onto RISC has imperfect sequence complementarity, the passenger (sense) strand is unwound leaving a mature miRNA bound to active RISC. The mature miRNA recognizes target sites (typically in the 3'-UTR) in the mRNA, leading to direct translational inhibition. Binding of miRNA to target mRNA may also lead to mRNA target degradation in processing (P)-bodies. Adapted from Kim and Rossi (2007).

## 1.2 siRNA as a therapeutic drug: challenges and limitations

SiRNA's high specificity for their target mRNA, potency and adaptability are favorable attributes. From the pharmaceutical industry's perspective, they are easy to manufacture, relatively cost-effective, compared to small molecule inhibitors, with limited cytotoxicity, and can be designed *in silico* with various stabilizing modifications to complement a specific mRNA of interest (Vaishnav *et al.*, 2010), and avoid the elicitation of off-target effects. Since transfected synthetic siRNA sequences can be used to induce RISC (Elbashir *et al.*, 2001a), the RNAi pathway can be potentially harnessed to downregulate the expression any disease-associated gene. Moreover, siRNA-mediated RNAi has become a routine tool for post-translational gene modification (Aagaard and Rossi, 2007, de Fougères *et al.*, 2007), as sequences are amenable to chemical modifications and improved pharmacologic profiles. Their high specificity may also allow targeting of specific alleles that differ from wild-type ones only by a few mutations. They are also less likely to interfere with gene regulation by endogenous miRNAs, as they enter the RNAi pathway later and are not produced by cleavage of longer dsRNA precursors by RISC (Grimm *et al.*, 2006, Rao *et al.*, 2009). Such is the potential of siRNA that within three years since its use as a gene expression modulator in cultured human cells was confirmed (Elbashir *et al.*, 2001a), antisense oligonucleotide-based therapeutics had entered clinical studies utilizing the endogenous RNA pathway as a therapeutic strategy. Many siRNA-based formulations are still being tested at various clinical phases (Summarized in Table 1.1). The only FDA-approved RNAi-based drugs as of 2014 are fomivirsen (marketed as Vitravene) for the treatment of cytomegalovirus retinitis (CMV) for HIV or immunocompromised patients via intraocular injection, and mipomersen (Kynamro) which targets mutated apolipoprotein B (apoB) in familial hypercholesterolemia. Both drugs were

developed by Isis Pharmaceuticals. It is interesting to observe that both FDA- approved drugs are composed of chemically stabilized single stranded antisense oligos; so far there have been no successful clinical trials with siRNA-based formulations.

Indeed, the development of siRNA-based therapies has been met with many challenges. As a drug that exploits a natural pathway to mediate its action, there is always the risk of saturating the RNAi pathway and disturbing the intracellular equilibrium. *In vivo* studies by Grimm *et al.* (2006) and Yi *et al.* (2005) have reported elevated, dose-dependent lethality in mice tested with 49 different short hairpin RNAs (shRNAs), due to liver toxicity. Their results suggested that the rate-limiting step to efficient RNAi was nuclear export by exportin-5, as introduced shRNAs had blocked the canonical miRNA pathway. Seminal work by Elbashir *et al.* (2001a), showed that transfection with oligonucleotide sequences shorter than 30 bp, did not induce any cytotoxicity. However, the introduction of longer ( $\geq 30$  bp) dsRNAs has been shown to activate the immune response by protein kinase R (PKR) - the first point of contact when viral dsRNA enters a mammalian cell. Similarly, transfection with GU-rich sequences such as the 5'-GUCCUCAA-3' 'danger motif' (Hornung *et al.*, 2005), leads to the induction of the interferon response through activation of pattern recognition receptors (PRRs), such as Toll-like receptors 3 (TLR3) on the cell surface, and TLR7/8 upon siRNA entry. TLR3 are transmembrane receptors that are found on the cell surface of dendritic cells, in the epithelium and endothelium (Zimmer *et al.*, 2011) and are able to 'sense' unmodified siRNA sequences  $\geq 23$ -33 bp in length, in a sequence-independent, but cell-type dependent manner (Kariko *et al.*, 2004, Reynolds *et al.*, 2006). Upon internalization, endocytic TLR7 and TLR8 expressed within dendritic cells, macrophages, or regulatory T cells (Kokkinopoulos *et al.*, 2005) can also sense internalized siRNA, in a sequence-dependent manner, favoring GU-rich or AU-rich

duplexes (Diebold *et al.*, 2006). Immune-mediated toxicity is thus elicited through the induction of interferon- $\alpha$  (IFN $\alpha$ ), IFN $\gamma$  and interleukin-12 (IL-12) (Robbins *et al.*, 2009). Another significant drawback is the unwanted side effects caused by less-than-perfect siRNA complementarity to their target sequences. Although Elbashir *et al.* (2001b) had reported a short, 7 bp ‘seed’ sequence as a targeting feature on cognate mRNA sufficed for RNAi, microarray screening analysis by Jackson *et al.* (2003) showed that a sequence as short as 11 bp could induce off-target effects with toxic phenotypes. A ‘seed’ sequence is necessary for mRNA recognition, as it places the target’s scissile phosphate group directly within the slicer site of AGO2 (Parker *et al.*, 2005)

Other issues with unformulated siRNAs as a therapeutic platform include poor serum stability, degradation by RNase A- type nucleases, poor bioavailability, a short half-life (Moschos *et al.*, 2007) or clotting effects (Bramsen and Kjems, 2012, Pecot *et al.*, 2011). In addition, compared to small molecule inhibitors (SMocs), the large size of siRNA molecules necessitate the development of stabilizing formulations that will allow the siRNA molecules to enter cells, without activating patrolling macrophages or the intracellular innate immunity protein repertoire, escape from endosomes, and retain potency and activity after loading onto the RNAi machinery *in vivo*. These hurdles, as well as the pharmacokinetic properties of siRNA, make this molecule a ‘difficult’ bioactive drug. In addition, its small size ( $\leq 10$  nm in diameter) poses serious challenges for systemic delivery, as it is rapidly cleared by the renal system; studies have shown that most of intravenously administered siRNA rapidly accumulates in the kidneys (Moschos *et al.*, 2007, van de Water *et al.*, 2006). Endosomal entrapment and subsequent degradation of internalized siRNA formulations by lysosomal nucleases, necessitate the development of fusogenic vectors that will become protonated upon acidification and allow the siRNA to

escape unharmed from endosomes. It is still unknown whether siRNA will ever become a universal ‘Superdrug’ that will revolutionize personalized medicine, but both academia and industry have been working hard at elucidating the key player in the puzzle: an efficacious delivery vector.

**Table 1.1 SiRNA-based therapeutics currently in clinical trials (2013).** Local delivery refers to ocular, pulmonary, pancreatic and colonic tissues; non-invasive patient administration. Targeted systemic delivery refers to intravenous (IV) uptake, or delivery by synthetic carriers. RSV, respiratory syncytial virus; TTR, transthyretin; VEGF, vascular endothelial growth factor; PLK1, polio-like kinase 1; PKN3, protein kinase N3, effector molecule in the phosphatidyl inositol 3 kinase (PI3K) metastatic pathway; miR-122, miRNA-122; HCV, Hepatitis C virus; SNALP, stable nucleic acid lipid particle; LNA, locked nucleic acid; AMD, age-related macular degeneration; DME, diabetic macular edema; TLR, Toll-like receptor. (Adapted from Zhou *et al.* (2013). Confirmed status April 2014)

Company	Lead Drug(s)	Target	Disease	Delivery	Clinical Trial Number, Trial Phase
<b>Alnylam</b>	ALN-RSV01	Nucleocapsid 'N' gene	RSV infection in lung transplant patients and prophylaxis in healthy patients	Local (nebulizer, intranasal spray)	NCT01065935, I/II (abandoned)
	ALN-VSP02	VEGF	Liver cancer	Local	NCT01158079, I
	ALN-TTR02	TTR	Hereditary amyloidosis	Systemic (SNALP)	NCT01961921 III (ongoing)
	ALN-TTRsc	TTR	Hereditary amyloidosis	Systemic (SNALP)	NCT01981837, II (ongoing)
	ALN-PCS	PCSK9	Severe hyper-cholesterolemia	Systemic (subcutaneous ESC-GalNAc conjugate)	NCT01437059, I (completed)
<b>Tekmira</b>	TKM-PLK1	PLK1	Solid cancer tumors	Systemic (SNALP)	NCT01437007, I/II (completed)
	TKM-EBOLA	Ebola viral infection	Ebola virus (biodefense); <i>Fast-tracked by the FDA, April 2014</i>	Systemic (subcutaneous LNA)	NCT01518881, I (terminated)
<b>Arrowhead/Silence</b>	Atu-027	PKN3	Advanced metastatic pancreatic adenocarcinoma	Systemic (Lipoplex; cationic liposomes)	NCT01808638, Ib/IIa (ongoing)
<b>Opko Health Inc</b>	Bevasiranib	VEGF receptor	AMD, DME	Local (Intravitreal)	NCT00306904II (terminated, low efficacy)
<b>Allergan/siRNA Therapeutics</b>	AGN-211745	VEGF receptor	AMD, DME	Local (Intravitreal)	NCT00363714, II (discontinued, off-target effects)
<b>Quark / Pfizer</b>	PF-655	Proangiogenic factor RTP801	AMD, DME	Local	NCT01445899, NCT00713518, I/II and II (terminated, TLR activation)

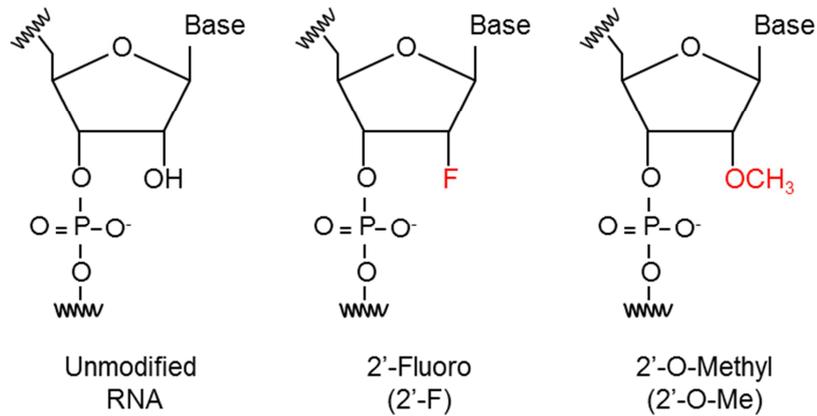
### 1.3 Engineering stabilizing modifications for siRNA

In addition to the above challenges in the development of siRNA as a therapeutic molecule, the elicitation of off-target effects as a reflection of the common RNAi pathway between miRNA and siRNA (Doench *et al.*, 2003) and the competition, and ultimate saturation, of the RNAi machinery by both cognate miRNAs and exogenously introduced siRNAs is still a major issue (Castanotto *et al.*, 2007, Vickers *et al.*, 2007). Chemical engineering of siRNA has overcome many of these limitations, either by modifying (i) the phosphodiester backbone, (ii) the ribose sugar, (iii) the nucleoside base, or (iv), the nucleotide structure itself, in order to maximize potency and minimize unwanted side-effects. Common backbone modifications include the phosphorothioate (PS) (Detzer *et al.*, 2008) and boranophosphate modifications (Hall *et al.*, 2006), which improve nuclease resistance, biodistribution and uptake by cells. Modifications in the furanose ring of the ribose involve substituting the 2'-OH group with either 2'-F or 2'-OMe structures that confer potency, specificity and a reduction in immunogenicity, with minimal effects on conformation and activity (Fig. 1.2). Finally, changes in the nucleoside base, most commonly Uracil, are used to increase base-pairing properties either alone or in conjunction with other stabilizing modifications (Bramsen and Kjems, 2012).

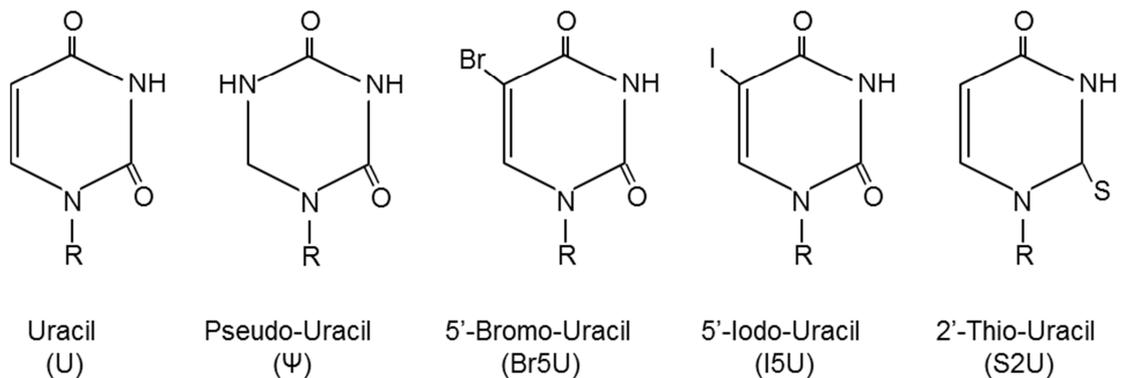
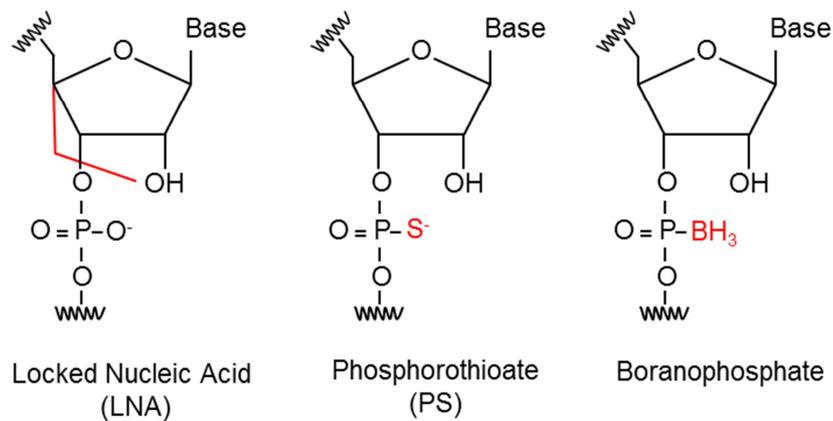
The site of modification on the double helix is critical, as chemical manipulation of the siRNA structure bias affects strand loading onto Argonaute proteins. As a general rule, each backbone modification decreases the melting temperature ( $T_m$ ) of siRNA duplexes, increasing their thermodynamic stability by 0.3-0.5°C per modification (Amarzguioui *et al.*, 2006, Harborth *et al.*, 2003). A difference in melting temperature ( $\Delta T_m$ ) of 3-5°C has been observed in locked nucleic acid (LNA) monomers (Koshkin and Wengel, 1998), nucleic acid analogues with enhanced thermodynamic stabilities. By 'locking' the C4 and

2'-OH groups in a ribose monomer, molecule rotation is restricted, thereby forcing the monomer to adopt an A-RNA structure (Kierzek *et al.*, 2009). Incorporation of LNA modifications into siRNA molecules has been shown to greatly siRNA delivery and specificity (Elmen *et al.*, 2004) and therefore provide attractive alternatives to siRNA oligos. Previous work by Chiu and Rana (2003), Schwarz *et al.* (2003), and more recently by Bramsen *et al.* (2009), has identified that siRNA is amenable to moderate, but not extensive, modifications or conjugations on the 3'-terminus of the antisense strand, or both 5'- and 3'- termini of the sense strand, with little influence on its activity-related properties. Many companies now offer custom-made siRNA sequences with proprietary modifications that can be used in gene-regulation studies, although the type and extent of such modifications are not always disclosed.

A.



B.



**Figure 1.2 Common chemical siRNA modifications.** **A.** Stabilising modifications in the ribose furanose ring (top panel) or in the phosphodiester backbone (lower panel). **B.** Side chain substitutions in Uracil, the most commonly stabilized base.

## 1.4 siRNA Delivery Strategies

However elegant siRNA modifications have been in increasing target specificity and stabilization against degradation, the current challenge in siRNA biotherapeutics development is targeted, efficient and non-immunogenic delivery to tissues of interest. Indeed, preservation of bioactivity, biodistribution, selectivity and target specificity for any given drug is no small feat. The development of either systemic or tissue-specific drug delivery vectors has been the bottleneck for many biotechnology companies, and the tombstone for some of big pharma's RNAi technology divisions (Novartis, for example, has decided to shut down its RNAi R&D operations, a decision 'driven by ongoing challenges with formulation and delivery [...]') (FierceBiotech, 2014) following Merck, Roche, Abbott and Pfizer in 2011.

The incorporation of specificity-enhancing, stabilising modifications have, to a great extent, minimised issues such serum degradation and off-target effects encountered with naked siRNA. Utilisation of naked siRNA in formulations is also problematic due to its highly charged anionic nature, which hinders uptake via the hydrophobic cell membrane. Formulation attempts include non-viral vectors to increase bioavailability and specificity, and to overcome the electrostatic barrier posed by the cell membrane. Such vectors involve lipid-based formulations, including neutral liposomes, cationic lipids and lipoplexes; stable nucleic acids (SNALPs), which are PEGylated lipid nanoparticles for systemic delivery, various polymers (chitosan nanoparticles, polyethyleneimine (PEI), cyclodextrin and dendrimers), each with its own advantages and limitations with respect to efficacy and toxicity. Lipid-based delivery strategies, especially cationic liposomes, have been used to deliver siRNA into mammalian cell cultures *in vitro* (Taetz *et al.*, 2009), but are of limited use in an *in vivo* setting as they fail to release cargo and elicit dose-dependent

inflammation (Spagnou *et al.*, 2004). Similarly, chitosan nanoparticles, derived from natural polysaccharides, lack endosomolytic escape properties, which limits their bioavailability even in an *in vitro* setting (Wu *et al.*, 2012). These findings further highlight the need of effective siRNA carriers that will: (i) be able to protect its cargo from nucleophilic/lysosomal degradation, (ii) be effectively internalized either specifically in tissues of interest, or be stable for systemic delivery, (iii) promote the cytoplasmic release of nucleotide cargo, (iv) exhibit high biological activity at low concentrations and finally (iv) possess a good biosafety profile for *in vivo* therapeutic approaches (Vaishnaw *et al.*, 2010).

## 1.5 Cell Penetrating Peptides

Cell penetrating peptides (CPPs); represent a class of therapeutically interesting vectors for targeted drug delivery. The classical definition of CPPs, or protein transduction domains (PTDs) is that of short cationic sequences (less than 30 amino acids long), rich in arginine (Arg, R), histidine (His, H), and lysine (Lys, K) residues. Their highly polar amino acid sequence allows them to cross the negatively charged lipid bilayer of the cell membrane, a key hurdle in the delivery of macromolecules into cells. Moreover, they have a low molecular weight (<500 Da) (Kerkis *et al.*, 2006). Their internalisation efficiency depends on the number of arginine (R) residues in the peptide backbone, with R6 and R8 synthetic repeats showing the highest, compared to other CPPs (Futaki *et al.*, 2001). Interactions with the cell membrane involve binding interactions between key polyanions on the external face of the cell membrane, such as heparan sulfate, heparin, polysialic acid, as well as nucleic acids, and CPP (Poon and Garipey, 2007).

The classification criteria for CPPs have greatly expanded in the past 10 years, and now involve different classes of molecules, that can span the cell membrane, including hydrophobic protein sequences, and tissue specific proteins identified by phage display (Milletti, 2012). Most cell penetrating peptides do not obey Lipinski's rule of five<sup>1</sup> (Lipinski *et al.*, 2001), and must therefore be examined on a case-by-case basis for efficacy.

CPPs can be categorised based on their origin or their classification (Table 1.2). Classical CPPs, such as the HIV transactivator of transcription (TAT), the Herpes simplex virus

---

<sup>1</sup> Lipinski's rule of 5: In a study of  $\geq 2000$  pharmacological agents, Lipinski concluded that a drug is more likely to be cell-permeant if  $MW \leq 500$ , lipophilicity ( $\text{Log}_p$ )  $\leq 5$ , number of H-bond donors  $\leq 5$  and number of H-bond acceptors  $\geq 10$ .

protein VP22, and the Antennapedia homeodomain from *Drosophila melanogaster*, are considered classical CPPs. These are naturally derived proteins, which obey the structural rules of amphipathicity or cationic charge and adopt an  $\alpha$ -helical structure upon interaction with the phospholipid membrane (Oehlke *et al.*, 1997, Scheller *et al.*, 2000). For example, HIV-Tat's ability to translocate through the cell membrane is conferred by its arginine residues whereas, Penetratin's ability (the third helix of the Antennapedia homeodomain), is based on its hydrophobic core (W6, F7) (Christiaens *et al.*, 2002, Derossi *et al.*, 1996, Magzoub *et al.*, 2001).

Two strategies have been used so far to conjugate siRNA cargo to CPPs; one is covalent, and the other non-covalent, via electrostatic interactions between protein and cargo. These are sufficient to hold the complexed siRNA-protein together and deliver it intracellularly (Eguchi *et al.*, 2009, Lundberg *et al.*, 2007).

Futaki *et al.* (2001) first demonstrated that poly-arginine sequences were able to transduce cargo intracellularly by virtue of bidentate hydrogen bonding between the guanidinium group in arginine and the phosphate backbone. Translocation efficiency was thus found to be directly proportional to the number of arginine residues within a cationic sequence (Tunnemann *et al.*, 2008). A number of either natural or synthetic sequences have since been designed, containing transducing motifs in tandem, or combinations thereof (Abes *et al.*, 2007, Meade and Dowdy, 2007) and tested in vitro with various cargoes, which have been either covalently or non-covalently conjugated to the peptides.

Covalent conjugation strategies include disulfide, amide, hydrazine and thiazolidine bonds between the 3' end of the sense strand and the C' or N' terminus of the protein (Chiu *et al.*, 2004, Davidson *et al.*, 2004, Detzer *et al.*, 2009, Moschos *et al.*, 2007, Muratovska and

Eccles, 2004) and have been used mostly in the delivery of oligonucleotide cargoes, such as morpholinos, peptide nucleic acids (PNA) (Fabani and Gait, 2008), steric block oligonucleotides (Abes *et al.*, 2008) or full-length proteins (Snyder and Dowdy, 2005), amongst others (Dietz and Bahr, 2004). Non-covalent strategies, which rely on amphipathic or cationic proteins with a polar and a non-polar domain, arranged either sequentially (i.e. within the primary sequence), or by conformation (once the protein has adopted its three-dimensional structure in solution) (Deshayes *et al.*, 2008), have become increasingly popular. Electrostatic interactions between the positively charged carrier and its anionic cargo suffice for efficient delivery, without necessitating the incorporation of chemical, and often structurally-modifying, conjugations. Simeoni *et al.* (2003), first illustrated the transduction of siRNA by MPG $\alpha$  via non-covalent conjugation *in vitro*, followed by reports with polyarginine (Kim *et al.*, 2006), HIV-Tat (Eguchi *et al.*, 2009) and the endosomolytic analogue of Penetratin, EB1 (Lundberg *et al.*, 2007).

Interest in CPP development as a drug delivery vector stems from their abilities to enter most, if not all, cell types non-specifically, including primary and difficult to transfect cell lines (Eguchi *et al.*, 2009). They also show rapid cellular uptake and do not rely on nucleic acid integration into the host cell genome (El-Andaloussi *et al.*, 2011, Sugita *et al.*, 2008), compared to viral vectors. The most commonly used CPPs are summarized in Table 1.2. Another class of CPPs are the stapled peptide CPPs. These were developed in an attempt to circumvent issues such as proteolytic degradations and renal clearance (Jenssen and Aspmo, 2008) and to provide enhanced conformation stability via  $\alpha,\alpha$ -disubstitution and macrocyclic bridge formation (Henchey *et al.*, 2008). By adopting such a hydrocarbon staple, proteins are induced to form  $\alpha$ -helices, with an increase in target affinity, thereby providing an alternative solution to inherently unstable CPPs (Verdine and Hilinski, 2012)

**Table 1.2** Examples of cell penetrating peptides that have shown efficient cargo transduction *in vitro*.  $\epsilon\text{NH}^{\alpha}$  represents four trifluoromethylquinoline-based derivatives via a succinylated lysine tree. Polar amino acids Arginine (R), red; Lysine (K), blue. (Laufer and Restle, 2008, Lukanowska *et al.*, 2013, Regberg *et al.*, 2012).

Cell penetrating peptide	Classification based on origin	Sequence	Reference
<b>AntpHD</b>	Classical	MGR <b>K</b> RGRQTY <b>R</b> YQT LELE <b>K</b> EFHF <b>N</b> RYL <b>R</b> <b>R</b> <b>R</b> R <b>I</b> E <b>I</b> A <b>H</b> A <b>L</b> C <b>L</b> T <b>E</b> R <b>Q</b> <b>I</b> <b>K</b> <b>I</b> WFQ <b>N</b> <b>R</b> <b>R</b> <b>M</b> <b>K</b> <b>W</b> <b>K</b> <b>K</b> <b>E</b> <b>N</b>	Joliot <i>et al.</i> (1991a)
<b>Penetratin</b>	Classical	<b>R</b> <b>Q</b> <b>I</b> <b>K</b> <b>I</b> <b>W</b> <b>F</b> <b>Q</b> <b>N</b> <b>R</b> <b>R</b> <b>M</b> <b>K</b> <b>W</b> <b>K</b> <b>K</b> <b>E</b> <b>N</b>	Derossi <i>et al.</i> (1994)
<b>VP22</b>	Classical	NAATAT <b>R</b> GRSAAS <b>R</b> P <b>T</b> <b>Q</b> <b>R</b> <b>P</b> <b>R</b> APARSAS <b>R</b> PR <b>R</b> <b>P</b> V <b>Q</b>	Elliott and O'Hare (1997)
<b>HIV-Tat</b>	Classical	<b>G</b> <b>R</b> <b>K</b> <b>K</b> <b>R</b> <b>R</b> <b>Q</b> <b>R</b> <b>R</b> <b>R</b> <b>Q</b> <b>C</b>	Dyson <i>et al.</i> (2004), Frankel and Pabo (1988)
<b>MPG<math>\alpha</math></b>	Classical	GALFLAFLAAALSLMG LWSQP <b>K</b> <b>K</b> <b>K</b> <b>R</b> <b>K</b> <b>V</b>	Simeoni <i>et al.</i> (2003)
<b>PepFect-6</b>	Chimeric	stearyl- AGYLLG <b>K</b> ( $\epsilon\text{NH}^{\alpha}$ )INL <b>K</b> ALAALAK <b>K</b> <b>K</b> IL-NH <sub>2</sub>	El-Andaloussi <i>et al.</i> (2011)
<b>TP10</b>	Chimeric	AGYLLG <b>K</b> INL <b>K</b> ALAAL AK <b>K</b> IL-NH <sub>2</sub>	El-Andaloussi <i>et al.</i> (2005)
<b>Oligo-Arginine</b>	Classical/ Synthetic	<b>R</b> <sub>(6-12)</sub>	Tunnemann <i>et al.</i> (2008)
<b>Antp-MEK1</b>	Sychnologic bioportide	<b>R</b> <b>Q</b> <b>I</b> <b>K</b> <b>I</b> <b>W</b> <b>F</b> <b>Q</b> <b>N</b> <b>R</b> <b>R</b> <b>M</b> <b>K</b> <b>W</b> <b>K</b> <b>K</b> <b>G</b> <b>M</b> <b>P</b> <b>K</b> <b>K</b> <b>K</b> <b>P</b> <b>T</b> <b>P</b> <b>I</b> <b>Q</b> <b>L</b> <b>N</b> <b>P</b>	Kelemen <i>et al.</i> (2002)
<b>Mastoparan</b>	Rhegnylogic bioportide	INL <b>K</b> ALAALAK <b>K</b> ILa	Higashijima <i>et al.</i> (1988), Pooga <i>et al.</i> (1998)
<b>CADY</b>	Synthetic	GL <b>W</b> <b>R</b> AL <b>W</b> <b>R</b> LL <b>R</b> SL <b>W</b> <b>R</b> LL <b>W</b> <b>R</b> <b>A</b>	Crombez <i>et al.</i> (2009)
<b>EB1</b>	Synthetic- endosomolytic	L <b>I</b> <b>R</b> <b>L</b> <b>W</b> <b>S</b> <b>H</b> <b>L</b> <b>I</b> <b>H</b> <b>I</b> <b>W</b> <b>F</b> <b>Q</b> <b>N</b> <b>R</b> <b>R</b> L <b>K</b> <b>W</b> <b>K</b> <b>K</b> <b>K</b> -amide	Lundberg <i>et al.</i> (2007)

## 1.6 HIV-Tat - a (brief) timeline

HIV's transactivator of transcription (Tat) was the first CPP to be discovered. Frankel and Pabo (1988), observed that exogenously added transcription-transactivating (Tat) protein from HIV-1 could enter HeLa cells unaided and localise in the nucleus. Mutational analyses by site-directed mutagenesis of polar amino acids with non-polar substitutions showed that the conserved basic domain (GRKKR), which acts as a nuclear localisation signal, was indispensable for its intracellular localisation (Hauber *et al.*, 1989). The full-length protein is 86 amino acids long, with a highly conserved cysteine region, which acts as the regulatory region, and a basic region (Tat<sup>47-58</sup>), rich in lysine and arginine residues (Arya *et al.*, 1985). Its function is to trans-activate transcription of the viral genome from the HIV long terminal repeat (LTR), one of the structural landmarks of the HIV ssRNA genome (Sodroski *et al.*, 1985). It does this by associating with the trans-acting response element (TAR) RNA, as well as improving translational efficiency (Marciniak *et al.*, 1990). Calnan *et al.* (1991) showed that the arginine-rich RNA-binding domain only adopts a helical structure upon binding dsRNA, and that this property is dependent on the overall charge density of Tat<sup>47-58</sup>, rather than sequence specificity. This makes Tat an RNA binding protein (Dingwall *et al.*, 1990), and it is this characteristic which has been exploited to effectively deliver covalently linked macromolecular cargoes such as full-length proteins (Schwarze *et al.*, 1999) and non-covalently linked nucleic acids with efficiency (Eguchi *et al.*, 2009).

The exact mechanism of how Tat enters cells was investigated by various groups at the time. Mann and Frankel (1991), reported a non-specific, endocytic-dependent mechanism of Rho-Tat internalisation in HeLa cells, by performing binding-kinetic time-course experiments at 37°C and 4°C. The latter abolished binding to the cell membrane, indicating

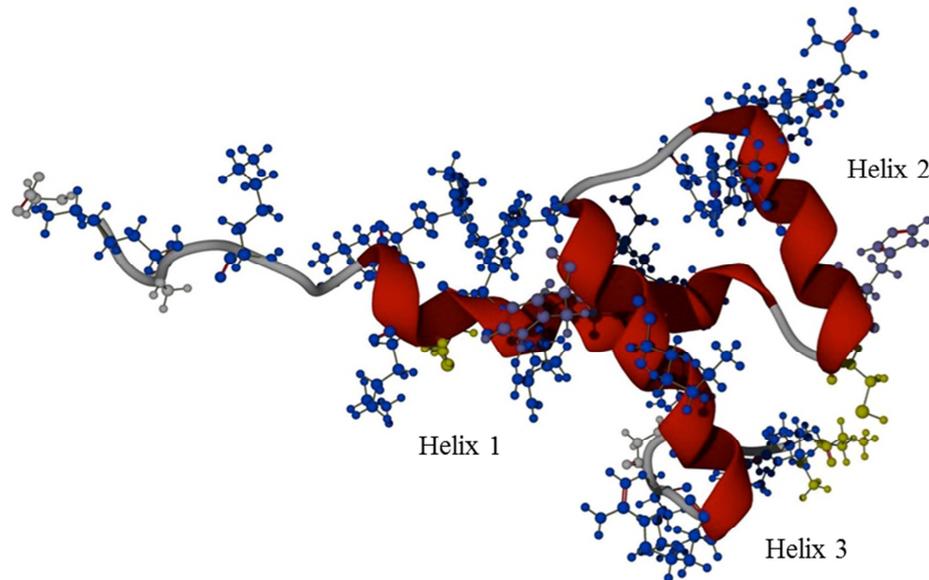
that an energy-dependent process may be involved in uptake. Interestingly, they reported that another, temperature-independent mechanism was involved in binding in a cutaneous T lymphocyte (H9) cell line, as treatment with trypsin, a serine protease which digests cell receptors, did not affect uptake. Dextran sulfate and heparin, polyanionic moieties on the cell surface, were identified as interacting molecules prior to internalisation (Mann and Frankel, 1991). Studies by Wadia *et al.* (2004) and Kaplan *et al.* (2005), described a macropinocytic method of Tat uptake in a number of cell lines, confirmed later by studies with fluorescein-labelled peptide (Fonseca *et al.*, 2009). Macropinocytosis is a non-endocytic type of bulk-transport, triggered when Tat interacts with HSPGs (Heparan sulfate proteoglycans) on the cell surface (Console *et al.*, 2003). Internalization occurs through large vesicles that pinch off the phospholipid membrane and form heterogeneous vesicles, which accumulate within the cell (Wadia *et al.*, 2004). The uptake, however of conjugated cargo, points towards an endocytic mode of uptake, depending on the molecular weight and transduced cell type, and typically resulting in endosomal sequestration (Fonseca *et al.*, 2009). More recently, the Dowdy group have developed a non-covalent, Tat-based siRNA delivery vector, which utilizes three Tat domains separated by a fusogenic hemagglutinin (HA) tag recombinantly expressed with a dsRNA binding motif (DRBM) from human PKR (Eguchi *et al.*, 2009). This recombinant protein, PTD-DRBD, was able to carry siRNA into a range of cell lines, including primary HUVEC cells, inducing potent and selective gene knockdown.

## 1.7 The Antennapedia homeodomain and Penetratin

Another classical CPP is the homeotic protein Antennapedia from *Drosophila melanogaster*. The Antennapedia (Antp) gene, a member of the Antennapedia Complex (ANTP-C) is a homeotic gene that encodes for DNA-binding Transcription Factors, which regulate segmental identity in the thorax (Lewis, 1978, Lewis et al., 1980). Dominant mutations in the Antp gene result in the transformation of antennae into legs (Wakimoto and Kaufman, 1981). Highest levels of gene transcription occur in the mesothorax, but not in the head during various stages of embryogenesis. Two promoters, P1 and P2, control the major protein coding region, composed of exons A, B, D & E (1512 nucleotides long) and C, D & E (1727 nucleotides long), respectively. Multiple AUG (start/Met) codons are present in exons A, B and C.

The Antennapedia homeoprotein is a proline and glutamine-rich, 378 amino acid (aa) protein (43 kDa). Its evolutionarily conserved homeodomain (AntpHD), 60 aa long, is found near the C-terminus and has been shown to bind DNA *in vitro* and *in vivo* (Desplan *et al.*, 1985). In 1988, Otting *et al* resolved the nuclear magnetic resonance (NMR) structure of the DNA-binding homeodomain as a helix-turn-helix motif (Otting *et al.*, 1988), akin to the structure of other prokaryotic regulatory proteins (Ohlendorf *et al.*, 1982). The seminal work by Alain Prochiantz's group identified that the Antennapedia homeodomain (AntpHD) adopts a tertiary structure comprised of 3  $\alpha$ -helices with a  $\beta$ -turn between the last two, which allows it to bind to cognate dsDNA sequence in target promoters (Joliot *et al.*, 1991a) (Fig. 1.3). Moreover, the group observed in 1991 that the evolutionarily conserved, 60 aa AntpHD could effectively translocate into nuclei and between neuronal cells thereby inducing differentiation and morphological changes. (Joliot *et al.*, 1991b). Through site-directed mutagenesis studies, the third helix of the protein (16

aa) was identified as the minimal sequence required for translocation, and was duly named ‘Penetratin’ (Pen) (Derossi *et al.*, 1994).



MGRKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTER**RQIKIWFQNRRM****KWKKEN**

Helix 1

Helix 2

Helix 3

**Figure 1.3 Secondary structure of the Antennapedia homeodomain protein (AntpHD).** The AntpHD exhibits a helix-turn-helix motif, common in cell penetrating proteins. Key amino acids responsible for its highly polar charge are shown in blue. Histidine residues, green; Methionine residues, grey. The sequence for the third helix, Penetratin, responsible for translocation, is shown in bold. Helical structures are underlined (Modelled with RSCB PDB Protein).

The 16 amino acid sequence –RQIKIWFQNRRMKKWKK- and its ability to translocate through biological membranes has been widely studied as a vector for the delivery of various cargoes, from chemical drugs, to proteins, and nucleic acids (Dietz and Bahr, 2004). A comparative study between Penetratin, Tat, transportan and polyarginine has demonstrated that unconjugated, rhodamine-labelled Penetratin exhibits negligible, cell type-independent cytotoxicity at concentrations of up to 100  $\mu$ M and a higher magnitude of uptake compared to Tat (Sugita *et al.*, 2008). Moreover, Penetratin failed to co-localize

with transferrin, a clathrin marker, but co-localized with cholera toxin, a lipid raft marker, indicating a lipid raft-dependent, but clathrin-independent mode of translocation (Jones *et al.*, 2005). Conversely, fluorescein-labelled Penetratin was found to be highly cytotoxic in HeLa and Jurkat cells, with a dose-dependent, cell type-independent mode of internalization, which was higher compared to Tat (Sugita *et al.*, 2008).

Thus far, only a handful of studies have investigated Penetratin as a siRNA carrier. The AntpHD has as yet, not been investigated for its potential siRNA transducing abilities, possibly due to its large molecular weight and dsDNA binding abilities. Penetratin was first covalently linked to siRNA, and applied to mammalian neurons, where it induced a gene knockdown effect (Davidson *et al.*, 2004). In addition, covalently linked siRNA against GFP induced a reduction in the expression of both stable and transient-expressing cells (Muratovska and Eccles, 2004). In *in vivo* lung studies, Pen-siRNA targeted against p38 MAPK was found to induce the innate immunity response (Moschos *et al.*, 2007). More recently, a folate-penetratin nanocomplex separated by a PEGylated linker has been used to deliver luciferase-specific siRNA, greatly decreasing its expression (Cheng and Saltzman, 2011). Some examples of cargoes successfully transduced by Penetratin are summarized in Table 1.3. Other CPPs have been covalently complexed with various forms of chemically stabilized dsRNAs, such as PNA or morpholinos (Abes *et al.*, 2008, Ivanova *et al.*, 2008), 2'-O-methyl phosphorothioate dsRNA (Hassane *et al.*, 2011), with uncharged oligonucleotide molecules covalently linked to a CPP moiety proving to be more potent and specific than direct conjugation to highly charged siRNA (Juliano *et al.*, 2008). It has been suggested that linking siRNA directly to amphipathic CPPs leads to charge neutralization, due to steric hindrance aggregation and loss of transducing abilities, which has consequently induced a shift to non-covalent conjugation approaches. This strategy for

siRNA conjugation includes delivery into a range of cell lines by the synthetic peptide MPG (Nguyen *et al.*, 2006, Simeoni *et al.*, 2003), successfully targeting Oct-4 (Zeineddine *et al.*, 2006) and cyclin B1 (Crombez *et al.*, 2007) *in vivo*. Moreover, Polyarginine (Kumar *et al.*, 2007), Penetratin (Lundberg *et al.*, 2007) and Tat (Meade and Dowdy, 2007) have also been utilized as siRNA carriers, successfully downregulating the expression of various targeted genes in a host of cell lines (Heitz *et al.*, 2009).

**Table 1.3** Notable examples of cargo successfully transduced by Penetratin. Adapted from Dietz and Bahr (2004).

<b>Cargo</b>	<b>Target/disease</b>	<b>Effect</b>	<b>Reference</b>
<b>doxorubicin</b>	anti-neoplastic drug delivery	6-fold increase in uptake through the BBB through <i>in situ</i> perfusion in rats	Rousselle <i>et al.</i> (2000), Rousselle <i>et al.</i> (2001)
<b>INK4a-derived peptide</b>	inhibition of pRB phosphorylation in p16 <sup>-/-</sup> cells	G1 arrest in p16-negative AsPC1 and BxPC3 pancreatic cell lines, human diploid fibroblasts	Bonfanti <i>et al.</i> (1997), Fahraeus <i>et al.</i> (1998), Fujimoto <i>et al.</i> (2000), Kato <i>et al.</i> (1998)
<b>p21</b>	anti-cancer treatment in p53 mutated cancers	growth inhibition in ovarian SKOV-3 and IGROV-1 cells; acute cytotoxicity in p53-mutated cells but not healthy ones in a SKOV-3 nude mouse model	Bonfanti <i>et al.</i> (1997), Kousparou <i>et al.</i> (2012)
<b>siRNA against p38 MAP kinase</b>	p38 mRNA	Reduction of p38 mRNA levels in a L929 cell line; Induction of innate immunity <i>in vivo</i>	Moschos <i>et al.</i> (2007)
<b>prAMP (proline rich antimicrobial peptides)</b>	inhibition of bacterial DNaK	increased the activity against the Gram-positive <i>Micrococcus luteus</i>	Hansen <i>et al.</i> (2012)
<b>PNA</b>	luciferase splice correction assay	β-globin splice correction and upregulation of the luciferase gene in HeLa-pLuc cells	Lundin <i>et al.</i> (2008)

## 1.8 Overcoming endosomal entrapment

Early studies on the uptake of CPPs on fixed cells *in vitro* had proposed a receptor and energy-independent mechanism of translocation, since uptake was observed at both 37°C and 4°C (Derossi *et al.*, 1996). Further studies in live cells, had attributed those observations to fixation artefacts (Lundberg *et al.*, 2003, Richard *et al.*, 2003), and proposed that CPPs enter cells by two distinct mechanisms, endocytosis (clathrin-mediated, caveolae-mediated and macropinocytosis) and direct translocation. The mechanism of uptake was found to depend on the presence or absence and type of cargo (Lundin *et al.*, 2008), temperature, primary protein sequence (Guterstam *et al.*, 2009, Jiao *et al.*, 2009) and concentration. Although the exact mechanisms of uptake are still elusive, most studies point towards initial bidentate interaction of the guanidinium group in arginine with sulfate atoms of heparin (Sakai and Matile, 2003) a member of the heparan sulfate proteoglycan family (HSPGs), and endocytosis-mediated uptake (Fuchs and Raines, 2004). The mode of uptake that follows varies with concentration and the attachment of cargo, with lower concentrations and macromolecules favoring an endocytic mechanism, whereas at concentrations above a membrane density-specific threshold, a shift to direct translocation is observed for a number of CPPs (Duchardt *et al.*, 2007, Kosuge *et al.*, 2008, Rydstrom *et al.*, 2011, Tunnemann *et al.*, 2008). For example, Penetratin has been found to favor direct translocation only at concentrations below 2 µM (Guterstam *et al.*, 2009). Lundin *et al.* (2008) compared the uptake mechanism of various cationic and amphipathic CPP-peptide nucleic acid (PNA) complexes, concluding that amphipathic CPPs were endocytosed in clathrin-coated vesicles, whereas cationic ones favored macropinocytosis. Once endocytosed, CPP-cargo complexes become trapped in endosomes and are eventually degraded; indeed, CPP concentrations necessary to elicit a biological effect are

usually  $\geq 10 \mu\text{M}$ , which may be too high for therapeutic formulations (Milletti, 2012). Attempts to bypass endosomal sequestration of CPPs coupled to macromolecular cargos have incorporated reagents such as chloroquine in a co-incubation step, as a buffer that stops the decrease in intra-endosomal pH, which may be relevant in an *in vitro* setting, but not *in vivo*, as cytotoxicity is elicited in a cell-type specific, and therefore tissue-specific, manner (Wadia *et al.*, 2004). Incorporation of the HA<sub>2</sub> subunit of the influenza hemagglutinin (HA) tag at the amino terminus of CPPs has been reported to aid fusogenicity (Eguchi *et al.*, 2009, Wadia *et al.*, 2004). HA<sub>2</sub> adopts an alpha-helical structure upon acidification in the endosome and proceeds to fuse with the endosomal membrane, allowing the release of viral nucleocapsid into the cytoplasm (Bullough *et al.*, 1994). CPP-HA<sub>2</sub> fusion proteins, when co-incubated with CPP-cargo in a stoichiometric ratio, have been reported to greatly enhance the delivery of nucleic acid/peptide cargo linked with Penetratin, Tat and transportan (El-Andaloussi *et al.*, 2005, El-Andaloussi *et al.*, 2006, Kaplan *et al.*, 2005, Wadia *et al.*, 2004). Another strategy involves the incorporation of histidine residues, which have been shown to disrupt endosomal membranes upon acidification via protonation of its imidazole ring (Midoux and Monsigny, 1999). A seminal study by Lundberg *et al* (2007), compared the efficiency of endosomal escape between EB1, a Penetratin analogue with Histidine insertions (LIRLWSHLIHIWFQN-RRLKWKKK), its parent peptide, Penetratin, as well as Penetratin fused to an HA<sub>2</sub> tag non-covalently linked to siRNA (Table 1.2). EB1 yielded a substantial decrease in luciferase activity compared to the latter two, demonstrating that unmodified penetratin-siRNA were effectively trapped within endosomes and could not mediate RNAi.

## 1.9 Double stranded RNA binding domains

Double stranded RNA (dsRNA) is ubiquitously involved in a myriad of pathogenic and non-pathogenic cellular processes. Viral genetic information is often in the form of dsRNA or ssRNA which, upon replication, form these double helical complexes. Moreover, there are cellular dsRNAs in the form of mRNA untranslated regions (UTRs), ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), as well as smaller precursors, such as those of short interfering RNAs (siRNA) and microRNAs (miRNA) (Masliah *et al.*, 2013). The diverse group of RNAs is recognized by a large superfamily of dsRNA binding proteins (DRBPs), which contain 1-5 dsRNA binding domains/motifs (DRBDs/DRBMs), zinc fingers, and sterile  $\alpha$ -motif (SAM) domains, amongst others (Masliah *et al.*, 2013). The double stranded RNA (dsRNA) binding domains represent conserved viral, eukaryotic and prokaryotic motifs within a large family of proteins that interact specifically with dsRNA. They are usually 65-70 amino acids in length (St Johnston *et al.*, 1992) and provide a vital means of regulating gene expression. Eukaryotic dsRNA binding proteins (DRBPs) include adenosine deaminase acting on RNA 1 (ADAR1) (Wang *et al.*, 2000), spermatid perinuclear RNA binding protein (SPNR) (Pires-daSilva *et al.*, 2001), DICER, nuclear factors associated with dsRNA 1 and 2 (NFAR 1 and 2) (Saunders *et al.*, 2001) and reside either in the nucleus or in the cytoplasm. The biological role of nuclear DRBPs is to regulate post-transcriptional control processes, such as RNA interference (RNAi), splicing, stability, transport as well as mRNA elongation and translation. They do so by binding the 5'- and/or 3'- UTR of mRNA molecules (Lee and Schedl, 2006). Cytoplasmic DRBPs include PKR, TAR RNA binding protein (TRBP), Protein activator of PKR (PACT) and the Staufen protein from *Drosophila melanogaster* (Table 1.4).

DRBDs are able to interact with dsRNA in a sequence independent manner, binding RNA sequences as short as 11 bp (Manche *et al.*, 1992). DRBD-containing proteins thereby interact primarily with A-form helical dsRNAs, since their minor groove is shallow and broad, allowing the formation of hydrogen bonds between the bridging oxygen atom in the ribose residue and the protein, but not ssRNAs, ssDNA or dsDNA (Fierro-Monti and Mathews, 2000, St Johnston *et al.*, 1992). Non-specific interactions between dsRNA and DRBDs occurs by binding the 2'-OH groups and the phosphodiester backbone of the double stranded helix, rather than specific nucleotides, pointing towards a shape-dependent, rather than sequence-dependent mechanism; crystallographic studies have shown that 11-16 base pair sequences are enough for protein-dsRNA binding (Ryter and Schultz, 1998). Although the DRBP family contains more than 500 members, the reason why some DRBPs have more than one DRBD has only been elucidated in the last 20 years; it has been suggested that different DRBDs within the same proteins bind dsRNA with different avidities, thereby stabilizing dsRNA-protein complex formation and mediating interactions with other DRBPs (Krovat and Jantsch, 1996). Indeed, several such protein-protein interactions have been identified, leading to heterodimerisation, autophosphorylation and activation, such as the interactions between human Protein Kinase R (PKR) and TRBP (Peters *et al.*, 2001), NFAR (Saunders *et al.*, 2001), SPNR (Coolidge and Patton, 2000) and PACT (Patel *et al.*, 2000).

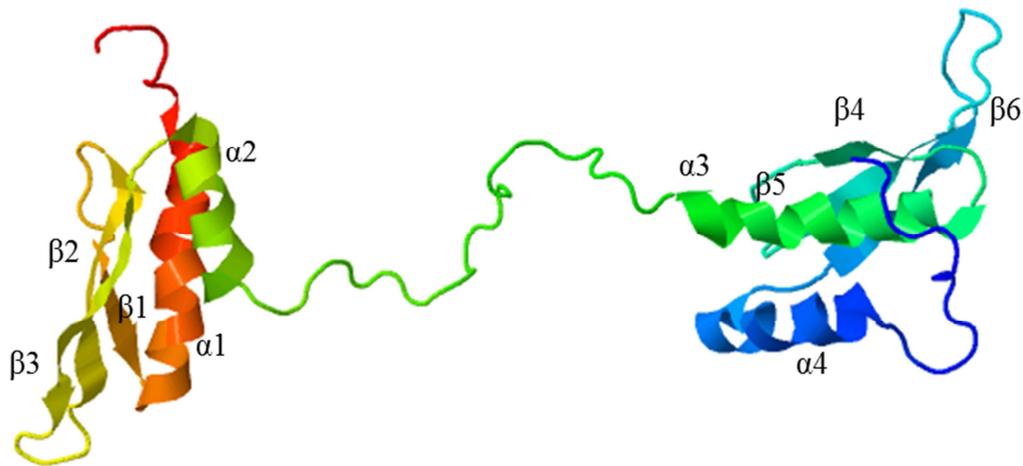
<b>Protein</b>	<b>Number of DRBMs</b>	<b>Localization</b>	<b>Organism</b>	<b>Molecular weight (kDa)</b>	<b>Cellular Function</b>	<b>Reference</b>
<b>PKR</b>	2	Cytoplasmic, 20% Nuclear	Human, rat	68	Antiviral defense, inhibition of protein synthesis	Galabru and Hovanessian (1987), Meurs <i>et al.</i> (1990) Levin and London (1978)
<b>NFAR 1, 2</b>	2	Nuclear	Human	90, 110	mRNA processing, PKR interaction	Saunders <i>et al.</i> (2001)
<b>TRBP</b>	3	Nuclear, Cytoplasmic	Human, mouse	40-50	Interaction with HIV transactivating region (TAR) RNA, PKR modulation	Gatignol <i>et al.</i> (1991), Park <i>et al.</i> (1994)
<b>ADAR 1,2,3</b>	3	Nuclear	Human, rat, mouse, <i>Xenopus</i>	150, 130, 130, 120	viral and brain mRNA editing (A→I) in the liver	O'Connell <i>et al.</i> (1995), Patterson <i>et al.</i> (1995), Chen <i>et al.</i> (2000)
<b>Staufen</b>	5	Nuclear, Cytoplasmic	<i>Drosophila</i> , human, rat, mouse	60, 65	Localization and mRNA translation	Ramos <i>et al.</i> (2000), Ferrandon <i>et al.</i> (1994)
<b>PACT</b>	3	Cytoplasmic	Human, mouse	35	Stress-induced PKR activation by autophosphorylation	Peters <i>et al.</i> (2001), Patel <i>et al.</i> (2000)
<b>RNaseIII / Drosha</b>	1, 2	Nuclear	<i>E. coli</i> , Human, <i>Drosophila</i>	160, 153	pre-mRNA processing, endoribonuclease activity	Nicholson (1996), Filippov <i>et al.</i> (2000)
<b>DICER</b>	1	Cytoplasmic	Human, mouse, <i>Drosophila</i>	220, 215, 210	Endoribonuclease and helicase activity in RNAi	Bernstein <i>et al.</i> (2001)

**Table 1.4** Examples and cellular functions of the members of the dsRNA binding protein family (DRBPs). DRBMs, dsRNA binding motifs. Adapted from Saunders and Barber (2003).

## 1.10 Human Protein Kinase R

The human isoform of Protein Kinase R (PKR) is encoded by a single gene on chromosome 2p21 and is a prominent member of the DRBPs. It has recently been evaluated in various reports for its dsRNA binding abilities, as part of engineered recombinant proteins for the delivery of therapeutically relevant siRNA (Eguchi *et al.*, 2009, Geoghegan *et al.*, 2012, Kim *et al.*, 2009a). It is a 551 aa long interferon-induced, dsRNA activated serine/threonine (Ser/Thr) kinase that plays a key role in the immune response against viral infections (Rice *et al.*, 1985, Samuel *et al.*, 1984), as well as in the regulation of signal transduction, apoptosis, cell proliferation and differentiation (Donze *et al.*, 1995, Wu and Kaufman, 1996). It therefore plays a critical function in allowing the cell to rapidly respond to external stimuli by shutting down its translational machinery (Merrick, 1992, Pain, 1996). In the case of viral infection, PKR recognizes and directly bind viral dsRNA via its two *N*-terminal tandem dsRNA binding motifs (DRBMs) in a sequential manner and induce the interferon response. This in turn allows the trans-phosphorylation of various Ser/Thr and Tyrosine residues (Ser83, Ser242, Thr88-90, Thr255, Thr258, Thr446 and Thr451) in its *C*-terminal kinase domain and subsequent homodimerization of the protein (Taylor *et al.*, 2001, Zhang *et al.*, 2001). Once the dsRNA substrate has been correctly inserted between the two DRBDs, activated PKR proceeds to phosphorylate the  $\alpha$ -subunit of eukaryotic translation initiation factor 2 (eIF2 $\alpha$ ), a rate limiting component of translation, which is docked within the catalytic domain. eIF2 $\alpha$  phosphorylation induces a dramatic arrest of protein translation and the induction of autophagy and ultimately the self-destruction of the infected host cell (Dabo and Meurs, 2012). PKR is able to bind most forms of dsRNA, including aptamers with bulges, internal and hairpin loops, and multistem junctions (Bevilacqua *et al.*, 1998), as

well as various viral dsRNAs, with imperfect Watson-Crick complementarity, as its two DRBMs adopt a dumb-bell shape with a canonical  $\alpha$ - $\beta$ - $\beta$ - $\beta$ - $\alpha$  topology (Dzananovic *et al.*, 2013) (Fig. 1.4).



**Figure 1.4** Structure of the two dsRNA binding motifs in human PKR that compose its dsRNA binding domain (DRBD). DRBM1 is shown on the left and DRBM2 on the right, linked by 22 amino acid loop. Adapted from Nanduri *et al.* (1998).

Mutational analysis studies by Nanduri *et al.* (1998) identified conserved binding sites within each DRBM, which were also confirmed by structural studies on the Xlrpba DRBM from *Xenopus laevis* (Ryter and Schultz, 1998). These critical residues correspond to three regions on each DRBM; namely regions  $\alpha$ 1 (N15-T16),  $\beta$ 1- $\beta$ 2 loop (P36-R39) and  $\beta$ 3- $\alpha$ 2 loop (R58-K64) in DRBM1, and regions  $\alpha$ 3 (N106-R107),  $\beta$ 4- $\beta$ 5 (H125-G130) and  $\beta$ 6- $\alpha$ 4 (S148-K154) in DRBM2 (Nanduri *et al.*, 2000).

Although PKR can non-specifically interact with as few as 15 base pairs (Bevilacqua and Cech, 1996, Schmedt *et al.*, 1995, Ucci *et al.*, 2007), a minimum of 30 bp is required for activation, homodimerization and the subsequent induction of immune signals, in response to exogenous dsRNA introduction into the cell (Lemaire *et al.*, 2008).

DRBM1<sup>9-77</sup> was found to be less conformationally restricted than DRBM2<sup>100-167</sup>, although similar in length (Nanduri *et al.*, 2000). This allows dsRNA binding with high affinity (dissociation constant,  $K_d = 3.8 \times 10^{-7}$ ) (Schmedt *et al.*, 1995). The binding affinity for dsRNA is increased 100-fold by the cooperative action of both tandem motifs ( $K_d = 4 \times 10^{-9}$ ) (Geoghegan *et al.*, 2012, Schmedt *et al.*, 1995).

A seminal report by Eguchi *et al.* (2009), expressed the DRBM1 from human PKR as a fusion protein with the HIV Tat protein (PTD; Tat-Tat-HA-Tat), creating an engineered PTD-DRBD siRNA delivery vector. They claimed that the DRBM1 (DRBD) from human PKR was sufficient in binding 21-23nt siRNA sequences at a 4:1 molar ratio, at 90° to the phosphate backbone and with a high enough avidity for intracellular transduction and post-translational gene silencing in a reporter cell line.

In another study Kim *et al.* (2009a) engineered a recombinant protein comprised of the entire DRBD<sup>1-177</sup> (DRBM1 and 2) with the fusogenic peptide KALA, and was able to downregulate the expression of GFP by delivering GFP-specific siRNA without endonuclease or serum degradation. A third study, by Geoghegan *et al.* (2012), evaluated the DRBD as a protein platform for siRNA delivery, and observed that the use of DRBM1 only was insufficient for stable binding of siRNA when fused to various different cell penetrating moieties (Penetratin, B2, Tat). The use of the entire DRBD, consisting of both DRBM1 and 2, however, yielded specific and stable dsRNA binding at low molar ratios. Intracellular delivery however was hindered due to the endosomal entrapment of the fusion protein-DRBD/siRNA complexes, which ultimately attenuated siRNA-induced gene knockdown.

## **1.11 PTP1B in Type 2 Diabetes: a potential target for siRNA-based intervention**

Type 2 Diabetes (T2DM) is a metabolic disorder that has reached epidemic proportions due to an increased prevalence of obesity and sedentary lifestyles. It accounts for 90% of all diabetes cases that affect 3 million people in the UK alone (DiabetesUK, 2013). Its core pathophysiological features include insulin resistance and insufficient  $\beta$ -cell insulin production, hyperglycemia and hyperinsulinemia, adipocytic lipolysis and increased glucose resorption in the kidneys (DeFronzo, 2009). Long-term complications include heart attacks, renal failure, stroke and retinopathy. Current monotherapies have various limitations, highlighting a need for the development of new pharmacologic agents (Carpino and Goodwin, 2010). Current therapies include various anti-hyperglycemic agents such as metformin, which lowers liver glucose output, but which elicits adverse side-effects such as lactic acidosis and renal damage, sulfonylureas, which increase insulin secretion, but whose adverse effects include hypoglycemia, glinides,  $\alpha$ -glucosidase inhibitors, which reduce the rate of polysaccharide breakdown, thiazolidinediones, that increase sensitivity to existing systemic glucose levels and insulin, which is still the 'golden standard' when it comes to lowering glucose levels (Mazzola, 2012). Other monotherapies include insulin secretagogues such as meglitinide, pioglitazone, or a dipeptidyl peptidase 4 (DPP-4) inhibitor (Irons and Minze, 2014). All monotherapies aim to by-pass diminished signaling by the Insulin Receptor (IR), characterized by aberrations in molecular pathways downstream of the IR which attenuate the signals mediated by the activation of the receptor upon ligand binding. These are caused by the de-phosphorylating action of a class of Protein Tyrosine Phosphatases (PTPs). This family of receptor-like tyrosine kinases plays an important role in the regulation of various signal transduction pathways.

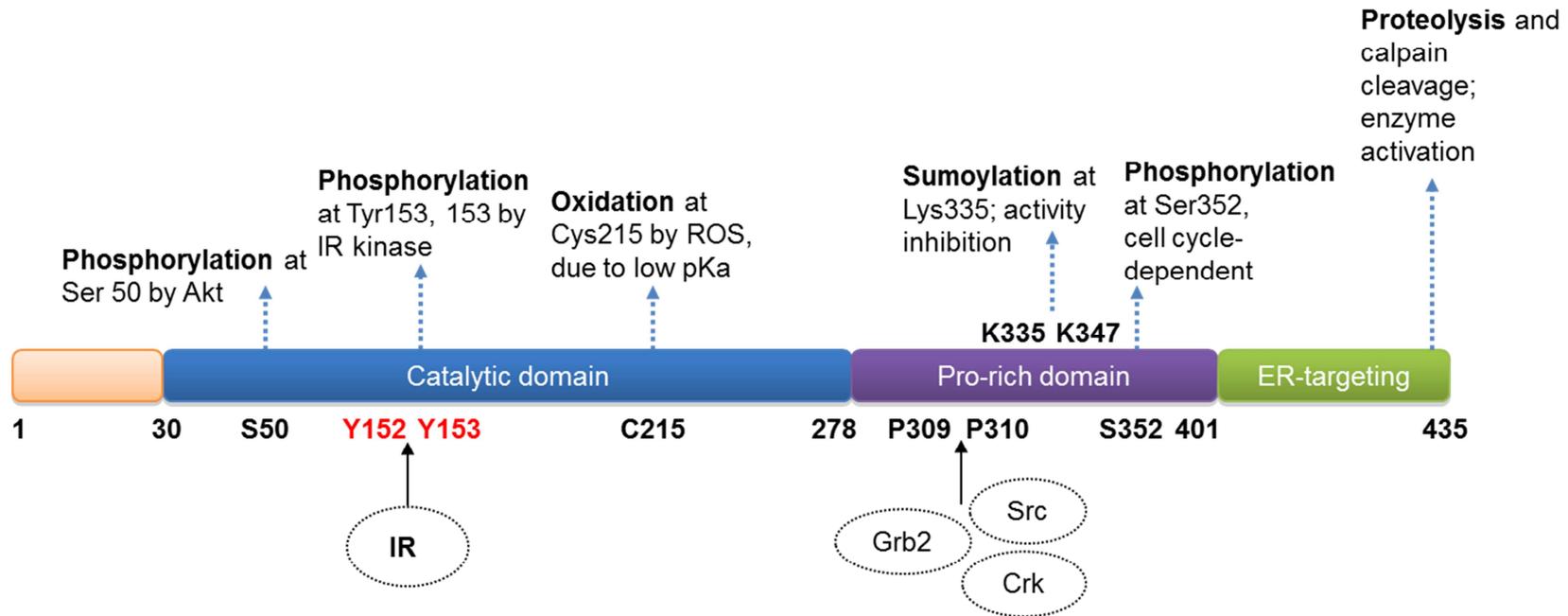
Signaling through the insulin receptor (IR) is activated when endogenous insulin, secreted by pancreatic  $\beta$  cells, binds to the two  $\alpha$ -chains of the extracellular domain of the receptor, triggering the activation of the intrinsic tyrosine kinase activity of its intracellular  $\beta$ -subunit tyrosine residues (Taniguchi *et al.*, 2006). Auto-phosphorylation of key tyrosine residues (Y1146, Y1150, Y1151), results in the recruitment and phosphorylation of insulin receptor substrates 1 and 2 (IRS1 and 2), allowing the regulatory subunit of phosphatidylinositol 3 kinase (PI3K) to dock. As PI3K is activated, the membrane phospholipid phosphatidylinositol 3,4,5-triphosphate (PIP3) is recruited and in turn, activates the second messengers 3-phosphoinositide-dependent protein kinase 1 and 2 (PDK1 and 2), ultimately activating protein kinase B (PKB)/Akt, a key mediator of the metabolic effects of insulin (Guo, 2014). Akt activation leads to glucose uptake by GLUT4, protein and glucose synthesis and gluconeogenesis (Taniguchi *et al.*, 2006). Growth and differentiation via the ras-MAPK pathway is also regulated by insulin, which elicits the GTP-dependent activation of p-21<sup>ras</sup>, stimulated by Grb-SOS (White, 2003). Ras-GTPase-activating protein (rasGAP) keeps ras in an inactive state by interacting with phosphorylated p62-Dok. The protein tyrosine phosphatase 1B (PTP1B) plays a key role in the feedback loop that dephosphorylates both the IR and p62 Dok, attenuating both PKB-mediated signaling and ras inactivation, thereby leading to a prolonged MAPK activation. PTP1B also attenuates Leptin signaling by dephosphorylating recruited Janus-2-kinase (JAK2) on the leptin receptor. Active JAK2 phosphorylates Signal Transducer and Activator of Transcription 3 (STAT3), which dimerizes and translocates to the nucleus to regulate gene expression. Abolishing signaling via JAK2 inhibits leptin feedback to the hypothalamus, which would normally elicit satiation and energy usage (Koren and Fantus, 2007). Insulin resistance occurs through perturbations in the phosphorylation status of signaling

molecules or increased degradation and decreased synthesis (Biddinger *et al.*, 2008), leading to a chronic elevated state of blood insulin and glucose levels.

PTP1B is encoded by the protein tyrosine phosphatase non-receptor type 1 gene (PTPN1) on chromosome 20q13 in humans – a region linked with insulin resistance, T2DM and obesity in human populations (Tsou and Bence, 2012). Although a ubiquitously expressed protein (Haj *et al.*, 2003), hepatic PTP1B has been implicated in the negative regulation of insulin signaling (Ahmad *et al.*, 1995). As a member of the large family of protein tyrosine phosphatases, it is involved in dephosphorylating key tyrosine residues on the IR and IRS proteins (Agouni *et al.*, 2011). As a member of the protein tyrosine phosphatase superfamily, it contains the highly conserved (I/V)HCXXGXXR(S/T) active site motif, and is regulated by phosphorylation within key residues in its catalytic site (Barford *et al.*, 1994) (Fig. 1.5).

Early studies in PTP1B knockout (PTP1B<sup>-/-</sup> KO) mice showed insulin hypersensitization with increased insulin receptor phosphorylation in liver and muscle tissue, compared to PTP1B<sup>+/+</sup> mice (Elchebly *et al.*, 1999). This was confirmed by later studies with KO mice that showed improved glucose homeostasis in muscle (Delibegovic *et al.*, 2007) and liver (Delibegovic *et al.*, 2009), and protection from diet-induced obesity (Klaman *et al.*, 2000). Moreover, studies in diabetic (*db/db*) and insulin-resistant (*ob/ob*) mice with PTP1B-targeting antisense oligonucleotides (ASOs) have shown a downregulation in the expression of genes normally associated with lipogenesis in adipose tissue and liver, improving glucose homeostasis (Waring *et al.*, 2003) and increasing insulin-stimulated PKB activation as well as reducing hyperinsulinemia (Zinker *et al.*, 2002). PTP1B-specific antisense oligonucleotide therapy in *ob/ob* hyperglycemic mice decreased mRNA levels by up to 50% in liver and adipose tissues (Rondinone *et al.*, 2002, Waring *et al.*, 2003, Zinker

*et al.*, 2002). In addition, Isis Pharmaceuticals are in Phase II studies with ISIS-PTP1B<sub>Rx</sub>, an antisense formulation that selectively targets PTP1B, for T2DM patients that are currently on metformin and show poor control of glucose levels (clinicaltrials.gov identifier: NCT01918865). No data has as yet been published by Isis, therefore the outcome of this work is still unclear. The development of potent, orally bioavailable inhibitors for PTP1B has so far been hindered by the 94% active site identity it shares with its closely related phosphatase T-cell PTP (TCPTP), encoded by the closely related tyrosine-protein phosphatase non-receptor type 2 gene (PTPN2) (Tiganis, 2013). Whereas PTP1B knockout improves insulin sensitivity, TCPTP knockout mice die within eight weeks after birth from anemia and systemic inflammation (Barr, 2010). The highly charged, conserved nature of its active site also poses a barrier to the development of small molecule inhibitors (Barr *et al.*, 2009). Taken together, evidence from antisense studies point towards a PTPN1 gene knockdown approach, since the high specificity of ASOs or siRNAs can be used to selectively target PTP1B, but not TCPTP in T2DM patients, as a novel therapeutic approach.



**Figure 1.5 Structural domains and PTP1B regulation.** Human PTP1B comprises of an *N*-terminal catalytic domain separated from the *C*-terminal endoplasmic reticulum (ER)-targeting domain by a proline (Pro)- rich domain. PTP1B is regulated via phosphorylation of tyrosine residues (Y152,Y153) by the insulin receptor (IR) kinase. Proline residues (P309-P310) interact with Src, Grb2, Crk and p130Cas. Sumoylation at lysine 335 (K335) inhibits its activity, whilst phosphorylation at serine 352 (S352) is cell-cycle dependent. Adapted from Yip *et al.* (2010).

## 1.12 Aims and Objectives

The cell penetrating properties of the CPP Penetratin in a non-cytotoxic or cell-specific manner has been well documented. The abilities of its parent protein, the AntpHD, have been less well studied, although its ability to carry cargo in the form of p21 DNA into p53-mutated tumors *in vivo* has been reported (Kousparou *et al.*, 2012). Utilizing the non-cytotoxic, cell penetrating AntpHD penetratin, as well as the endosomolytic analogue EB1 previously developed by Lundberg *et al.* (2007), fused to the dsRNA binding domain(s) from human PKR, as in the studies by Eguchi *et al.* (2009) and Geoghegan *et al.* (2012) this project aims to clone, purify and characterize various recombinant proteins and investigate whether they can provide an efficient, non-cytotoxic platform for the delivery of therapeutically relevant siRNA *in vitro*.

The first part of this project specifically aimed to:

1. Clone, express and purify several AntpHD-DRBM1 and Pen-DRBM1 fusion proteins. These would be produced from a synthetic template comprised of the cDNA sequence for the Antennapedia homeodomain and the dsRNA binding motif 1 from human PKR separated by a flexible glycine/serine linker.. Purified proteins were characterized by proteomic and mass-spectrometry analyses.
2. Determine whether purified proteins retained their transducing properties, and their potential intracellular localization in a HEK293 and HepG2 cell lines by fluorescence studies in both live and fixed cells.
3. Determine whether purified proteins retained their dsRNA binding potentials and subsequently assess their complex formation efficiencies.

4. Assess whether proteins can transduce the cell membrane when non-covalently complexed with cargo siRNA and induce GFP-specific gene knockdown in a HEK293-dEGFP reporter cell line.

The primary sequence of Penetratin and the AntpHD indicated that the fusion proteins may have lacked endosomolytic abilities and were likely to be sequestered into endosomes should uptake involve endocytosis. Moreover, the affinity of the entire DRBD (DRBMx2) compared to the DRBM1 only was expected to yield stronger siRNA binding.

The second part of this project therefore aimed to:

1. Clone and purify several EB1-DRBMx2 and DRBMx2-EB1 constructs
2. Purify the constructs Tat-DRBMx2, Penetratin-DRBMx2 and DRBMx2 used in a report by Geoghegan *et al.* (2012) and compare siRNA transduction between HEK293 and HepG2 cells.
3. Assess the efficiency of siRNA-induced knockdown in a HEK293-dEGFP reporter cell and compare any differences in knockdown efficiency to the work done by Geoghegan *et al.* (2012) in an HPRT-expressing reporter cell line.
4. Attempt to assess PTP1B knockdown at the protein level in HepG2 cells following treatment with recombinant proteins non-covalently conjugated to PTPN1-specific siRNA.

Chapter 2. Characterisation of  
recombinant proteins comprised of the  
AntpHD or Penetratin and DRBM1  
from human PKR

## 2.1 Introduction

In order to generate a recombinant protein that can deliver therapeutically relevant siRNA *in vitro*, several constructs were designed to incorporate two different functional domains separated by a flexible linker. The cell-penetrating moiety was comprised of the entire Antennapedia homeodomain (AntpHD) from *Drosophila melanogaster* or its third helix, Penetratin (Pen), whilst its dsRNA binding moiety comprised of the dsRNA Binding Motif 1 (DRBM1) from the human Protein Kinase R (PKR) (Appendix I). The Antennapedia homeodomain has been found to effectively translocate through the plasma membrane of eukaryotic cells and is able to mediate transduction of cargo molecules ranging from small proteins to antisense oligonucleotides (Derossi *et al.*, 1996, Duchardt *et al.*, 2007). Penetratin, the third helix in the homeodomain protein is the minimal sequence required for translocation (Joliot *et al.*, 1991a).

PKR is a human protein involved in the immune response by recognizing foreign dsRNA that has been introduced by a virus into the cell via its double-stranded RNA binding domain (DRBD) (Rice *et al.*, 1985, Samuel *et al.*, 1984). This domain is made up of two distinct motifs (Ryter and Schultz, 1998), DRBM1 and 2. Previous research by Heinicke *et al.* (2009) has shown that the first motif (DRBM1) binds dsRNA with higher avidity than the second motif (DRBM2). A study by Eguchi *et al.* (2009) has utilised a recombinant protein (PTD-DRBD), comprised of the first DRBM fused to an HIV Tat peptide, to effectively mask the large negative charge of siRNA and prevent PTD inactivation, thereby enhancing the cellular uptake of attached siRNA. The potential of DRBM1 fused to a different CPP sequence (in this case the Antennapedia homeodomain and Penetratin) for siRNA delivery, is investigated here.

For this project's purposes, the cDNA sequence corresponding to the AntpHD Isoform 1 (canonical sequence) or Penetratin, was identified in GenBank (accession number M14496.1), and corresponded to the amino acid sequence of the homeobox (UniProt ID: P02833-1 [amino acids 297-356]). Sequence specific primers were designed for the specific amplification of its third helix, Penetratin, during subcloning by PCR. The cDNA sequence for the DRBM1 from human PKR was obtained from Genbank (accession number M35663.1), corresponding to amino acids 9-77 (Uniprot ID: P19525-1). These two domains were separated by a small, flexible linker (-GGGGSGGGGS-; [G<sub>4</sub>S]<sub>2</sub>) which was comprised of solubility-promoting, small amino acids (Arai *et al.*, 2001, Trinh *et al.*, 2004) to yield several constructs encoding for bifunctional proteins with the hypothetical properties of efficient siRNA binding and translocation through plasma membranes. A synthetic cDNA template corresponding to the AntpHD-linker-DRBM1 sequence supplied in a pBluescriptII vector was synthesized commercially (Epoch Biosciences, Missouri City, TX, USA).

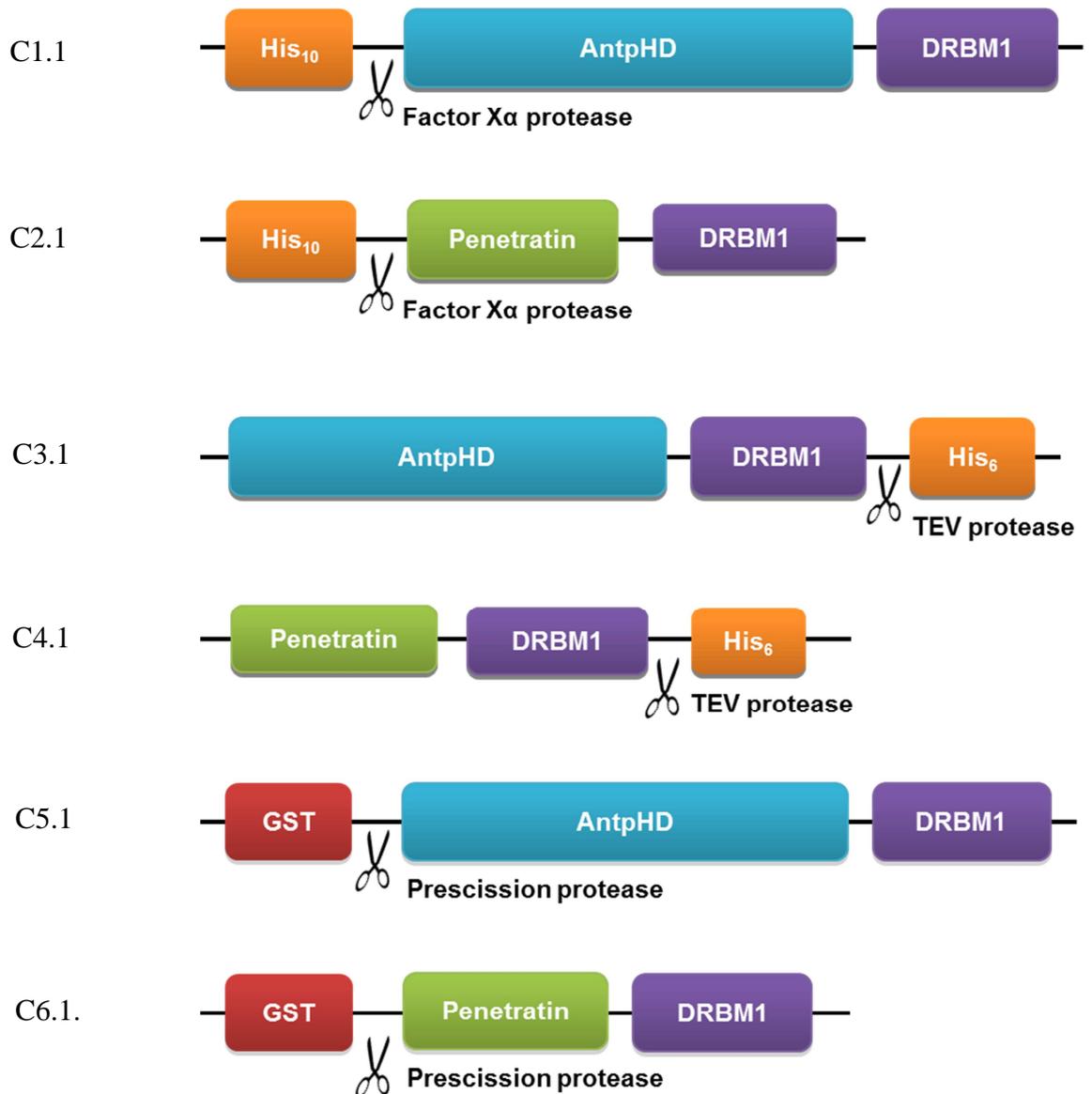
In order to facilitate expression and purification of fusion proteins, oligonucleotide primers were designed such that constructs could be PCR amplified and inserted by restriction enzyme cloning into a modified pET32-a, or pGEX-6P-2 vector (Appendix II). This approach enabled the generation of several constructs with either an *N*-terminal or *C*-terminal affinity tags and different protease cleavage sites (as shown in Fig. 2.1), and was used in order to maximise the likelihood of identity of highly expressed, soluble fusion proteins. This was done in order to provide a 5'-His<sub>10</sub> tag from the vector with a Factor X $\alpha$  protease cleavage tag, or a 3'-end His<sub>6</sub> tag with a TEV protease cleavage tag to each construct. Poly(His) and GST affinity tags were introduced to the *in silico* design of proteins in anticipation of increased solubility, efficiency of refolding, improved yield, and

minimisation of proteolytic degradation upon purification (Waugh, 2005). The terminus at which the affinity tag is attached is critical to the aforementioned parameters, however, the effects of each solubility tag at each of the two protein termini can only be assessed case-by-case (Arnau *et al.*, 2006a, Arnau *et al.*, 2006b). Subcloning the synthetic template into a pGEX-6P-2 vector plasmid provides a 5'-end GST tag with a downstream human Rhinovirus 3C Protease tag (commercially sold as Prescission Protease (PP), GE Healthcare) for its removal following purification (Fig. 2.1). The poly(His) tags allows purification by Immobilised Metal Affinity Chromatography (IMAC) over a Nickel nitroloacetic (Ni-NTA) column, whereas the GST tag allows purification by affinity chromatography over a Glutathione Sepharose 4b (GS4B) column. All sequences were assembled using Vector NTI Advance v11.5.2 software (Invitrogen, Paisley, UK) to construct six novel recombinant proteins:

- C1.1: His<sub>10</sub> – AntpHD-DRBM1 with a Factor X $\alpha$  cleavage site
- C2.1: His<sub>10</sub> – Pen-DRBM1 with a Factor X $\alpha$  cleavage site
- C3.1: AntpHD-DRBM1 –His<sub>6</sub> with a TEV protease cleavage site
- C4.1: Pen-DRBM1 – His<sub>6</sub> with a TEV protease cleavage site
- C5.1: GST – AntpHD-DRBM1 with a Prescission protease cleavage site
- C6.1: GST – Pen-DRBM1 with a Prescission protease cleavage site

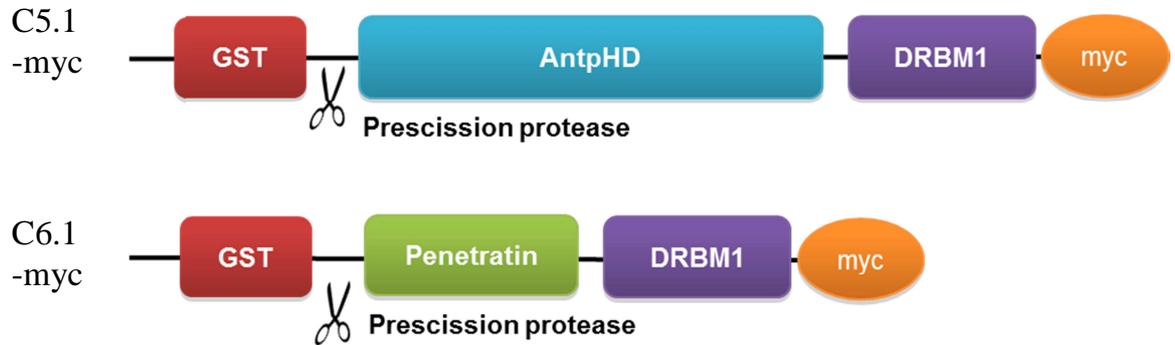
A c-myc epitope tag (as a sequence corresponding to the amino acid sequence (-EQKLISEEDLN)) was also added to the C-terminus of C5.1 and C6.1 by insert-specific primers using C5.1 as a template (Fig. 2.2) to assist in localisation studies and immunoblotting.

**Construct Name:**



**Figure 2.1** Schematic of the fusion protein constructs in a modified pET32-a vector (C1.1-C4.1) or a pGEX-6P-2 vector (C5.1-C6.1). They are comprised of either: the Antennapedia homeodomain (AntpHD); Penetratin; the dsRNA binding motif 1 of human protein kinase R (DRBM1) and either an N-terminal 10xHistidine tag (His<sub>10</sub>), a C-terminal 6xHistidine tag (His<sub>6</sub>) or an N-terminal Glutathione-S-transferase (GST) tag. Protease cleavage sites for removal of affinity tags are indicated (✂). TEV: Tobacco Etch virus protease; Precission Protease: Rhinovirus 3C protease. Construct domains are separated by a flexible (G4S)<sub>2</sub> linker.

**Construct Name:**

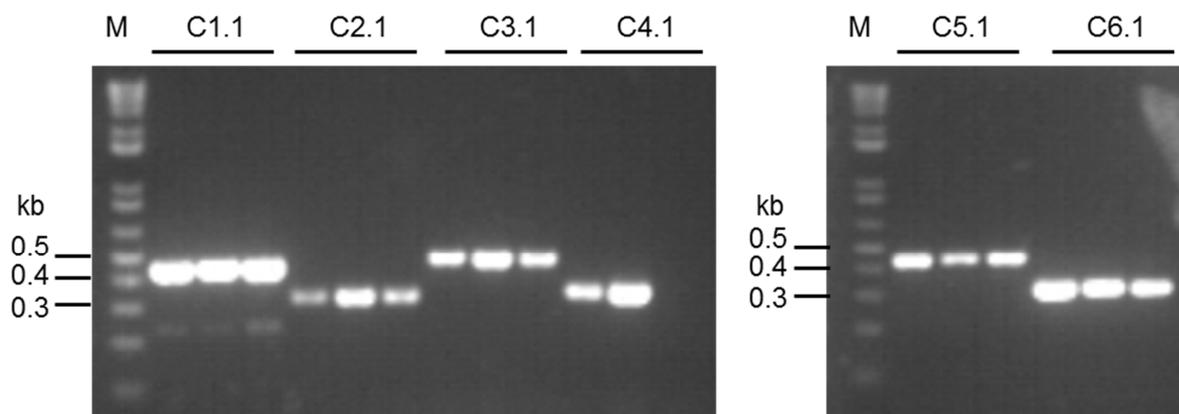


**Figure 2.2** Schematic of the myc-tagged recombinant proteins with an *N*-terminus Glutathione S-transferase (GST) tag in a pGEX-6P-2 vector. They are comprised of either: the Antennapedia homeodomain (AntpHD); Penetratin; the dsRNA binding motif 1 of human protein kinase R (DRBM1); a GST affinity tag and a c-myc epitope tag with a Precision protease cleavage site (indicated as ✂). Construct domains are separated by a flexible (G<sub>4</sub>S)<sub>2</sub> linker.

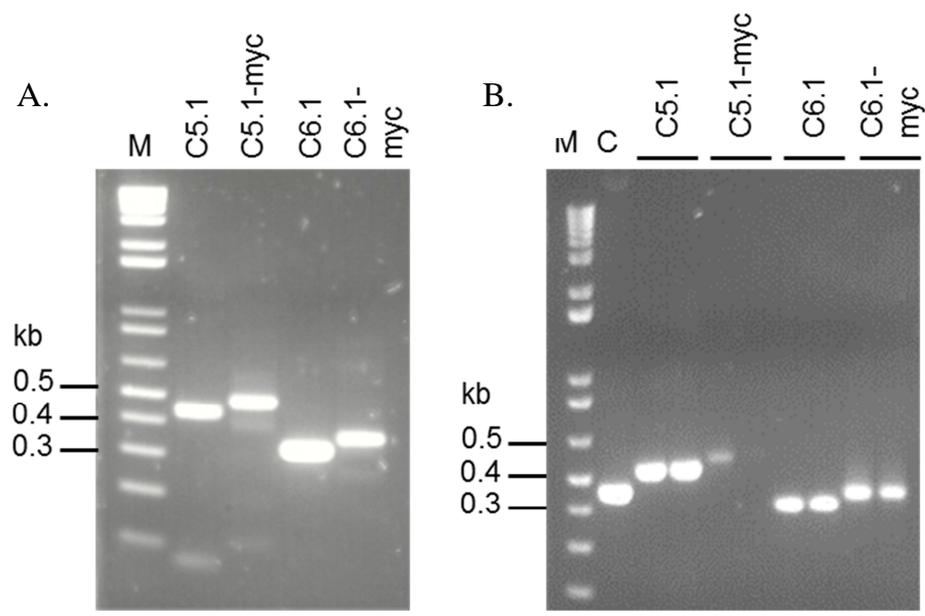
## 2.2 Results

### 2.2.1 Subcloning the recombinant constructs into expression vectors

Constructs were amplified by PCR with a Platinum *pfx* polymerase (Invitrogen, Paisley, UK), ligated into their corresponding plasmid vectors. Purified insert/vector DNA was used to transform the propagation strain XL1-Blue. Three colonies from each plate were screened directly for correct insert sizes by colony PCR. C1.1-C4.1 were analysed using insert-specific forward and reverse primers, while C5.1 and C6.1 were amplified with a vector-specific forward primer and an insert-specific reverse primer (Appendix I) and analysed by agarose gel electrophoresis. Expected sizes for inserts were 495 bp (C.1), 369 bp (C2.1), 480 bp (C3.1), 351 bp (C4.1), 408 bp (C5.1) and 339 bp (C6.1) and corresponded to the bands on the gels (Fig 2.3). Sequences were confirmed by sequencing (GATC Biotech, London, UK) and analysed with Sequence Scanner v1.0 (Applied Biosystems) and ExPASy Translate (Swiss Institute of Bioinformatics, Switzerland). C5.1 was used as a template for the PCR amplification of the c-myc epitope tag at the 3'-end of pGEX-6P-2.C5.1/C6.1 with insert-specific reverse primers. Amplified fragments were used to transform XL1-Blue cells and band sizes were confirmed by colony PCR, agarose gel electrophoresis and sequencing (Fig. 2.4).



**Figure 2.3** Insert screening by colony PCR following ligation of insert in pET32-a (C1.1-C4.1) or pGEX-6P-2 (C5.1-C6.1). Expected size of bands for His<sub>10</sub>-AntpHD-DRBM1 (C1.1), 0.49 kb; His<sub>10</sub>-Pen-DRBM1 (C2.1), 0.37 kb. AntpHD-DRBM1-His<sub>6</sub> (C3.1), 0.48 kb; and Pen-DRBM1-His<sub>6</sub> (C4.1), 0.35 kb. Bands at 0.4 kb and 0.34 kb correspond to AntpHD-DRBM1 (C5.1) and Pen-DRBM1 (C6.1) in pGEX-6P-2, respectively. M: 1 kb Marker. Gels were visualized under a UVP transilluminator.



**Figure 2.4** PCR amplification of GST - tagged C5.1 and C6.1 with a c-myc epitope tag. **A.** PCR amplification of C5.1, C6.1, C5.1-myc and C6.1-myc. **B.** Colony screening by PCR of XL1-blue colonies following transformation with pGEX-6P-2-ligated inserts. C: Negative control template DNA (C5.1). M: 1 kb Marker. Gels were visualized under a UVP transilluminator.

### **2.2.2 Recombinant protein expression on a 50ml scale and purification by affinity chromatography**

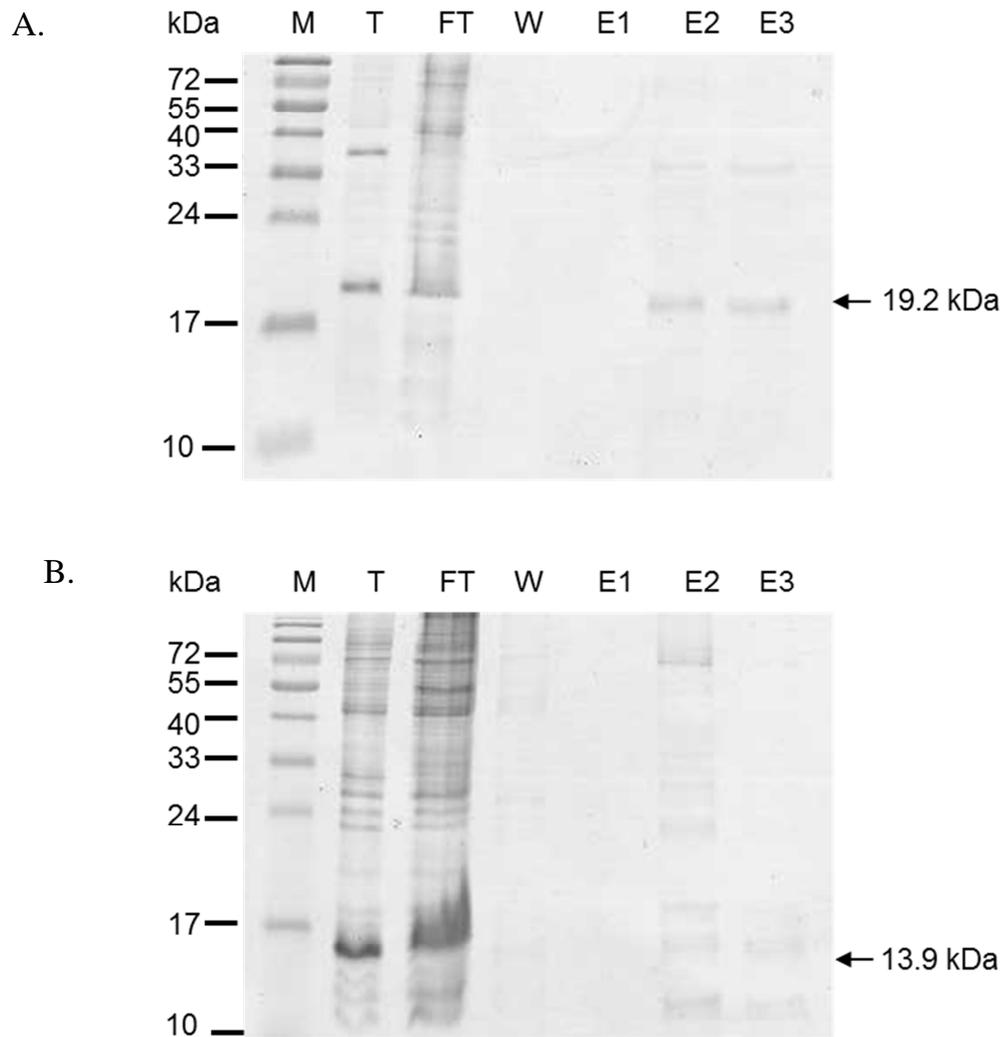
Sequenced constructs generated in section 2.2.1 were used to transform BL21(DE3)pRARE cells in LB broth containing 100 µg/ml ampicillin and 34µg/ml chloramphenicol. Protein production was induced with Isopropyl β-D-1-thiogalactopyranoside (IPTG) and His-tagged proteins were purified by Immobilised Metal Affinity Chromatography (IMAC) on Nickel-Nitriloacetic acid (Ni-NTA) columns charged with 0.1M Ni<sup>2+</sup> under native conditions. GST-tagged proteins were purified over a Glutathione sepharose 4B (GS4B) column. The total, flowthrough, wash and eluted fractions were collected and analysed by polyacrylamide gel electrophoresis (SDS-PAGE).

His<sub>10</sub>-tagged C1.1 (His<sub>10</sub>-AntpHD-DRBM1) was found to be mostly insoluble and expressed at low levels by the expression host, as shown by the weak bands in the total (T) and flowthrough (FT) fractions, especially when compared to C2.1 (His<sub>10</sub>-Pen-DRBM1) (Fig. 2.5). C-terminally His<sub>6</sub>-tagged C3.1 was also insoluble, compared to C4.1, although similarly expressed at low levels (Fig. 2.6). The presence of protein in the eluted fractions of C2.1 and C4.1 (E1-E3) warranted further investigation by upscaling to purifications from 0.5L culture purifications.

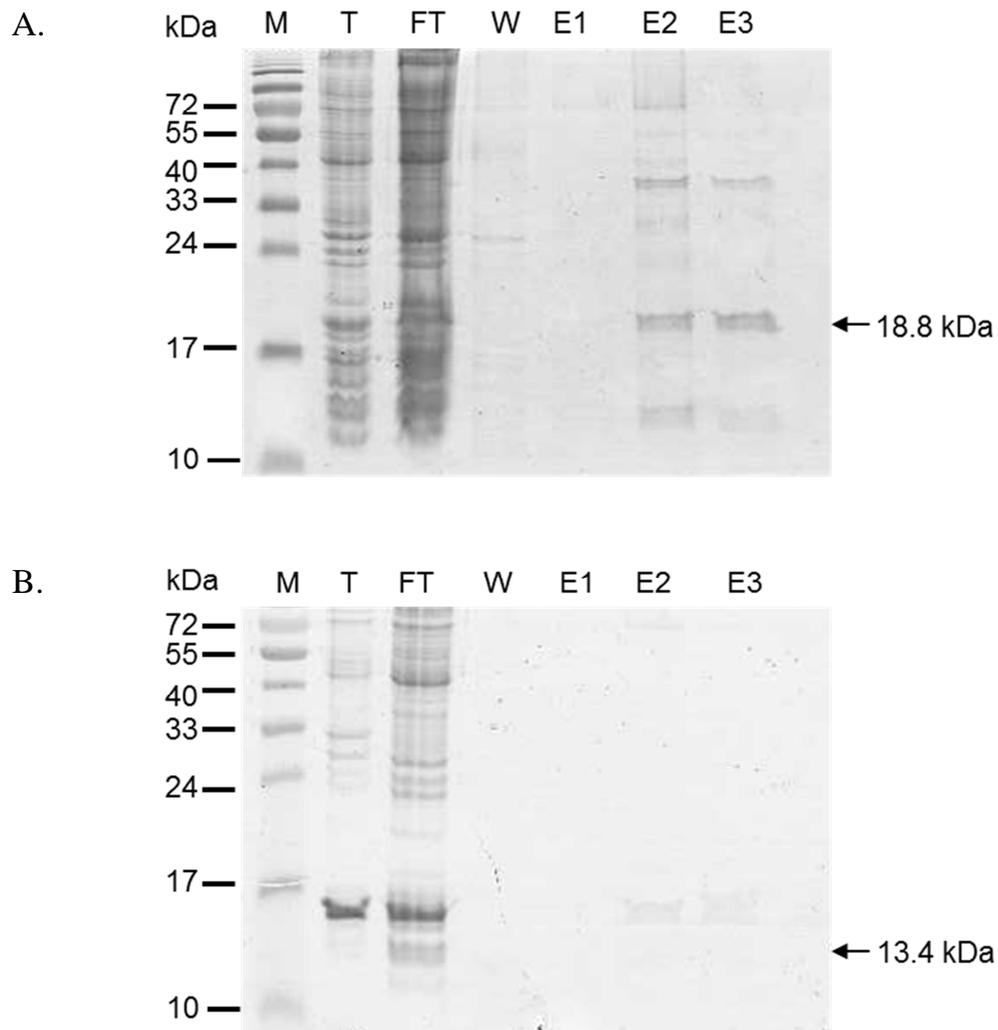
In contrast, GST-tagged proteins were expressed at higher levels compared to the His-tagged proteins, as indicated by the difference in band intensities in the total fractions during SDS-PAGE analysis. In this instance, GST-AntpHD-DRBM1 (C5.1) is efficiently eluted from the Glutathione sepharose 4B (GS4B) column, compared to GST-Pen-DRBM1 (C6.1) (Fig. 2.7 samples E1-E2). Isolation of the protein in the eluted fractions by competitive displacement with 10 mM reduced glutathione provided the foundations for further expression studies and the use of C5.1 as the template for the PCR amplification of

the c-myc epitope tag at the C-terminus of these proteins. Myc-tagged GST-AntpHD-DRBM1 (C5.1-myc) and GST-Pen-DRBM1-myc (C6.1-myc) were expressed at high levels (Fig. 2.8), although faint bands by SDS-PAGE analysis in the eluted fractions of C6.1-myc indicated that the protein was isolated at very low levels and that it is mostly expressed as an insoluble protein.

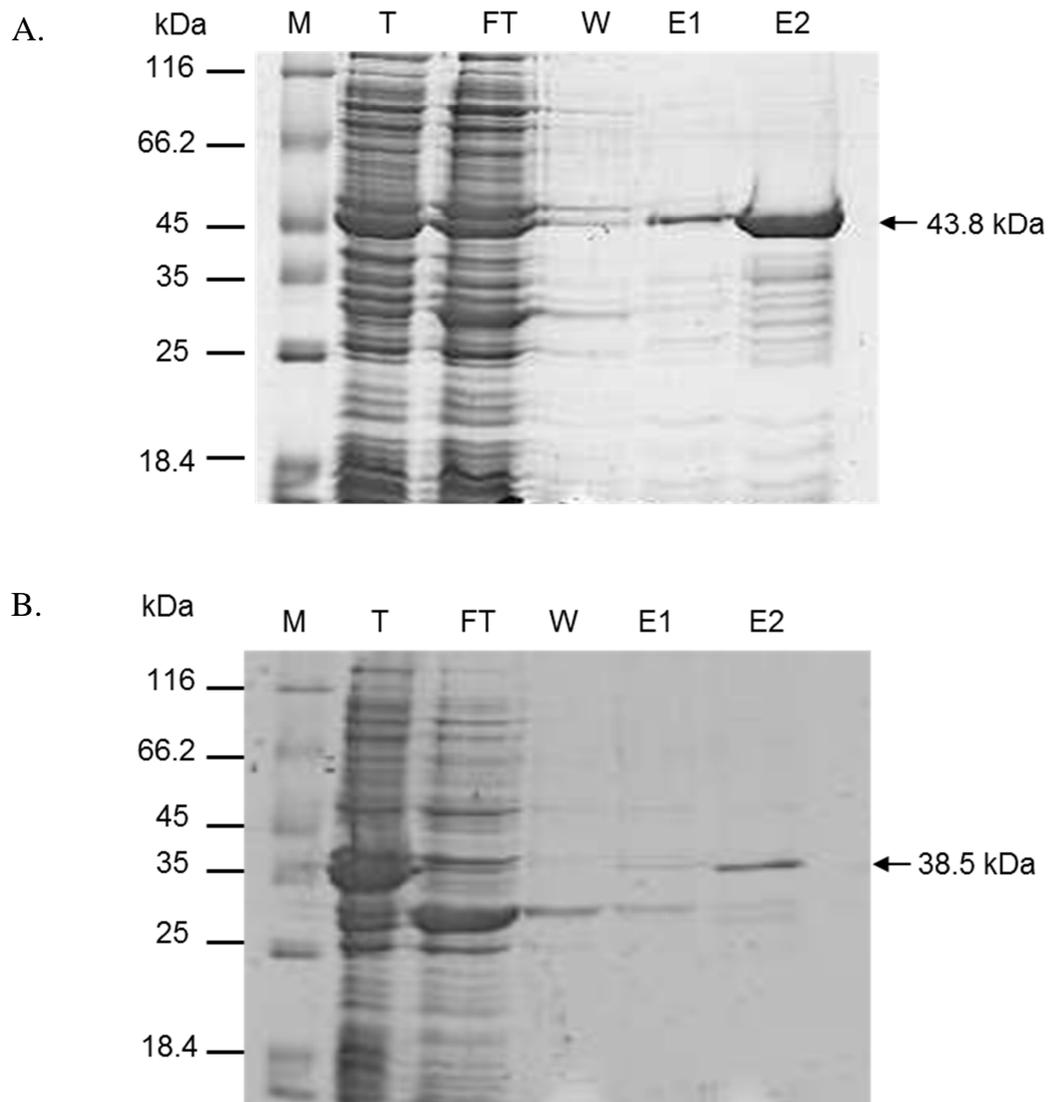
The presence of Pen-DRBM1 constructs (C2.1 and C4.1) mostly in the flowthrough fractions during purification over Ni-NTA or GS4B columns indicates that it is not sufficiently binding to the column and that methods for purification need further optimisation. To investigate whether increased yields of eluted C2.1 and C4.1 could be obtained from larger scale cultures under the same conditions, these proteins were purified from 0.5L cultures.



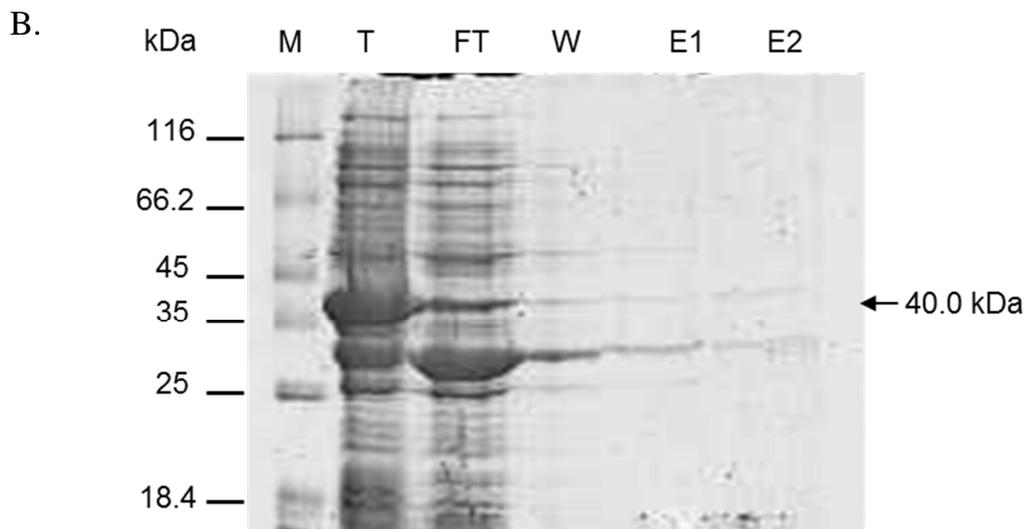
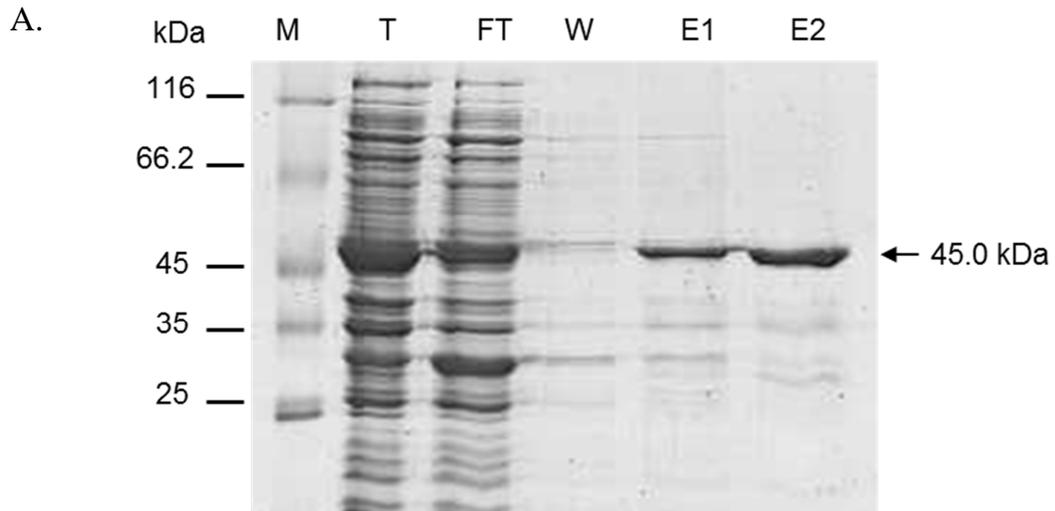
**Figure 2.5 Small-scale expression analysis of His<sub>10</sub> constructs C1.1 and C2.1 by SDS-PAGE.** **A.** Construct C1.1; **B.** Construct C2.1. 20  $\mu$ l samples were analysed by SDS-PAGE on a 12% w/v acrylamide gel. Fractions from the Ni-NTA purification are shown: T, Total fraction (in 6 M Urea); FT, Flowthrough; W, Wash, E1-E3, Eluates 1-3. M, molecular weight marker. Arrows indicate the expected molecular weight of fusion proteins in kDa. 15  $\mu$ l sample loaded per well. The gel was stained with Coomassie blue (0.5% w/v) and visualized on a BioRad scanner.



**Figure 2.6 Small-scale expression analysis of His<sub>6</sub> constructs C3.1 and C4.1 by SDS-PAGE.** **A.** Construct C3.1; **B.** Construct C4.1. 20 µl samples were analysed by SDS-PAGE on a 12% w/v acrylamide gel. Fractions from the Ni-NTA purification are shown: T, Total fraction (in 6 M Urea); FT, Flowthrough; W, Wash, E1-E3, Eluates 1-3. M, molecular weight marker. Arrows indicate the expected molecular weight of fusion proteins in kDa. 15 µl sample loaded per well The gel was stained with Coomassie blue (0.5% w/v) and visualized on a BioRad scanner.



**Figure 2.7 Small-scale expression analysis of GST-fusion constructs C5.1 and C6.1 by SDS-PAGE. A.** Construct C5.1; **B.** Construct C6.1. 20  $\mu$ l samples were analysed by 12% SDS-PAGE on a 12% w/v acrylamide gel. Fractions from the Ni-NTA purification are shown; T, Total fraction (in 6 M urea); FT, flowthrough; W, Wash, E1-E2, Eluates 1-2. M, molecular weight marker (kDa). 20 $\mu$ l sample loaded per well The gel was stained with Coomassie blue (0.5% w/v) and visualized on a Biorad scanner.



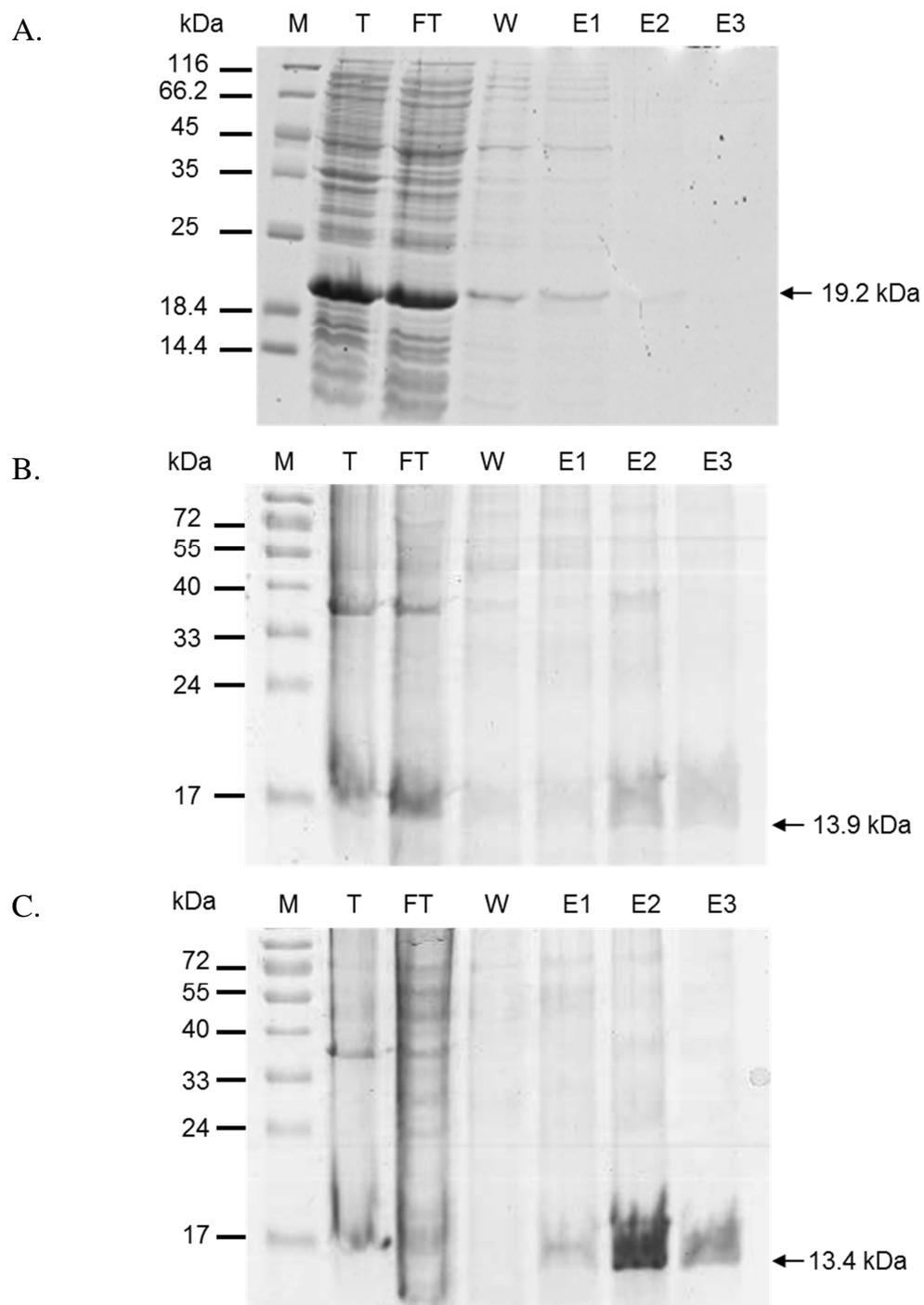
**Figure 2.8 Small-scale expression analysis of GST-fusion constructs C5.1-myc and C6.1-myc by SDS-PAGE. A. Construct C5.1-myc; B. Construct C6.1-myc.** 20  $\mu$ l samples were analysed by 12% SDS-PAGE on a 12% w/v acrylamide gel. Fractions from the Ni-NTA purification are shown; T, Total fraction (in 6 M urea); FT, flowthrough; W, Wash, E1-E2, Eluates 1-2. M, molecular weight marker (kDa). 20  $\mu$ l sample loaded per well The gel was stained with Coomassie blue (0.5% w/v) and visualized on a Biorad scanner.

### **2.2.3 Poly(His)- tagged protein purifications from 500 ml *E.coli* cultures under native conditions**

Following small scale expression tests on 50ml cultures, expression testing on a 500ml scale aimed to estimate His-tagged protein concentrations at each step of the purification process and to provide some insight as to the role of each tag with regards to solubility or expression enhancement. Prior His-tagged protein analysis had indicated that AntpHD-DRBM1 (C1.1 and C3.1) was both insoluble and expressed at low levels, an observation supported by purifications on 50 ml and 500 ml scales under the same conditions for C1.1 (Fig. 2.9 A). Larger scale cultures for C3.1 were not pursued due to the presence of contaminating nucleic acids in the flowthrough fractions.

Although SDS-PAGE gel band intensities signified that C1.1 was well expressed, it was mostly found in the Total (crude lysate) and the Flowthrough (cleared lysate) fractions and did not bind to the Ni-NTA column efficiently. Conversely, His<sub>10</sub>-Pen-DRBM1 (C2.1), which was expressed as a soluble fraction on a 50ml scale, was expressed at low levels and at low purities under native conditions. There was little protein isolated in the eluted fractions (Fig. 2.9 B).

C-terminus His<sub>6</sub>-tagged Pen-DRBM1 (C4.1) was expressed at higher levels than its N-terminally-tagged counterpart and was successfully eluted from the column. However, the banding pattern with vertical streaking in the eluted fraction lanes (E1-E3) pointed towards nucleic acid contamination during purification (Fig. 2.9 C) and therefore necessitated the addition of an extra purification step over an anion exchange column.



**Figure 2.9 Expression analysis of C1.1, C2.1 and C4.1 from 0.5L cultures by SDS-PAGE. A.** Construct C1.1; **B.** Construct C2.1; **C.** Construct C4.1. Samples were analysed by SDS-PAGE on a 12% w/v polyacrylamide gel. Fractions from the Ni-NTA purification are shown; T, Total fraction (in 6 M urea); FT, flowthrough; W, Wash, E1-E3, Eluates 1-3. M, molecular weight marker (kDa). 15  $\mu$ l sample loaded per well. The gel was stained with Coomassie blue (0.5% w/v) and visualized on a Biorad scanner.

#### 2.2.4 GST-tagged protein purifications from 500 ml *E.coli* cultures

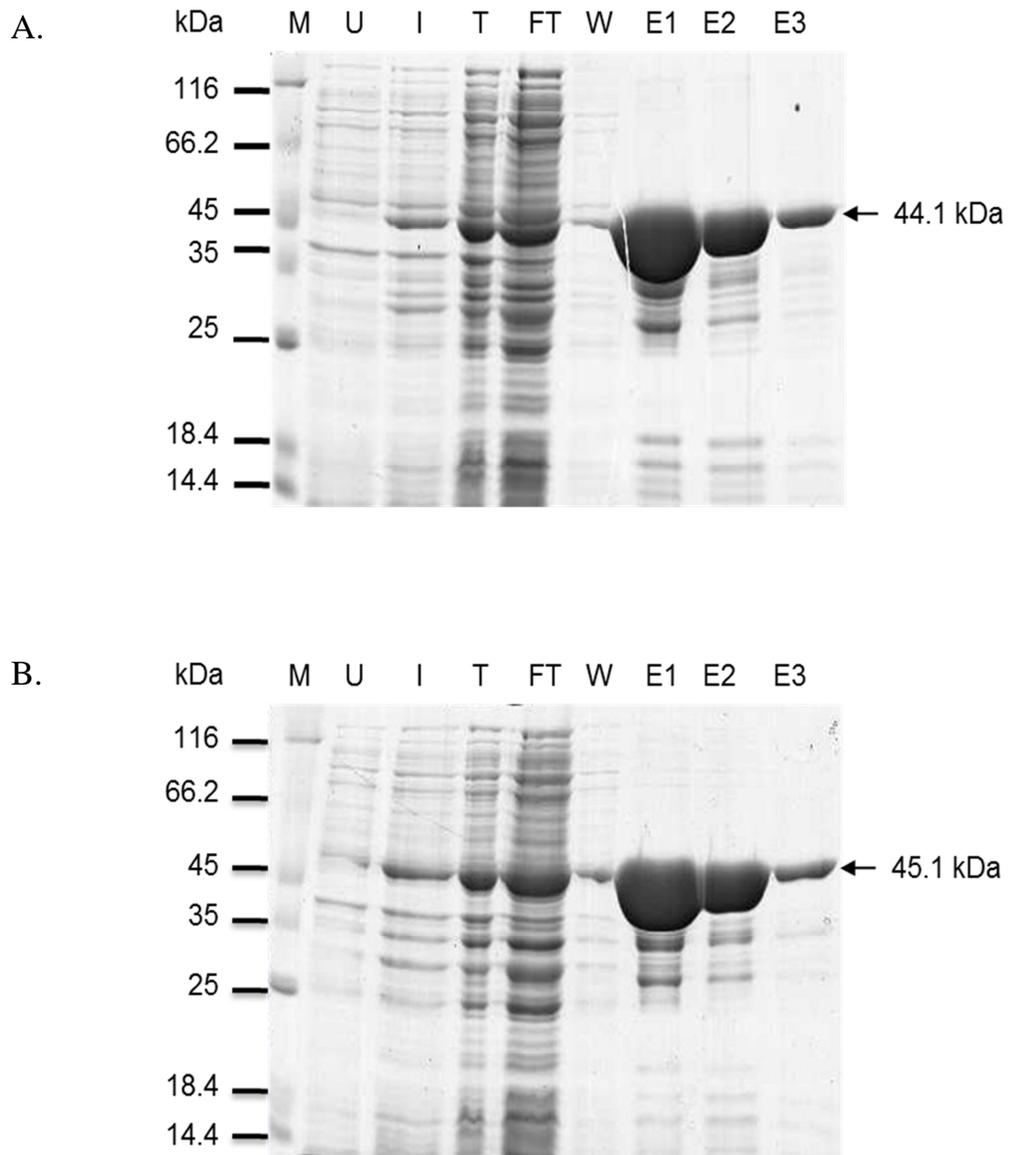
Having successfully purified GST-tagged C5.1, C5.1-myc, C6.1 and C6.1-myc, from 50ml *E.coli* cultures, protein purification was scaled to 500 ml. Both C5.1 and C5.1-myc were isolated with 80% purities following batch binding and gravity flow affinity purification over a GS4B column (Fig. 2.10). The induction of protein expression from BL21(DE3)pRARE with 0.5 mM IPTG was evident in the induced (I) fraction lanes (Fig. 2.10 and 2.11, Lane I) compared to the uninduced fraction lanes (U). In the case of C5.1 and C5.1-myc, there was little protein lost in the flowthrough (FT) and wash (W) fractions. Both proteins were successfully eluted from the GS4B column (E1-E3) by competitive displacement with 10 mM reduced glutathione in GST elution buffer (containing 500 mM NaCl and 10 mM DTT). Pooled eluate concentration was analysed by Nanodrop at  $A_{280}$  using each protein's predicted extinction coefficient ( $\epsilon$ ) and molecular weight (MW) (Appendix I). As a method, measuring the spectrophotometric absorbance of protein at  $A_{280}$  by Nanodrop provides reliable results with 98% accuracy, provided that samples are purified. Measurements are done based on the number of tryptophan and tyrosine residues and disulphide (cysteine-cysteine) bonds. Absorptivity is calculated based on Beer's Law (List of formulae, p18). In the case of uncharacterised proteins, however, concentration assessment by Nanodrop may not be as accurate as the Bicinchoninic acid (BCA) assay, although it eliminates the necessity for standard curve interpolation.

Pooled eluate concentrations for C5.1 and C5.1-myc were 3.98 mg/ml (with a final yield of 11.7 mg) and 3.53 mg/ml (with a final yield of 10.2 mg), respectively (Table 2.1). In contrast, most GST-tagged C6.1 and C6.1-myc, were lost in the FT fractions (Fig. 2.11 A and B). Soluble C6.1 and C6.1-myc protein bound to the GS4B column were eluted, although with lower purities and yields in comparison to C5.1 and C5.1-myc. Pooled eluate

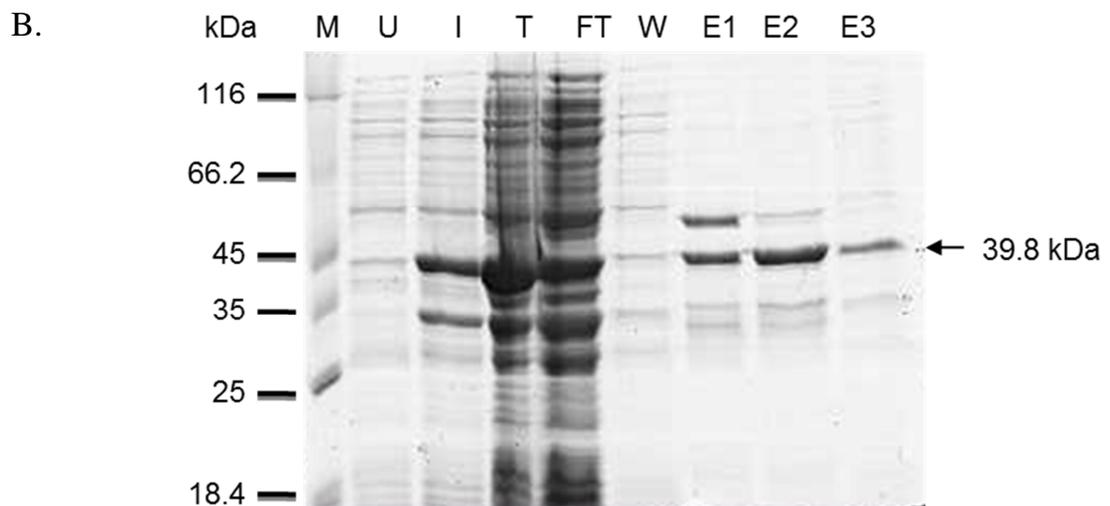
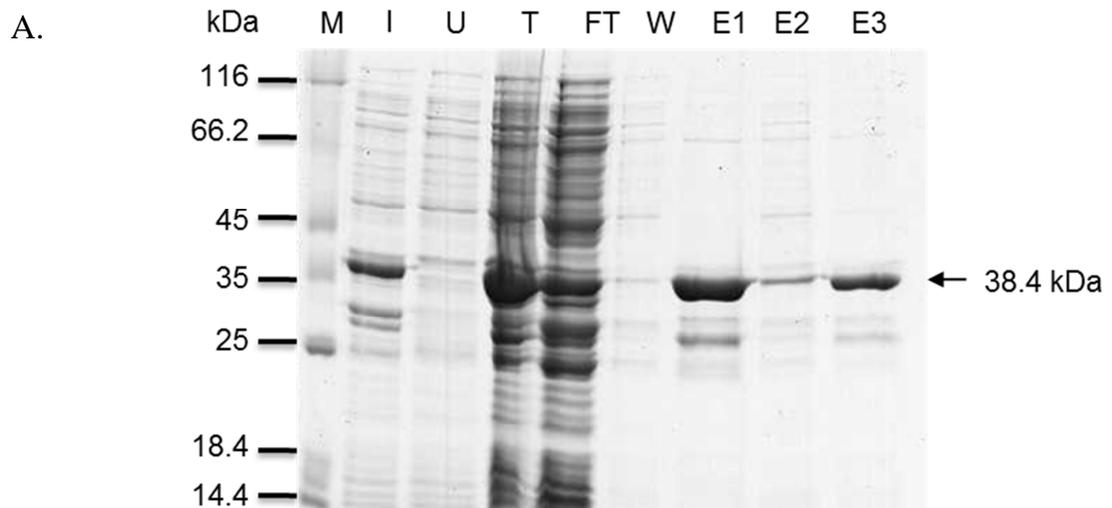
concentrations for C6.1 and C6.1-myc were 0.90 and 0.95 mg/ml, respectively. In order to cleave the GST tag with PP at a 1:5 molar ratio to total protein, according to the protocol developed by Barr *et al.* (2009), pooled eluates were buffer exchanged first into equilibration buffer (containing 200 mM NaCl and 2 mM DTT) to remove glutathione ions and decrease the molar concentrations of DTT and salt. This was done as they were incompatible with enzymatic activity at 4°C. The subsequent decrease in the ionic strength of the buffer resulted in a marked decrease in protein yields: C5.1 decreased by 19.4 %, C5.1-myc by 24.8%, C6.1 by 57.4% and C6.1-myc by 62.6% (Table 2.1).

Further yield losses were observed upon tag cleavage by PP at a 1:5 ratio to protein in GST equilibration buffer. Protein desalting into 1xPBS resulted in further yield losses; C5.1 decreased by 51.8%, C5.1-myc by 14.9%, C6.1 by 46.6% and C6.1-myc by 57.4%, compared to yields in equilibration buffer before tag cleavage (Fig. 2.12 and 2.13).

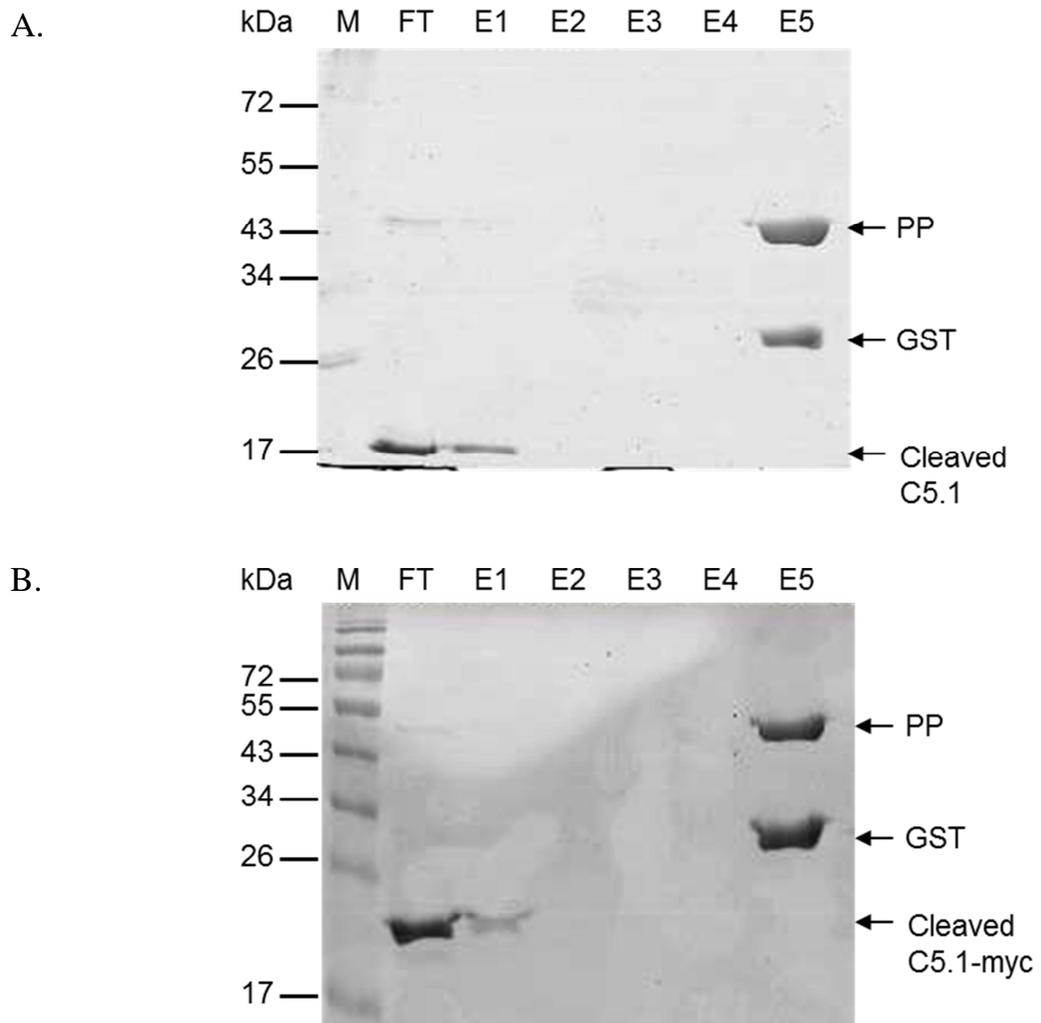
Our multi-step purification strategy (purification, buffer exchange, protease cleavage and desalting) resulted in overall losses of 61.1%, 36.0%, 77.2% and 84.1% for C5.1, C5.1-myc, C6.1 and C6.1-myc, respectively (Table 2.1). Final yields for C5.1, C5.1-myc, C6.1 and C6.1-myc were 6.6 mg, 4.6 mg, 0.4 mg and 0.4 mg, respectively.



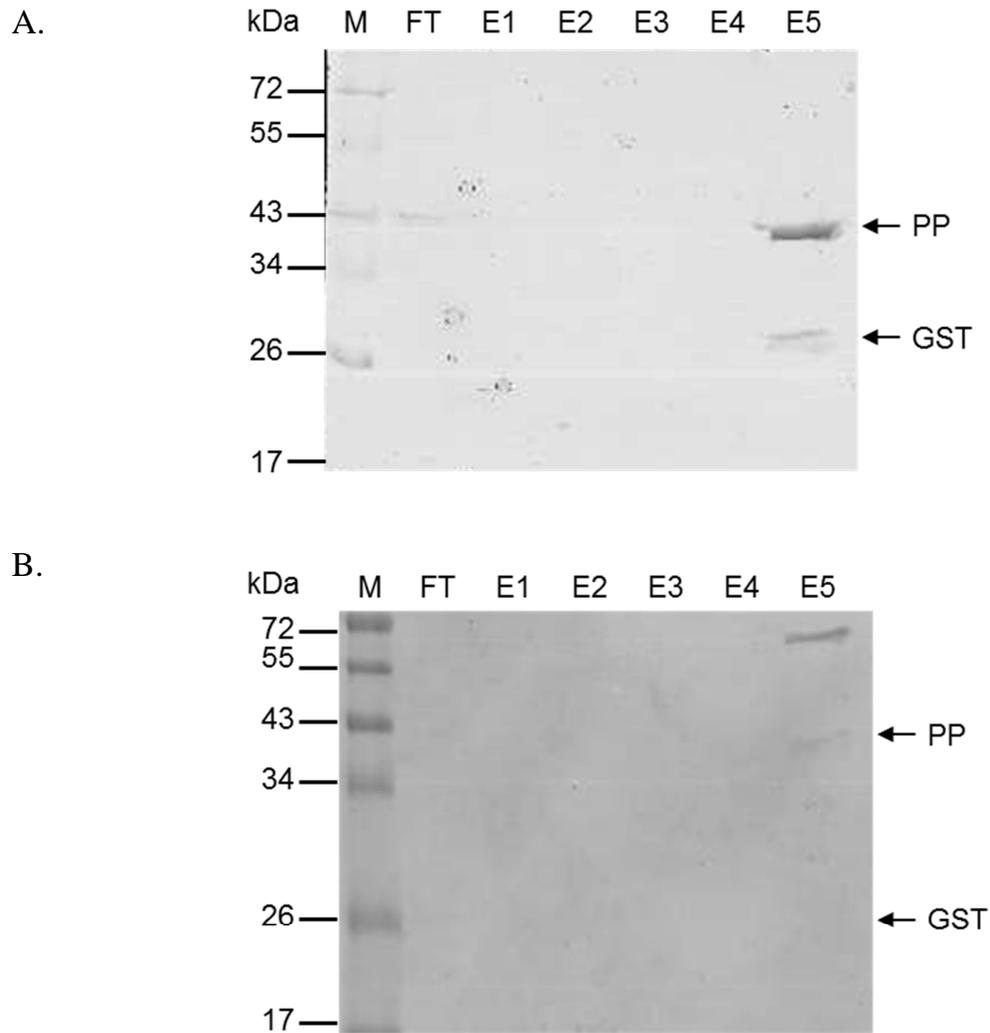
**Figure 2.10 Expression analysis of GST-tagged constructs C5.1 and C5.1-myc from 0.5L cultures by SDS-PAGE.** **A.** Construct C5.1; **B.** Construct C5.1-myc. Samples were analysed by SDS-PAGE on a 12% w/v acrylamide gel. Fractions from the GS4B purification are shown: U, Uninduced fraction; I, Induced fraction T, Total fraction (in 6 M Urea); FT, Flowthrough; W, Wash, E1-E3, Eluates 1-3. M, molecular weight marker. Arrows indicate the expected molecular weight of fusion proteins in kDa. 20  $\mu$ l sample loaded per well. The gel was stained with Coomassie blue (0.5% w/v) and visualized on a BioRad scanner. n=1.



**Figure 2.11 Expression analysis of GST-tagged constructs C6.1 and C6.1-myc from 0.5L cultures by SDS-PAGE. A. Construct C6.1; B. Construct C6.1-myc.** Samples were analysed by SDS-PAGE on a 12% w/v acrylamide gel. Fractions from the GS4B purification are shown: U, Uninduced fraction; I, Induced fraction T, Total fraction (in 6M Urea); FT, Flowthrough; W, Wash, E1-E3, Eluates 1-3. M, molecular weight marker. Arrows indicate the expected molecular weight of fusion proteins in kDa. 15  $\mu$ l sample loaded per well. The gel was stained with Coomassie blue (0.5% w/v) and visualized on a BioRad scanner. n=1.



**Figure 2.12 0.5L expression analysis of C5.1 and C5.1-myc in 1xPBS following GST-tag cleavage, by SDS-PAGE. A. Construct C5.1; B. Construct C5.1-myc.** Samples were analysed by SDS-PAGE on a 12% w/v acrylamide gel. Fractions from the GS4B purification and tag cleavage by Precission protease are shown: FT, Flowthrough; E1-E4, Eluates 1-4, E5, Eluate 5. M, molecular weight marker. Arrows indicate proteins present in each fraction. PP, Precission protease. 20  $\mu$ l sample loaded per well. Gels were stained with Coomassie blue (0.5% w/v) and visualized on a BioRad scanner. n=1.



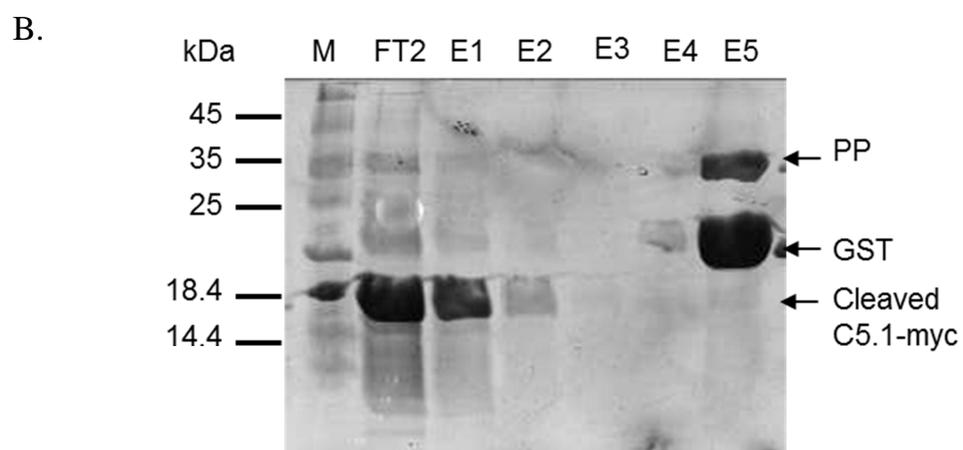
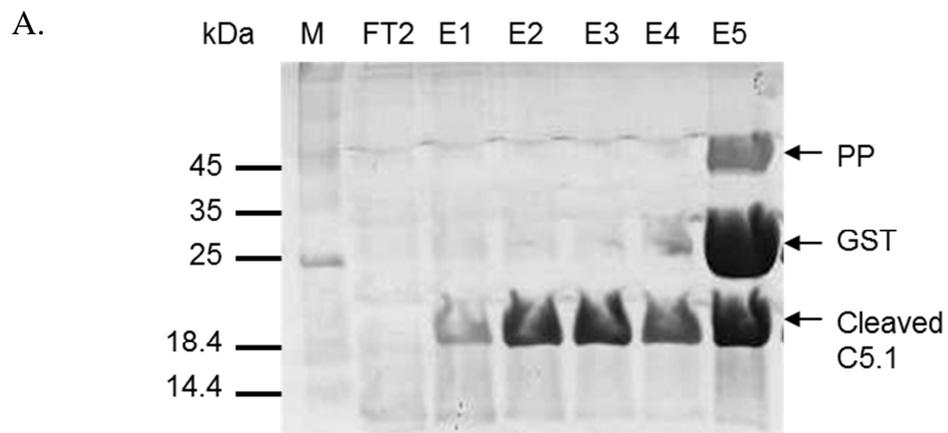
**Figure 2.13 0.5L expression analysis of C6.1 and C6.1-myc in 1xPBS following GST-tag cleavage, by SDS-PAGE. A. Construct C6.1; B. Construct C6.1-myc.** Samples were analysed by SDS-PAGE on a 12% w/v acrylamide gel. Fractions from the GS4B purification and tag cleavage by Precission protease are shown: FT, Flowthrough; E1-E4, Eluates 1-4, E5, Eluate 5. M, molecular weight marker. Arrows indicate proteins present in each fraction. PP, Precission protease. 20  $\mu$ l sample loaded per well. The gel was stained with Coomassie blue (0.5% w/v) and visualized on a BioRad scanner. n=1.

**Table 2.1** Yield analysis of GST- tagged fusion proteins obtained during purification from 0.5L *E. coli* cultures. Yield represents the total protein found in eluted samples at each step of the purification process. % change represents change in yield between each purification step. Overall % change represents the change in yield from the point of elution to the point of buffer-exchange into 1xPBS after GST tag removal by Prescission protease. n=1.

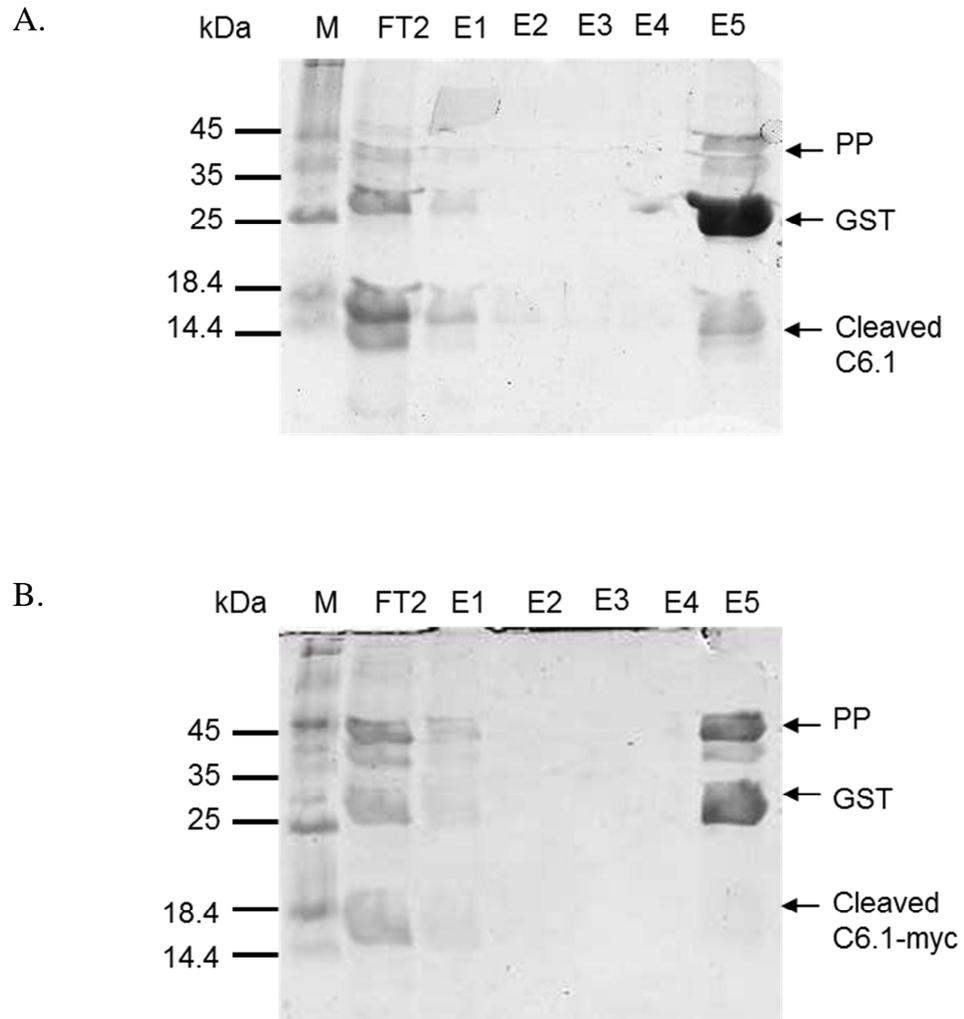
<b>Protein</b>	<b>Yield in elution buffer (mg)</b>	<b>Yield following buffer-exchange into in equilibration buffer (mg)</b>	<b>Change in yield (%)</b>	<b>Yield in 1xPBS following tag cleavage by PP and desalting (mg)</b>	<b>Change in yield (%)</b>	<b>Overall % change (Step 1 → 3)</b>
<b>C5.1</b>	11.7	9.4	-19.4	6.6	-51.8	-61.1
<b>C5.1-myc</b>	10.2	7.7	-24.8	4.6	-14.9	-36.0
<b>C6.1</b>	1.8	0.8	-57.4	0.4	-46.6	-77.2
<b>C6.1-myc</b>	2.5	1.0	-62.6	0.4	-57.4	-84.1

### **2.2.5 Large scale purifications of GST-tagged C5.1, C6.1-myc, C6.1 and C6.1-myc**

Loss of protein yields from 500 ml purifications indicated that the process of purifying GST-tagged proteins needed optimization. The long purification protocol adopted with 0.5L cultures, comprised of a sequential buffer exchange to achieve GST tag cleavage and desalting, had proved detrimental to protein yields. C5.1 and C5.1-myc, which had exhibited fewer % losses, were up-scaled to 4L cultures and purified, while C6.1 and C6.1-myc from 3L cultures. The purification strategy was adjusted to reflect changes in volumes and the slow kinetics between the GST tag and the resin. To minimise yield losses, a batch-binding strategy with PP and on-column cleavage was adopted. Cleared lysates and resin were batch-bound for at least one hour at 4°C before they were applied to the column; the sepharose matrix with bound protein was incubated with PP at a 5:1 ratio of enzyme to estimated protein yield. Following GST tag removal, proteins were eluted from the column by gravity flow affinity chromatography into GST binding buffer. A final elution step with glutathione-containing elution buffer released the protease, cleaved tag and any uncleaved protein that remained on the column (Fig. 2.14 and 2.15, Lane E5). Fractions containing purified protein (FT2, E1-E4) were pooled and buffer exchanged into 1xPBS. SDS-PAGE analysis indicated that C5.1 was isolated mostly in fractions E1-E4, with very little protein being lost from the column in the FT2 fraction, while C5.1-myc, C6.1 and C6.1-myc were isolated mostly in the FT2-E2 fractions. Uncleaved C5.1 and C6.1 are also visible in the E5 fraction. SDS-PAGE results are representative of 2 independent experiments (n=2). Final yields in 1xPBS for C5.1, C5.1-myc from 4L cultures were 35.5 mg and 64.35 mg, respectively, while final yields for C6.1 and C6.1myc from 3L cultures were 5.3 mg and 3.9 mg, respectively (Table 2.2).



**Figure 2.14 4L expression analysis of C5.1 and C5.1-myc following GST-tag cleavage, by SDS-PAGE. A.** Construct C5.1; **B.** Construct C5.1-myc. Samples were analysed by SDS-PAGE on a 12% w/v acrylamide gel. Fractions from the GS4B purification and tag cleavage by Prescission protease are shown: FT2, Flowthrough 2; E1-E4, Eluates 1-4, E5, Eluate 5. M, molecular weight marker. Arrows indicate proteins present in each fraction. PP, Prescission protease. 20  $\mu$ l sample loaded per well. The gel was stained with Coomassie blue (0.5% w/v) and visualized on a BioRad scanner. n=2.



**Figure 2.15 3L expression analysis of C6.1 and C6.1-myc following GST-tag cleavage, by SDS-PAGE. A. Construct C6.1; B. Construct C6.1-myc.** Samples were analysed by SDS-PAGE on a 12% w/v acrylamide gel. Fractions from the GS4B purification are shown: FT2, Flowthrough 2; E1-E4, Eluates 1-4, E5, Eluate 5. M, molecular weight marker. Arrows indicate proteins in the E5 fraction. PP, Precision protease. 20  $\mu$ l sample loaded per well. The gel was stained with Coomassie blue (0.5% w/v) and visualized on a BioRad scanner. n=2.

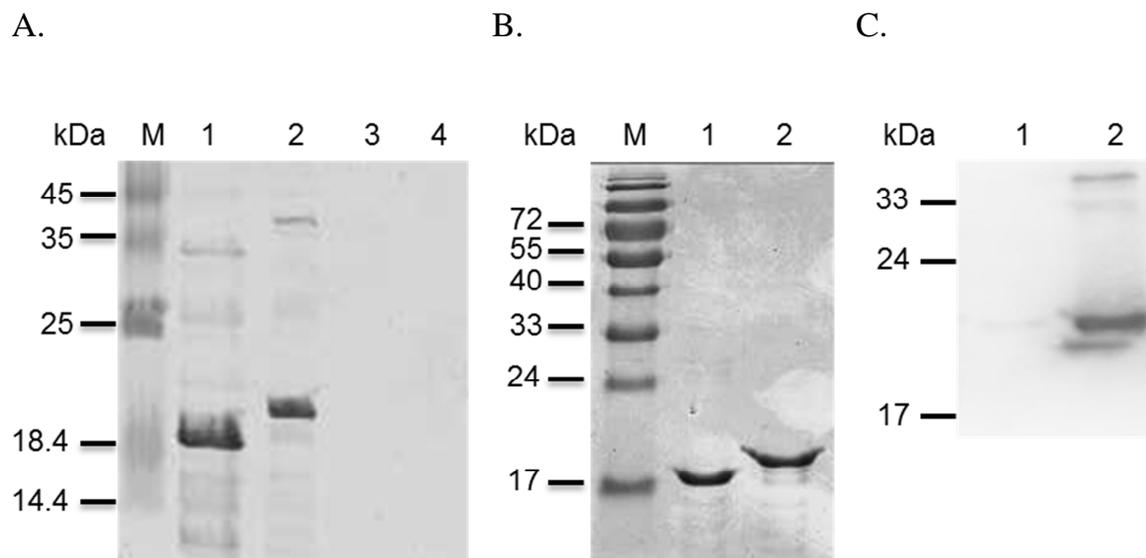
**Table 2.2 Large-scale protein purification yields for C5.1, C5.1-myc, C6.1 and C6.1-myc.** The GST tag was cleaved by Prescission protease and eluates were concentrated in an Amicon Ultra-15 column before buffer exchange into 1xPBS/10% (v/v) glycerol. Results are representative of 2 independent experiments (n=2).

Protein	Protein yields (mg) or concentrations (mg/ml) after each purification step					
	Eluates after GST cleavage in GST Binding buffer (mg)	Concentration in GST binding buffer (mg)	% change in Yield	Buffer Exchange into 1xPBS /10% (v/v) Glycerol (mg)	% change in Yield	% Change in Overall Yield
<b>C5.1</b>	45.1	36.4	-19.30%	35.5	-2.6%	-21.3%
<b>C5.1-myc</b>	118.6	97.4	-17.89%	64.4	-33.9%	-45.7%
<b>C6.1</b>	5.0	7.5	+51.21%	5.3	-29.7%	+6.3%
<b>C6.1-myc</b>	8.8	3.0	-65.90%	3.9	+22.1%	-65.9%

### **2.2.6 C5.1 and C5.1-myc stability assessment and detection by western blotting**

In order to assess the stability of proteins prior to further *in vitro* studies, purified C5.1 and C5.1-myc in PBS were characterised by SDS-PAGE before and after thawing. Stability was assessed by centrifugation at 16,100 x *g* and by immunoblot with an anti-myc antibody (Fig 2.16). The difference in band intensity before and after centrifugation indicated that at least some C5.1 and C5.1-myc protein had precipitated, even when equal amounts of protein were loaded on an SDS-PAGE gel (20 µg).

C6.1 and C6.1-myc were also assessed by SDS-PAGE following purification, but appeared to have precipitated following buffer exchange into PBS/10% (v/v) glycerol (Fig 2.16 A). SDS-PAGE and western blot analysis (Fig. 2.16 C) of C5.1 and C5.1-myc with an anti-myc antibody was done to confirm expression and integrity of the myc-epitope tag, as well as to validate the anti-myc antibody prior to subsequent immunofluorescence studies with C5.1-myc. Final yields were 33.8 mg, 36.6 mg, 0.3 mg and 0.7 mg for C5.1, C5.1-myc, C6.1 and C6.1-myc, respectively (Table 2.3).



**Figure 2.16 Stability analysis by centrifugation and immunoblot of purified C5.1, C5.1-myc, C6.1 and C6.1-myc in 1xPBS.** **A.** Proteins C5.1 (Lane 1) and C5.1-myc (Lane 2) remain soluble in 1xPBS whilst C6.1 (Lane 3) and C6.1-myc (Lane 4) precipitate following buffer exchange. **B.** C5.1 and C5.1-myc (Lanes 1-2) centrifuged at 16,100 x g for 30 minutes at 4°C remain stable with little loss of protein. 20 µg protein were analysed on a 12% w/v SDS-PAGE gel. M; Marker. **C.** Immunoblot analysis on 10 µg C5.1 (Lane 1) and C5.1-myc (Lane 2) with an anti-myc primary antibody and an HRP-conjugated secondary antibody. 20 µg total protein loaded. Bands were detected by chemiluminescence on a UVP Auto Chemi Darkroom Imaging system.

**Table 2.3 Final concentrations and yields of cleaved proteins.** Yields were obtained following purification of proteins from large scale cultures in 1xPBS, following centrifugation at 4°C and 16,100 x g.

Protein	Concentration (mg/ml)	Final yield (mg)
C5.1	1.13	33.8
C5.1-myc	1.22	36.6
C6.1	0.075	0.30
C6.1-myc	0.175	0.70

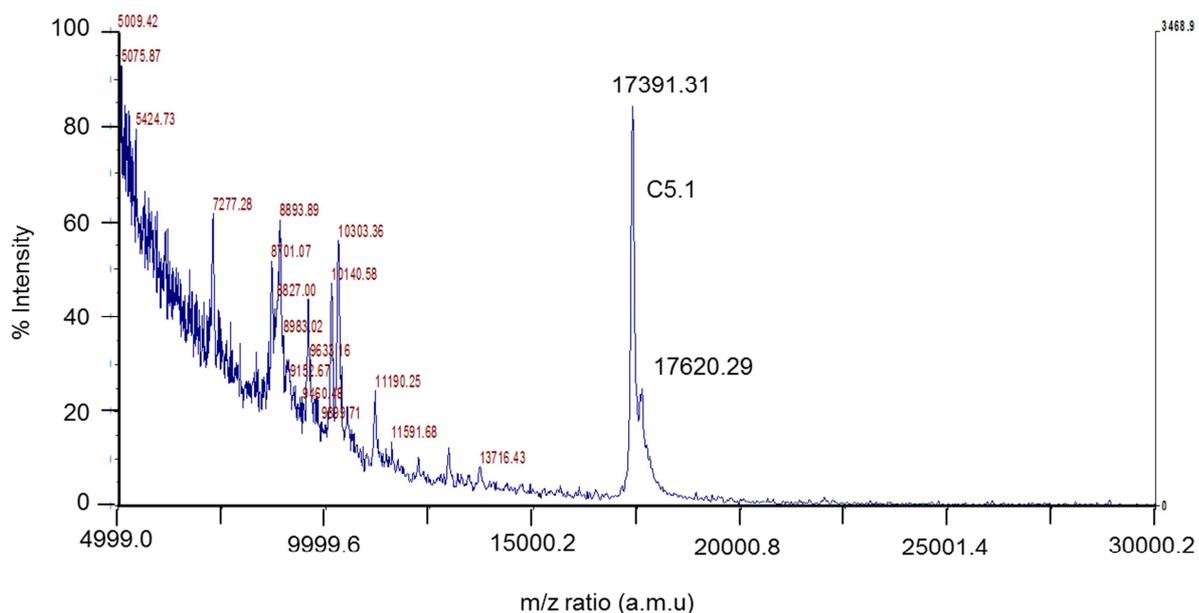
### 2.2.7 Mass and purity determination for C5.1 and C5.1-myc by Mass Spectroscopy

Knowledge of the exact molecular mass and biochemical parameters of a potential therapeutic molecule is crucial for pre-clinical evaluations *in vitro*. Before proceeding with localisation experiments in HEK293 cells, Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight Mass Spectrometry (MALDI-TOF) analysis was performed after positive ionization of the 10 µg C5.1 in formic acid. The mass to charge ratio ( $m/z$ ) spectrum showed a dominant peak at 17,391 atomic mass units (amu) (Fig. 2.17), in good accordance to the 17,401 Da value predicted by bioinformatic analysis carried out by the ExPASy ProtParam online tool (Gasteiger E., 2005). A larger peak at  $m/z$  17,620 indicated the formation of a protein species that was approximately 229 Da larger than the predominant species. It is not known whether this species the result of post-translational modifications or nucleic acid contamination.

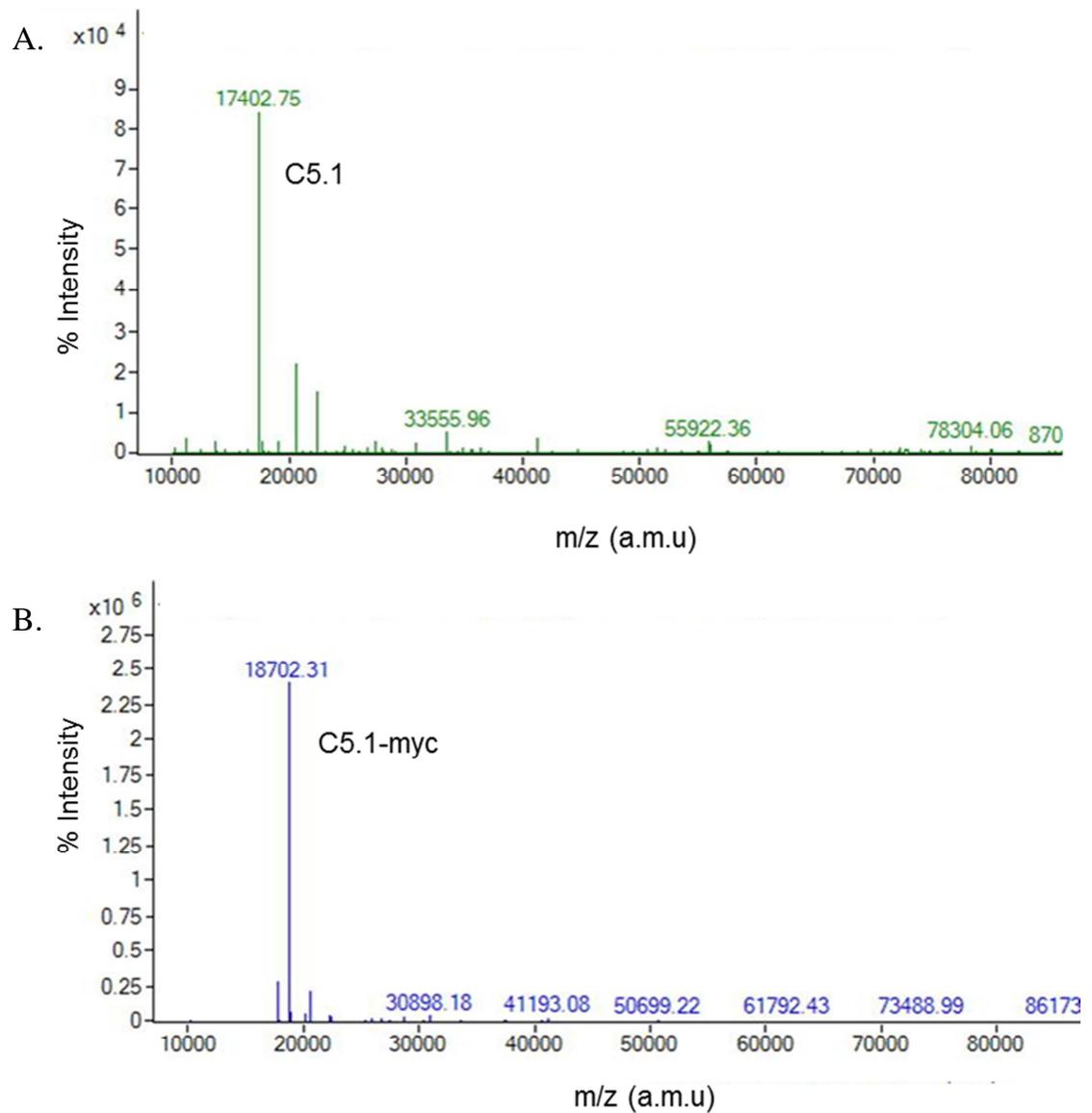
The intact protein mass of C5.1 and C5.1-myc was further assessed by Electrospray Ionisation-Time-of-Flight Mass Spectrometry (ESI-QTOF MS). Electrospray ionisation is known as a "soft" ionisation method as the sample is ionised by the addition (or removal) of a proton, with very little extra energy remaining to cause fragmentation of the sample ions and has a mass accuracy of 99.9%. Each peak represents the intact protein molecule carrying a different number of charges (in this case,  $H^+$ ). The deconvoluted spectrum of C5.1 in positive ionisation mode produced a dominant peak at 17,402.75 amu, consistent with our predicted MW of 17,401 Da with few, larger impurities (Fig. 2.18 A).

Consistent with a predicted MW of 18,701 Da, the  $m/z$  spectrum for C5.1-myc produced a dominant peak at 18,702.31 amu with smaller impurities, which could be attributed to phosphate groups in the buffer (Fig. 2.18 B).

MALDI-TOF was performed at the Proteomics facility, UCL, while ESI-QTOF was performed at the Structural Genomics Consortium (SGC), University of Oxford.



**Figure 2.17 Intact C5.1 molecular weight assessed by MALDI-TOF.** The dominant ion peak shows the actual mass of C5.1 to be 17,391 Da, in good accordance with the expected molecular mass of 17,401 Da. The m/z ratio is the mass-to-charge ratio measured in a.m.u (atomic mass units).



**Figure 2.18 Deconvoluted ESI-QTOF spectra for C5.1 (A) and C5.1-myc (B).** Purified proteins were positively ionised in formic acid and assessed for molecular weight and purity. **A.** The dominant ion peak at 17,402.75 Da represents the actual mass of C5.1 (expected molecular mass, 17,401 Da). **B.** C5.1-myc shows a dominant peak at 18,702.31 Da (expected molecular mass of 18,701 Da). The m/z ratio is the mass-to-charge ratio measured in a.m.u (atomic mass units).

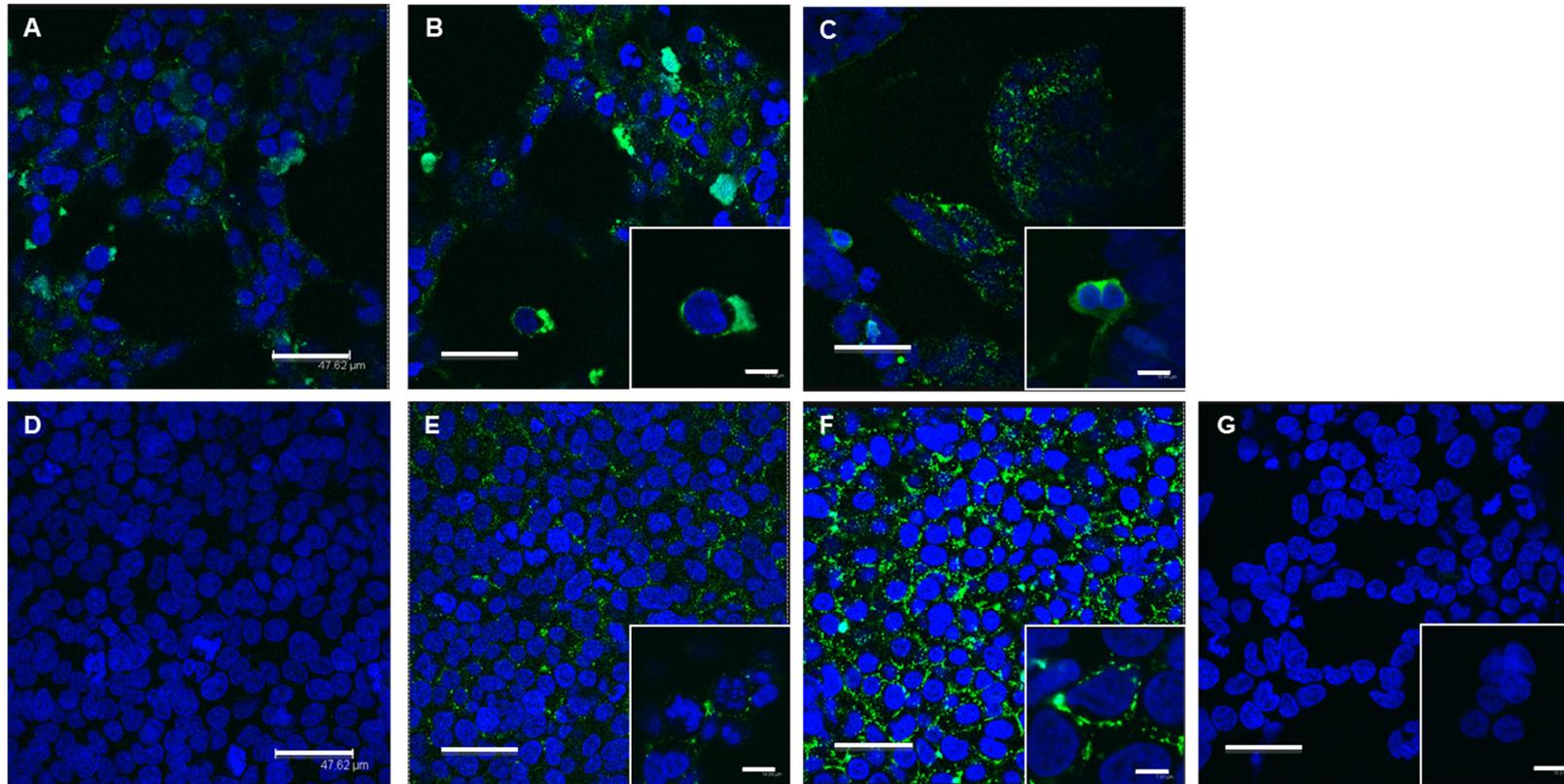
### **2.2.8 C5.1-myc shows a dose-dependent intracellular localisation in HEK293 and HepG2 cells**

Following protein characterisation, it was important to ascertain whether C5.1-myc has the potential for intracellular transduction. Localisation studies by indirect immunofluorescence with an anti-myc antibody were therefore performed. These subsequently allowed the determination of the protein's intracellular localisation upon internalisation. For this purpose, human embryonic kidney 293 (HEK293) cells were used; their size (20-30  $\mu\text{m}$ ), division time (70% confluency within 48 hours of passaging), robust morphology, as well as ease of manipulation as a transfection host, make them an excellent mammalian cell vector for translocation/transduction experiments (Thomas and Smart, 2005). The human hepatocellular carcinoma (HepG2) cell line on the other hand, is an epithelial cell line which endogenously expresses protein tyrosine phosphatase 1B (PTP1B). PTP1B dephosphorylates the insulin receptor (IR), thereby negatively regulating IR signalling and is a potential target for T2D (Asante-Appiah and Kennedy, 2003, Cheng *et al.*, 2002, Drake and Posner, 1998, Goldstein *et al.*, 1998). This cell line has also been successfully used in other cell penetrating peptide studies (El-Andaloussi *et al.*, 2011).

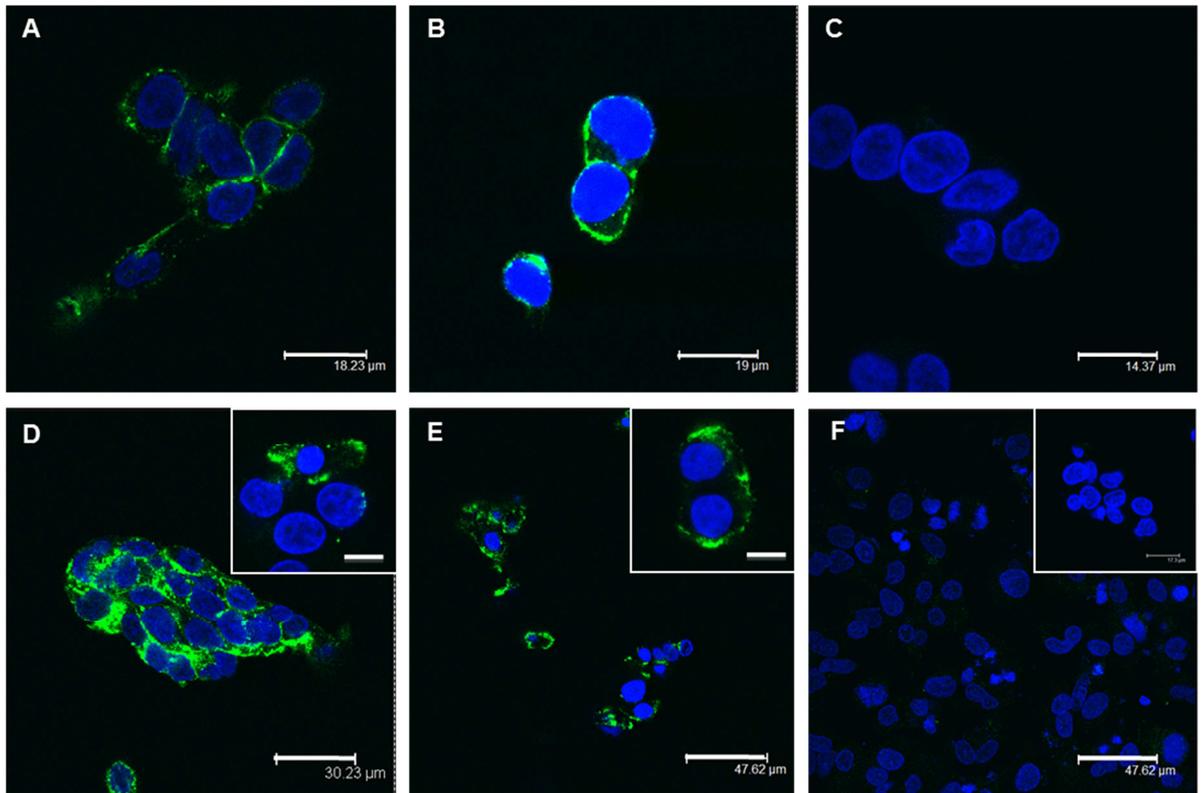
Purified C5.1-myc was incubated with seeded, adherent cells for either 2 or 24 hours at 3 different concentrations. The protein's c-myc epitope tag was detected with an anti-myc antibody (clone 4A6, Millipore) and a fluorescein thiocyanate (FITC) - labelled secondary antibody following fixation/permeabilisation with 4% Paraformaldehyde/0.2% Triton-X. Nuclei were stained with ToPro3. Finally, cells were visualized with a Leica SP2 confocal microscope. Indirect immunofluorescence of C5.1-myc in HEK293 cells showed a dose-dependent and time-dependent translocation at both timepoints (Fig 2.19, enlarged version in Appendix I.4).

Poor translocation was observed after 2-hour incubation even at the lowest molar concentration assayed (0.1  $\mu\text{M}$ ) (Fig. 2.19 A), as the protein at this point was localized primarily on the exterior face of the plasma membrane. At 1.0  $\mu\text{M}$ , the protein was seen to translocate into cells more efficiently without altering cell morphology, with a mixture of diffuse and punctate distributions (Fig. 2.19 B). Similar distribution patterns were observed at 10.0  $\mu\text{M}$  concentrations, with cell morphology remaining unaffected following treatment, compared to the mock control (Fig. 2.19 C). Conversely, there was little evidence of translocation at the 0.1  $\mu\text{M}$  concentration after 24-hour incubation (Fig. 2.19 D), possibly due to the low protein concentration. Protein internalization at this time-point appeared to be dose-dependent, with translocation and increased extra-cellular aggregation of the protein increasing between 1.0  $\mu\text{M}$  and 10.0  $\mu\text{M}$  (Fig. 2.19 E-F). Entrapment into endosomal vesicles upon transduction was also observed, which appeared to be dose-dependent at both time-points.

Qualitative observations from indirect immunofluorescence experiments with C5.1-myc at 0.1 and 1.0  $\mu\text{M}$  concentrations in fixed HepG2 cells after 2 and 24-hour incubations also pointed towards a dose-dependent and time-dependent uptake mechanism. After a 2-hour incubation, C5.1-myc at 0.1  $\mu\text{M}$  concentration localized primarily on the outer membrane fold (Fig. 2.20 A). Internalized protein distribution was mostly diffuse at both 0.1 and 1.0  $\mu\text{M}$  concentrations, with some evidence of endosomal entrapment (Fig. 2.20 A-B). At 24 hours, the protein was effectively translocating into cells at both 0.1  $\mu\text{M}$  and 1.0  $\mu\text{M}$  concentrations with evidence of perinuclear endosomal entrapment, typical of late endosome localization (Fig. 2.20 D-E).



**Figure 2.19** Intracellular localisation of C5.1-myc in fixed HEK293 cells is time-dependent and dose-dependent. Cells were incubated with 0.1  $\mu\text{M}$ , 1.0  $\mu\text{M}$  or 10.0  $\mu\text{M}$  protein for either 2 or 24 hours, permeabilised, and fixed. Protein was detected with an anti-myc primary antibody and a FITC-labelled secondary antibody. Cells were visualized with a TCS SP2 Leica confocal microscope under a 63x objective. Nuclei were counterstained with Topro3. **A.** 0.1  $\mu\text{M}$ , 2 hours; **B.** 1.0  $\mu\text{M}$ , 2 hours; **C.** 10.0  $\mu\text{M}$ , 2 hours. **D.** 0.1  $\mu\text{M}$ , 24 hours; **E.** 1.0  $\mu\text{M}$ , 24 hours; **F.** 10.0  $\mu\text{M}$ , 24 hours. **G.** Mock control (PBS), 24 hours. Scale bar: 47.6  $\mu\text{m}$ . Insets show magnified cells on a 10  $\mu\text{m}$  scale.



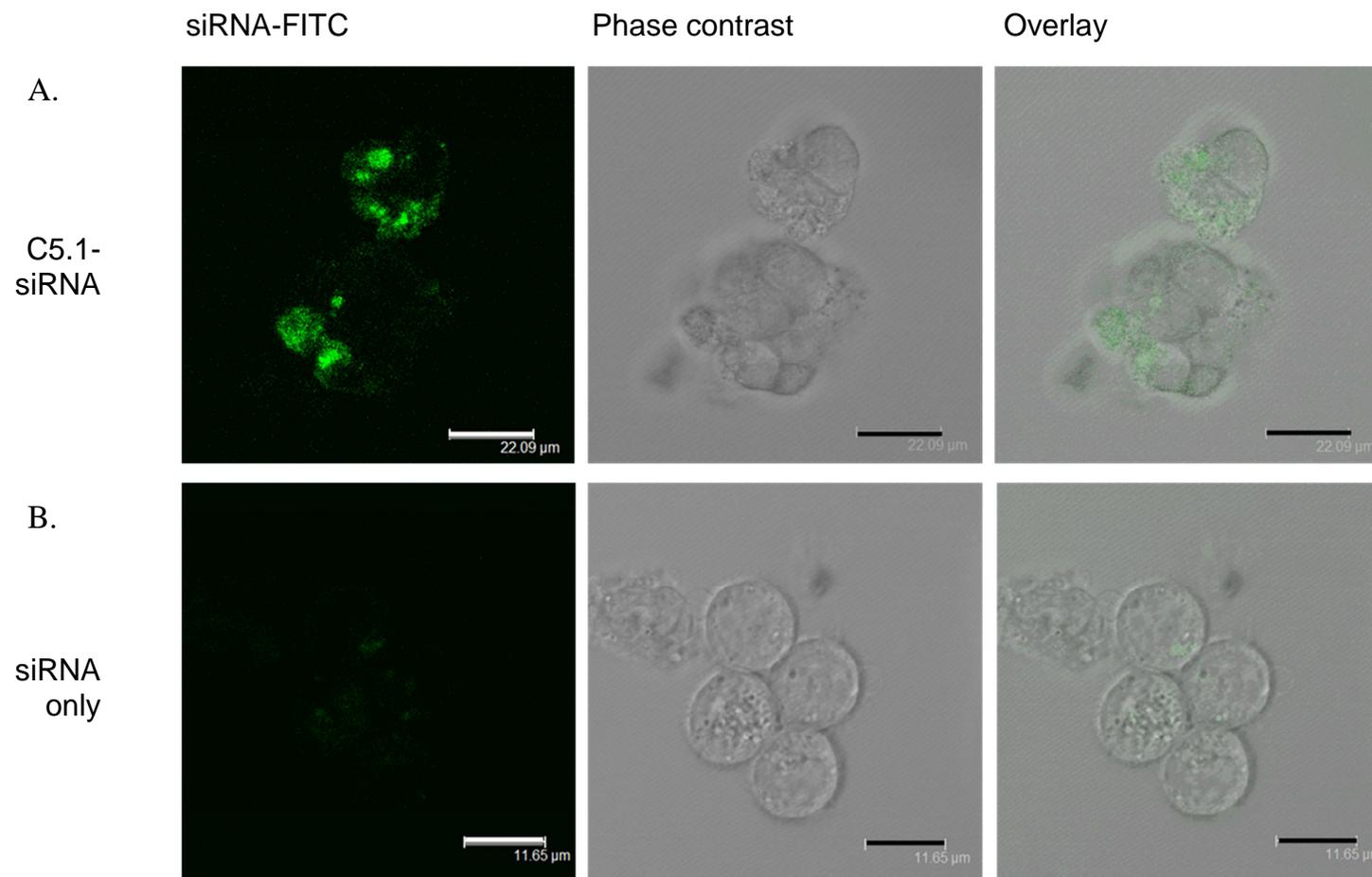
**Figure 2.20 Intracellular localization of C5.1-myc in fixed HepG2 cells is time-dependent and dose-dependent.** Cells were incubated with 0.1  $\mu\text{M}$ , 1.0  $\mu\text{M}$  or 10.0  $\mu\text{M}$  protein for either 2 or 24 hours, permeabilized, and fixed. Protein was detected with an anti-myc primary antibody and a FITC-labelled secondary antibody. Cells were visualized with a TCS SP2 Leica confocal microscope under a 63x objective. Nuclei were counterstained with Topro3. **A.** 0.1  $\mu\text{M}$ , 2 hours; **B.** 1.0  $\mu\text{M}$ , 2 hours; **C.** Mock control (PBS), 2 hours. **D.** 0.1  $\mu\text{M}$ , 24 hours; **E.** 1.0  $\mu\text{M}$ , 24 hours; **F.** Mock control (PBS), 24 hours. Scale bar: 15  $\mu\text{m}$  (upper panels) and 47.6  $\mu\text{m}$  (lower panels) . Insets show magnified cells on a 10  $\mu\text{m}$  scale.

### **2.2.9 Assessing the potential of C5.1 to deliver siRNA in live HepG2 cells with a lectin counterstain**

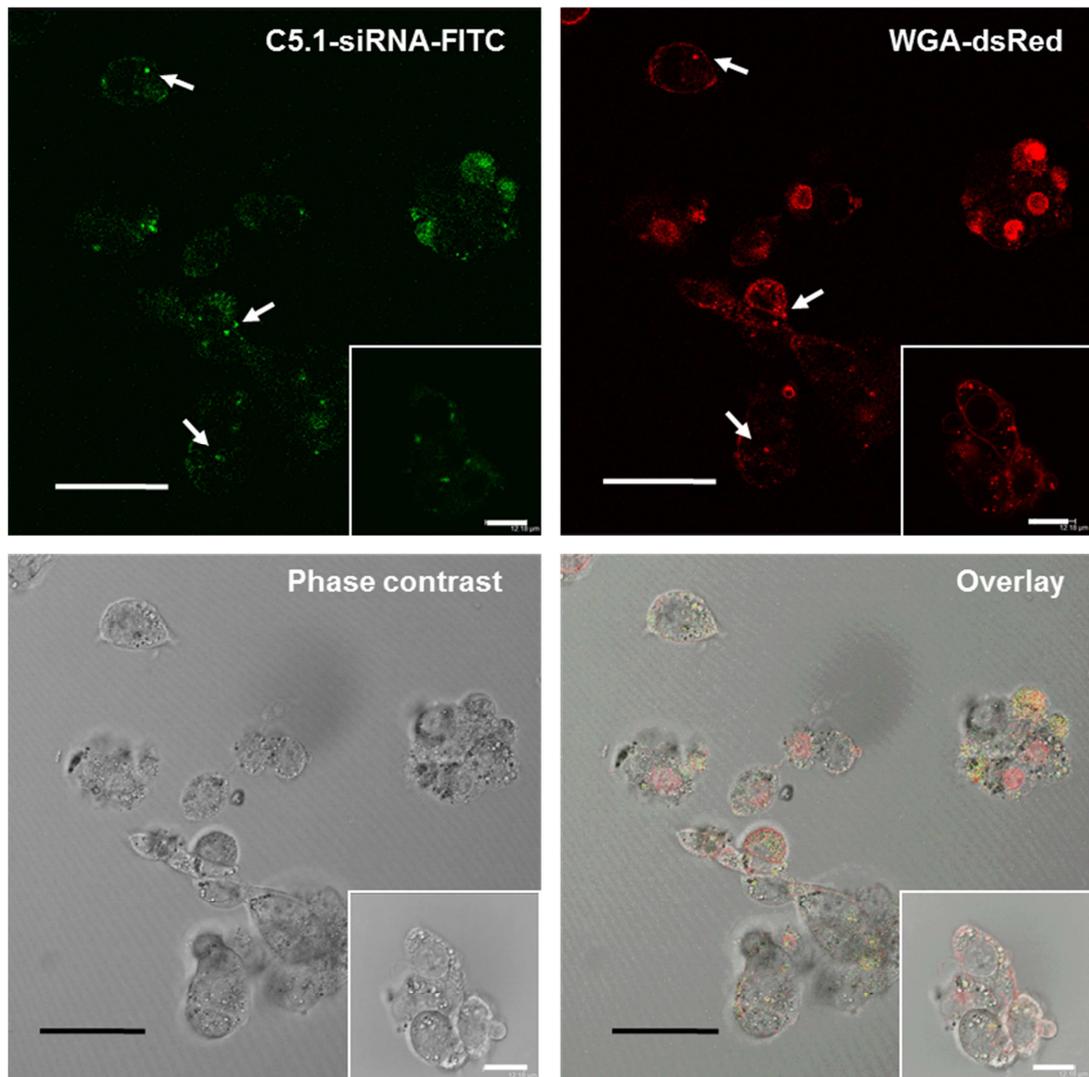
In order to address the issue of potential artefacts in fixed HepG2 cells and the potential of C5.1 to effectively complex with and deliver siRNA intracellularly, a non-specific siRNA oligomer fluorescently labelled with fluorescein (FITC) (Block-iT Fluorescent oligo, Life Technologies) was complexed at a 1:50 molar ratio with C5.1. Cells were incubated with complexes for one hour in the absence of a membrane counterstain (Fig. 2.21) and for two hours with a wheat germ agglutinin (WGA)-dsRed counterstain. WGA-dsRed is a 38 kDa cationic lectin which selectively binds sialic acids and *N*-acetylglucosaminyl residues on the cell membrane and can therefore be used counterstaining live cell localisation assays (Wright *et al.*, 1984a, Wright *et al.*, 1984b). Assuming 98% siRNA encapsulation (van Asbeck *et al.*, 2013), C5.1 localized intracellularly following a one hour incubation with HepG2 cells (Fig. 2.21 A), compared to the siRNA control (Fig. 2.21 B). The somewhat aggregated distribution around the cell membrane necessitated the use of a membrane-specific stain.

The lack of a membrane counterstain was addressed using WGA-dsRed at a final concentration of 1mg/ml for 30 minutes (Fig. 2.22). Confocal analysis of live HepG2 cells indicated that WGA-dsRed localization was not limited to the membrane and that instead, intracellular localisation occurred within distinct vesicular structures, possibly endosomes. Complexed C5.1-siRNA-FITC was observed in the cytoplasmic compartment with a mostly diffuse distribution interspersed with vesicles, thereby demonstrating C5.1-mediated siRNA transduction. Some overlap in the FITC-dsRed signals within cells pointed towards a common uptake pathway.

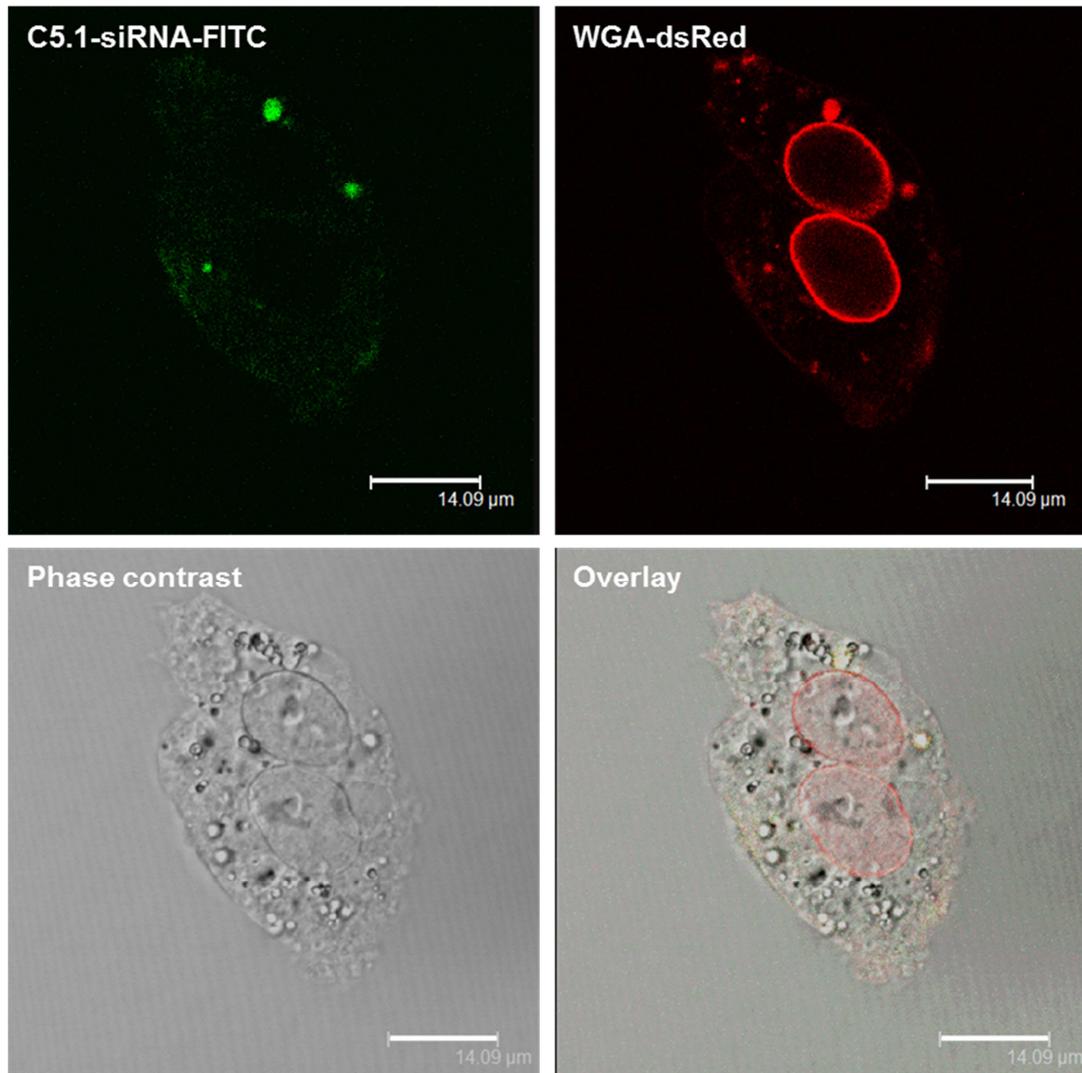
When cells were counterstained with 2.5 mg/ml WGA-dsRed for 10 minutes, the lectin localized primarily in the perinuclear region and in distinct punctae within the cytoplasm (Fig. 2.23). Conversely, C5.1-siRNA uptake resulted in a diffuse cytoplasmic distribution characterised by the absence of endosomal structures, which pointed towards a non-endocytic mechanism of uptake (Fig. 2.23).



**Figure 2.21** C5.1-siRNA localization in live HepG2 cells in the absence of a membrane counterstain. C5.1 was co-incubated with 30 pmol siRNA-fluorescein (FITC) at a 50:1 molar ratio for one hour and applied to HepG2 cells to a final siRNA concentration of 50 nM. **A.** C5.1-siRNA-FITC, **B.** siRNA only. Scale bar, 22  $\mu$ M (upper panels) and 11.7  $\mu$ M (lower panels).



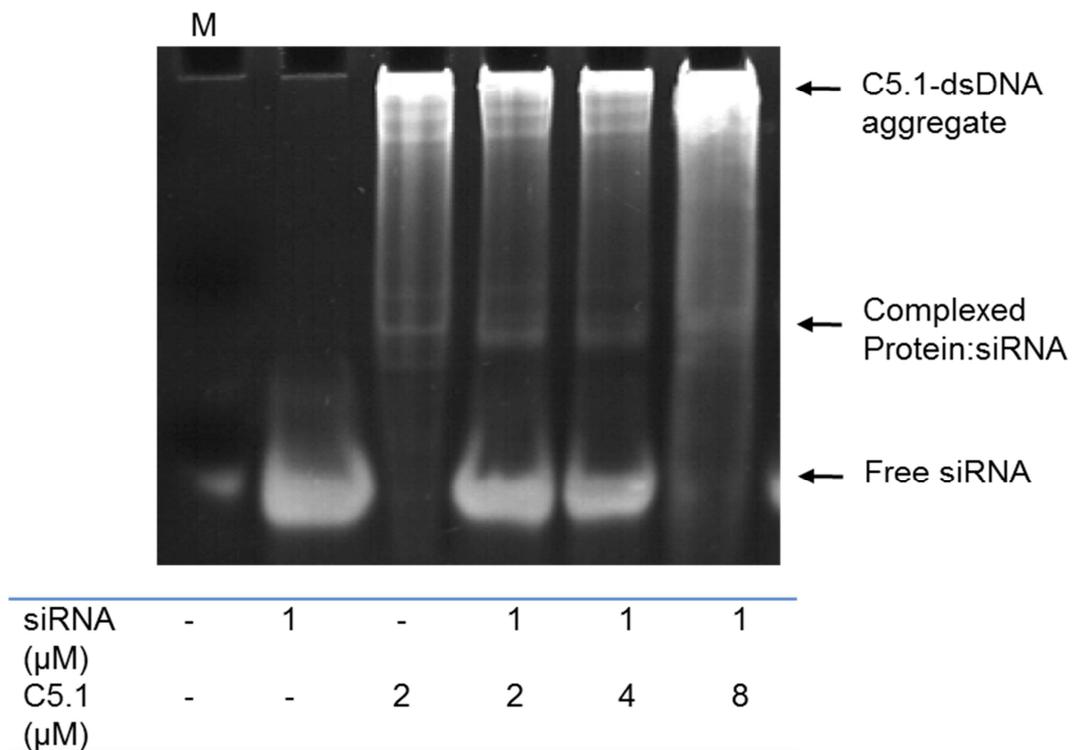
**Figure 2.22 Intracellular localization of C5.1-siRNA in live HepG2 cells counterstained with 1 mg/ml wheat germ agglutinin (WGA)-dsRed for 30 minutes.** C5.1-siRNA complexes were allowed to form by co-incubation and incubated with HepG2 cells for one hour at 37°C. WGA-dsRed, used as a counterstain localizes within distinct vesicular structures. Cells were visualized with a TCS SP2 Leica confocal microscope. Scale bar, 47.2  $\mu$ M and 12  $\mu$ M (insets).



**Figure 2.23** Intracellular localization of C5.1-siRNA-FITC in live HepG2 counter-stained with 2.5 mg/ml wheat germ agglutinin (WGA)-dsRed for 10 minutes. C5.1-siRNA complexes were allowed to form by co-incubation and incubated with HepG2 cells for two hours at 37°C. WGA-dsRed, used as a counterstain, localizes mainly around the nucleus. Cells were visualized with a TCS SP2 Leica confocal microscope. Scale bar, 14 μM.

### **2.2.10 C5.1 shows diminished siRNA binding through the DRBM1**

The transducing abilities of C5.1-myc were assessed *in vitro*, with the protein exhibiting a time-dependent and dose-dependent translocation into HEK293 and HepG2 cells. It seemed essential to also assess the dsRNA binding abilities conferred to the fusion protein by the C-terminal DRBM1 moiety. The Electrophoretic Mobility Shift Assay (EMSA) is a sensitive method to assess protein-nucleic acid interaction by observing a gel shift on either a highly concentrated agarose gel (typically 3-5% w/v) or a dilute native PAGE gel (typically 5-6% w/v polyacrylamide). It is based on the premise that the electrophoretic mobility shift of protein-bound nucleic acid is less than that of free nucleic acid (Hellman and Fried, 2007). Eguchi *et al.* (2009), was able to show that the DRBM1 domain from human PKR in the PTD-DRBD fusion protein could bind 21-23nt dsRNA sequences in a sequence independent manner, with high avidity at a 4:1 ratio. To test whether these results could be replicated with C5.1, excess protein was allowed to form a complex with 1  $\mu$ M siRNA (Stealth siRNA, Invitrogen, Paisley, UK) at various molar ratios, for 30 minutes at 4°C and assessed by EMSA on a 6% native PAGE gel. The gel was stained with Ethidium bromide (EtBr) and analysed under UV light. At a 2:1 and 4:1 and 8:1 molar ratios, only a partial gel shift was observed, indicative of insufficient binding by the DRBM1. The protein-siRNA complexes aggregated at the top of the wells and failed to migrate towards the anode. Fluorescence under UV light was expected only for Ethidium bromide-stained nucleic acid; instead, the negative control protein (C5.1 only) was seen to fluoresce (Fig. 2.24). Coupled with electrophoretic mobility results by Geoghegan *et al.* (2012) and recombinant proteins comprised of the DRBM1 coupled to various CPPs, new constructs, encompassing the tandem DRBMx2 domain were therefore designed and purified.



**Figure 2.24 Electrophoretic mobility shift assay with C5.1 and siRNA at various molar ratios.** Complexes were allowed to form in PBS at 4°C and analysed by polyacrylamide gel electrophoresis with a 6% native PAGE gel in 0.5x TBE buffer. n=3.

## 2.3 Discussion

Utilization of the endogenous RNAi pathway as a post-translational gene modification strategy has been greatly investigated as a routine tool to regulate gene expression. *In vivo*, RNAi has the potential as a therapeutic intervention in genetic diseases, where single nucleotide polymorphisms in mutant allele transcripts exist (Whitehead *et al.*, 2009, Wen and Meng, 2014). Moreover, RNAi has therapeutic potential in viral diseases, such as Hepatitis C, whose single stranded RNA genome that encodes for a polyprotein is an attractive target (Aagaard and Rossi, 2007). The major hurdle in the utilization of RNAi is systemic and targeted delivery, as siRNA's intrinsic physicochemical properties (large anionic charge, susceptibility to degradation) prevent it from crossing the cellular membrane readily, and thus render it a poor pharmaceutical candidate. Viral siRNA vectors are unsuitable for siRNA delivery, due to concerns over immunogenicity and insertional mutagenesis whereas lipid-based strategies *in vivo* have shown rapid liver clearance and lack of tissue specificity. Cell penetrating peptides, such as the Antennapedia homeodomain (AntpHD) and penetratin, represent a class of protein-based delivery vectors, whose low cytotoxicities and versatile cargo-transduction have the potential to be suitable delivery vector alternatives. Although the AntpHD has been investigated as a dsDNA vector to restore p21 expression in p53-mutated cancers, with encouraging results *in vivo* (Kousparou *et al.*, 2012), its potential as a siRNA vector has not been investigated. Penetratin, on the other hand, has been covalently linked to siRNA (Davidson *et al.*, 2004, Muratovska and Eccles, 2004) inducing potent gene knockdown. Prior to this study, non-covalent siRNA conjugation strategies with the AntpHD have not been investigated, whereas this strategy with Penetratin has only been reported by Lundberg *et al.* (2007) and Geoghegan *et al.* (2012) while a seminal study by Eguchi *et al.*

(2009) had induced potent gene knockdown in a variety of adherent and primary cell lines using a recombinant construct comprised of the Tat domain fused to the DRBM1 (PTD-DRBD; Tat-Tat-HA-Tat).

This chapter therefore aimed to investigate whether the AntpHD and Penetratin, expressed as recombinant proteins with the DRBM1 from human PKR, could be used as siRNA vectors, and whether the non-covalent strategy of complex formation would produce similar results to the study by Eguchi *et al.* (2009) in a HEK293 and HepG2 cell lines.

We have chosen to express these vectors as recombinant proteins from a bacterial cell system following heterologous protein induction by IPTG, similar to the methodology adopted by Eguchi *et al.* (2009). In this case, it was important to obtain soluble protein under native conditions by recombinant methods and affinity chromatography, as refolding insoluble structures back to their precise, three-dimensional shape is often a lengthy process and often leads to decreased yields (Arakawa *et al.*, 2007). Affinity tags, such as the GST or poly(His) tags often enhance the solubility of recombinant proteins and aid purification under native or denaturing conditions (Waugh, 2005). Although His<sub>6</sub>-tags have been used to purify proteins at high yields and purities, longer (His<sub>10</sub>) tags have also been used successfully (Grishammer and Tucker, 1997). Constructs obtained either an *N*-terminal or *C*-terminal Poly(His) tag (~1 kDa) (Bornhorst and Falke, 2000) by insertion in a modified pET32a vector, or Glutathione S-transferase (GST) tag (26 kDa) (Smith and Johnson, 1988) by insertion into a pGEX-6P-2 vector, for affinity chromatography purification. The purification methodology adopted allowed the assessment of a) the expression state of proteins when purified under native conditions from 50 ml *E. coli* cultures and b) the relative yield when cultures were upscaled to 500 ml.

Expression tests on 50ml BL21DE3pRARE cultures transformed with pET32-a/insert plasmids indicated that the terminus at which the tag was attached, did not increase the solubility of expressed proteins. Although *N*-terminal poly(His) tags have been reported to enhance soluble protein expression compared to *C*-terminal ones (Busso *et al.*, 2005, Dyson *et al.*, 2004), in this case, both His<sub>10</sub>-AntpHD-DRBM1 (C1.1) and AntpHD-DRBM1-His<sub>6</sub> (C3.1) were expressed as insoluble fractions (as seen in the total fractions, and not in the eluates). The low levels of expression may signify that these proteins are toxic to the host cell, as expression of heterologous genes rich in codons for arginine (R), isoleucine (I), leucine (L) and proline (P), such as C1.1-C4.1, have been implicated in translational stalling, causing protein expression to either fail or proceed at very low levels (Francis and Page, 2001).

Similarly, His<sub>10</sub>-Pen-DRBM1 (C2.1) and Pen-DRBM1-His<sub>6</sub> (C4.1) were mostly present in the flowthrough and total fractions, with minimal amounts of protein successfully eluted from the columns, although higher expression levels may signify a lesser degree of toxicity.

Considering that recombinant proteins  $\leq 60$  kDa are usually expressed as soluble fractions by BL21 *E.coli* strains (Graslund *et al.*, 2008a, Graslund *et al.*, 2008b), construct size may not be the cause of insoluble expression. Francis and Page (2001), brought forward the hypothesis that the proportion of a protein that forms  $\alpha$ -helices has positive effects on refolding and soluble expression. The lack of soluble expression, then, could be explained by bioinformatic analysis of the secondary sequence. Consistent with Francis and Paige's observations, the AntpHD-based proteins (C1.1 and C3.1), adopt mostly a coiled conformation (67.7% and 76.1%, respectively), with only 28% and 20% of each protein in an  $\alpha$ -helical conformation. The presence of disordered coils, may have promoted

aggregation (Francis and Page, 2001). Conversely, C2.1 and C4.1 conformations were predominantly helical (~42%), which may explain the higher levels of protein found in the cleared lysate (flowthrough) fraction. Further studies with C1.1 on a 0.5L scale confirmed that this protein was expressed as an insoluble fraction. On the same scale, C2.1 and C4.1 were successfully purified; however, vertical streaking following analysis by SDS-PAGE signified nucleic acid contamination. This could be attributed to the absence of *DNaseI* in the lysis buffer, which would have digested any unsheared DNA following cell lysis by sonication. The incorporation of a DE52 anion exchange step to remove endogenous nucleic acid would have therefore been beneficial, and was incorporated as a final purification step following GST-tagged protein purification. Furthermore, these results indicated that poly(His)-tagged proteins were expressed as insoluble aggregates and needed to be purified under denaturing conditions. It would be interesting to investigate whether the conformation adopted in solution by these proteins, would have an effect on their uptake efficiency by cells *in vitro*.

The GST-tag (Smith and Johnson, 1988) was added to the *N*-terminus of both constructs by ligation into a pGEX-6P-2 plasmid to yield C5.1 and C6.1. The incorporation of a c-myc epitope tag at the *C*-terminus to yield C5.1-myc and C6.1-myc was done for downstream immunofluorescence studies. Small scale expression testing showed that the GST-tag enhanced soluble expression of C5.1 and C6.1, compared to their His-tagged counterparts, consistent with previous observations by Hammarstrom *et al.* (2002).

The higher expression levels for all GST-tagged constructs also signified that the GST-tag enhanced the expression of fusion proteins and not just their solubility. In 0.5L

purifications, C5.1 was successfully eluted with 80% purity, confirming that in this case, the GST tag promoted solubility and effective purification of these fusion proteins.

The GST-tag was cleaved from pooled eluates by the proteolytic action of Prescission protease (PP). PP is a human rhinovirus 3C protease-GST chimera, which recognizes the –LEVLFG/GP- internal sequence downstream of the tag and cleaves between the glutamate and glycine residues. Tag removal was facilitated while the enzyme was immobilized on the GS4B column in a buffer formulation that lacked reduced glutathione. This necessitated the buffer exchange of purified proteins into an equilibration buffer that did not include reduced glutathione. This process resulted in a step-wise decrease of protein yields, which signified that they were unstable in equilibration buffer. C6.1 and C6.1-myc decreased by 77.2% and 84.1%, respectively, with 0.40 mg as a final yield in PBS/10% glycerol. Step-wise yield calculations indicated that the lengthy process of batch-binding, purification, buffer exchange, proteolytic tag cleavage and a final exchange into PBS caused both proteins to precipitate, possibly as a result of changes in ionic strength, buffer composition and increased handling.

Their lower levels of expression and instability during manipulation, meant that an alternative strategy, such as solid-phase synthesis, would have been advantageous for these proteins at this stage. Indeed, most studies in the last 10 years have used chemically synthesized proteins, and most commonly proteins prepared by solid-phase synthesis, via an Fmoc (9H-fluoren-9-ylmethoxycarbonyl) (El-Andaloussi *et al.*, 2011) or Tbc (tert-butoxycarbonyl) protection strategy (El-Andaloussi *et al.*, 2007b, Jones *et al.*, 2005). This strategy, whereby the peptide chain is sequentially coupled step-wise on a solid support, allows the high-grade purification of proteins without loss of stability or degradation (Hu, 2009) and the covalent conjugation of cargo (Jarver *et al.*, 2012, Coursindel *et al.*, 2012).

It is however, very costly, and synthesis is limited to an average amino acid length of 36 amino acids. This strategy would therefore be beneficial only in the development of constructs based on the shorter, 16 amino acid Penetratin constructs, rather than the longer AntpHD. Another alternative, would have been to purify these constructs by IMAC followed by fast protein liquid chromatography (FPLC) (GE Healthcare) on an Akta purification system, utilizing a linear gradient during buffer exchange, rather than a manual buffer exchange process, similar to the successful methodology adopted by Eguchi *et al.* (2009).

In order to obtain sufficient amounts of all GST-tagged proteins for downstream assays, bacterial cultures were scaled up to 4L for C5.1 and C5.1-myc, and 3L for C6.1 and C6.1-myc. To minimize the steps involved, a batch-binding strategy with Prescission protease was adopted, prior to purification by affinity chromatography. Subsequent concentration resulted in C5.1, C5.1-myc and C6.1-myc precipitating to various extents; only the concentration of C6.1 increased by 51% during this step. Yields decreased in the order of C6.1-myc>C5.1>C6.1>C5.1-myc. Overall, C5.1 decreased by 21.3%, compared to 61.1% previously. C6.1 showed increased stability, whereas C6.1-myc increased by 6.3% and finally, yields for C6.1-myc also showed improvement, with less protein precipitating. These data suggest that a streamlined purification strategy involving batch-binding and on-column cleavage helps minimize protein losses due to precipitation during extended handling, and that the addition of glycerol can aid protein solubility by preventing aggregation.

The presence of the myc epitope was confirmed by immunoblotting with an anti-myc primary antibody. Assessment of protein stability by centrifugation at 16,100 x g and 4°C indicated that C5.1 was relatively stable, with little loss in yield, whilst C5.1-myc yield was almost halved. C6.1 and C6.1-myc on the other hand completely precipitated, which

confirmed that they were unstable in PBS/10% (v/v) glycerol. The intact molecular weights of C5.1 and C5.1-myc were assessed by MALDI-TOF and ESI-QTOF mass spectroscopy, and confirmed that expected molecular weights were in good accordance with actual molecular weights, with small contaminating ions attributed to phosphate groups in the buffer or nucleic acids. As ESI is a 'soft' ionization method, it allows the precise interrogation of con-covalent complexes (Akashi *et al.*, 2005, Sundqvist *et al.*, 2005), such as the ones between Tat and TAR RNA and Tat-TAR RNA complexes and the antibiotic neomycin (Sannes-Lowery *et al.*, 1997). It would therefore have allowed the interrogation of both the protein, and non-covalent protein-nucleic acid complexes with respect to the stoichiometry and the dissociation constants between the protein and endogenous nucleic acid.

Localization studies in fixed HEK293 and HepG2 cells by indirect immunofluorescence utilized the C-terminal c-myc tag of C5.1. Indirect immunofluorescence experiments in fixed HEK293 and HepG2 cells with C5.1-myc, indicated a dose-dependent and time-dependent mechanism of uptake, as well as periplasmic membrane aggregation at high concentrations (10  $\mu\text{M}$ ). Even though a large AntpHD-Rab3 (93 aa) chimera has been successfully used in transduction studies in chick embryo myoblasts (Perez *et al.*, 1992), the distribution of C5.1-myc (161 aa) pointed towards extra-cellular aggregation, indicative of auto-association on the outer leaflet of the cell membrane in HEK293 cells (Eiriksdottir *et al.*, 2010). This observation could be attributed a size-specific molecular weight cut-off point for CPP entry through cell membranes, and may have further implications for the delivery of macromolecular cargo by the homeodomain. The punctate distribution of internalized protein at 0.1  $\mu\text{M}$  and 1.0  $\mu\text{M}$  concentrations support uptake by an endocytic mechanism at both timepoints, consistent with observations by Jiao *et al.*

(2009) whereas at 10  $\mu\text{M}$ , the diffuse distribution throughout the cytoplasm support endocytosis and direct translocation after two hours in HEK293 cells, consistent with MALDI-TOF analysis of penetratin at the same molar concentration in CHO cells (Alves *et al.*, 2011). Diminished uptake after 24 hours indicated that uptake was time-dependent, consistent with findings in other studies with Penetratin by Dom *et al.* (2003), possibly due to protein degradation outside of the protective cellular environment, or sequestration into endosomes and subsequent degradation.

Conversely, in HepG2 cells, C5.1-myc was internalized more efficiently after 24 hours, rather than after two hours. This difference further reinforces the notion that uptake is both time-dependent and cell-type dependent, with endosomal entrapment and aggregation occurring at higher micromolar concentrations. Formation of early endosomes (EE) occurs when molecules span the cell membrane and enter the cytosol, by a primarily endocytic pathway; trapped molecules are then degraded as the endosome matures to form late endosomes (LE), 250-1000 nm in diameter and an intra-lumen pH range of 4.9-6.0 (Huotari and Helenius, 2011), which ultimately leads to degradation.

Although other groups have provided conflicting reports with regards to the molecular mechanisms of cell entry with Penetratin, (Alves *et al.*, 2011, Drin *et al.*, 2003, Duchardt *et al.*, 2007), it is clear that differences in time-dependent transduction efficiencies between cell lines are caused by differences with regards to binding ligands on the cell surface, such as the proportion of sialic acid and proteoglycans present. It would be interesting to quantitate the amounts of C5.1 or C5.1-myc by spectroscopy, as this would provide new insight as to its cell penetrating efficiency and variations between different time-points, concentrations and cell lines, with potential implications as to its use as a transduction vector in different tissues.

Protein detection with a primary antibody against the c-myc epitope and a FITC-labeled secondary antibody has been widely used in the past, producing results with enhanced resolution by laser scanning confocal microscopy. However, indirect immunofluorescence, involving cell fixation with paraformaldehyde and permeabilization with a non-ionic surfactant such as Triton X-100, is a process that has been widely implicated in artefactual redistribution of protein around the plasma membrane as a result of membrane perturbations (Green *et al.*, 2003, Richard *et al.*, 2003)

Subsequently, most current methods of protein detection in living cells involve the covalent linking of fluorescent probes in the form of organic dyes (fluorescein isothiocyanate (FITC, 389 Da), rhodamine (tetramethyl rhodamine isothiocyanate (TRITC, 479 Da) or biological fluorophores, such as GFP (~27 kDa), amongst others (Cazes, 2001, Chalfie *et al.*, 1994). A major drawback of covalently linked probes for fluorimetry lies in their large molecular weight or possible alteration of biochemical characteristics, especially when used in conjunction with bifunctional fusion proteins such as C5.1. The methodology developed in this study therefore attempted to circumvent these issues whilst assessing the siRNA-binding efficiency of the DRBM1 moiety via non-covalent interactions. Complexes were allowed to form by co-incubation at a 50:1 ratio of protein to siRNA-fluorescein (FITC) (Eguchi *et al.*, 2009, Geoghegan *et al.*, 2012) in Optimem and applied to adhered HepG2 cells. This molar ratio has been previously identified as the optimal ratio of ~98% siRNA encapsulation by any CPP, irrespective of charge or peptide mass (van Asbeck *et al.*, 2013). Moreover, HepG2 cells have been reported to express *N*-linked sialic acid on their cell surface, which makes this cell line suitable for live localization assays (Wu *et al.*, 2006) as previous reports have identified this  $\alpha$ -2, 8-polysialic acid as the membrane ligand for AntpHD (Bloch-Gallego *et al.*, 1993, Joliot *et*

*al.*, 1991b). Similarly, wheat germ agglutinin (WGA)-dsRed, a cationic lectin labeled with red fluorescent Alexa Fluor 594, selectively binds *N*-acetylneuraminic and poly-sialic acids on the cell surface and is routinely used as a cell membrane marker. Initially, we wanted to investigate whether WGA-dsRed could be used as a live cell membrane counterstain. Unexpectedly, incubating cells with WGA at 37°C for 30 minutes following C5.1-siRNA treatment for two hours, resulted in cellular uptake into peripheral endosomes, consistent with observations by Raub *et al.* (1990). We were curious then, to assess whether WGA-dsRed uptake following C5.1-siRNA treatment was inhibited, or whether WGA-dsRed competed with C5.1 for the same plasma membrane binding sites, and by consequence, whether this would result in diminished uptake of WGA-dsRed. Similar intracellular localization within endosomal structures were observed, which indicated that a common endocytic uptake mechanism may occur, although the majority of C5.1-siRNA appeared to translocate directly into the cells. Finally, this observation confirmed that C5.1 could effectively bind and transduce siRNA cargo intracellularly at a 50:1 molar excess, although results are only preliminary. Again, this observation would have to be repeated using quantitative methods such as flow cytometry to assess the amounts of internalized protein-siRNA complexes in the absence or presence of WGA-dsRed, or by MALDI-TOF spectrometry, as previously shown by Alves *et al.* (2011). Geoghegan *et al.* (2012) reported that the DRBM1 from human PKR did not bind siRNA and thus did not induce a complete gel shift at 4:1 molar ratios to siRNA, as previously reported by Eguchi *et al.* (2009). His group had achieved complete encapsulation of siRNA with DRBM1 at 4:1 ratios, which in fact had been the result of non-specific interactions of the Tat-Tat-HA-Tat penetrating moiety of PTD-DRBD.

Results obtained using C5.1-siRNA in EMSA assays (n=3) are in agreement to those of Geoghegan *et al*, as siRNA binding by the DRBM1 was incomplete even at 8:1 molar ratios. In this case, aggregate formation at the top of the wells in the presence of EtBr without any nucleic acid bound to the siRNA and failure of C5.1 to induce a gel shift when complexed with siRNA was observed. This finding may be a consequence of the native purification strategy employed, which, in the absence of a *DNase* treatment, resulted in co-purification of endogenous nucleic acid. Consequently, nucleic acid contamination may have saturated the protein and inhibited efficient dsRNA binding, but illustrates how C5.1 can efficiently bind nucleic acid. This observation was not surprising considering that the biological function of the homeoprotein as part of the *Drosophila* homeobox is to bind DNA.

In conclusion, increased aggregation and decreased siRNA binding affinity at low molar ratios render the C5.1 protein an inefficient siRNA delivery vector, at least in this study. In addition to binding siRNA, a vector protein suitable for delivering cargo to the cytoplasm must also facilitate its release from endosomal vesicles that target it for degradation. The fact that C5.1, in lieu of a fusogenic tag, becomes sequestered in endosomes upon uptake, affects the efficient cytoplasmic delivery of siRNA. In light of Geoghegan's findings with Penetratin-DRBM1, proteins comprised of the DRBM1 were deemed insufficient siRNA vectors and new constructs were designed and purified, comprising of the endosomolytic EB1 penetratin analogue, fused to both tandem domains from human PKR. However, the fact that C5.1 localized intracellularly even in the presence of contaminating nucleic acid, leaves open the possibility that C5.1 may be used as dsRNA delivery vector, provided that a streamlined purification strategy is further developed to incorporate a fusogenic tag,

allow the efficient purification from contaminating nucleic acid, and efficiently refold or elute the protein into a suitable buffer that stabilizes it.

# Chapter 3. Characterisation of endosomolytic, cell penetrating, dsRNA-binding fusion proteins

### 3.1 Introduction

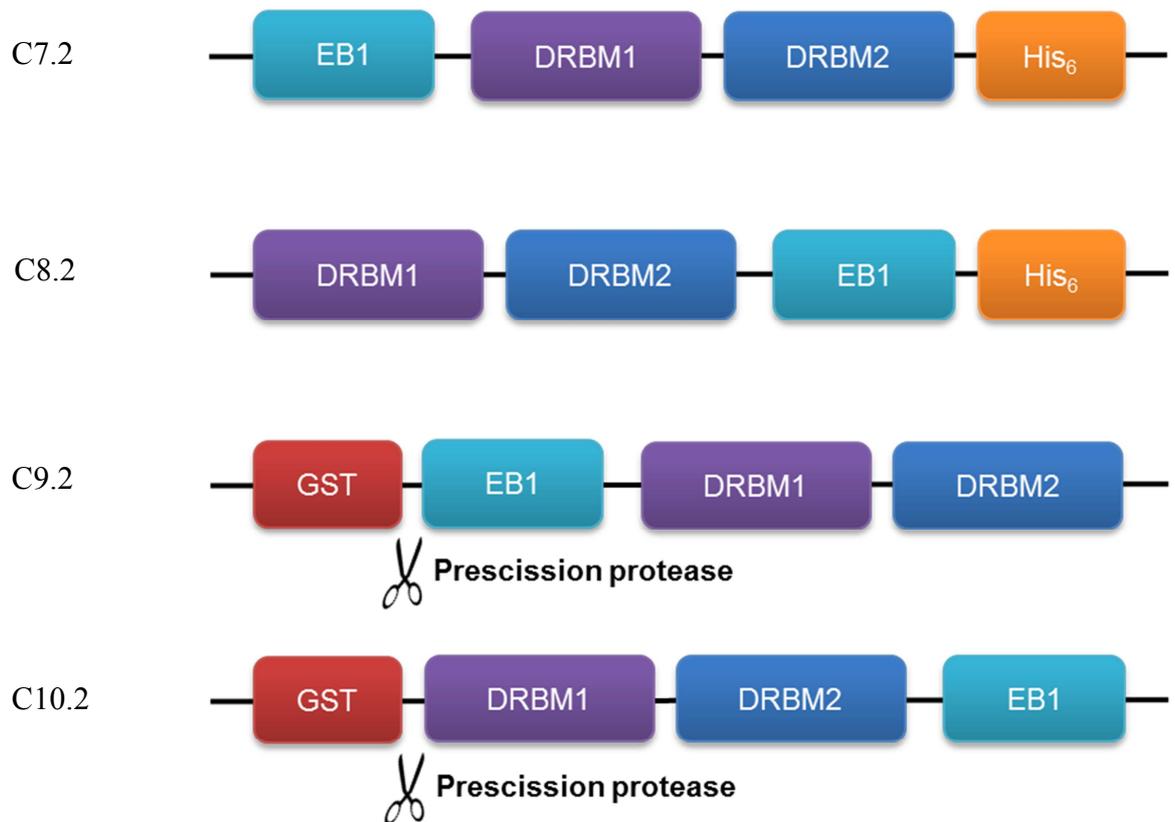
Endosomal entrapment is an intracellular process triggered when uptake favours an endocytic mechanism, rather than direct translocation (Duchardt *et al.*, 2007, Duchardt *et al.*, 2009, Kosuge *et al.*, 2008, Tunnemann *et al.*, 2006, Tunnemann *et al.*, 2008). Cell-penetrating protein uptake via endocytosis has been reported to occur at low micromolar concentrations as a membrane composition-dependent mechanism. Endosomal entrapment and subsequent concomitant degradation in lysosomes, seems to be a limiting step in the delivery of cell penetrating proteins and their attached cargo. Various strategies have been employed to promote endosomal escape, such as the use of sucrose, calcium and chloroquine, or the incorporation of a fusogenic hemagglutinin (HA) tag within the protein sequence (Eguchi *et al.*, 2009); their use however, is severely limited in an *in vivo* setting. Another strategy is the use of the ‘proton sponge’ effect, whereby histidine residues, with an imidazole group and protonatable secondary amine groups with a pKa ~6.0, are incorporated in the primary amino acid sequence to act as a proton sponge in acidic endosomes (Midoux *et al.*, 2009). As the endosomes mature, membrane-bound ATPase pumps recruit cytosolic protons, leading to a decrease in pH. Incorporated histidines resist this acidification. As proton influx increases, so does the influx of water, leading to a buildup of osmotic pressure, subsequent rupture of endosomes and cargo release (Boussif *et al.*, 1995, Pack *et al.*, 2005). Moreover, a protein’s secondary structure is critical for endosomolysis, with an alpha helical structure adopted by primary amphipathic and arginine rich proteins, such as Penetratin, promoting protonation (Biggin and Sansom, 1999, Varkouhi *et al.*, 2011).

An elegant study by Lundberg *et al.* (2007) modified the amino acid sequence of Penetratin, incorporating a short *N*-terminal amino acid sequence that promoted an alpha

helix upon protonation in the endosome and histidine residues to assist in the formation of a proton sponge. This Penetratin analogue, EB1, was then used to non-covalently bind siRNA and induce luciferase gene silencing *in vitro*. It displayed high binding efficiency at low molar ratios, and increased knockdown effects compared to its parent peptide.

In another seminal study by Geoghegan *et al*, recombinant proteins were comprised of a cell penetrating peptide moiety (Penetratin, HIV Tat) fused to the tandem dsRNA binding domains (DRBMx2). These fusion proteins exhibited efficient siRNA binding and gene-specific knockdown *in vitro*. The DRBD (herein DRBMx2) binds siRNA with high avidity ( $K_d = 4 \times 10^{-9}$  mol/L) (Chapter 4 Introduction), compared to DRBM1 or DRBM2 alone (Manche *et al.*, 1992, Patel and Sen, 1992, Schmedt *et al.*, 1995, Tian and Mathews, 2001). In this chapter, a set of siRNA carrier proteins were designed with the aim of incorporating domains capable of binding siRNA with high affinity, traversing the cell membrane and releasing siRNA cargo from endosomes. The endosomolytic EB1 Penetratin analogue developed by Lundberg *et al.* (2007) was therefore expressed as a recombinant protein with the tandem dsRNA binding domain from human PKR (DRBMx2) with either GST or His<sub>6</sub> affinity tags, and purified under either native or denaturing conditions to obtain refolded protein (Fig. 3.1). To assess any effects on binding efficiency or endosomolysis, the EB1 sequence was placed at both the *N*-terminal position of the resulting protein, and the *C*-terminus.

**Construct name:**



**Figure 3.1** Schematic of the EB1-based fusion proteins cloned into a modified pET32-a or pGEX-6P-2 vector. They are comprised of either: EB1, the endosomal analogue of Penetratin; the entire DRBD of human PKR comprised of dsRNA binding motifs 1 and 2 (DRBM1-DRBM2); a C-terminal 6xHistidine tag (His<sub>6</sub>) or an N-terminal Glutathione-S-transferase (GST) tag with a downstream Precission protease cleavage site.

## 3.2 Results

### 3.2.1 Subcloning EB1-DRBMx2 constructs into plasmid vectors

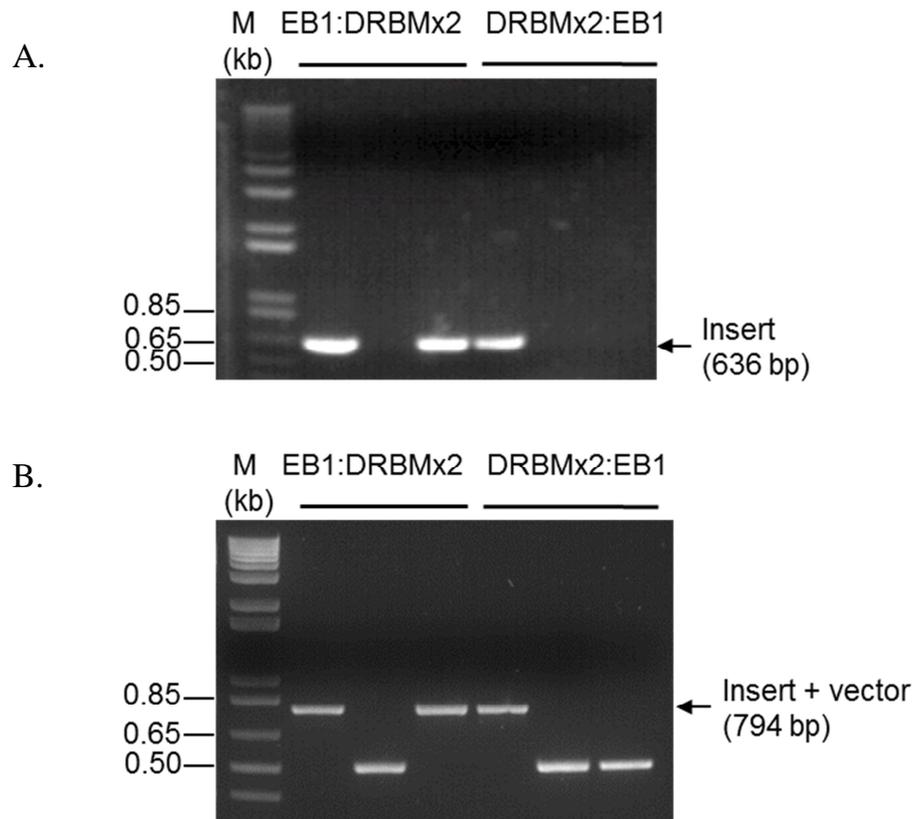
Sequences for EB1-DRBMx2 and DRBMx2-EB1 were obtained by reverse translating the amino acid sequence of EB1 fused to the nucleotide sequence of the entire DRBD, separated by the original linker sequence (G<sub>4</sub>S)<sub>2</sub> in a synthetic plasmid with a pUC backbone (636 bp) (Epoch Biosciences). PCR-amplified fragments were digested with *Bam*HI/*Xho*I and subcloned into pGEX-6P-2 (C9.2 and C10.2); similarly, to obtain a C-terminal His<sub>6</sub> tag, constructs digested with *Nde*I/*Xho*I were subcloned into a modified pET32-a plasmid vector (C7.2 and C8.2) (Fig. 3.1).

The XL1-Blue *E.coli* strain was transformed and colonies were screened directly for inserts by colony PCR. Constructs ligated into pGEX-6P-2 were analysed by colony PCR with insert-specific forward and reverse primers (Fig. 3.2 A) and vector-specific forward and reverse primers, which produced fragments at 794 bp, as expected (Fig. 3.2 B). The slightly larger amplicon size corresponded to the PCR amplification of pGEX-6P-2 vector sequences. Colonies corresponding to the 'insert + vector' bands (794 bp) were sequenced with a pGEX-5' forward primer at GATC Biotech (London) and sequences were analysed with Sequence Scanner v1.0 (Applied Biosystems) and ExPASy Translate (Gasteiger E., 2005). C9.2 and C10.2 sequences that did not contain any mutations were used as templates to replace the *Bam*HI site with an *Nde*I site by PCR.

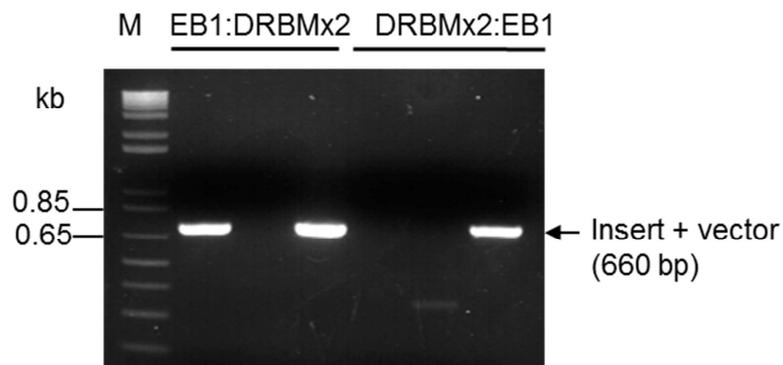
*Nde*I/*Xho*I digested fragments, ligated into a modified pET32-a plasmid vector were used to transform XL1-Blue cells. His-tagged proteins were purified by denaturing Immobilised Metal Affinity Chromatography (IMAC) on a Cobalt chloride (CoCl<sub>2</sub>)

nitrioloacetic (Co-NTA) column, whereas GST-tagged proteins were purified by affinity chromatography on a Glutathione Sepharose 4B (GS4B) matrix.

Direct colony screening by PCR using a T7 (vector-specific) forward primer and a gene-specific reverse primer yielded ~660 bp bands when analysed by agarose gel electrophoresis, which were attributed to primer amplification of a short vector sequence by the forward primer (Fig. 3.3). Screened colonies were sequenced with a T7 minus1 universal primer, as before, and used to transform BL21(DE3)pRARE.



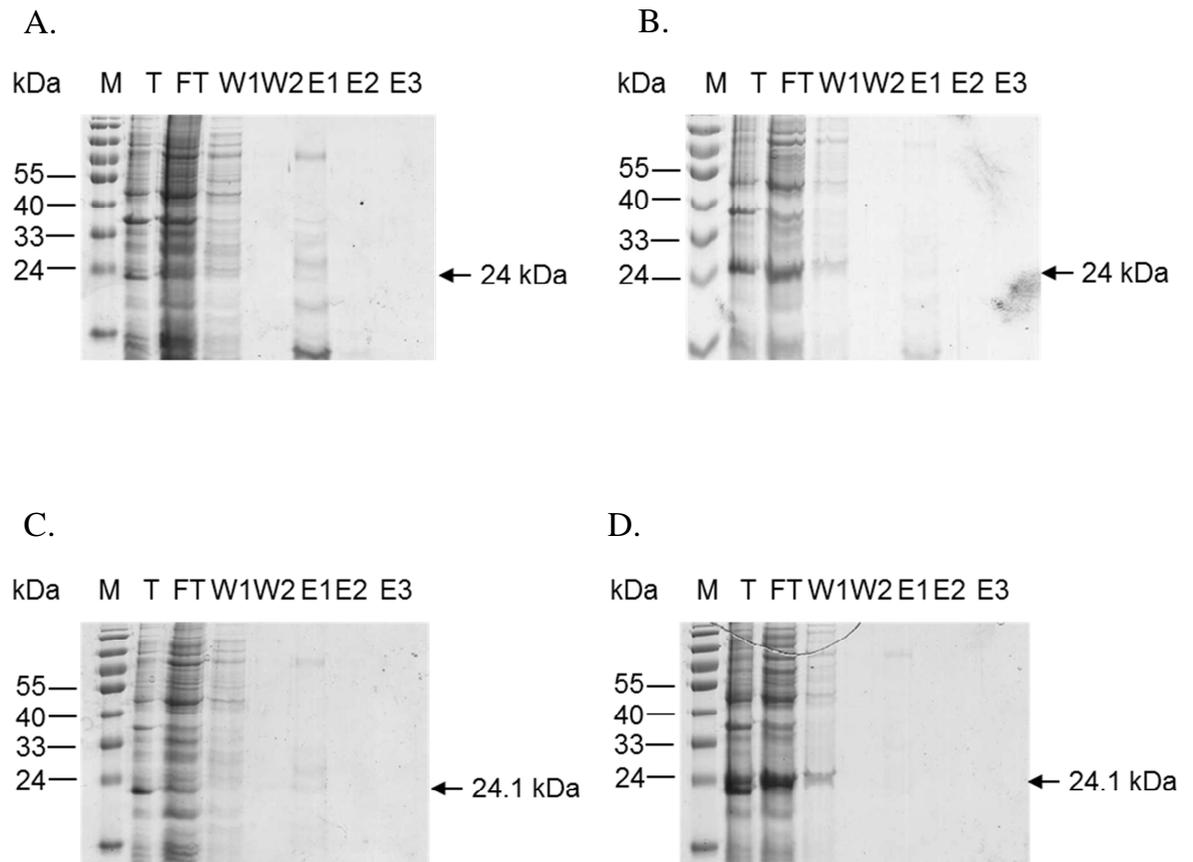
**Figure 3.2** Insert screening by colony PCR following ligation in pGEX-6P-2. **A.** Colony PCR with gene-specific primers. **B.** Colony PCR with vector specific primers. EB1-DRBMx2, C7.2; DRBMx2:EB1, C8.2. Expected size of bands; ~0.8 kb. M, 1 kb marker.



**Figure 3.3** Insert screening by colony PCR following ligation in pET32-a, by a vector-specific forward primer and a gene-specific reverse primer. EB1-DRBMx2, C9.2; DRBMx2:EB1, C10.2. Expected size of bands; ~0.67 kb. M, 1 kb marker.

### **3.2.2 Small-scale expression optimization of His-tagged C7.2 and C8.2**

Since GST-tagged endosomolytic fusion proteins were expressed as insoluble fractions, their His-tagged counterparts were expressed from 50 ml bacterial cultures and purified under native conditions by IMAC on a Cobalt nitriloacetic sepharose matrix (Co-NTA). This provided an estimate of both levels of expression and associated protein solubility, before upscaling to 1 L cultures. Protein expression was induced with 0.5 mM IPTG for 15 hours at either 18°C or 37°C; cultures were pelleted, lysed by sonication, purified and fractions were analysed by SDS-PAGE electrophoresis. Incubation of cultures at 18°C upon induction resulted in low expression of both His<sub>6</sub>-C7.2 and His<sub>6</sub>-C8.2 (Fig. 3.4 A and C, Total fractions), compared to protein levels expressed at 37°C (Fig. 3.4 B and D). Protein was lost in the flowthrough fraction, with no protein eluted from the column after incubations at either temperature, indicating that most of the protein was insoluble. Moreover, vertical streaking in bands from flowthrough fractions (Fig. 3.4 B and D) indicated nucleic acid contamination or degradation. This necessitated purification by denaturing IMAC.



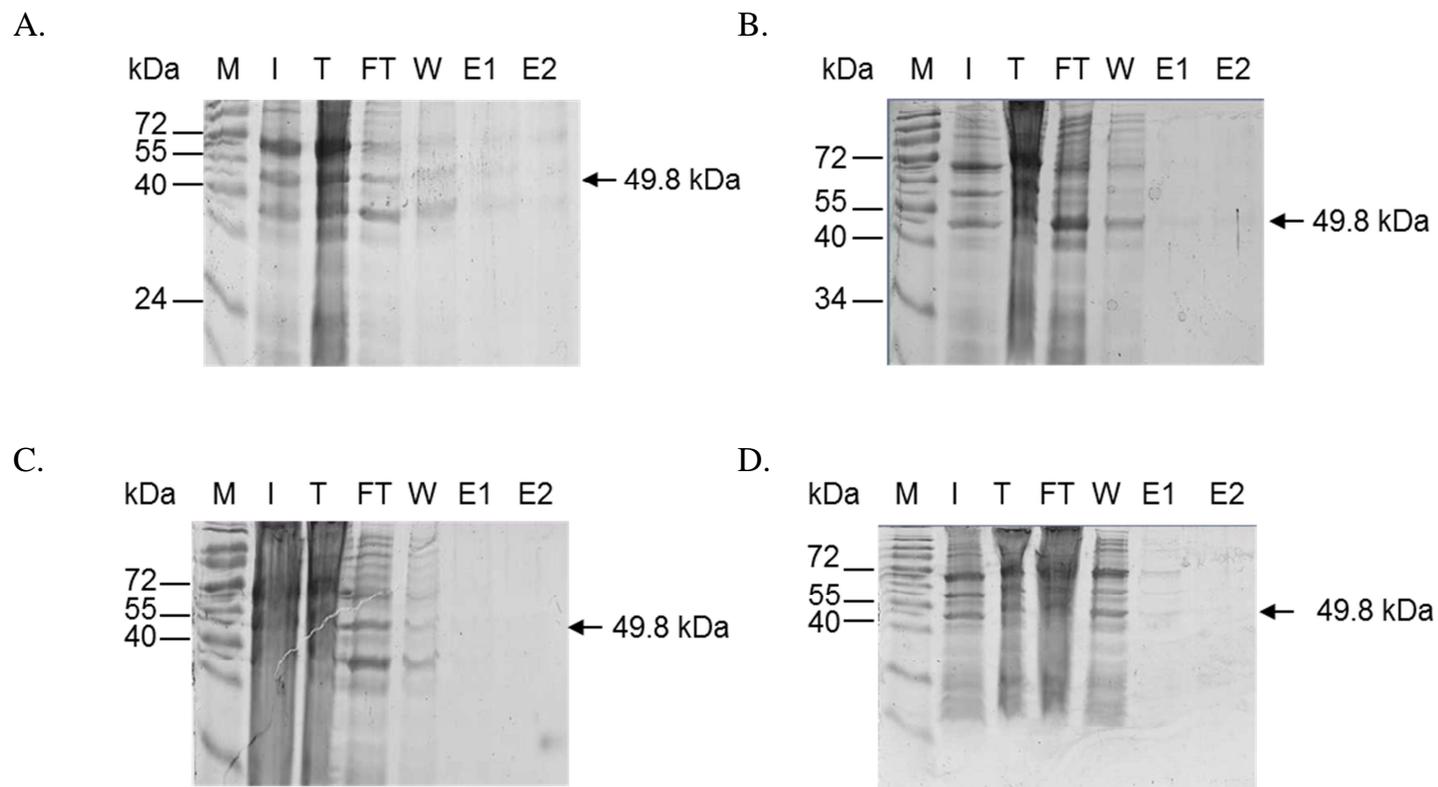
**Figure 3.4 Optimization of His<sub>6</sub>-tagged C7.2 and C8.2 protein expression following incubation at 18°C and 37°C.** Proteins were purified from 50ml cultures by affinity chromatography by immobilized metal ion chromatography over a Co-NTA matrix and analysed on a 12% w/v SDS-PAGE gel. **A.** His<sub>6</sub>-C7.2 incubated at 18°C, **B.** His<sub>6</sub>-C7.2; 37°C, **C.** His<sub>6</sub>-C8.2; 18°C, **D.** His<sub>6</sub>-C8.2; 37°C. Expected molecular weights (MW): His<sub>6</sub>-C7.2, 24 kDa; His<sub>6</sub>-C8.2, 24.1 kDa. T, Total (crude) lysate; FT, Flowthrough (soluble) fraction; W, Wash; E1-E3, Eluates 1-3. M, Marker. Gels were stained with 0.5% Coomassie and visualized under a UVP transilluminator.

### 3.2.3 Small-scale expression optimization of GST-tagged C9.2 and C10.2

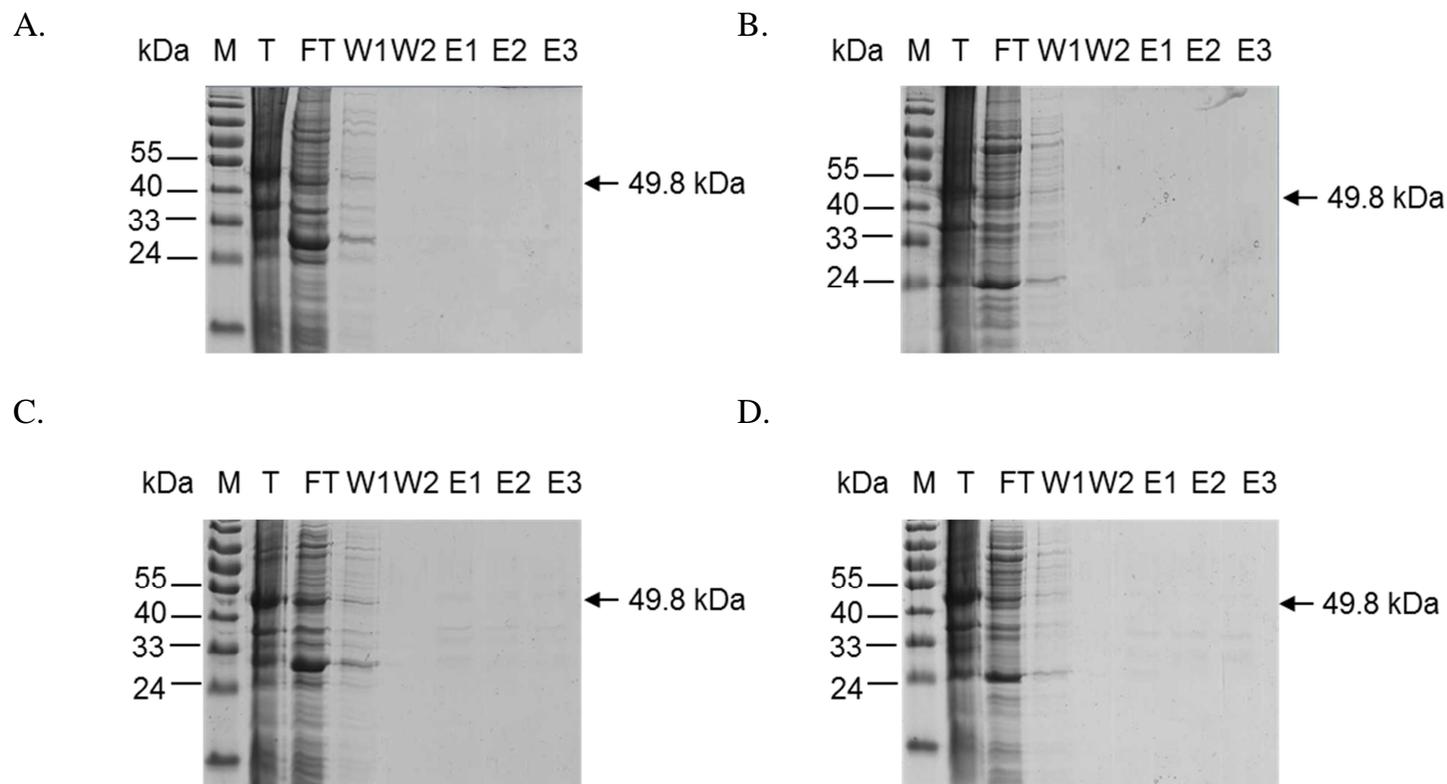
Plasmid DNA from sequence-verified constructs C9.2 and C10.2 in the pGEX-6P-2 vector were used to transform the expression strain BL21(DE3)pRARE. In order to optimise soluble expression from bacterial cultures, two critical conditions for recombinant protein expression were assessed: temperature and Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) concentration. Induction of heterologous protein expression from the lac operon occurs with a 0.1-1.0 mM IPTG concentration (Tolia and Joshua-Tor, 2006). GST-tagged protein expression from 50 ml bacterial cultures was therefore optimised by varying IPTG concentrations at 0.5 mM and 1.0 mM concentrations once O.D<sub>600</sub> had reached ~0.7. Proteins were then purified under native conditions on a GS4B column. IPTG concentration had no effect on the overall levels of C9.2 expression (Fig 3.5 A and B), however, it appeared to increase soluble expression (seen in the flowthrough fractions) at 1.0 mM concentration in comparison to 0.5 mM. No protein was observed in the eluted fractions.

Low expression levels were observed for C10.2 following induction with either 0.5 mM or 1.0 mM IPTG, with protein seen in the flowthrough fractions, but not eluted (Fig. 3.5 C and D). Since protein levels appeared unaffected by the concentration of IPTG, the induction temperature was varied for both proteins. C9.2 and C10.2 were induced with 0.5 mM IPTG and incubated for 15 hours at either 18°C or 37°C, followed by lysate clearing and purification on a glutathione agarose matrix as before. C9.2 incubated at 18°C was expressed at low levels as inclusion bodies; protein was visible in the crude lysate (Fig. 3.6 A, sample 'T'), but not in the flowthrough (FT), wash (W) or any of the eluted fractions (E1-E2). Several contaminating proteins were observed in the FT; similar observations were made when the protein was purified from cultures incubated at 37°C (Fig. 3.6 B).

C10.2 was mostly expressed as insoluble protein, present in the Total (T) fraction, with little protein remaining in the flowthrough (FT) fraction, after incubation at both 18°C and 37°C (Fig 3.6 C and D). Interestingly, a ~28 kDa protein co-purified in the FT fractions of both C9.2 and C10.2 from 18°C (Fig. 3.6 A and C), but a slightly smaller protein (~24 kDa) was evident in the same fractions from 37°C culture purifications (Fig. 3.6 B and D). No proteins were eluted successfully from the glutathione sepharose matrix under any of the expression parameters tested.



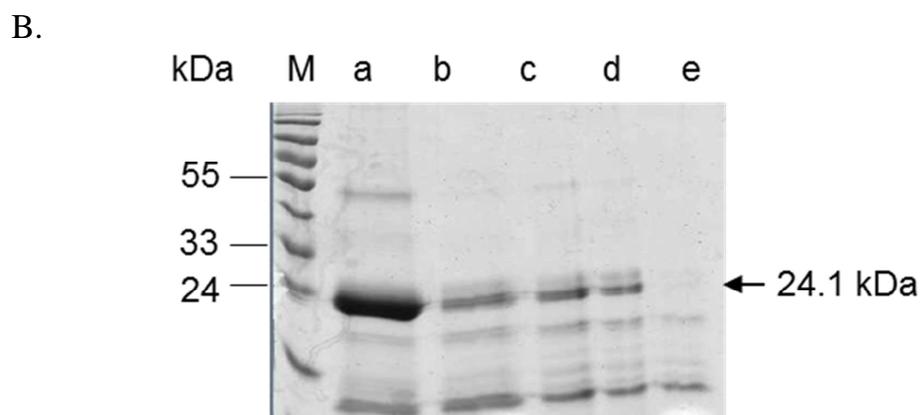
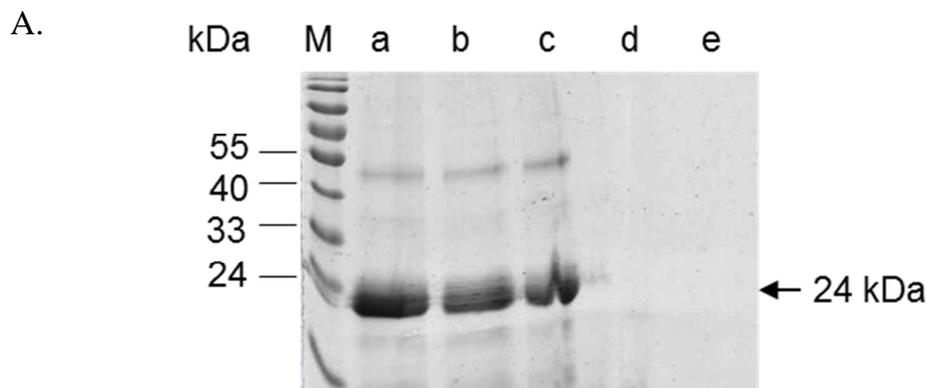
**Figure 3.5 Optimization of GST-tagged C9.2 and C10.2 protein expression following induction with 0.5 mM and 1.0 mM IPTG. Proteins were purified from 50ml cultures by affinity chromatography over a GS4B column and analysed on a 12% w/v SDS-PAGE gel. A. GST-C9.2, induction with 0.5 mM IPTG, B. GST-C9.2; 1.0 mM IPTG, C. GST-C10.2; 0.5 mM IPTG, D. GST-C10.2; 1.0 mM IPTG. Expected molecular weight (MW), 49.8 kDa. I, Induced fraction; T, Total (crude) lysate; FT, Flowthrough (soluble) fraction; W, Wash; E1-E2, Eluates 1-2. M, Marker. Gels were stained with 0.5% (v/v) Coomassie and visualized under a UVP transilluminator. n=1.**



**Figure 3.6 Optimization of GST-tagged C9.2 and C10.2 protein expression following induction incubation at 18°C or 37°C.** Proteins were purified from 50ml cultures by affinity chromatography over a glutathione sepharose (GS4B) matrix and analysed on a 12% w/v SDS-PAGE gel. **A.** GST-C9.2 incubated at 18°C; **B.** GST-C9.2 incubated at 37°C; **C.** GST-C10.2 incubated at 18°C; **D.** GST-C10.2 incubated at 37°C. Expected molecular weight (MW), 49.8 kDa. T, Total (crude) lysate; FT, Flowthrough (soluble) fraction; W, Wash; E1-E3, Eluates 1-3. M, Marker. Gels were stained with 0.5% (v/v) Coomassie and visualized under a UVP transilluminator. n=1.

### 3.2.4 Large scale purifications of C7.2 and C8.2 under denaturing conditions

The expression of C7.2 and C8.2 necessitated purification under denaturing conditions and subsequent refolding. In order to obtain C7.2 and C8.2 in sufficient amounts for *in vitro* assays, proteins were purified from 1L BL21(DE3)pRARE cultures induced with 0.5 mM IPTG at 37°C for 15 hours. Proteins were eluted from the Co-NTA column in His-elution buffer (pH 7.4) containing 6 M Guanidine-HCl (Gu-HCl) and high imidazole (300 mM) and salt concentrations (500mM NaCl). Eluates were then pooled and refolded into 5L refolding buffer (with a lower imidazole concentration, 20 mM) and buffer-exchanged into 5L His-storage buffer. The stability of dialysate fractions at each step of the refolding/buffer exchange process was assessed by centrifugation at 16,100 x g and 4°C and SDS-PAGE. Proteins were successfully eluted from the Co-NTA column, and refolded (Fig. 3.7 A and B; (a)). However, excessive precipitation of C8.2 was observed after centrifugation, with little loss of C7.2 (Fig. 3.7 A and B; (b)). A second dialysis step was included to buffer exchange proteins into His-storage buffer. Visible precipitation within the dialysis tubing prompted the increase of NaCl concentration from 50 mM to 150 mM and the addition of 10% (v/v) glycerol as a stabilizing excipient, which prevented further aggregation (Fig. 3.7 A and B; (c)). Upon centrifugation, there was complete aggregation of C7.2, while C8.2 remained stable (Fig. 3.7 A and B; (d)); however upon final analysis of collected samples, C8.2 appeared to have precipitated, as it did not yield any bands on the gel at the expected molecular weight (Fig. 3.7 A and B; (e)). Final yields for C7.2 and C8.2 were 0.7 mg and 0.9 mg, respectively.



**Figure 3.7 Analysis of purified C7.2 and C8.2 by SDS-PAGE at various stages during refolding and buffer exchange. Fractions from 1L cultures were analysed by 12% w/v SDS-PAGE. A. C7.2; B. C8.2. Lanes (a) purified protein in His-refolding buffer; (b) purified protein in refolding buffer following centrifugation at 16,100 x g; (c) refolded protein following buffer exchange into His-storage buffer/150 mM NaCl/10% (v/v) glycerol; (d) refolded protein in His-storage buffer/150mM NaCl/10% (v/v) glycerol following centrifugation at 16,100 x g and (e) final protein fraction. M, Marker. n=1.**

### 3.3 Discussion

Experimental evidence from a number of studies support the hypothesis that cell-penetrating peptide uptake mechanisms vary between direct translocation and endocytosis depending on protein cargo and the size of complexes (Morris *et al.*, 2001, Simeoni *et al.*, 2003, van Asbeck *et al.*, 2013). Due to the complexity of factors involved in uptake pathways, there is no direct correlation between complex formation ability and intracellular delivery of bioactive siRNA cargo. Endosomal escape is therefore a rate-limiting step for the delivery of macromolecules into cells; the inclusion of a fusogenic moiety has been greatly investigated. Previous studies by Wadia *et al.* (2004), for example, had coupled a fusogenic HA2 tag to a Tat-Cre peptide, which exhibited increased transduction efficiency compared to Tat-Cre alone. A non-covalent strategy utilising a Penetratin analogue with histidine modifications (EB1), managed to deliver siRNA with higher efficiency than Penetratin-HA<sub>2</sub> and Penetratin alone (Lundberg *et al.*, 2007). C5.1, comprised of the Antennapedia homeodomain, lacks the ability to escape from endosomes by virtue of its primary structure. In light of Lundberg's findings, we wanted to investigate whether the Penetratin analogue EB1, expressed as recombinant protein with the entire DRBD (DRBMx2) from human PKR, would show enhanced siRNA binding and transduction efficiencies. Moreover, we aimed to compare its theoretical properties to those of C5.1, which bound siRNA with low affinity, and co-purified with endogenous nucleic acids. This chapter aimed to clone, purify and characterize two novel proteins comprised of EB1 as a fusion protein with DRBM1 and 2 from human PKR (DRBMx2), or its counterpart in tandem (DRBMx2-EB1) as a non-covalent vector for siRNA delivery. Since protein stability is a critical parameter that may only be determined empirically, purification of EB1-based proteins was attempted under both native and denaturing conditions. Two

parameters, IPTG concentration and incubation temperature, were considered. Since the *tac* promoter in pGEX-6P-2 and the T7 promoter in pET32-a are under the control of IPTG, varying its concentration can affect both the rate of expression and their solubility, by displacing the *lac* repressor encoded by the *lacI<sup>q</sup>* gene and inducing the overexpression of protein (Amersham, 2013). Purifications of GST-tagged C9.2 and C10.2 under native conditions from BL21(DE3)pRARE cultures transformed with the pGEX-6P-2-insert plasmids, revealed that both proteins were expressed at low levels, which were unaffected by the concentration of IPTG used to induce heterologous protein production through the *lacI<sup>q</sup>* system. C9.2 was expressed at marginally higher levels with 1.0 mM IPTG, rather than with 0.5 mM. Since a lower IPTG concentration could increase soluble protein production albeit with decreased yield (by allowing the host to fold the protein properly) (Sorensen and Mortensen, 2005), the lower concentration was chosen for induction temperature optimization. Transformed host cells were allowed to grow at 37°C until an optical density (OD<sub>600</sub>) ~0.6-0.7 was attained, which represents the logarithmic growth phase of bacteria in batch cultures, and 0.5 mM IPTG was added to the inoculum. The temperature was either kept constant or decreased to 18°C for 15 hours, inducing protein production through the *tac* promoter. It has been demonstrated previously that temperature is a critical factor affecting soluble protein production (Wang, 2005), with lower temperatures favouring expression into the cytoplasm. The often, direct correlation between aggregation and temperature has been attributed to the partial elimination of heat-shock proteins produced at lower temperatures (Chesshyre and Hipkiss, 1989, Schein, 1989). It was therefore expected that the expression levels of C9.2 and C10.2 would be increased by incubation at 18°C, rather than at 37°C. Observations from 50 ml-scale purifications, did not support this notion; low levels of insoluble protein were observed at

both temperatures assessed for both proteins. This may be attributed to the large molecular weight of both proteins (49.8 kDa); a large scale examination by Dyson *et al.* (2004) of 95 recombinant proteins identified molecular weight as a critical parameter for soluble expression, with larger proteins (~40 kDa) needing a solubility enhancing tag, such as thioredoxin, for soluble expression. A 31.5 kDa co-purifying protein in the FT fraction may correspond to  $\beta$ -lactamase co-purifying, a by-product of expression via IPTG induction (Amersham, 2013).

In summary, since purifications of GST-tagged constructs were unsuccessful, His<sub>6</sub>-tagged C7.2 and C8.2 were purified, first under native conditions and then in the presence of 6 M Guanidinium hydrochloride (Gu-HCl) as a chaotrope.

The addition of Gu-HCl in purification formulations for insoluble proteins, is beneficial in order to solubilize the protein. Gu-HCl must then be removed by dialysis or rapid dilution to allow the protein to return it to its native state (Maxwell *et al.*, 2003). To increase the binding selectivity while decreasing the amount of co-purifying contaminants, nitrilotriacetic agarose was precharged with Co<sup>2+</sup> ions as 0.1 M CoCl<sub>2</sub>, rather than Ni<sup>2+</sup>. Temperature optimization assays at 18°C and 37°C on a 50 ml scale, indicated that both proteins were expressed at higher levels at 37°C, rather than 18°C. Protein was mostly present in the crude lysate (total fraction) with some protein seen in the cleared lysate (flowthrough), signifying that protein must be solubilized prior to purification. Vertical streaking in the SDS-PAGE gel for C7.2 fractions indicated co-purification of nucleic acid or partial degradation, as these preliminary screening tests did not involve *DNase* treatment or protease inhibitors. Since no protein was eluted from the column under native conditions, purifications were attempted under denaturing conditions. The addition of lysozyme, MgCl<sub>2</sub>, benzonase and protease inhibitor cocktail in the His-resuspension and

His-binding buffers was necessary to prevent nucleic acid contamination and proteolytic degradation. Resuspension of the pellet in 1.1x His-binding buffer with a high ionic strength (0.55M NaCl) with 6.66M Gu-HCl, and 2mM  $\beta$ -mercaptoethanol (a potent disulfide bond and RNase denaturant) and subsequent purification by gravity flow affinity chromatography allowed the successful isolation of C7.2 and C8.2 in elution buffer. Sequential dialysis was done to remove any  $\beta$ -mercaptoethanol and Gu-HCl ions from the eluted proteins upon refolding. Proteins were dialysed into His-refolding buffer of high ionic strength (500 mM NaCl) in order to prevent aggregation and 20 mM imidazole, which resulted in C10.2 precipitating. This could be attributed to improper folding, whilst C9.2 remained stable, indicating a greater proportion of properly folded protein.

Both proteins were lost during a second dialysis step into storage buffer. Aggregation at this point could be explained by the decrease of ionic strength (150 mM NaCl), which causes non-specific aggregation and subsequent precipitation of proteins even in the presence of 10% glycerol as a stabilising excipient.

In theory, these novel constructs have great potential as non-covalent vectors of siRNA delivery vectors. The EB1 proteins had exhibited effective siRNA complex formation via EMSA assays at 5:1 molar excess to siRNA, compared to Penetratin, which shifted siRNA at 10:1 molar ratios, using the non-covalent co-incubation strategy as used by Geoghegan *et al.* (2012), Eguchi *et al.* (2009) and Lundberg *et al.* (2007) Moreover, EB1 proteins at 25:1 molar ratios to siRNA *in vitro* showed effective luciferase knockdown following a 36 hour treatment in a transiently transfected HeLa-pGL3 cell line, even in the absence of a dsRNA binding moiety (Lundberg *et al.*, 2007). The siRNA binding properties of EB1 could have been further potentiated by incorporating the DRBMx2. This in turn may have resulted in an effective, non-covalent siRNA carrier with endosomolytic abilities.

The unsuccessful purification of recombinant proteins however, did not allow for downstream analysis; this chapter shows that the purification strategy adopted in this instance, under both native and denaturing conditions, allowed the purification of protein, which was however lost upon dialysis.

Obtaining functional, pure and native-state protein from bacterial expression systems has been found to depend on solubilisation and refolding (Tsumoto *et al.*, 2003). According to Maxwell *et al.* (2003). With one-step dialysis, as was done in this chapter, the concentration of the denaturant is decreased as the rate of refolding into the native step increases; this sometimes results in nonspecific aggregation, even if the protein concentration remains near-constant within the dialysis tubing. The exchange of a buffer containing 6 M Gu-HCl to one with none (0 M Gu-HCl), then, may have forced proteins into a rigid structures, which, if improperly folded, cannot revert back to the native state. It may have been beneficial to adopt a step-wise dialysis process, which would sequentially refold proteins by manual exchange of buffer from one of high denaturant concentration to one of medium and finally to one of low concentration, thereby controlling the equilibrium at each stage (Tsumoto *et al.*, 1998).

The possibility of buffer contamination despite one's best intentions during dialysis is also another factor that affects refolded protein yields, as reported by Maachupalli-Reddy *et al.* (1997); these include nucleic acid and contaminating proteins that may have co-purified along with the target protein. This is consistent with our observations with C7.2 and C8.2 proteins in refolding buffer (Fig. 3.7, lane (a)). Purification optimization to isolate the target protein with 80-95% purity, prior to refolding, is therefore critical to efficient refolding.

Finally, buffer exchange via a size-exclusion chromatography approach that incorporates a flow rate suitable for protein folding kinetics may be a suitable alternative to dialysis. The column matrix (such as a G25 column), equilibrated in an intermediate denaturant concentration buffer, may prevent the formation aggregated species, via hydrogen or hydrophobic interactions between the column and the refolded protein. This approach has been successfully used for the purification of interleukin-6 (IL-6), and is primarily beneficial for proteins with disulphide bonds (Cys-Cys bonds) (Ejima *et al.*, 1999).

Chapter 4. Characterisation of fusion  
proteins containing both dsRNA  
binding domains from PKR.

## 4.1 Introduction

The sequence-independent, efficient dsRNA binding properties of human PKR are conferred by its regulatory binding domain (DRBD). The domain is comprised of two tandem motifs, DRBM1(amino acids 6-79) and DRBM2 (amino acids 96-169), which adopt a dumb-bell shape upon interaction with dsRNA; this allows DRBM2 to ‘release’ the kinase domain, shifting the entire PKR to an open conformation and inducing PKR autophosphorylation and activation (Nanduri *et al.*, 1998, Nanduri *et al.*, 2000).

Although both DRBMs bind dsRNA, DRBM1 ( $K_d = 2.8 \times 10^{-7}$  M) has a higher affinity compared to DRBM2, potentiated by the cooperative action of both tandem motifs ( $K_d = 4 \times 10^{-9}$  M) (Patel and Sen, 1992, Schmedt *et al.*, 1995, Tian and Mathews, 2001).

In 2012, Geoghegan *et al.*, reported that cell penetrating protein fusions with the dsRNA binding motif 1 (DRBM1) from human PKR bind siRNA with low affinity, requiring a 20-60-fold molar excess for *in vitro* knockdown as opposed to the 4:1 ratio as previously reported with PTD-DRBD (Tat-Tat-HA-Tat-DRBM1) (Eguchi *et al.*, 2009). The 4:1 ratio reported was the result of non-specific interactions of Tat with siRNA. The group purified several constructs comprised of a cell penetrating moiety and both dsRNA binding domains (DRBMx2) and used them to assess siRNA binding via electrophoretic mobility shift assays (EMSA) and gene knockdown.

Although the utilization of the DRBMx2 yielded stable and specific siRNA binding compared to DRBM1 alone, these fusion proteins (Tat-DRBMx2 and Pen-DRBMx2) failed to mediate gene-specific knockdown when targeting the endogenous gene hypoxanthine-guanine phosphoribosyl transferase (HPRT) in HeLa cells. This lack of gene expression knockdown was attributed to the endosomal entrapment of the protein-siRNA

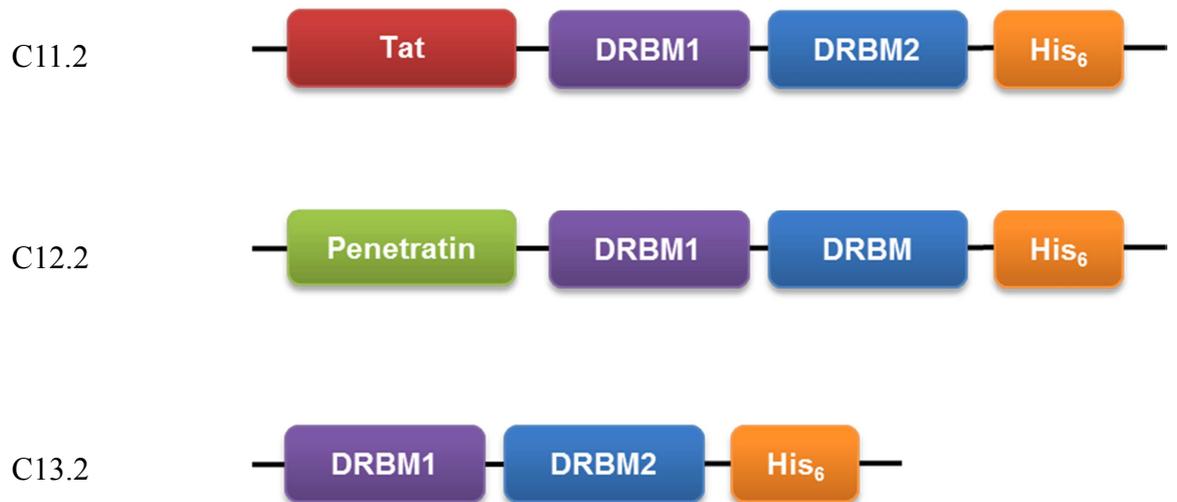
cargo, even though efficient complex formation and transduction *in vitro* was observed. Pen-DRBMx2, yielded 60% HPRT knockdown in the presence of chloroquine, an endosomolytic agent (El-Sayed *et al.*, 2009, Xiao and Samulski, 2012).

Considering that the binding efficiency of Pen-DRBMx2 had not been investigated by the group via EMSA assays, we set out to investigate the molar ratio necessary to induce a gel shift. Moreover, the transduction efficiency and cytotoxicity of Penetratin and Tat have been reported to be cell-line specific (Sugita *et al.*, 2008). Although Penetratin exhibits enhanced cellular transduction in the HeLa cell line compared to Tat (Mueller *et al.*, 2008), it has also been reported to induce cytotoxicity (Sugita *et al.*, 2008). With these findings in mind, we set out to investigate whether Pen-DRBMx2 and Tat-DRBMx2 may be more effective in a HepG2 cell line, targeting the endogenous expression of PTP1B, and in a HEK293-dEGFP reporter cell line. Previous electrophoretic mobility shift assay experiments with C5.1, comprised of the Antennapedia homeodomain (AntpHD) and the dsRNA binding motif 1 from human PKR, revealed that C5.1 co-purified with endogenous nucleic acid, and that at the molar ratios tested, dsRNA was not complexed by the protein as previously reported (Eguchi *et al.*, 2009). These observations were in good accordance with those made by Geoghegan *et al.* with DRBM1.

Plasmids encoding for His<sub>6</sub>-Tat-dsRNA binding motifs 1 and 2 (DRBMx2) (C11.2), His<sub>6</sub>-Penetratin-DRBMx2 (C12.2) and His<sub>6</sub>-DRBMx2 (C13.2) were obtained (a kind gift from Professor Davidson, University of Iowa) (Fig 4.1). The siRNA oligos used in this study were also longer and lacked 2'-OH overhangs, compared to those used in by Geoghegan *et al.*

Plasmids were used for protein expression studies from BL21(DE3)pRARE cultures and purifications under denaturing conditions, followed by refolding by dialysis. Purified proteins were used to assess cytotoxicity in HEK293 and HepG2 cells. In order to quantitatively assess uptake and gene knockdown by flow cytometry analysis, a HEK293 reporter cell line stably expressing destabilised enhanced green fluorescent protein (dEGFP) was created. The pd1EGFP-N1 plasmid (Clontech, Saint-Germain-en-Laye, France), used to transfect HEK293 cells, encodes the destabilised variant of the wild-type GFP protein from *Aequorea victoria* fused to a mouse ornithine decarboxylase (MODC<sup>422-461</sup>) with a PEST amino acid sequence which destabilises the protein, with a half-life of ~2 hours, allowing knockdown analysis at 24 or 48 hours after treatment. It also encodes for a neomycin resistance gene, allowing the specific selection of dEGFP-expressing cells from a polyclonal culture.

**Construct Name:**



**Figure 4.1** Schematic of fusion proteins comprised of the HIV Tat; Penetratin; DRBM1 and DRBM2; His<sub>6</sub> affinity tag. C11.2, Tat-DRBMx2-His<sub>6</sub>; C12.2, Pen-DRBMx2-His<sub>6</sub>; C13.2, DRBMx2-His<sub>6</sub>.

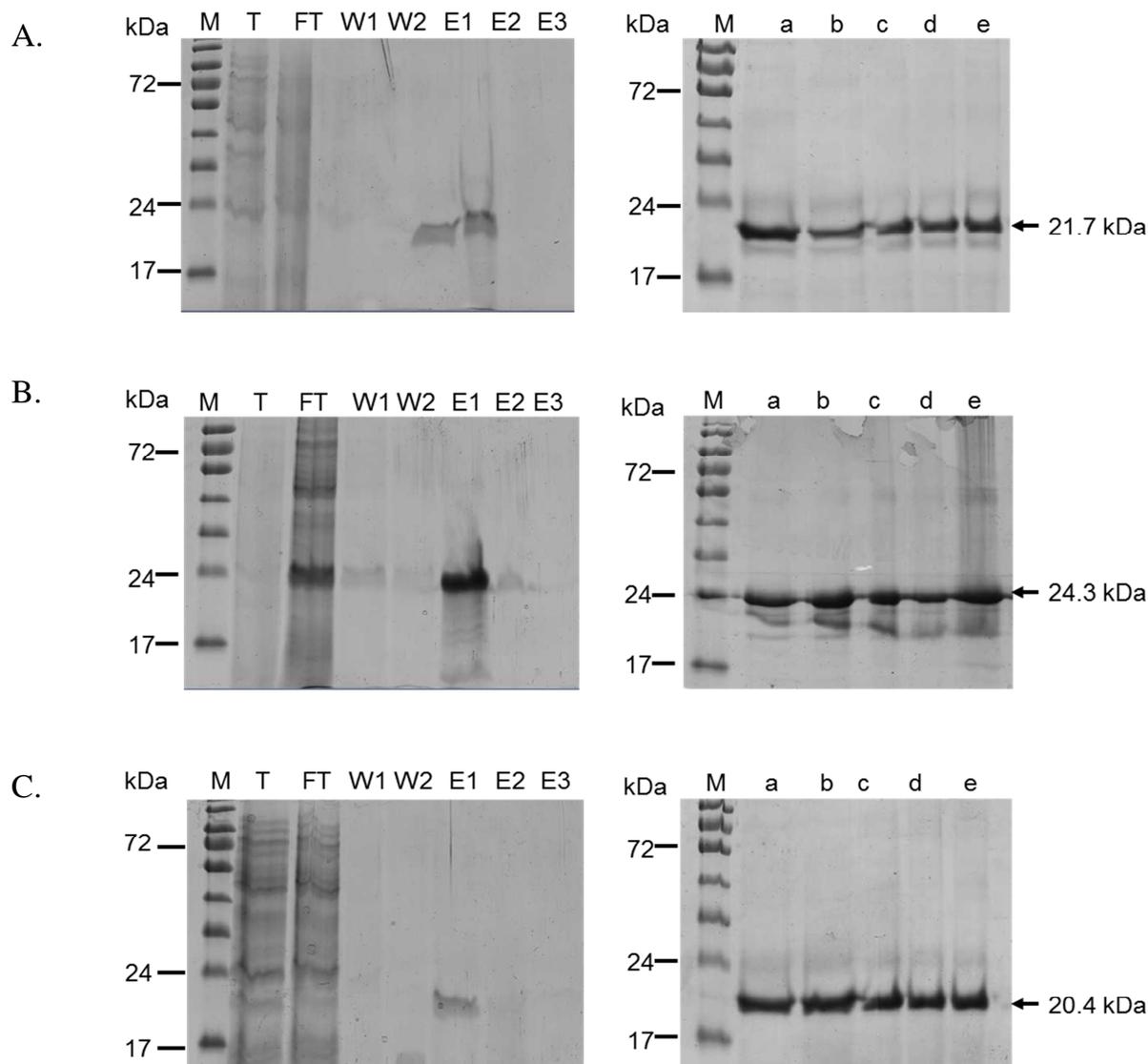
## 4.2 Results

### 4.2.1 Purification and Refolding of C11.2, C12.2 and C13.2

Recombinant C11.2, C12.2 and C13.2 were purified under denaturing conditions, as purified by Geoghegan *et al.* (2012) with some modifications with regards to the refolding strategy. C11.2 (Fig. 4.2 A) was successfully eluted in the E1-E2 fractions, while C12.2 and C13.2 were present in the first eluate only (E1), with little protein loss in the FT fractions (Fig 4.2 B-C). Eluates were sequentially dialysed first into 5L His-refolding buffer to eliminate  $\text{Gu}^+$  ions and then into 5L His-storage buffer in order to refold the proteins into their native state. Stability was assessed by centrifugation (Fig 4.2). All three proteins were isolated with ~90% purities (Fig. 4.2 A, C) except C12.2 (Penetratin-DRBM2), which showed some degradation or contamination with a smaller co-purifying protein (Fig. 4.2 B). Concentrations of final fractions in storage buffer were analysed by Nanodrop at  $A_{280}$  using the expected molecular weights (MW, kDa) and extinction coefficients ( $\epsilon$ ) (Appendix I). The final yields for C11.2 (expected molecular weight (MW): 21.7 kDa), C12.2 (expected MW: 24.3kDa) and 10.2 (expected MW: 20.4 kDa) were isolated with yields of 0.9 mg, 4.3 mg and 1.5 mg, respectively (Table 4.1).

**Table 4.1 Final yields for purified proteins C11.2-C13.2 from 0.5 L cultures.** Yields were calculated following purification under denaturing conditions and refolding by sequential dialysis.

Protein	Protein concentration (mg/ml)	Final yield of refolded proteins from 0.5L purifications (mg)
C11.2	0.2	0.9
C12.2	0.7	4.3
C13.2	0.2	1.5



**Figure 4.2 Purification of C11.2, C12.2 and C13.2 from 0.5L bacterial cultures under denaturing conditions.** Left panels, purification fractions under denaturing conditions; right panels, eluted fractions at various stages of the refolding process. **A.** C11.2; **B.** C12.2; **C.** C13.2. Samples were analysed on a 12% w/v SDS-PAGE gel. Purification fractions analysed; T, Total; FT, Flowthrough; W1-W2, Washes 1-2; E1-E3, Eluates 1-3 and dialysate fractions; (a) pre-spin, in refolding buffer; (b) post-spin at 16,100 x g, in refolding buffer; (c) pre-spin, in storage buffer; (d) pre-spin, in storage buffer /10% (v/v) glycerol and (e) post-spin at 16,100 x g, in storage buffer /10% (v/v) glycerol. Arrows indicate expected size in kDa. M, Marker.

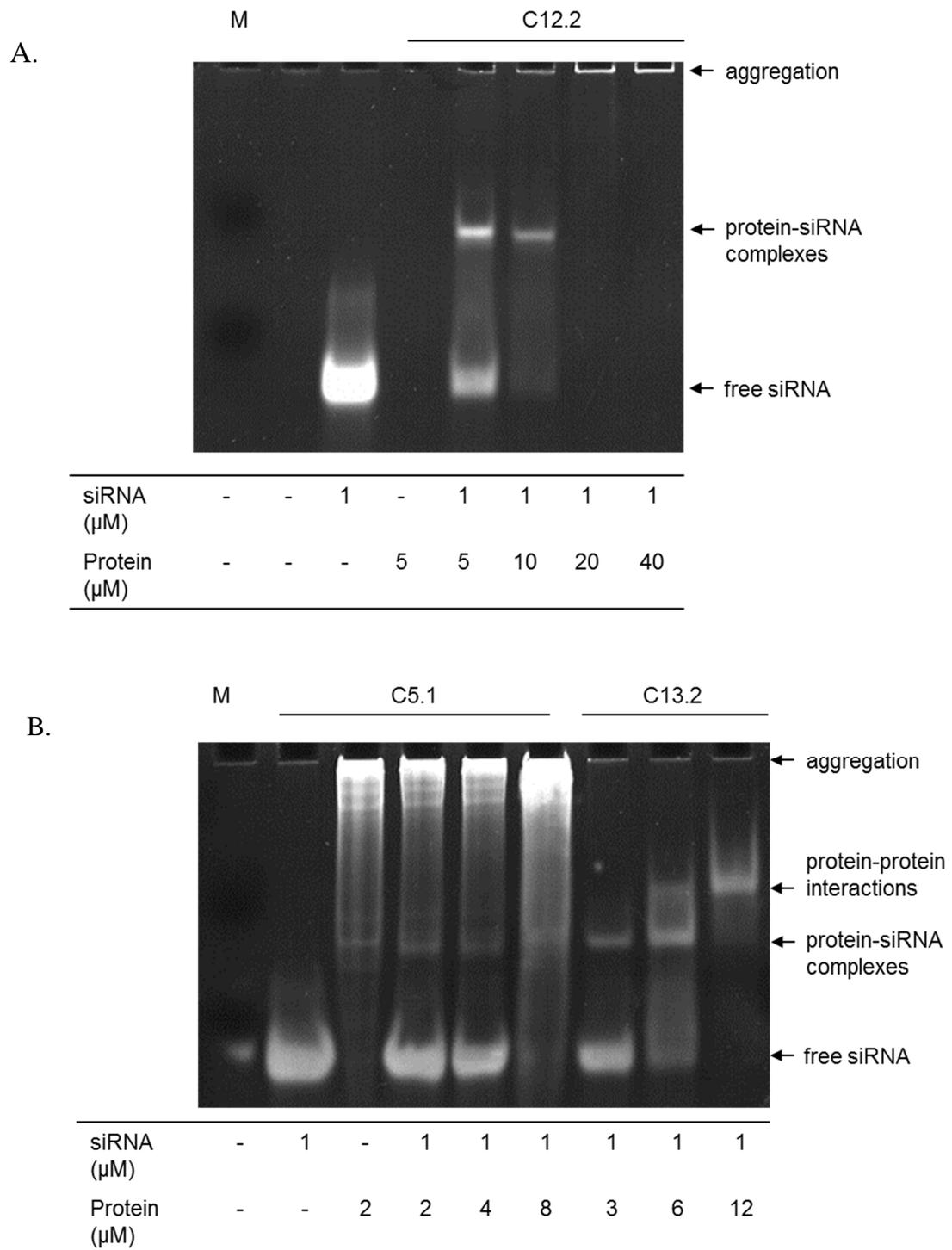
## 4.2.2 Functional analysis of purified C11.2, C12.2 and C13.2

### 4.2.2.1 Assessment of siRNA binding abilities by EMSA

The molar concentrations required by C12.2 and C13.2 for effective siRNA complex formation, were assessed by EMSA and compared to the molar ratios tested with C5.1 and dsRNA. Prior analysis by Geoghegan *et al.* (2012) had revealed that Tat-DRBMx2 (C11.2) bound siRNA with lower affinity than PTD-DRBD. In our case, these results could not be replicated due to the low molar concentration of purified C11.2 (3.65  $\mu$ M).

Excess C12.2, upon complex formation with siRNA at 5:1, 10:1, 20:1 and 40:1, produced a gel shift at the 5:1 molar ratio, with some free siRNA remaining. At a 10:1 molar ratio, a complete gel shift was observed, followed by protein aggregation and minimal migration at the 20:1 and 40:1 molar ratios (Fig 4.3 A).

C5.1, C12.2 and C13.2 in molar excess were allowed to complex with 1.0  $\mu$ M siRNA at various molar ratios. C13.2 showed a partial gel shift at a 3:1 molar ratio, indicating incomplete binding by the C13.2 protein. At a 6:1 molar ratio, further complex formation was observed, as indicated by the increased 'shift' intensity on the gel. A third shift, or 'supershift' was also observed at this molar ratio. Unexpectedly, a 'supershift' band was also evident at a 12:1 molar ratio; free siRNA was not observed on the gel. C13.2 was therefore able to bind dsRNA with high specificity and avidity at both 6:1 and 12:1 molar ratios. By comparison, C5.1 showed nucleic acid contamination and aggregation (Fig 4.3 B).



**Figure 4.3 Electrophoretic mobility shift assays with C12.2 (A) or C5.1 and C13.2 (B) complexed with siRNA at various molar ratios. A.** C12.2-siRNA. **B.** Comparison of the DRBM1 (C5.1) and DRBMx2 (C13.2). Proteins and siRNA were incubated at 4°C for 30 minutes in PBS and analysed by polyacrylamide gel electrophoresis on a 6% (w/v) native PAGE gel in 0.5x TBE buffer. The third ‘supershift’ band seen with C13.2 indicates protein-protein interactions at higher molar ratios.

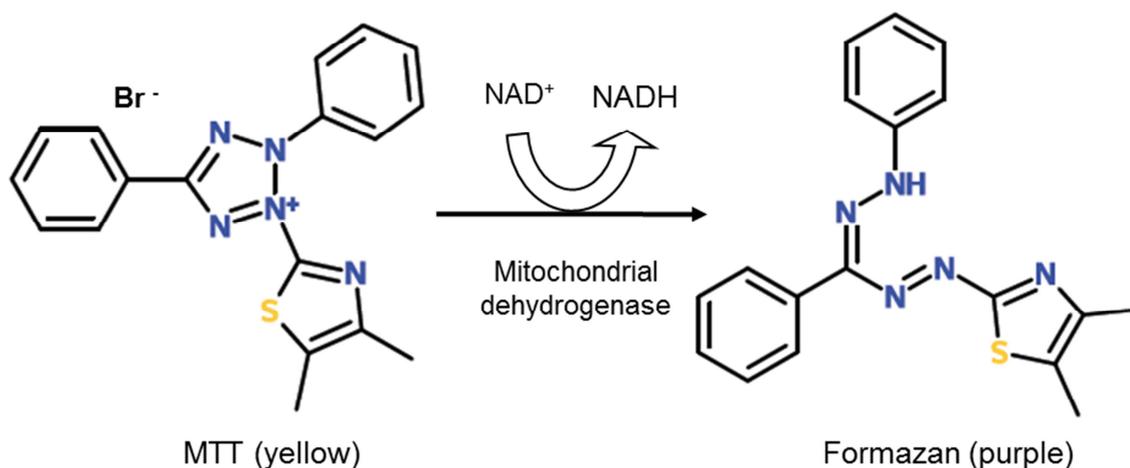
#### **4.2.2.2 Viability assays with purified C5.1, C11.2, C12.2 and C13.2 in HEK293 and HepG2 cells.**

Having investigated the siRNA-binding capacities of C5.1, C12.2 and C13.2 by EMSA, the potential for cytotoxicity or other effects on cell viability were investigated via the MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assay in HEK293 and HepG2 cells (Mosmann, 1983).

Tetrazolium salts are composed of a positively charged tetrazole ring core with four nitrogen atoms, surrounded by three aromatic phenyl groups and counterbalanced with one negatively charged sulfonate group (Fig. 4.4) (Berridge *et al.*, 2005). The tetrazole ring is reduced by NAD(P)H-dependent oxidoreductases in the mitochondria of metabolically active (and therefore viable) cells following uptake via changes in the plasma membrane potential (Bernas and Dobrucki, 2002, Bernas and Dobrucki, 2000), forming brightly coloured formazan crystals. Cell viability, can thus be directly assessed by dissolving formazan crystals in an organic solvent such as DMSO, and measuring the absorbance at 540-570nm (Mo *et al.*, 2012).

Since the reduction of MTT to formazan correlates with metabolic activity in actively dividing cells, decreased  $A_{570}$  is expected in non-dividing, or quiescent cells. HEK293 cells were seeded in 96-well microplates and allowed to adhere prior to treatment with 0.1, 0.3, 1.0 or 3.3  $\mu\text{M}$  C5.1, C11.2, C12.2, C13.2 or mock-treated control for 24 hours in low serum media (Optimem). Since most knockdown assays *in vitro* are performed with siRNA at a final concentration of 50 nM and a protein:siRNA ratio of 50:1, this concentration range provided some insight as to any potential cytotoxicity effects prior to loading with siRNA. Cytotoxicity was assessed by adding MTT to a final concentration of 0.45 mg/ml

and measuring absorbance at  $A_{570}$ . Background signal was measured  $A_{640}$ , and values were subtracted from their respective  $A_{570}$  values to remove background fluorescence caused by non-specific interactions between the dye and media (Riss TL, 2013). Results shown represent average % viability compared to mock-treated control (Optimem).



**Figure 4.4** Schematic of the chemical reaction of MTT reduction to formazan, catalyzed by mitochondrial NADH oxidoreductase. MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide; NADH, NADH dehydrogenase;  $\text{NAD}^+$ , oxidised dehydrogenase.  $\text{Br}^-$ , Bromide ions.

In HEK293 (Fig. 4.5 A), C5.1, C11.2, C12.2 and C13.2 increased cell viability, or at least mitochondrial activity at 0.1, 0.3 and 1.0  $\mu\text{M}$  concentrations, compared to the mock-treated control. At 3.3  $\mu\text{M}$ , C11.2 exhibited dose-dependent cytotoxicity, with only ~60% cell viability following a 24 hour treatment. However, pairwise comparisons between the 0.1, 0.3 and 1  $\mu\text{M}$  concentrations against the 3.3  $\mu\text{M}$  concentration of C12.2 were not significant ( $P > 0.05$ ), indicating that C12.2 does not elicit dose-dependent cytotoxicity at the concentrations tested. For C11.2, similar pairwise comparisons between concentrations were significant ( $P < 0.02$ ), indicating that changes in cell viability are dose-dependent. Similarly, C13.3 showed dose-dependent changes between the 0.1, 0.3 and 3.3  $\mu\text{M}$

concentrations, but not between the 1.0 and 3.3  $\mu\text{M}$  concentrations. Differences in cell viability due to protein concentration were not significant for C5.1 (Appendix II). Taken together, these results indicate that at the concentration range tested, cytotoxicity increases in the order of  $\text{C12.2} \leq \text{C5.1} < \text{C13.2} < \text{C11.2}$  in HEK293 cells ( $*p < 0.05$ ) (Fig 4.5 A). Results represent three independent experiments ( $n=3$ ). Since no cytotoxicity was evident in HEK293 cells at the lowest molar concentration, it was omitted from HepG2 viability assays.

Similar effects on cell viability were observed in HepG2 cells treated with 0.1-3.3  $\mu\text{M}$  protein, with C11.2 eliciting the highest cytotoxicity at all three concentrations when compared to C12.2 (Fig. 4.5 B,  $^+p < 0.05$ ). The increase in cell viability observed with 1.0  $\mu\text{M}$  treatment with C11.2, C12.2 and C13.2 was not significant at  $P \leq 0.05$ .

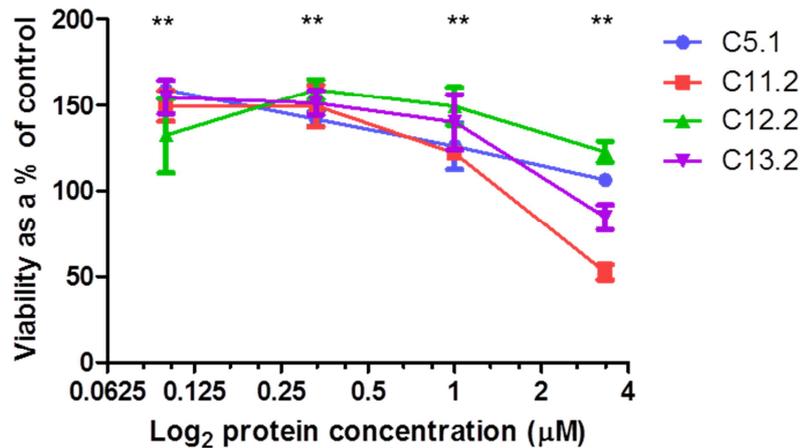
Pairwise comparisons between the different concentrations for each protein, were not significant. Differences between each protein at each concentration were not significant in HepG2 assays, either ( $P > 0.05$ ).

Overall, effects on cell viability were cell-type dependent; however protein-specific results were significant between C11.2 and C12.2, indicating that overall, C12.2 is less cytotoxic than C11.2, at least in HepG2 cells. Results for HepG2 cells represent three independent experiments with biological replicates in triplicate ( $n=9$ ).

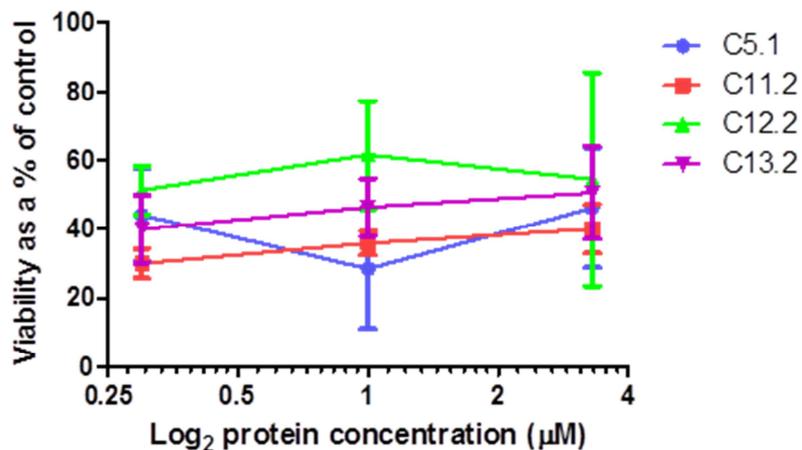
Mean values were calculated for each treatment, normalised against the untreated control cells and expressed as %. Error bars reflect  $\pm$  standard deviation values ( $\pm$  SD) for HEK293 cells ( $n=3$ ), and  $\pm$  standard error of the mean ( $\pm$  SEM) for HepG2 cells ( $n=9$ ). Statistical analysis was done with a 2-way Anova, and statistical significance was established with p-values set at 0.05 for pairwise comparisons, followed by Dunnet's post-

hoc test. In conclusion, statistically significant, dose-dependent cytotoxicity was observed in HEK293 cells, but not HepG2 cells, which signifies that cell-type specific effects on mitochondrial activity are exerted by these proteins.

A.



B.



**Figure 4.5 Cell viability in HEK293 and HepG2 cells following a 24 hour incubation with increasing concentrations of C5.1, C11.2, C12.2 and C13.2 is cell-line dependent.** Cells were incubated with increasing molar concentrations of proteins for 24 hours and viability was assessed by the MTT reduction assay. **A.** HEK293 cell viability with 0.1, 0.3, 1.0 and 3.3 µM proteins. n=3; error bars ±SD. **B.** HepG2 cell viability with 0.3, 1.0 and 3.3 µM proteins. n=9; error bars ±SEM. Statistical analysis was done by 2-way Anova followed by Dunnet's post-hoc test of each protein concentration against the highest concentration. \*P<0.05, \*\*P<0.01. The mock-treated (vehicle) control was used to normalise values.

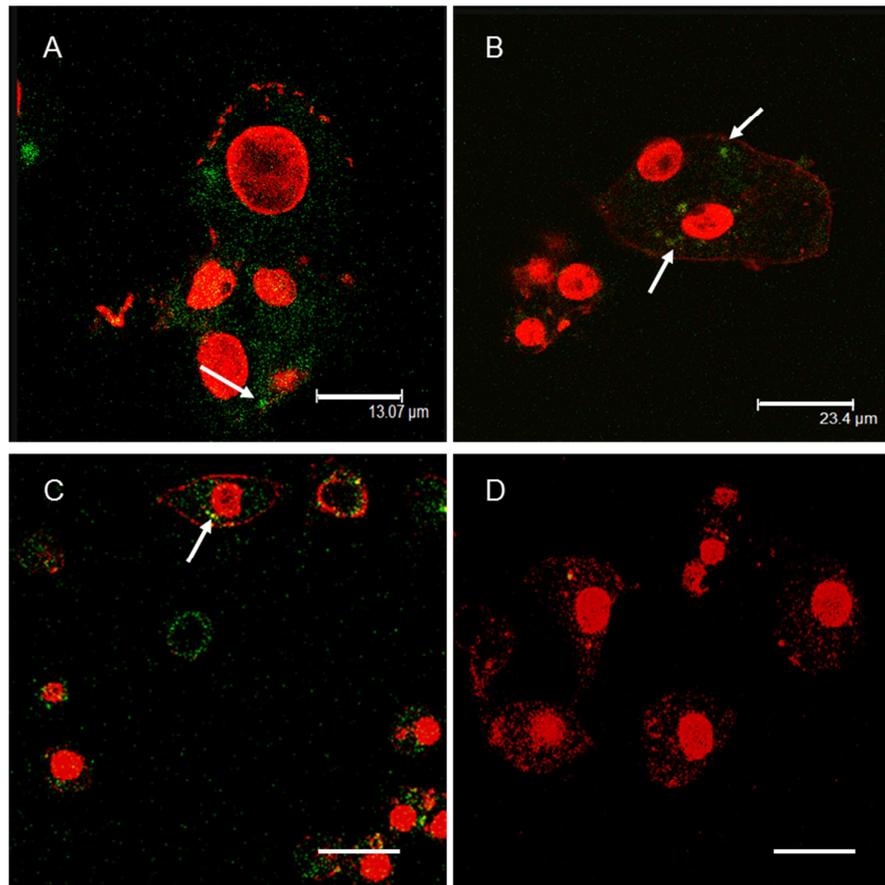
#### **4.2.2.3 Qualitative assessment of protein-siRNA complex uptake in live HepG2 cells with C5.1, C12.2 and fluorescent siRNA**

Electrophoretic mobility shift assays with C5.1, C12.2 and C13.2 complexed with siRNA showed that DRBMx2 binds siRNA at lower molar ratios than DRBM1 alone (Fig. 4.3 B).

In order to compare siRNA transduction efficacy between C5.1 and C12.2, proteins were co-incubated with 60 pmol siRNA-fluorescein at a 50:1 molar ratio (to give a final siRNA concentration =100 nM) and applied to adhered HepG2 cells for 2 hours. WGA-dsRed was used as a counterstain, localising primarily in the membrane and the nucleus at a concentration of 2.5 mg/ml (Fig. 4.6 D).

Assuming that ~98 % siRNA is complexed with protein at this molar ratio (van Asbeck *et al.*, 2013), C12.2 delivered siRNA to the cytoplasm of adherent HepG2 cells (Fig. 4.6 A), with comparable efficiency to the lipofectamine-transfected control (Fig. 4.6 C).

The punctate distribution points towards an endocytic mechanism of uptake, consistent with results by Geoghegan *et al.* (2012). C5.1-siRNA, on the other hand, exhibited poor intracellular localization with a diffuse distribution interspersed by distinct punctae, which may be endosome-entrapped (Fig. 4.6 B).

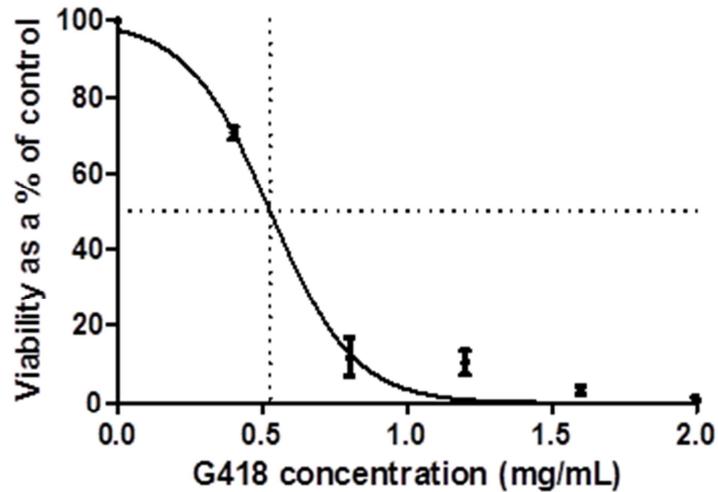


**Figure 4.6** Intracellular localisation of C5.1 and C12.2 complexed with 100 nM siRNA-fluorescein in live HepG2 cells. Protein-siRNA complexes were allowed to form at a 50:1 molar ratio and incubated with cells for two hours at 37°C. WGA-dsRed was used as a counterstain, translocating to the nucleus. **A.** C12.2-siRNA-fluorescein; **B.** C5.1-siRNA-fluorescein; **C.** Lipofectamine RNAiMax-siRNA-fluorescein and **D.** Mock-treated control. Arrows indicate endosome-entrapped siRNA. Cells were visualized with a TCS SP2 Leica confocal microscope. Scale bar: 47.6 µm.

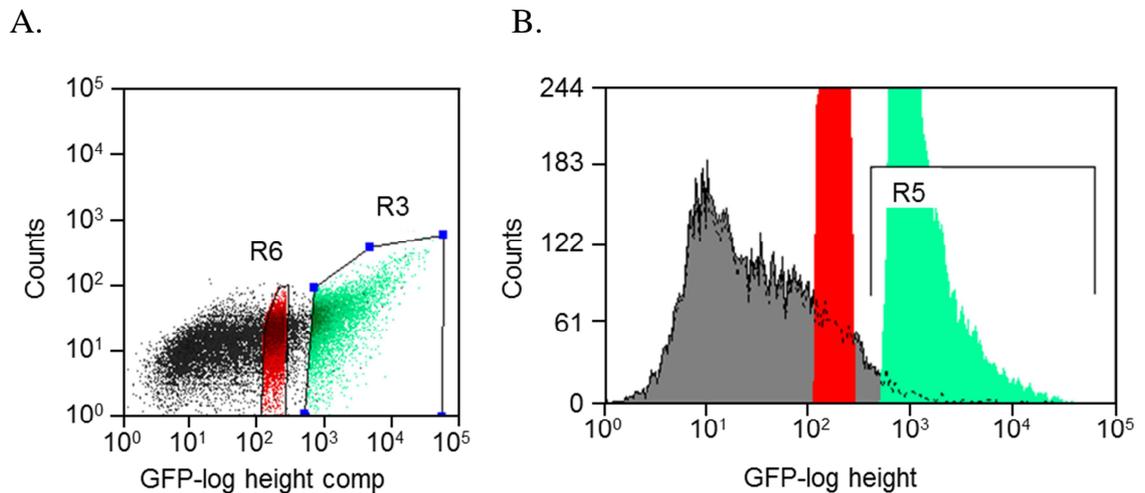
#### 4.2.2.4 Generation of a HEK293 reporter cell line

A HEK293 reporter cell line stably expressing destabilized enhanced green fluorescent protein (dEGFP) was created in order to assess whether fusion protein-mediated siRNA transduction could effectively induce transient knockdown of protein expression by RNAi. Cells were transfected with pd1EGFP-N1 plasmid DNA. To determine the optimal antibiotic selection concentration a kill curve was generated following a dose-response assay with 0.4-2 mg/ml G418 for 10 days. Lethal dose (LD) was observed at 2 mg/ml, with decreased viability ( $\leq 15\%$ ) observed at doses  $\geq 0.8$  mg/ml (Fig. 4.7). A sub-lethal dosage of 1 mg/ml G418 was therefore used as the effective antibiotic concentration for a further two weeks. Fluorescence activated cell sorting analysis (FACS) indicated that a range of dEGFP levels (measured on a FITC  $\log_{10}$  scale) was stably expressed after auto-fluorescence levels were normalised with untransfected HEK293 cells. Event limits were set at  $1 \times 10^4$  cells. First, a population gate was set to encompass the majority of live, healthy cells based on the forward scatter (FSC, indicating cell size) and side scatter (SSC, indicating granularity) axes. Then, cell populations at GFP log height  $\sim 10^2$  and  $\sim 10^3$  were considered to be 'Medium' and 'Highly' – expressing populations and gated (R6 and R3 regions, respectively). Highly-expressing cells (R3) represented only 2.75% of the total viable cell population, whereas the 'medium' expressing cells (R6) made up 12.19% of the population (Fig. 4.8 A). The cell population (gated as R5) represented cells that expressed the protein at high ( $10^3$ ) levels. This subpopulation equated to only 4.35% of cells (Fig. 4.8 B). Cells were analysed by Fluorescence-activated cell sorting (FACS) following stable selection with 1.0 mg/ml G418 into a 'Medium'-expressing and a 'Highly'-expressing population with specific expression profiles (Fig 4.9) Following sorting into distinct populations on the basis of their dEGFP expression, the maintenance dosage was decreased

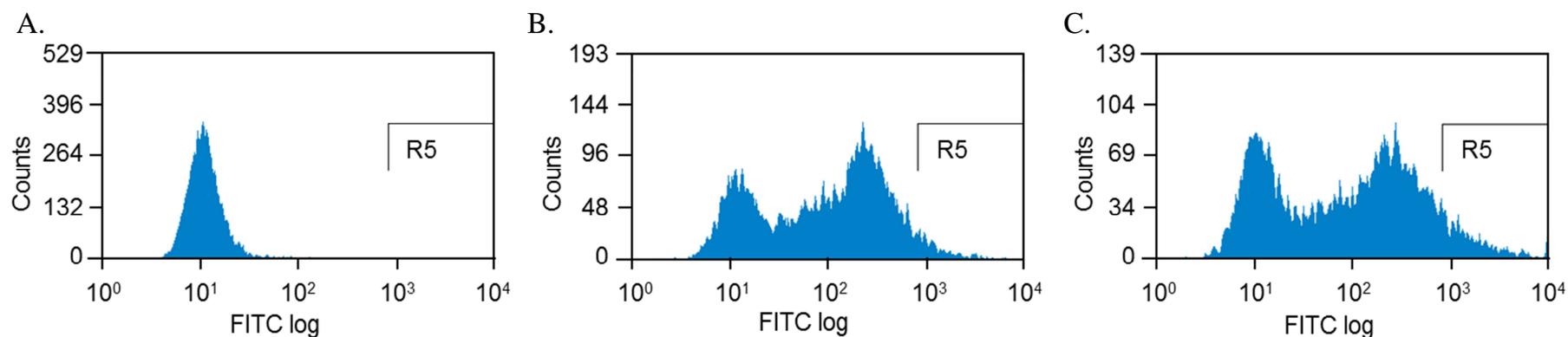
to 0.5 mg/ml, which shifted both population expression levels to a lower expression profile (Fig. 4.10). FACS analysis was done at the Institute of Child Health, University College London (ICH, UCL).



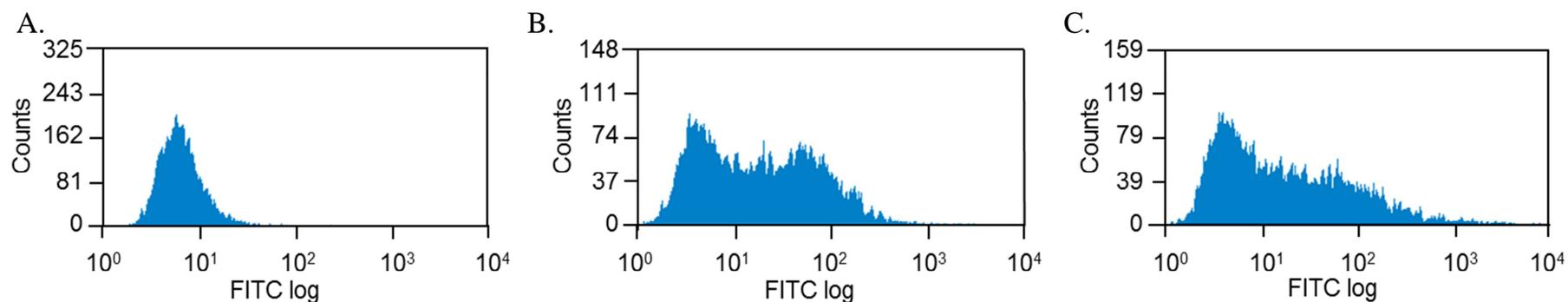
**Figure 4.7** Cell viability of HEK293-dEGFP cells treated for 10 days with increasing concentrations of G418 following transformation with pEGFP-N1. Stable clones were counted as a % of untransfected cells by the trypan-blue assay.



**Figure 4.8** FACS analysis of stable HEK293-dEGFP clones. **A.** 2-D dot plot of  $1 \times 10^4$  HEK293 cells stably transfected with pd1-EGFP-N1. R6, 'medium' expressing cells (GFP log  $\sim 10^2$ ); R3, highly expressing cells (GFP log  $\geq 10^3$ ); R6, medium-expressing viable HEK293-dEGFP cells. **B.** Histogram of gated cell populations. R5, highly expressing viable HEK293-dEGFP cell population.



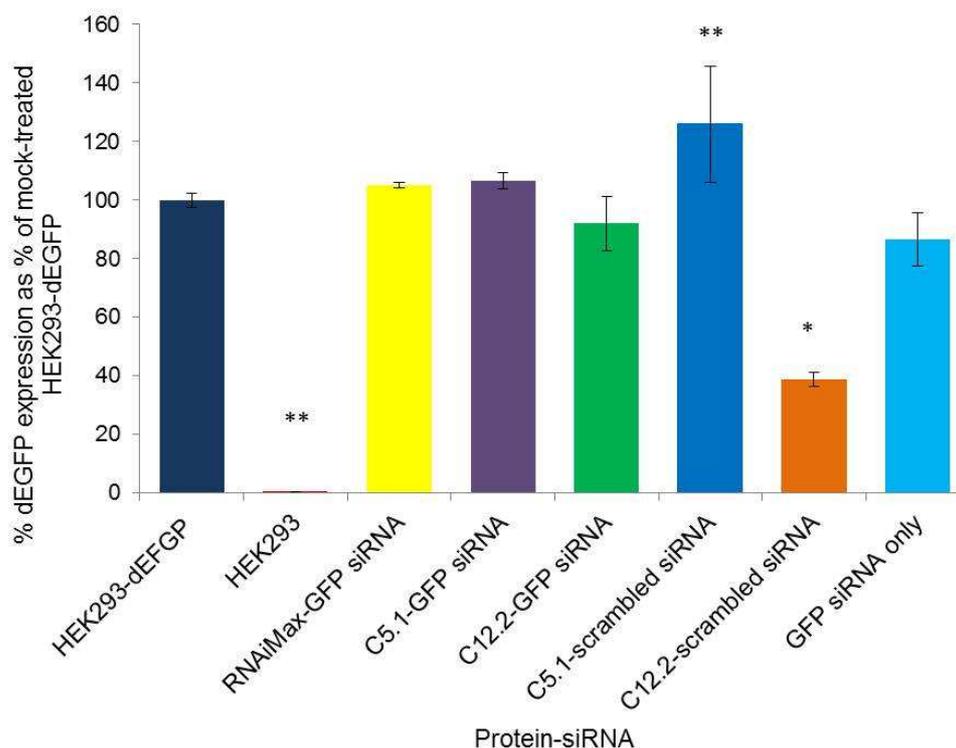
**Figure 4.9** FACS histogram analysis of stably transfected HEK293-dEGFP cell populations selected with 1.0 mg/ml G418. **A.** Untransfected HEK293 cells. **B.** Medium-expressing HEK293-dEGFP cell population (FITC log  $10^2$ ). **C.** Highly-expressing HEK293-dEGFP cell population (FITC  $\geq \log 10^3$ ). R5, highly expressing population within A, B and C. Total cell count in histogram analysis,  $7 \times 10^4$  cells.



**Figure 4.10** FACS histogram analysis of stably transfected HEK293-dEGFP cell populations maintained with 0.5 mg/ml G418 shows a shift to a lower dEGFP phenotype. **A.** Untransfected HEK293 cells. **B.** Medium-expressing HEK293-dEGFP cell population (FITC log  $\geq 10^2$ ). **C.** Highly-expressing HEK293-dEGFP cell population (FITC  $\geq \log 10^3$ ). R5, highly expressing gated population. Total cell count in histogram analysis,  $7 \times 10^4$  cells.

#### 4.2.2.5 Transient knockdown assays in a HEK293-dEGFP reporter cell line

In order to establish whether cell penetrating fusion proteins could effectively deliver siRNA into living cells and induce potent and specific gene knockdown, C5.1 and C12.2 were complexed with either 10 nM or 100 nM final siRNA concentrations in low-serum media (Optimem).  $2.5 \times 10^5$  HEK293-dEGFP cells were seeded and adherent cells were transduced with complexes for 5 hours. Gene knockdown was observed by flow cytometry after either 24 or 48 hours treatment, with propidium iodide (PI) as a dead cell counterstain. 10 nM siRNA as a final concentration is the minimum concentration that is able to induce knockdown *in vitro*, and therefore served as a starting point for knockdown assays (Invitrogen, 2007). Since C5.1 and C12.2 lack the capability to escape from endosomes upon internalisation, they become sequestered into vesicles following uptake, thereby limiting the amount of siRNA cargo that is loaded onto the RNAi machinery (Geoghegan *et al.*, 2012). To enhance endosomal escape, cells were treated with 50  $\mu$ M chloroquine in complete growth media following transduction, according to the method from Geoghegan *et al.* (2012). Lipofectamine RNAiMax was used as a positive control in order to assess the efficiency of knockdown using this siRNA concentration. Flow cytometry analysis on  $1 \times 10^4$  HEK293-dEGFP cells in the presence of PI, showed that C5.1-siRNA failed to induce GFP expression knockdown actually increasing dEGFP expression levels (%). No knockdown was observed with RNAiMax-transfected siRNA (Fig. 4.11). Interestingly, C12.2-siRNA produced a modest GFP expression knockdown, whereas siRNA only treatment decreased the levels of dEGFP expression by ~20%. Although live cells, that have taken up PI at a  $\log_{10} \geq 10^2$ , were gated, the cytotoxic or non-specific results of this analysis cannot be excluded completely.



**Figure 4.11 dEGFP expression analysis following treatment with 10 nM siRNA complexed with C5.1 or C12.2 for 48 hours.** Cells were treated with either C5.1 or C12.2 complexed with 10 nM GFP-specific siRNA at 50:1 molar ratios.  $1 \times 10^4$  HEK293-dEGFP cells were analysed for GFP fluorescence by flow cytometry in the presence of propidium iodide. Mean values for each treatment were normalised against the mock-treated HEK293-dEGFP control and expressed as a %. Error bars represent  $\pm$ SD, n=3. Statistical analysis was done by 1-way Anova, followed by Dunnet's post-hoc test against the siRNA control. \*P<0.05, \*\*P<0.01.

None of the Lipofectamine-siRNA or protein-siRNA treatments, mediated any statistically significant knockdown. This indicated that the siRNA concentration was insufficient for the number of cells seeded. The assay was therefore repeated with 100 nM siRNA. To limit potential cytotoxicity elicited by a high siRNA concentration, protein-siRNA transductions were assessed 24 hours following treatment, rather than after 48.

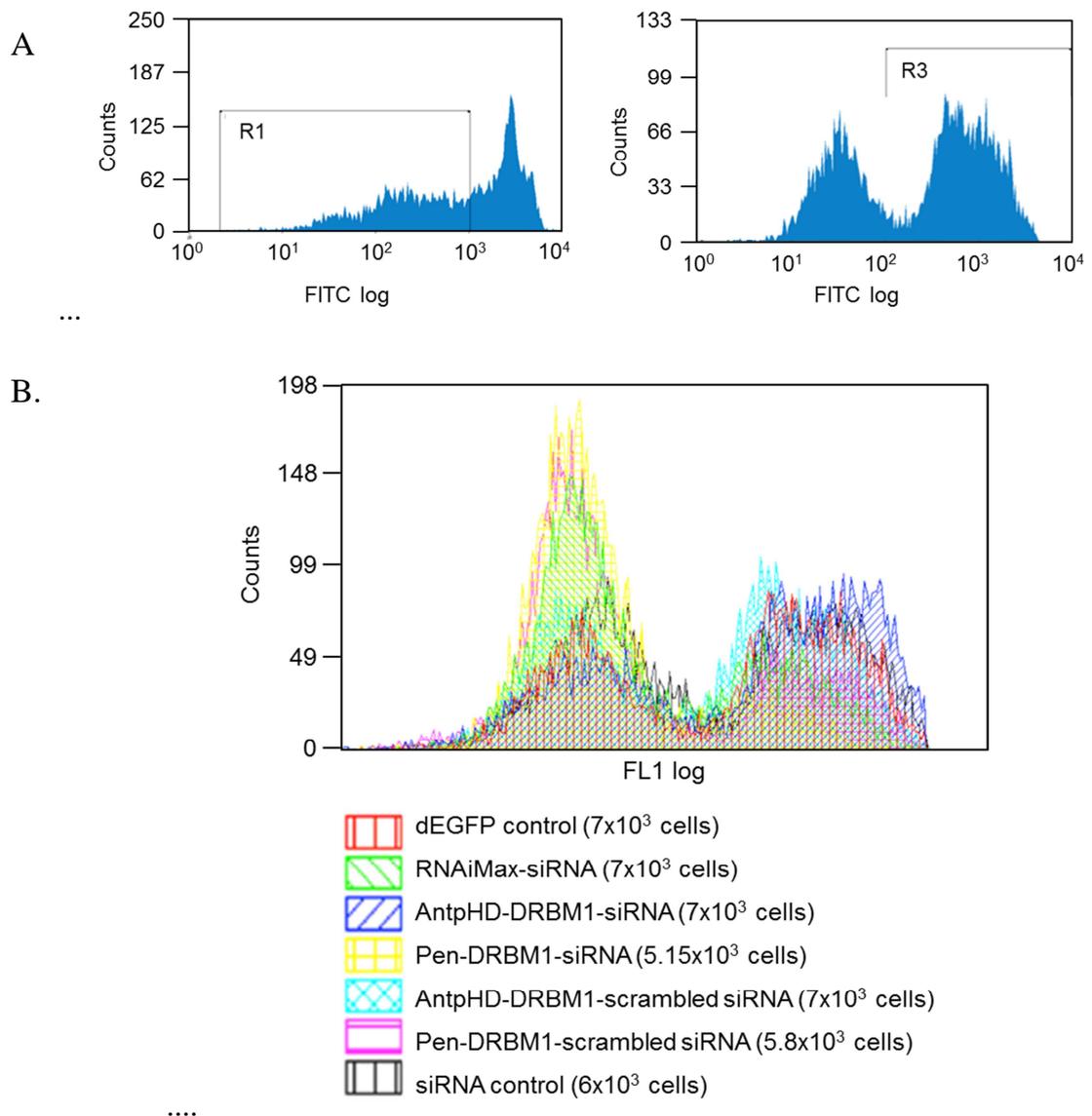
A population sample of  $1 \times 10^4$  cells was analysed by flow cytometry with a FITC/PI filter. Untreated HEK293 cells were used to normalise against autofluorescence (FITC  $\log_{10}$

$\leq 10^3$ ); this cell population was excluded from further analysis. Cells were stained with 0.5 mg/ml PI with a  $\log_{10}$  value of  $\geq 10^3$  were also gated and excluded from analysis (R1), therefore only live, highly expressing (FITC  $\log_{10} \geq 10^3$ ) were considered when assessing for knockdown efficiency (R3) (Fig. 4.13). Mock-treated HEK293-dEGFP cells were used as the normalisation control (*i.e* set to 100%). All sample % histogram values were averaged and normalised against the untreated control. Two distinct subpopulations exist within the HEK293-dEGFP reporter cell line analysed, although the selection marker G418 was used at very high concentrations (1.8mg/ml) to ensure high stable plasmid expression (FITC  $\log_{10} \geq 10^3$ ). This is a normal distribution of fluorescence within a polyclonal stable cell line; a subpopulation has lost the dEGFP plasmid but is still resistant to G418 and another is expressing the pd1-EGFP-N1 plasmid at high levels.

As expected, mock-treated HEK293 cells showed minimal fluorescence, whilst the cell population transfected with Lipofectamine RNAiMax showed a 40% decrease in dEGFP expression after 24 hours. C5.1-siRNA and C5.1-scrambled siRNA control failed to induce any dEGFP knockdown. Interestingly, C12.2-siRNA treatment showed a marked decrease ( $\approx 60\%$ ) in dEGFP expression – although treatment with scrambled, non-specific siRNA also showed a comparable decrease. siRNA-only treatment in the absence of protein or transfection agent elicited a modest, yet significant (20%) decrease in dEGFP expression, which was unexpected, as negatively charged siRNA molecules are usually excluded from cells by virtue of the negatively charged membrane bilayer (Fig.4.13). This observation cannot be attributed to cytotoxic effects since cells expressing PI at  $\geq \log 10^3$  values (and were therefore considered dead), were omitted from further analysis.

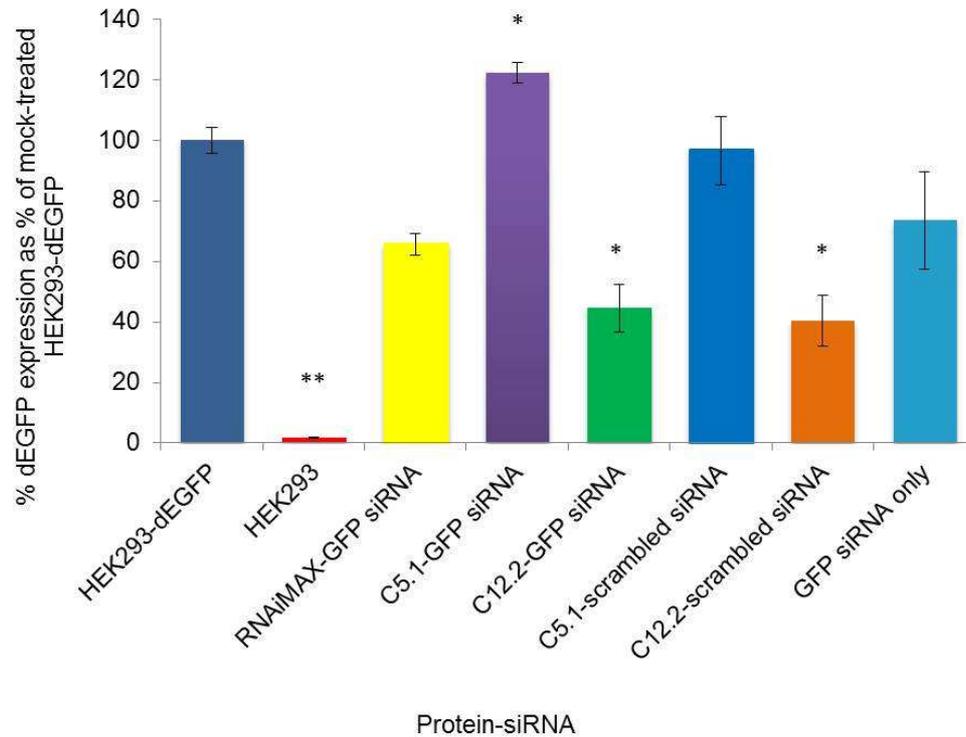
Live (Fig. 4.12 A, R1), highly expressing cells were gated (Fig. 4.12 A, R3) and used to normalize cell populations prior to univariate analysis of overlaid treatment averages (Fig.

4.12 B). A decrease in dEGFP expression is shown as a shift in fluorescence, and was observed for cells treated with C12.2-siRNA (yellow), C12.2-scrambled siRNA control (pink) and lipofectamine-siRNA (green), in comparison to expression profiles observed with the mock-treated control (red), C5.1-siRNA (blue) and siRNA only (black).



**Figure 4.12 FACS analysis of a HEK293-dEGFP cell line following treatment with protein-siRNA complexes for 24 hours. A.** Histogram analysis of HEK293-dEGFP cells showing the gated subpopulations of the cells prior to dEGFP knockdown analysis. R1 (left panel): Live cell subpopulation of untreated HEK293-dEGFP cells ( $\text{PI } \log_{10} \leq 10^3$ ). R3 (right panel): Highly expressing, live, untreated HEK293-dEGFP subpopulation after superimposition of the R1 gate. **B.** Univariate analysis showing overlaid histograms indicative of GFP knockdown. FL1: Fluorescence channel (FITC).  $n=3$ . Subpopulations

showing a decrease in dEGFP expression are shown shifted to the left; the highly expressing HEK293-dEGFP population is shown on the right of the FL1 channel axis.



**Figure 4.13 dEGFP expression analysis following treatment with 100 nM siRNA complexed with C5.1 or C12.2 for 24 hours.** Cells were treated with either C5.1 or C12.2 complexed with 100 nM GFP-specific siRNA at 50:1 molar ratios.  $1 \times 10^4$  HEK293-dEGFP cells were analysed by flow cytometry in the presence of propidium iodide. Mean values for each treatment were normalised against the mock-treated HEK293-dEGFP control and expressed as a %. Error bars represent  $\pm$ SD, n=3. Statistical analysis was done by 1-way Anova, followed by Dunnet's HSD test against the siRNA control. \*P<0.05, \*\*P<0.01.

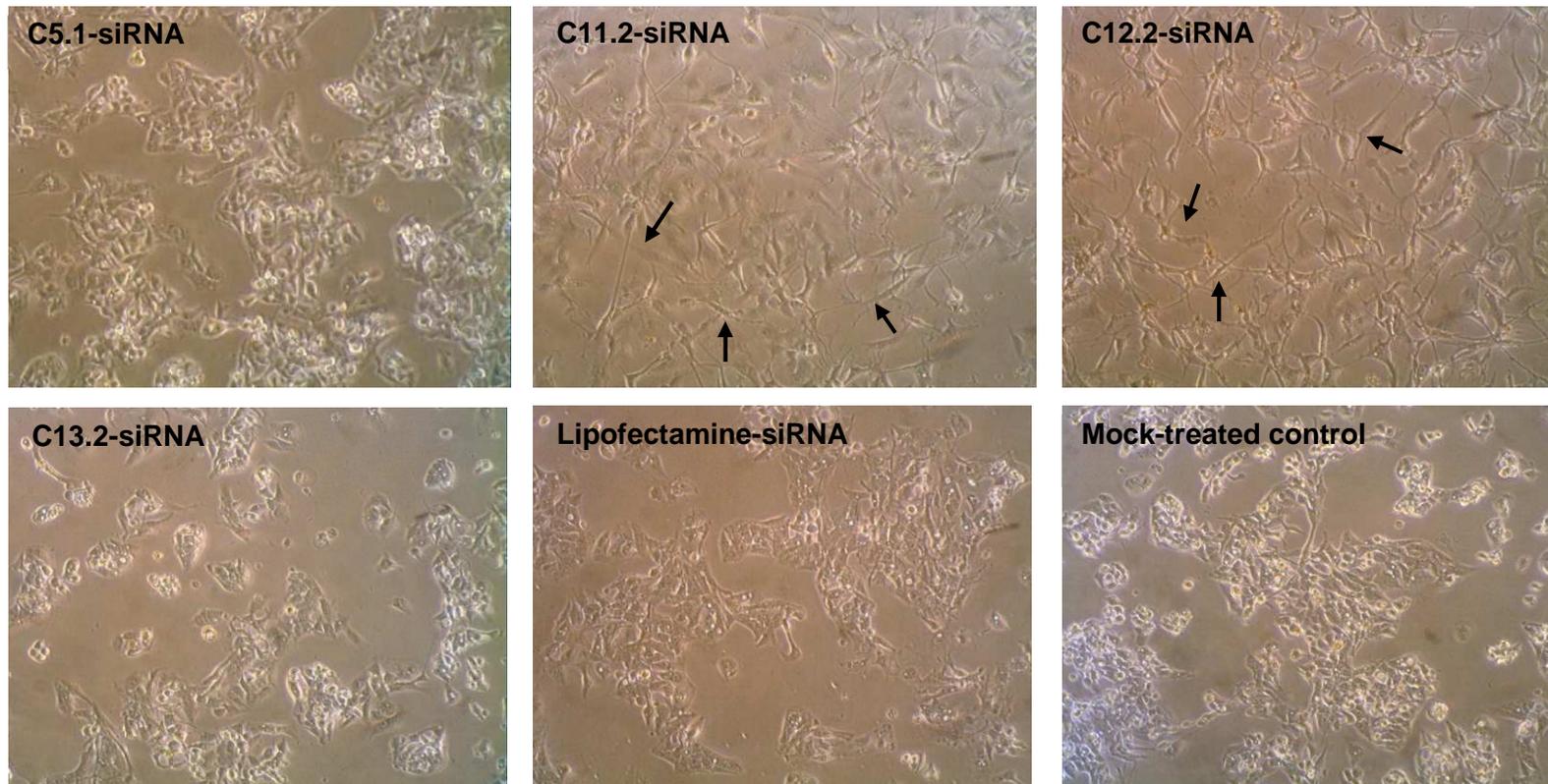
#### 4.2.2.6 Protein –induced morphology changes in HepG2 cells

Potential effects on the morphology of HepG2 cells were briefly assessed prior to proof-of-principle experiments with protein-siRNA PTP1B knockdown. One of the intracellular targets for siRNA-induced knockdown in this study was the protein tyrosine phosphatase 1B (PTP1B) protein, which is endogenously expressed in HepG2 cells and is a therapeutic target for type 2 diabetes (T2D).

Lipofectamine RNAiMax and mock-treated HepG2 cells were used as controls. The siRNA duplex used corresponded to a 21-23nt sequence in the PTPN1 gene, which encodes for PTP1B (Life Technologies). Protein-siRNA complexes were formed with 25 nM siRNA at a 1:50 molar ratio to proteins (Fig. 4.14). Cells were treated for 5 hours at 37°C in low-serum media and incubated for 72 hours (Johnson *et al.*, 2010, Lu *et al.*, 2008). Cell morphology was assessed by phase-contrast microscopy, following treatment. Interestingly, C11.2-siRNA and C12.2-siRNA induced marked morphological changes in HepG2 cells, causing a shift from the typical epithelial phenotype associated with hepatocellular carcinoma cell lines, to a fibroblastic one, generally associated with connective tissue. Morphological changes were visible as elongated, bipolar apices, rather than the regular polygonal structures that HepG2 cells adopt.

C5.1-siRNA, C13.2-siRNA and Lipofectamine RNAiMAX-siRNA complexes did not cause any morphological or colony structure changes, compared to the mock-treated control. None of the proteins appeared to induce cytotoxicity at the molar concentration tested with the siRNA (1.25  $\mu$ M). These preliminary, qualitative results did not include siRNA-only and protein-only treatments, which would need to be included in subsequent

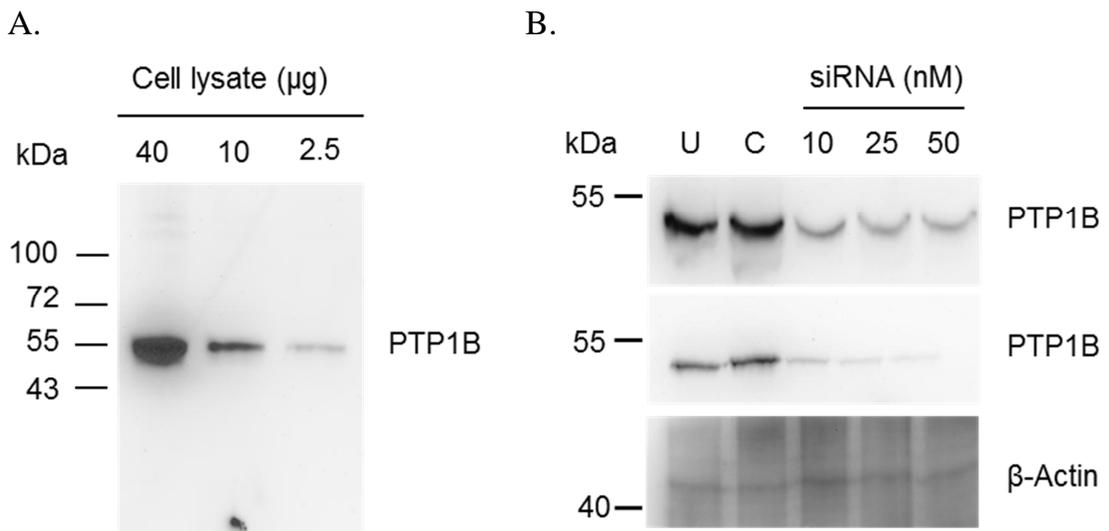
assays in order to assess whether morphological changes were induced by the protein-siRNA complexes, or the protein itself.



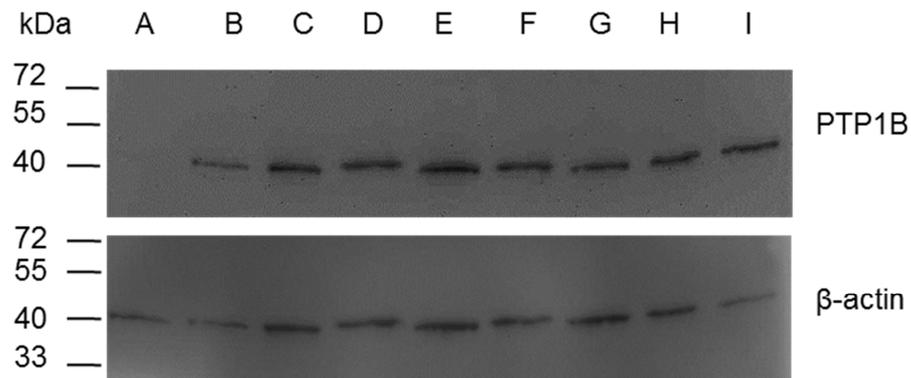
**Figure 4.14** Phase-contrast microscopy of HepG2 cells transduced with 25 nM PTPN1 siRNA complexed with proteins at a 50:1 molar ratio for 72 hours. Cells were incubated with protein-siRNA complexes for 5 hours before media were changed and incubated at 37°C for 72 hours before analysis by phase-contrast microscopy. Arrows represent apical structure formation induced by protein treatment. Mock-treated control, OptiMem only.

#### 4.2.2.7 Protein-siRNA assessment of PTP1B knockdown

C5.1, C11.2, C12.2 and C13.2 fusion proteins had been previously complexed with 25 nM PTPN1 siRNA and incubated HepG2 cells, with C11.2-siRNA and C12.2-siRNA had induced morphological changes in HepG2 cells. It was still unknown, though, whether a) these complexes had indeed transduced the cells and b) whether complexed PTPN1-specific siRNA had induced a transient knockdown in the expression levels of protein tyrosine phosphatase 1B (PTP1B). Since HepG2 endogenously expresses PTP1B, knockdown assays were performed, and protein levels were assessed by immunoblotting with a PTPN1-specific antibody (Fig. 4.15 A and B). Western blotting with an anti-PTP1B antibody (BD Biosciences, Oxford, UK) established antibody specificity at a 1:2500 dilution and 10 µg as the minimum amount of crude lysate protein resolved (Fig. 4.15 A). Preliminary transfections using lipofectamine and two siRNA oligonucleotide sequences (VHS204190 and VHS204191) (Methods), corresponding to the PTPN1 gene, showed a dose-dependent PTP1B knockdown effect with 10, 25 and 50 nM siRNA at 72 hours post-transfection (Fig 4.15 B). The minimum effective concentration required to induce complete PTP1B expression knockdown for both siRNA sequences was 25 nM. Lipofectamine-transfected siRNA induced a complete knockdown of PTP1B expression (Fig. 4.16, lane A) By contrast, none of the fusion proteins complexed with 25 nM siRNA induced any detectable protein expression knockdown at 50:1 molar ratios 72 hours post-transduction, even when 15 µg crude protein lysate samples were loaded (Fig. 4.16, Lanes B-I).



**Figure 4.15 HepG2 cell lysate immunoblot analysis with an  $\alpha$ -PTP1B antibody. A.** Lysate amount optimization from untreated HepG2 cells for detection by immunoblotting. **B.** PTP1B knockdown efficiency analysis following transfection with lipofectamine RNAiMax and two anti-PTPN1 siRNAs at various molar concentrations. Top panel, siRNA oligo 1 (VHS204190); lower panel, siRNA oligo 2 (VHS204191).  $\beta$ -Actin was used as a loading control. 15  $\mu$ g protein loaded. PTP1B, 50 kDa;  $\beta$ -Actin, 42 kDa.



**Figure 4.16 Immunoblot analysis on HepG2 cell lysates following PTP1B knockdown for 72 hours. A.** Lipofectamine RNAiMax-siRNA; B. Lipofectamine RNAiMax-scrambled control siRNA; C. C5.1-siRNA; D. C12.2-siRNA; E. C11.2-siRNA; F. C13.2-siRNA; G. C9.2-siRNA; H. C10.2-siRNA and I. Untransfected control.  $\beta$ -Actin was used as a loading control. 15  $\mu$ g protein loaded. PTP1B, 50 kDa;  $\beta$ -Actin, 42 kDa.

### 4.3 Discussion

In this chapter, we aimed to purify and characterise recombinant proteins that contain the entire dsRNA binding domain (DRBMx2) from human PKR (Fig. 4.1), with the aim of assessing their siRNA binding efficiency by electrophoretic mobility shift assays; their ability to deliver GFP-specific siRNA to a dEGFP-expressing HEK293 reporter cell line; potential cytotoxicity by the MTT assay, and finally, the proteins' potential to induce PTP1B expression knockdown, evident at the protein level, by immunoblotting.

Previous research from Geoghegan *et al.*, had purified Tat-DRBMx2 (C11.2), Penetratin-DRBMx2 (C12.2) and DRBMx2 (C13.2) under denaturing conditions from bacterial cultures by low pressure chromatography. Purification and refolding were done on-column, and proteins were eluted, concentrated and buffer exchanged into PBS/10% glycerol (Geoghegan *et al.*, 2012). The approach used in this project, involved purification under denaturing conditions and sequential dialysis to refold proteins into their native state. Buffer-exchange to remove  $\text{Gu}^+$  ions and imidazole, resulted to the near-complete loss of C11.2 (final molar concentration = 3.65  $\mu\text{M}$ ). This protein was thereby excluded from these analyses. To eliminate the possibility that incorrectly folded protein had been purified, each step of the dialysis process was followed by a centrifugation step that eliminated aggregated protein as a precipitate. Increased precipitation with each dialysis step was observed for all proteins, even with the addition of 10% glycerol as an excipient, (Arakawa and Timasheff, 1983, Arakawa and Timasheff, 1984b, Arakawa and Timasheff, 1984a, Arakawa *et al.*, 2007, Frokjaer and Otzen, 2005, Gerlsma, 1968, Jorgensen *et al.*, 2009, Taneja and Ahmad, 1994). Although the process of dialysis is effective in the elimination of unwanted ions from the final buffer formulation, it appears to leave innately unstable proteins, such as C11.2 susceptible to degradation due to exogenous factors, such

as variations in temperature during handling, and slow buffer exchange kinetics. An alternative purification strategy, such as an ion exchange column, may have eliminated stability-dependent aggregation.

EMSA assays with C12.2 and C13.2 yielded effective siRNA complex formation at low molar ratios with little or no aggregation, in comparison to C5.1. Initial EMSA assays with fluorescent siRNA and no EtBr staining were unsuccessful, subsequent experiments utilised an unconjugated siRNA oligonucleotide (n=2). It would be expected that both C12.2 and C13.2, since they contain the same dsRNA binding domains, would bind siRNA with the same efficiency. SiRNA binding by C12.2 yielded a complete shift at a 10:1 molar ratio, whereas C13.2 showed a complete shift at a 6:1. This finding contradicts the 2:1 molar ratio observed by Geoghegan *et al.* (2012). Since the methodology adopted was the same, this discrepancy could be attributed to structural differences between the siRNA molecules used due to the different stabilising modifications; in our case, a 25-mer siRNA with no overhangs was used, with other proprietary sense strand modifications, whereas the study by Geoghegan *et al.* utilised 21-mer siRNAs. Although the sequence length is not expected to affect binding efficiency (Bycroft *et al.*, 1995, Kharrat *et al.*, 1995, Ryter and Schultz, 1998), (Dzananovic *et al.*, 2013), these discrepancies may be attributed to the lack of 2'-OH overhangs, or other proprietary modifications in the siRNA used. Indeed, crystal structure studies on the highly conserved DRBD from *Xenopus laevis*, showed that RNA-protein interactions occur through hydrogen bonding with 2'-OH groups present in the minor groove, in A-form dsRNA (Ryter and Schultz, 1998). A non-equivalence of dsRNA binding by the DRBMx2 occurs with nucleoside modifications or the incorporation of stabilising G-U wobble pairs (Nallagatla *et al.*, 2008, Nallagatla and Bevilacqua, 2008),

that ultimately change the overall structure of the A-form helix (Masquida and Westhof, 2000, Weeks and Crothers, 1993).

MTT assay results pointed towards dose-dependent as well as cell-line specific effects of unconjugated proteins on cell viability. In HEK293 cells, relative cytotoxicities at 3.3  $\mu\text{M}$  were highest for C11.2 and C13.2, with C5.1 and C12.2 eliciting the lowest (n=3). Although C11.2 also showed the greatest effects on cell viability at the 3.3  $\mu\text{M}$  concentration in HepG2 cells, statistical analysis of these assays (n=9) showed no significance at  $p\text{-values}\leq 0.05$  (Appendix II.2), indicating that there is no correlation between our concentration range and mitochondrial activity in this cell line.

By comparison, pair-wise comparisons between C11.2 and C13.2, or between proteins and concentrations in HEK293 cells were significant at  $p\leq 0.05$ , indicating that cytotoxicity is indeed dose-dependent, at least for these two proteins. Cell viability differences between C5.1 and C12.2, were not statistically significant (Appendix II.2). The observation of cell line-specific effects on cell viability is consistent with previous research with Penetratin by Sugita *et al.*, 2008. Their group had observed Penetratin-induced cytotoxicity in HeLa, but not HaCat and A431 cells at concentrations  $\geq 100 \mu\text{M}$ , with negligible cytotoxicities below that concentration (Sugita *et al.*, 2008), with negligible cytotoxicity at 50  $\mu\text{M}$  was observed in CHO cells (El-Andaloussi *et al.*, 2007a). Our results obtained from HEK293 cell assays point towards dose-dependent cytotoxicity at a much lower concentration range than previously tested (El-Andaloussi *et al.*, 2007a, Mueller *et al.*, 2008, Sugita *et al.*, 2008). It is still unknown why mitochondrial activity is affected by various concentrations of proteins, although overestimations of mitochondrial activity (and by extension, cell numbers), are a common limitation of this assay (Wang *et al.*, 2010). Also unknown are the causes behind cell line susceptibilities. It would be interesting to replicate these

findings using another colorimetric assay, such as the WST-8 assay. The WST-8 assay involves the extracellular, rather than intracellular reduction of water-soluble tetrazolium salts, and thus does not require the solubilisation step with DMSO, minimizes cytotoxicity, can be colorimetrically assayed and provides better accuracy with regards to cell viability (Berridge *et al.*, 2005). Moreover, our observations reinforce the notion that cell-type specific effects should be considered when preparing protein-based therapeutic formulations.

Morphology-specific effects, especially with non-covalently conjugated siRNA cargo, have not been previously reported for HepG2 cells. Here, we have shown that a 72-hour incubation of HepG2 cells with C11.2 and C12.2 complexed with 25 nM siRNA elicit distinct morphological changes in HepG2 cells, which were not observed by similar treatment with Lipofectamine-siRNA, C13.3-siRNA or C5.1-siRNA. Again, it is unclear as to why Tat-DRBMx2 (C11.2) and Pen-DRBMx2 (C12.2), but not DRBMx2 (C13.2) or AntpHD-DRBMx2 (C5.1), have caused these changes. Both proteins appear to shift the morphology of epithelial HepG2 cells from a well-defined, cubical morphology, to a more fibroblastic one with apical structures extending in a bipolar manner. Assuming that complexes are indeed taken up by the cells, it would be interesting to assess potential changes on the actin cytoskeleton by phalloidin staining. Holm *et al* (2011) investigated the effects of 20  $\mu$ M fluorescently labelled Penetratin and its retro-inverso (RI) enantiomer with the same amino acid sequence in an epithelial mammary gland adenocarcinoma cell line (MDA-MB-231) (Holm *et al.*, 2011). Although Penetratin did not elicit any changes to the arrangement of the actin cytoskeleton, RI-Penetratin caused drastic changes in the adhesive actin bundles, producing a similar phenotype to the HepG2 cells treated with C11.2 and C12.2. Absence of morphological changes with Lipofectamine-siRNA treatment

signify that siRNA is not the causative agent, however, in the absence of a siRNA-only and protein-only controls, these qualitative results can only be construed as preliminary and in need of further investigation.

Localisation assays in live HepG2 cells with a WGA-dsRed counterstain showed enhanced intracellular localisation of C12.2-fluorescent siRNA complexes, compared to C5.1-fluorescent siRNA and lipofectamine-transfected cells. The punctate distribution observed is consistent with previous work with unconjugated Penetratin in HEK293, HeLa, MDCK and Cos-7 cells (Mueller *et al.*, 2008), presumably by associating with heparan sulfate proteoglycans on the cell surface (Console *et al.*, 2003, Marty *et al.*, 2004, Tyagi *et al.*, 2001). Since these are qualitative results, the transduction efficiency cannot be quantified, and may be attributed to an enhanced complex formation efficiency of the DRBMx2 in C12.2 compared to the DRBM1 in C5.1, consistent with observations by Geoghegan *et al.* (2012). An alternative explanation would be due to the lack of co-purified nucleic acids.

Since C5.1 was able to transduce siRNA intracellularly, this indicates that the protein is still able to bind siRNA, and transduce it into cells, consistent with observation by Eguchi *et al.*, albeit with lower affinity than intact DRBMx2 (C13.2). This observation could be explained by the mode of dsRNA binding by PKR; although DRBM1 can effectively bind dsRNA, DRBM2 is required to stabilise the overall complex (Dabo and Meurs, 2012).

To quantify C5.1 and C12.2-mediated siRNA transduction, a dEGFP-expressing HEK293 reporter cell line was developed from a polyclonal culture transformed with the destabilised variant of GFP from *Aquorea victoria*. Stable clones were selected with a high dose of geneticin (G418), due to the generation of the kill-curve after plasmid transfection, and not before. However, fluorescence-activated cell sorting (FACS) of stable clones

revealed two distinct sub-populations, with the majority of cells showing stable plasmid integration, suitable for further assays. The actual percentage of cells expressing dEGFP at high levels ( $\text{FITC log}_{10} \leq 10^3$ ) may be attributed to the expression of many plasmid copies within each cell. Decreasing the maintenance dose shifted GFP expression levels and decreased the number of highly expressing clones. Further attempts to increase the proportion of highly expressing cells by increasing the antibiotic concentration did not change the expression profile, with a large number of cells (67%) still showing stable plasmid integration.

Eguchi *et al* (2009) had reported robust dGFP expression knockdown in a lung adenocarcinoma reporter cell line (H1299-dGFP/dsRed) with PTD-DRBD following treatment for 24 hours, that was sustained after 48 hours. Treatment of HEK293-dEGFP cells with either C5.1 or C12.2-siRNA at a final siRNA concentration of 10 nM did not elicit any effects on dEGFP expression; neither did lipofectamine-transfected siRNA, which indicated that the siRNA concentration was insufficient to induce knockdown. Considering previous observations of C5.1 aggregating, (*i.e.* associating with, but not entering the cell) at the plasma membrane of fixed HepG2 cells, and the work of Mueller *et al*, (2008), who reported that fewer than 10% of CPPs in general are effectively internalised, the possibility that an insufficient number of protein-siRNA complexes were successfully internalised remains. Since C12.2 (Geoghegan *et al*, 2012) and C5.1 lack the ability escape from endosomes upon internalisation, an incubation step for 24 hours in complete GM supplemented with 50  $\mu\text{M}$  chloroquine, a known endosomolytic agent, was performed (Geoghegan *et al.*, 2012). In following assays ( $n=3$ ), the final concentration of siRNA was subsequently increased to 100 nM and knockdown was assessed after 24, rather than 48 hours. Non-covalent complexes with C12.2-siRNA were able to decrease

dEGFP expression significantly, which indicated that complexes were transduced into cells and siRNA was able to induce RNAi. However, the same observations occurred with the non-specific scrambled control, which was unexpected. The possibility that a non-specific decrease in dEGFP expression is caused by a decrease in viable cell numbers was taken into consideration by staining cells with PI (0.5 mg/ml) prior to analysis and gating cells that had appeared to be alive. However, further work is needed to assess any non-specific knockdown effects that may be caused by the scrambled control siRNA.

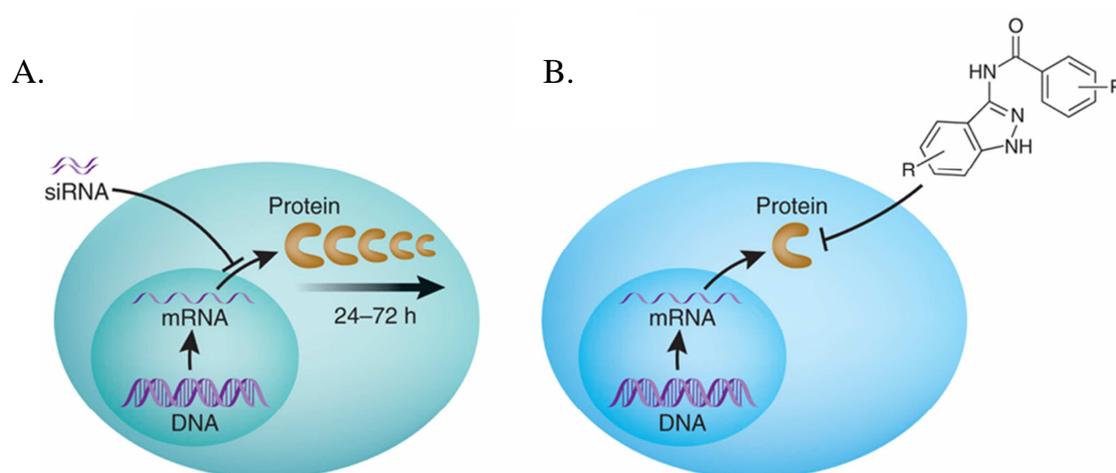
C5.1-siRNA on the other hand, failed to mediate any dEGFP knockdown, indicating that the DRBM1 does not bind the siRNA sufficiently, consistent with previous observations from EMSA assays. In contrast, Eguchi *et al.*, had reported robust and specific reporter gene knockdown with the PTD-DRBM1 protein non-covalently loaded with siRNA; their results however, were obtained with 400 nM siRNA. Moreover, their observations were attributed to non-specific interactions with the Tat moieties within the protein, which aided siRNA complex formation (Calnan *et al.*, 1991, Geoghegan *et al.*, 2012, Mujeeb *et al.*, 1994).

As a proof-of-concept experiment aimed to validate whether any of the recombinant proteins were able to knockdown the expression of PTP1B in HepG2 cells, all purified proteins were complexed with 25 nM siRNA and used to treat cells for 72 hours. This RNA concentration was optimised with lipofectamine transfections (n=2), with evidence at the protein level assessed by western blot. Other optimisation assays on untreated whole cell lysates with an anti-PTP1B antibody identified 10 µg as the minimum amount of crude lysate protein resolved by immunoblotting. Whole cell lysate analysis following knockdown was therefore done with 15µg crude protein. Apart from the lipofectamine-treated control, none of the recombinant protein-siRNA complexes induced knockdown. It

would be interesting to investigate potential changes in putative primary mRNA transcript levels instead, by RT-PCR (Reverse Transcription Polymerase Chain Reaction), which would provide more accurate information as to the efficiency of these proteins to mediate RNAi.

# Chapter 5. Conclusions and future directions

RNAi-based therapies have the potential to revolutionize personal medicine by specifically targeting aberrant gene expression - the cause of a cornucopia of chronic, debilitating and fatal diseases. In theory, RNAi based therapies target primary mRNA transcripts before they are translated into aberrant or misfolded protein, in a potent, specific, universal and non-cytotoxic manner, as opposed to small molecule inhibitors (SMocs) that target the protein product from affected DNA following transcription and translation from its putative mRNA (Fig. 5.1).



**Figure 5.1** Modes of action of the attenuation of protein activity. **A.** Translation inhibition by siRNA-mediated RNAi; **B.** Inhibition of protein activity by small molecule inhibitors (SMocs). Adapted from Weiss *et al.* (2007).

An important issue for the development of these therapies lies in their delivery; as Meade and Dowdy wrote, ‘delivery of nucleic acid therapies is the 800-pound gorilla in the room’ (Meade and Dowdy, 2009). Current nanoparticle-based strategies that include cationic motifs, whether these are polymers, peptides or liposomal formulations, aim to increase *in vivo* circulation time, reduce renal clearance and avoid the stimulation of the interferon response (de Fougerolles *et al.*, 2007, Vaishnav *et al.*, 2010). The most commonly used approach thus far was the covalent conjugation between carrier and cargo, however the

large anionic charge of the siRNA backbone is amenable to non-covalent approaches which may be able to release cargo intracellularly upon internalisation.

In light of recent developments in the non-covalent delivery of siRNA by incorporating a dsRNA binding domain sequence with a cell-penetrating peptide (CPP, in this instance the HIV-Tat protein) (Eguchi *et al.*, 2009), that exhibited potent and specific gene knockdown in a number of cell lines, we set out to investigate whether the Antennapedia homeodomain, a transduction domain part of a much larger protein complex from *D. melanogaster*, could potentially be an efficacious siRNA delivery vector. To aid complex formation with siRNA, we recombinantly expressed the AntpHD, as well as Penetratin, with a dsRNA binding domain from human PKR – a strategy which has proved highly effective in inducing potent and specific gene knockdown in recent studies by Eguchi *et al.* (2009) and Geoghegan *et al.* (2012).

The development of peptide-based delivery vectors, may seem ideal in theory – after all, cell penetrating peptides, or transduction domains, have been well documented in the last 20 years and indeed pose a beautiful alternative to viral vectors and toxic liposomal formulations. They are small (15-30 kDa) in size, are able to traverse the cell membrane with generally negligent cytotoxicity, and their generally cationic charge allows for the attachment of anionic cargo. In practice, their Achilles' heel lies in a formulation that keeps the protein both stable and active following purification from a heterologous system. Other potential issues include endosomal entrapment following internalisation, serum instability of siRNA cargo, and degradation upon systemic administration (Gooding *et al.*, 2012).

In the context of siRNA delivery, proteins to be used as vectors ought to be devoid of any nucleic acid contamination during purification, to be properly folded and retain their biological activity; thus purified through an optimised strategy, and finally be stable in a formulation that will allow complex formation with siRNA. Lastly, they should be able to bind their cargo with high affinity, transduce it intracellularly and release it into the cytoplasm upon protein-cargo-induced release from endosomal vesicles.

We have designed various DRBM1 or DRBMx2-based fusion proteins with various cell penetrating moieties, cloned them, and by studying biochemical parameters such as their molecular weight, pI and primary structure, we have purified them under both native and denaturing conditions from a bacterial cell system in suitable buffers for their isolation. Although similar with respect to their lengths, pIs and overall charge, there was no universality in the purification process. Indeed, from our observations using two well-characterised affinity tags (the poly(His) and GST tags), differences in levels of expression, solubility, stability and final yields were immense. Aggregation was by far the most common problem encountered, with yield losses that prevented some proteins, such as the Pen-DRBMx1 (C6.1 and C6.1-myc), EB1-DRBMx2 (C7.2, C9.2), DRBMx2-EB1 (C8.2, C10.2) and Tat-DRBM1 (C11.2), from being isolated for further analysis. Indeed, the optimisation of critical parameters (such as culture temperature, IPTG concentration for induction, expression time, media choice and the choice of purification strategy) is necessary for streamlined protein production. Solid-phase synthesis is at the moment, prohibitive to large-scale protein production, due to costs. Time and money permitting, the purification strategies adopted in this project would have been further optimised to include a wider variety of purification methods, such as ion exchange chromatography, the use of DNA-binding protein affinity chromatography, a wider choice of buffer pH and the

development of an efficient refolding strategy that minimized aggregation and allowed for the isolation of increased yields of properly folded protein. Although care was taken to address the innate instability of our proteins with buffers of high ionic strength and stabilising excipients (Frokjaer and Otzen, 2005, Jorgensen *et al.*, 2009, Wang, 2005), further work is needed in order to successfully generate a comprehensive purification strategy, as critical parameters need to be addressed on a case-by-case basis (Arakawa *et al.*, 2007). Further purification steps may include ion exchange chromatography via fast protein exclusion chromatography (HPLC) to remove any bound nucleic acid, as well as size exclusion chromatography to remove smaller contaminating protein species from the final purified product.

Results obtained with the Antennapedia homeodomain as an siRNA carrier in this thesis highlighted the biological activity of the AntpHD as a DNA binding protein by nature (Muller *et al.*, 1988). This conclusion was drawn upon evaluation of the C5.1 construct binding with siRNA yet co-purifying with endogenous DNA. This was an important observation that limited its use as a siRNA carrier in this project and one which should be addressed with DNase digestion during future purifications. Immunofluorescence and live cell localization assays did confirm its transducing abilities in both cell lines tested, however, its inefficient siRNA complex formation as demonstrated by EMSA assays, and the failure to induce any RNAi-induced knockdown in our reporter cell line, signified that this particular fusion protein is not an effective delivery vector. It would have been interesting to investigate the outcomes of these assays following DNA-free purification of these constructs. Endosomal entrapment upon cell entry could be attributed to the lack of endosomolytic histidines within primary amino acid sequence of C5.1 and C5.1-myc. This could be addressed by incorporating a pH-sensitive linker that improves the chances of

successful release of the protein and its cargo from endosomes upon internalisation, as demonstrated by June *et al.* (2010) and Chen *et al.* (2010). Fusion proteins containing the EB1 peptide sequence were designed to address the issue of endosomolytic entrapment; although these proteins were successfully isolated, they were subsequently lost during refolding by dialysis. With respect to these constructs, an optimised refolding strategy that limits yield losses during the last stages of refolding and storage may be beneficial; alternatives to sequential dialysis, as outlined in Chapter 3, may have allowed further evaluation of these fusion proteins.

It would have been interesting to compare the siRNA binding efficiency of these proteins in comparison to the EB1 construct by Lundberg *et al.* (2007); the incorporation of the entire DRBD may have increased the siRNA-binding capabilities, while the EB1 moiety would have addressed the issue of endosomal sequestration.

Our results with fusion proteins C5.1 (AntpHD-DRBM1), C11.2 (Tat-DRBMx2), C12.2 (Penetratin-DRBMx2) and C13.2 (DRBMx2) further reinforced the notion of cell-type specific cytotoxicity for all the proteins tested, although C12.2 elicited significantly less cytotoxicity than C11.2 in both cell lines. Moreover, the successful delivery of siRNA and the induction of statistically significant GFP expression knockdown by C12.2 in a reporter cell line, leaving the prospect of further developing Penetratin-based, rather than Tat-based recombinant proteins, open to further investigation.

C11.2-siRNA and C12.2-siRNA induced morphological changes in HepG2 cells, an observation previously unreported. Further work is necessary to understand the molecular mechanisms behind the morphological changes observed in HepG2 cells following treatment with C11.2 and C12.2, as our results were only preliminary and addressed only

phenotypic changes. Are they signs of toxicity? Is the cellular phenotype associated with pre-apoptotic or pro-proliferative pathways? It has been shown that the AntpHD induces neuronal differentiation (Cosgaya *et al.*, 1998, Joliot *et al.*, 1991a), but its effects in adherent cell lines are still unknown. Does uptake by mammalian cells induce any changes in endogenous gene expression? These are interesting questions that need to be answered if further development of these peptides is to be pursued.

# Chapter 6. Materials and Methods

## 6.1 Materials

### 6.1.1 Bacterial Strains

**Table 6.1** *E.coli* strains and associated genotypes

Strain	Genotype	Source
XL1-Blue	recA1, endA1, gyrA96, thi-1, hsdR17 supE44, relA1, lac [F' proAB lacIqZΔM15 Tn10 (Tet <sup>r</sup> )]	Stratagene (A. Markiv, U. of Westminster)
BL21(DE3)pRARE	OmpT <sup>-</sup> , Ion <sup>-</sup> , F <sup>-</sup> dcm ompT hsdS(r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) gal λ(DE3)	SGC, Oxford

### 6.1.2 Plasmids

**Table 6.2** Plasmid vectors and inserts

Name	Vector	Insert	Affinity Tag	Cloning sites	Proteolytic enzyme-Tag removal
<b>C1.1</b>	pET32-a	AntpHD-DRBM1	His <sub>10</sub>	<i>NcoI/XhoI</i>	Factor Xα
<b>C2.1</b>	pET32-a	Pen-DRBM1	His <sub>10</sub>	<i>NcoI/XhoI</i>	Factor Xα
<b>C3.1</b>	pET32-a	AntpHD-DRBM1	His <sub>6</sub>	<i>NdeI/XhoI</i>	TEV
<b>C4.1</b>	pET32-a	Pen-DRBM1	His <sub>6</sub>	<i>NdeI/XhoI</i>	TEV
<b>C5.1</b>	pGEX-6P-2	AntpHD-DRBM1	GST	<i>BamHI/XhoI</i>	Prescission protease
<b>C6.1</b>	pGEX-6P-2	Pen-DRBM1	GST	<i>BamHI/XhoI</i>	Prescission protease
<b>C5.1-myc</b>	pGEX-6P-2	AntpHD-DRBM1-myc	GST	<i>BamHI/XhoI</i>	Prescission protease
<b>C6.1-myc</b>	pGEX-6P-2	Pen-DRBM1-myc	GST	<i>BamHI/XhoI</i>	Prescission protease
<b>C7.2</b>	pGEX-6P-2	EB1-DRBMx2	GST	<i>BamHI/XhoI</i>	Prescission protease
<b>C8.2</b>	pGEX-6P-2	DRBMx2-EB1	GST	<i>BamHI/XhoI</i>	Prescission protease
<b>C9.2</b>	pET32-a	EB1-DRBMx2	His <sub>6</sub>	<i>NdeI/XhoI</i>	-
<b>C10.2</b>	pET32-a	DRBMx2-EB1	His <sub>6</sub>	<i>NdeI/XhoI</i>	-
<b>C11.2</b>	pET11-d	Tat-DRBMx2	His <sub>6</sub>	-	-
<b>C12.2</b>	pET11-d	Pen-DRBMx2	His <sub>6</sub>	-	-
<b>C13.2</b>	pET11-d	DRBMx2	His <sub>6</sub>	-	-

### 6.1.3 Cloning primers

**Table 6.3 Primer list for PCR amplification of constructs C1.1-C6.1**

<b>Construct</b>	<b>Primers (F&amp;R)</b>	<b>Forward primer sequence (5'→3')</b>	<b>Reverse primer sequence (5'→3')</b>
<b>C1.1</b>	TR7.1	GATCCCATGGGCCGC	GATCCTCGAGCTACGCTTTTT
	TR7.2	AAACGG	TG TTCAGAATTTCCACG
<b>C2.1</b>	TR7.3	GATCCATGGGCCGCC	GATCCTCGAGCTACGCTTTTT
	TR7.4	AGATTAAAATTTGGT TTC	TG TTCAGAATTTCC
<b>C3.1</b>	TR7.5	GATCCATATGGGCCG	GATCCTCGAGGGATTGGAAGT
	TR7.6	CAAACGCGG	ACAGGTTCTCGGTCGCTTTTT TG TTCAGAATTTCC
<b>C4.1</b>	TR7.7	GATCCATATGCGCCA	GATCCTCGAGGGATTGGAAGT
	TR7.6	GATTAAAATTTGGT TC	ACAGGTTCTCGGTCGCTTTTT TG TTCAGAATTTCC
<b>C5.1</b>	TR7.8	GATCGGATCCGGCC	GATCCTCGAGTCAACTAACTG
	5&6_R	GCAAACGCGG	CCTTCTTTTCCTTCAGAATTTC CACCGCCAGTTTCG
<b>C6.1</b>	TR7.10	GATCGGATCCCGCCA	GATCCTCGAGTCAACTAACTG
	5&6_R	GATTAAAATTTGGT TC	CCTTCTTTTCCTTCAGAATTTC CACCGCCAGTTTCG
<b>C5.1-myc</b>	TR7.8	GATCGGATCCGGCC	GATCCTCGAGTCAATTCAGAT
	5&6_myc	GCAAACGCGG	CCTCTTCTGAGATGAGTTTTT GTTCACTAACTGCCTTCTTTTC CTTCAGAATTTCCACCGCCAG TTTCG
<b>C6.1-myc</b>	TR7.10	GATCGGATCCCGCCA	GATCCTCGAGTCAATTCAGAT
	5&6_myc	GATTAAAATTTGGT TC	CCTCTTCTGAGATGAGTTTTT GTTCACTAACTGCCTTCTTTTC CTTCAGAATTTCCACCGCCAG TTTCG

**Table 6.4 Primer list for PCR amplification of constructs C7.2-C10.2**

<b>Construct</b>	<b>Forward primer sequence (5'→3')</b>	<b>Reverse primer sequence (5'→3')</b>
<b>C7.2</b>	GATCCATATGCTGATTCG CCTGTGGAGC	GATCCTCGAGGTCAGATTTCACTGA GGTTTCTTCT
<b>C8.2</b>	GATCCATATGGCTGGTGA TCTTTCAGCAGG	GATCCTCGAGTTTTTTTTTCCATTCA GGCG
<b>C9.2</b>	GATCCATATGCTGATTCG CCTGTGGAGC	GATCCTCGAGGTCAGATTTCACTGA GGTTTCTTCT
<b>C10.2</b>	GATCCATATGGCTGGTGA TCTTTCAGCAGG	GATCCTCGAGTTTTTTTTTCCATTCA GGCG

## 6.1.4 Sequencing primers

**Table 6.5 Sequencing primer list**

<b>Construct</b>	<b>Primer</b>	<b>Primer sequence (5'→3')</b>
<b>C1.1</b>	T7minus1	AATACGACTCACTATAGGG
<b>C2.1</b>	T7minus1	AATACGACTCACTATAGGG
<b>C3.1</b>	T7minus1	AATACGACTCACTATAGGG
<b>C4.1</b>	T7minus1	AATACGACTCACTATAGGG
<b>C5.1</b>	pGEX5-FP	AACGTATTGAAGCTATCCC
<b>C6.1</b>	pGEX5-FP	AACGTATTGAAGCTATCCC
<b>C5.1-myc</b>	5&6_ myc	GATCGGATCCGGCCGCAAACGCGGGATCCTCG AGTCAATTCAGATCCTCTTCTGAGATGAGTTTTT GTTCACTAACTGCCTTCTTTTCCTTCAGAATTC CACCGCCAGTTTCG
<b>C6.1-myc</b>	5&6_ myc	GATCGGATCCCGCCAGATTA AAAATTTGGTTTCG ATCCTCGAGTCAATTCAGATCCTCTTCTGAGAT GAGTTTTTGTTCCTAACTGCCTTCTTTTCCTTC AGAATTTCCACCGCCAGTTTCG
<b>C7.2</b>	T7minus1	AATACGACTCACTATAGGG
<b>C8.2</b>	T7minus1	AATACGACTCACTATAGGG
<b>C9.2</b>	pGEX5-FP	AACGTATTGAAGCTATCCC
<b>C10.2</b>	pGEX5-FP	AACGTATTGAAGCTATCCC
<b>C11.2</b>	pUC-FP	GCCAGTGAATTCGAGCTCGG
<b>C12.2</b>	pUC-FP	GCCAGTGAATTCGAGCTCGG
<b>C13.2</b>	pUC-FP	GCCAGTGAATTCGAGCTCGG

## 6.1.5 Protein purification buffers, media, solutions

### 6.1.5.1 General Media and Antibiotics

XL1-Blue and BL21(DE3)pRare cells were grown in Luria Bertani broth: 1% (w/v)

Tryptone, 0.5% (w/v) NaCl, 0.5% (w/v) yeast extract.

**Table 6.6 Antibiotics used in bacterial cell cultures or in cell culture**

Name	Stock concentration (w/v)	Working concentration	Source
<b>Ampicillin</b>	100 mg/ml	100µg/ml	Sigma
<b>Chloramphenicol</b>	34 mg/ml in 80% EtOH	34 µg/ml	Sigma
<b>Kanamycin</b>	10 mg/ml	10 µg/ml	Sigma
<b>Geneticin (G418)</b>	50 mg/ml	0.4-2 mg/ml	Sigma

**Table 6.7 General buffers and solutions<sup>2</sup>**

Name	Composition
<b>10x SDS running buffer</b>	250 mM Tris-base, 2.5 M glycine, 1% (w/v) SDS, in ddH <sub>2</sub> O, pH 8.3
<b>Coomassie total protein stain</b>	0.05% -0.1% (w/v) Coomassie Brilliant blue, 50% (v/v) Methanol, 10% (v/v) Acetic acid in ddH <sub>2</sub> O
<b>Destaining solution</b>	7% (v/v) Acetic acid in ddH <sub>2</sub> O
<b>5x Transfer buffer</b>	960 mM Glycine, 120 mM Tris-base in ddH <sub>2</sub> O
<b>1x Transfer buffer</b>	20% (v/v) 5x Transfer buffer, 10% (v/v) Methanol
<b>10x TBS</b>	200 mM Tris-base 1.5 M NaCl, (pH 7.6) in ddH <sub>2</sub> O
<b>TBS-T</b>	10% (v/v) 10x TBS, 0.05% (v/v) Tween-20 in ddH <sub>2</sub> O

<sup>2</sup> General molecular biology reagents were from Fisher Scientific (Loughborough, UK) unless otherwise stated.

**Table 6.8 Reagents used in Polyacrylamide gel electrophoresis (SDS-PAGE)**

Reagent	12 % Resolving gel(x2) (ml)	15 % Resolving gel (x2) (ml)	4% Stacking gel (x2) (ml)
ddH <sub>2</sub> O	4.9	4.15	3.6
1.5 M Tris-base pH 8.8	2.6	2.6	-
0.5 M Tris-base, pH 6.8	-	-	0.625
10% (w/v) sodium dodecyl sulfate	0.1	0.1	0.05
40% (w/v) Bis-Acrylamide	3.0	3.75	0.5
10% (v/v) ammonium persulfate	0.1	0.1	0.05
TEMED	0.02	0.02	0.01

**6.1.5.2 Protein purification buffers****Table 6.9 Protein purification buffers used in IMAC under native conditions**

Name	Composition
His-Binding buffer	50 mM HEPES, 500 mM NaCl, 5 mM Imidazole, 5% (v/v) Glycerol
His-Wash buffer	50 mM HEPES, 500 mM NaCl, 30 mM Imidazole, 5% (v/v) Glycerol
His-Elution buffer	50 mM HEPES, 500 mM NaCl, 250 mM Imidazole, 5% (v/v) Glycerol

**Table 6.10 Protein purification buffers used in GST affinity chromatography**

Name	Composition
GST-Binding buffer	50mM Tris-HCl, 500 mM NaCl, 10 mM DTT, 5% (v/v) Glycerol, pH 7.5
GST-Wash buffer	50mM Tris-HCl, 500 mM NaCl, 10 mM DTT, 5% (v/v) Glycerol, pH 7.5
GST-Elution buffer	50mM Tris-HCl, 500 mM NaCl, 10 mM DTT, 5% (v/v) Glycerol, 10 mM reduced glutathione pH 8.0,
GST-Equilibration (cleavage) buffer	50mM Tris-HCl, 200 mM NaCl, 2 mM DTT, 5% (v/v) Glycerol, pH 7.5
GST-Storage buffer	PBS, 10% (v/v) Glycerol, pH 7.4

**Table 6.11 Protein purification buffers used in IMAC under denaturing conditions**

<b>Name</b>	<b>Composition</b>
<b>His-Resuspension buffer</b>	50 mM sodium phosphate (pH 7.4)
<b>Additives (per 0.5L)</b>	1 tablet complete mini protease inhibitor (without EDTA), 1200 U benzonase, 30,000 U lysozyme, 2 mM MgCl <sub>2</sub>
<b>1.1x His-Guanidine binding buffer</b>	50 mM sodium phosphate, 550 mM NaCl, 22.2 mM imidazole, 6.66 M Guanidine-HCl, 2 mM β-mercaptoethanol (pH 7.4)
<b>His-Wash buffer</b>	50 mM sodium phosphate, 550 mM NaCl, 22.2 mM imidazole, 6.66 M Guanidine-HCl, 2 mM β-mercaptoethanol (pH 7.4)
<b>His-Elution buffer</b>	50 mM sodium phosphate, 500 mM NaCl, 150 mM imidazole (pH 7.4)
<b>His-Refolding buffer</b>	50 mM sodium phosphate (pH 7.4), 500 mM NaCl, 20 mM imidazole (pH 7.4)
<b>Storage buffer</b>	25 mM Hepes, 150 mM NaCl, 10% glycerol (pH 7.4)

## **6.2 Methods**

### **6.2.1 Plasmids**

The AntpHD-DRBM1 plasmid was obtained from Epoch Biosciences (Missouri City, TX, USA). The pd1-deGFP-N1 plasmid was kindly provided by Drs Chang and Moore (University of Pittsburgh). Modified pET32-a lacking the thioredoxin (trx) tag was a kind gift from Dr Markiv (University of Westminster). EB1-DRBMx2 and DRBMx2-EB1 were from Genscript (Piscataway, NJ, USA). The Tat-DRBMx2, Pen-DRBMx2 and DRBMx2 plasmids were a kind gift from Prof. Beverly (University of Iowa).

### **6.2.2 Competent *E. coli* strain preparation**

XL1-Blue and BL21(DE3)pRARE cells were grown in 200 ml fresh LB broth inoculated with 10 ml starter cultures in the absence of antibiotics at 37°C and 250 rpm in a shaking incubator, until an O.D<sub>595</sub> ~0.4 was reached. The flask was then chilled on ice for 30 minutes and 50 ml were aliquoted into pre-chilled sterile centrifuge tubes. Cells were harvested by centrifugation for 10 minutes at 4,500 rpm at 4°C. Supernatants were discarded and each pellet was resuspended in 12.5 ml 0.1 M MgCl<sub>2</sub>. Resuspended cells were centrifuged again as before and the pellets were resuspended in 25 ml 0.1 M CaCl<sub>2</sub>. Cells were incubated on ice for 30 minutes, centrifuged again and the supernatant was discarded. Pellets were resuspended in 700 µl 0.1 M CaCl<sub>2</sub> and 300 µl 50% glycerol, before being stored as 50 µl aliquots at -80°C.

### **6.2.3 cDNA Cloning by PCR**

For C1.1-C6.1, a synthetic cDNA template encoding for the entire Antennapedia homeodomain (AntpHD) linked to the dsRNA binding domain 1 (DRBM1) from human PKR by a small flexible linker ((G<sub>4</sub>S)<sub>2</sub>), was obtained as an *EcoRV* fragment inserted into

a modified pBSK plasmid lacking its MCS (Epoch Biosciences, (Missouri City, TX, USA) and used as the template for the construction of 8 new constructs, encoding for either the AntpHD-DRBM1 or Penetratin-DRBM1. AntpHD-DRBM1 and Pen-DRBM1 were cloned by PCR using Platinum *Pfx* DNA Polymerase (Invitrogen, Paisley, UK) as *NcoI/XhoI* fragments with an *N*-terminus His<sub>10</sub> tag and a Factor-X $\alpha$  cleavage site for the tag's subsequent removal following purification (His<sub>10</sub>-HQEA-AntpHD/Pen-DRBM1, where -HQEA- corresponds to the Factor X $\alpha$  protease cleavage site). Similarly, they were also cloned as *NdeI/XhoI* fragments with a *C*-terminus His<sub>6</sub> tag separated from the fusion peptide by a TEV protease cleavage site (AntpHD/Pen-DRBM1-ENLYFQ(G/S)-His<sub>6</sub>, where ENLYFQ(G/S)- corresponds to the TEV cleavage site). An *N*-terminally GST tagged construct was generated by inserting the PCR product into a vector pGEX-6P-2 (GE Healthcare, Uppsala, Sweden). Both AntpHD-DRBM1 and Pen-DRBM1 were cloned as *BamHI/XhoI* fragments with an *N*-terminus Prescission protease (PP) cleavage site (-LEVLFQGP-). The c-myc epitope tag (-EQKLISEEDLN-) was added to the *C*-terminus of GST-AntpHD-DRBM1 and GST-Pen-DRBM1 using the amplified GST-AntpHD-DRBM1 sequence as a template. 10  $\mu$ M of custom-designed oligonucleotide primers were used for this purpose on an MJ Mini Thermal Cycler (Biorad, Hertfordshire, UK). The cDNA templates for EB1-DRBMx2 and DRBMx2-EB1 were obtained as synthetic plasmids with a pUC57 backbone from Genscript (Piscataway, NJ, USA) and sequenced with a universal pUC57 forward and reverse primer. *NdeI/XhoI* or *BamHI/XhoI* cleavage sites inserted by PCR and amplified fragments were cloned into a modified pET32-a and pGEX-6P-2, respectively. Thermocycler reaction conditions were optimised for each construct by varying the annealing temperature of the primers and the absence or addition of enhancer, accordingly.

#### **6.2.4 Cloning into plasmid vectors**

Purified target sequences with PCR-inserted restriction sites were double-digested with 1-2 units of appropriate restriction endonucleases (REs) (New England Biolabs, Herts, UK) in a final volume of 50 µl prior to ligation with plasmid vectors. *N*-terminus His<sub>10</sub>-tagged AntpHD-DRBM1 and Pen-DRBM1 were double-digested with *NcoI/XhoI* for 60 minutes at 37°C while constructs cloned with a *C*-terminus His<sub>6</sub>-tag were digested with *NdeI/XhoI*. A modified pET32-a vector plasmid, was also digested with either *NcoI/XhoI* or *NdeI/XhoI* enzymes to create appropriate sticky ends for ligation. Constructs to be inserted into pGEX-6P-2 (GE Healthcare, Uppsala, Sweden) were digested with *BamHI/XhoI* under the same conditions as before. 2 µg of each plasmid vector were also digested with the same REs. Ligations were done with a T4 DNA ligase as part of a Rapid Ligation kit (Roche, Sussex, UK) at insert-to-vector molar ratios of 3:1 and 5:1, respectively, for 2 hours at 25°C, using a linearised plasmid vector (minimum 50 ng) and digested insert. Their concentration and purities assessed by a Nanodrop 1000 micro-volume spectrophotometer (Thermo Scientific, DE, USA) at 260 nm and 280 nm.

#### **6.2.5 Gel extraction and ligation into plasmid vectors**

For gels bands to be blind-excised and purified, samples were run in duplicate on the same gel, and the gel was halved prior to visualization under UV light, to avoid inserting mutations.

Digested vector plasmids were isolated by gel blind-excision, solubilised and eluted in 50 µl EB buffer using a QiaQuick Gel Extraction kit (Qiagen, Crawley, UK) according to the manufacturer's specifications. Digested PCR fragments were ligated into either a modified pET32-a vector or a pGEX-6P-2 vector using T4 DNA ligase, respectively, as before, for

5-30 minutes at 25°C. The resulting ligation mixes were used to transform competent XL1-Blue *E. coli*, which were then aseptically plated on amp<sup>+</sup> (100 µg/ml) plates overnight at 37°C. Similarly, Tat-DRBMx2-His<sub>6</sub>, Pen-DRBMx2-His<sub>6</sub> and DRBMx2-His<sub>6</sub> plasmids ligated with pET11-d were used to transform competent XL1-Blue cells as before.

### **6.2.6 Transformation of competent cells**

50 µl competent *E. coli* cells were thawed on ice for 5 minutes. 1-2 µl plasmid DNA (1-2 µg/µl) was added, and cells were incubated on ice for 5 minutes. For transformation, cells were heat-shocked for 60 seconds at 42°C and allowed to recuperate on ice for 5 minutes. 250 µl SOC buffer was added and cells were incubated in a shaking incubator at 37°C and 250 rpm for 60 minutes. Transformed cells were plated on agar plates containing the antibiotic corresponding to the plasmid's resistance marker, and allowed to grow at 37°C for 16 hours. For plasmid propagation, the *E. coli* propagation strain XL1-Blue was used. For protein expression, the BL21(DE3)pRARE strain was used, which is chloramphenicol resistant. It contains plasmids with genes encoding for rare codon tRNAs (*argU*, *argW*, *glyT*, *ileX*, *leuW*, *metT*, *proL*, *thrT*, *thrU* and *tyrU*) for the rare codons AGA and AGG (Arginine), GGA (Glycine), AUA (Isoleucine), CUA (Leucine), AUG (Methionine), CCC (Proline), ACA and ACC (Threonine) and UAC (Tyrosine), as well as reduced mRNA degradation. The DE3 lysogen encodes for the T7 polymerase, which is used to induce transcription and protein expression from the T7 promoter in *E. coli*.

### **6.2.7 Plasmid propagation**

To obtain stock DNA, 1-1.5 µl sequenced DNA (≥50 ng/µl) was used to transform 50 µl of competent XL1-Blue *E. coli*, which were then plated on ampicillin-containing agar plates (100 µg/ml) and incubated overnight at 37°C. 10 ml LB broth starter cultures supplemented

with 100 µg/ml ampicillin were inoculated with 1 colony from each plate and incubated overnight at 37°C and 250 rpm on a shaking incubator. DNA was purified from the centrifuged cell pellet using a QiaPrep Spin Miniprep kit into 50.0 µl Elution buffer (EB), according to the manufacturer's specifications (Qiagen, Crawley, UK).

### **6.2.8 Agarose gel electrophoresis**

The correct size of amplicons was confirmed by agarose gel electrophoresis on 2% gels in 1xTAE Buffer. PCR-amplified sequence bands corresponding to the correct size of insert were pooled together, and treated with *DpnI* endonuclease (recognition site 5'-Gm<sup>6</sup>ATC-3') to remove methylated/hemimethylated bases on the template DNA strand (as DNA isolated from most *E.coli* strains is *dam* methylated). PCR fragments were recovered with a QiaQuick PCR purification kit (Qiagen, Crawley, UK) in 50.0 µl EB buffer (10 mM Tris-Cl, pH 8.5) according to the manufacturer's specifications. Their concentration and purity were assessed at 260 nm and 280 nm.

2% Agarose gels were prepared by dissolving 1.2 g of Agarose (Fisher Scientific, Loughborough, UK) in 60 ml 1x TAE buffer (40 mM Tris, 20 mM Acetic acid, 1 mM EDTA) diluted from 50x stock (2 M Tris-base, 0.5 M Disodium EDTA (pH 8.0), 1M acetic acid) and heating in a microwave until all solid had dissolved . Prior to pouring, 0.6 µl SYBR Safe DNA staining dye (Life Technologies, CA, US) was added and the gel was allowed to solidify in its gel tank. 20-25 µl samples with 6x loading dye and 5 µl marker (0.4 µl of 1 kb PLUS Marker (Invitrogen, Paisley, UK) diluted into 4.0 µl 1xTAE buffer and 1.0 µl 6x loading dye) were loaded. Gels were run at 100V for 60 minutes.

### **6.2.9 DNA purifications**

Following transformation of XL1-Blue cells and plating on agar plates containing the appropriate antibiotic, one colony was used to inoculate 10 ml Luria-Bertani (LB) broth (Fisher Scientific, Loughborough, UK) supplemented with ampicillin. Cells were allowed to grow overnight at 37°C and 250 rpm on a shaking incubator for a maximum of 16 hours. Cells were then pelleted by centrifugation at 4,500 rpm for 10 minutes, and purified from silica membrane columns as per the manufacturer's instructions using a QiaPrep spin Miniprep kit (Qiagen, Crawley, UK). DNA was eluted in 50 µl EB buffer. Concentrations and purity were assessed using a Nanodrop 1000 (Thermo Scientific, DE, USA) at 260 nm and 280 nm.

### **6.2.10 Validation of target sequence amplification by colony PCR**

To confirm that transformed XL1-Blue cells contained the gene of interest, and not empty vector plasmids, three colonies from each plate were used to inoculate 3 x 10 µl ddH<sub>2</sub>O, and colony PCR was performed on 1.0 µl template DNA from a diluted colony using Biotaq Red DNA polymerase (Bioline, London, UK) with 10 µM gene-specific forward and reverse primers. Amplified fragments were analysed on a 2% agarose gel electrophoresis at 100V for 35 minutes and visualized under UV light for correct size. Colony mixes with amplicons showing correct insert sizes were then used to inoculate 10 ml Luria-Bertani (LB) broth starter cultures supplemented with 100 µg/ml ampicillin overnight at 37°C and 250 rpm for a maximum of 16 hours. The next day, cells were pelleted by centrifugation at 4,500 rpm on a bench-top centrifuge for 10 minutes at 25°C. Plasmid DNA was extracted and purified from the harvested pellet via alkaline lysis using a Qiaprep Spin Miniprep kit on a bench-top centrifuge (Qiagen, Crawley, UK), according to the manufacturer's specifications and eluted into 50 µl EB buffer. Sample concentration

was assessed with a Nanodrop, before storing at -20°C for future use. To confirm that transformed cells contained the fusion gene of interest and not template DNA, 1.0 µg miniprep DNA from each construct was digested with *NdeI/XhoI*, *NcoI/XhoI* or *BamHI/XhoI*, according to its amplified restriction sites, and analysed by agarose gel electrophoresis. Finally, 100 ng/µl purified DNA from each sample cloned into pET32-a was sequenced using 5 µM T7-minus1 primer corresponding to the T7 promoter. In the case of pGEX-6P-2-ligated inserts, a pGEX5' universal forward primer was used (GATC Biotech, London, UK). Chromatogram files were analysed on Finch TV software (v.1.4.0, Geospiza Inc, US).

#### **6.2.11 Small scale purifications of C1.1-C4.1 by immobilized metal ion chromatography (IMAC) under native conditions**

Purified plasmids from XL1-Blue containing constructs of interest were used to transform the expression strain BL21(DE3)pRARE. Cells were plated on LB Miller agar plates (Fisher, Loughborough, UK) containing 34 µg/ml chloramphenicol and 100 µg/ml ampicillin and incubated overnight at 37°C. The next day, 10 ml LB Lennox broth (Fisher Scientific, Loughborough, UK) supplemented with Chl/Amp were aseptically inoculated with one colony from each plate and incubated at 37°C on a rotating shaker (250 rpm) for a maximum of 16 hours. To assess soluble protein expression on a 50 ml scale, 1 ml from 10 ml starter cultures was used to inoculate 50 ml LB broth supplemented with Chl/Amp and incubated on a Innova 43 rotating shaker (New Brunswick Biosciences, CT, USA) at 37°C and 250 rpm to an O.D<sub>600</sub> ~0.3, after which temperature was decreased to 18°C until O.D<sub>600</sub>~0.7. Protein production was induced with 0.5 mM IPTG at 18°C overnight. Cultures were pelleted by centrifugation at 4,500 rpm, re-suspended in 5 ml His-Binding buffer on ice and mechanically lysed by sonication for 5 minutes at 10 second intervals

followed by a 30 second rest period. 100 µl supernatant was collected as the 'Total' fraction (in 2x v/v 6 M Urea). Lysate was cleared by centrifugation at 13,000 rpm and 4°C, and His-tagged constructs were purified under native conditions by IMAC over a nickel-nitrilotriacetic acid (Nickel-NTA) matrix. Proteins were then purified by gravity filtration through a 0.2 ml b.v of IMAC Sepharose 6 Fast Flow resin (GE Healthcare, Uppsala, Sweden) that had been primed with 5x column volumes (c.v) of 0.2 M NiSO<sub>4</sub>•6H<sub>2</sub>O and equilibrated with 5 ml His-Binding buffer at 4°C. The soluble fraction was collected ('Flowthrough'). The resin was then washed with 5 ml His-Wash buffer and the flowthrough was collected as 'Wash'. Finally, the proteins were eluted from the metal ion matrix with 2x 200 µl Elution buffer as 'Eluates 1 and 2'. Samples were analysed by the method of Laemmli (1970) on 12% running and 4% stacking 0.75 mm polyacrylamide gels at 160-200 V for 45 minutes using a Mini-Protean Tetra Cell (BioRad, Herfordshire, UK) and stained with 0.5% (v/v) Coomassie Blue and destained with 7% (v/v) acetic acid.

#### **6.2.12 Small scale purifications of GST-tagged proteins by affinity chromatography**

GST-tagged peptides from 50 ml bacterial cultures were purified over glutathione-sepharose columns using a 0.4 ml b.v. of Glutathione Sepharose 4B (GS4B) resin (GE Healthcare, Uppsala, Sweden) pre-equilibrated with 5 ml GST-Binding buffer. The cell pellet was lysed by sonication (73mm probe, power 40%, 10 second cycles) and the lysate was cleared by centrifugation at 16,000 x g and 4°C. Cleared lysate was applied to the columns and collected as 'Flowthrough' by gravity flow. Columns were washed with 5 ml GST-binding buffer and the flowthrough was collected as 'Wash'. Peptides were eluted 2x with 200 µl GST-elution buffer . All samples were analysed by SDS-PAGE as before.

### **6.2.13 Large scale purifications of GST-tagged C5.1, C6.1, C5.1-myc and C6.1-myc**

Glycerol stocks from BL21(DE3)pRare were used to inoculate 10 ml  $\text{Chl}^+/\text{Amp}^+$  LB starter cultures overnight, which were then transferred to 1L LB broth in 3L Erlenmeyer flasks ( $\text{Chl}^+/\text{Amp}^+$ ). Cultures were grown to an  $\text{O.D}_{600} \sim 0.3$  at  $37^\circ\text{C}$  and then at  $18^\circ\text{C}$  until  $\text{O.D}_{600} \sim 0.7$  before induction with 0.5 mM IPTG overnight on a rotating shaker at  $18^\circ\text{C}$  (180 rpm). The next day, cultures were transferred to 800 ml rotor flasks and spun for 30 minutes at 4,500 rpm. The supernatant was discarded, and the pellet was again spun in 50 ml centrifuge tubes (Fisherbrand, Loughborough, UK) for 20 minutes at  $4^\circ\text{C}$ . Pellets were resuspended in a total volume of 25 ml GST-Binding buffer and sonicated for 10 minutes on ice (73 mm probe, power 40%, 10 second cycles). A 100  $\mu\text{l}$  aliquot in 200  $\mu\text{l}$  6 M Urea was collected as the 'Total' fraction. Lysates were cleared in 80 ml round-bottom flasks by centrifugation at  $24,000 \times g$  at  $4^\circ\text{C}$  for 30 minutes.

For a bed volume (b.v) of 4 ml, 5 ml Glutathione Sepharose 4B (GS4B) resin in 20 % (v/v) ethanol (GE Healthcare, Uppsala, Sweden) was centrifuged at  $500 \times g$  and resuspended in 5x b.v GST-binding buffer and spun again. The resin was resuspended in 4 ml (1x b.v) GST-binding buffer and kept on ice.

For batch-binding, cleared lysate was applied to the resin and incubated on a rotating mixer for at 2 hours at  $4^\circ\text{C}$  and centrifuged at  $500 \times g$  for 15 minutes or until the supernatant was clear. The supernatant was discarded, while the resin was resuspended in 40 ml GST-binding buffer, and sequentially transferred as 10 ml aliquots to a 20 ml glass column and collected as 'Flowthrough 1'. The column was then washed 1x with 40 ml GST-binding buffer ('Wash 1'), and 3x with 10 ml, which were then pooled ('Wash 2'). For on-column cleavage of the GST tag, Prescission protease in GST binding buffer was applied to the

resin at a 1:5 enzyme:protein ratio in a sealed column overnight on a rotating mixer at 4°C. The next day, the column was mounted and the flowthrough ('Flowthrough 2') was collected. Cleaved protein was eluted from the column sequentially with 4x 8 ml GST-binding buffer ('Eluate 1-4'). A final elution step with 10 ml GST-Elution buffer released cleaved GST, Prescission protease and any uncleaved protein. All samples were assessed for purity and concentration by Nanodrop, and analysed on either 12% or 15% polyacrylamide gels at 180V for 45 minutes. Samples containing the cleaved protein of interest were pooled together (FT2, E1-E4, or a combination of the above) and concentrated over an Amicon Ultra-15 Centrifugal Filter device (3K MWCO) (Millipore, MA, USA) and a fixed-angle rotor at 5,000 x g, according to the manufacturer's specifications. Concentrations were assessed by Nanodrop and concentrated samples were buffer exchanged into either 10% (v/v) Glycerol/PBS or 1xPBS in 5 ml Zeba Desalt Spin columns (Thermo Scientific, DE, USA) and analysed by Nanodrop using the expected molecular weight (MW) and extinction coefficient ( $\epsilon$ ) and SDS-PAGE followed by Coomassie staining or Western blotting. Proteins were aliquoted into 1 ml and 200  $\mu$ l stocks and stored at -80°C for future use.

#### **6.2.14 Denaturing cobalt IMAC purification of His-tagged proteins**

1L bacterial cultures for EB1-DRBMx2-His<sub>6</sub> (C7.2) and DRBMx2-EB1-His<sub>6</sub> (C8.2) and 0.5L cultures for Tat-DRBMx2-His<sub>6</sub> (C11.2), Pen-DRBMx2-His<sub>6</sub> (C12.2) and DRBMx2-His<sub>6</sub> (C13.2) were grown at either 18°C or 37°C. C7.2 and C8.2 were grown for 8 hours, whereas C11.2-C13.2 for 4 hours on a shaking incubator. The following reagents were used on 0.5L cultures, and were doubled for 1L cultures. Cells were pelleted by centrifugation as before. For 0.5 L cultures, pellets were resuspended in 5 ml His-resuspension buffer and quick-frozen at -80°C for 15 minutes. Cells were then thawed and

incubated with 1 Complete Mini Protease Inhibitor tablet (Roche, Manheim, Germany), 1200 U benzonase, 30,000 U rLysozyme, 2 mM MgCl<sub>2</sub> on ice for 30 minutes. 10 ml of 1.1x binding buffer was added and cells were lysed by sonication for 7 minutes over ice (73 mm probe, power 40%, 20 second cycles, with 60 seconds rest). 35 ml chilled 1.1x binding buffer was subsequently added with 3 mM β-Mercaptoethanol for a final volume of 50 ml. Lysates were cleared by centrifugation at 16,000 x g for 40 minutes at 4°C. The supernatant was filtered through a 0.45 μM Millex PVDF filter (Millipore, MA, USA) and protein was purified on nitriloacetic acid columns pre-charged with 0.1 M CoCl<sub>2</sub> and pre-equilibrated with 5 b.v of elution. Following elution, proteins were sequentially dialysed in a SnakeSkin Dialysis tubing (Thermo Scientific, DE, US) into 5 L refolding buffer and then into 5 L storage buffer. Dialysed protein was briefly centrifuged at 13,000 x g for 2 minutes to remove any precipitated protein, aliquoted and stored at -80°C.

#### **6.2.15 Trichloroacetic Acid – Ethanol (TCA-EtOH) Protein Purification**

Samples from large-scale (0.5L) purifications under denaturing conditions were isolated in buffers containing 6-6.66 M Gu-HCl and 2 mM β-mercaptoethanol. Gu-HCl, unlike Urea, is not soluble in SDS and cannot be resolved by SDS-PAGE. To remove guanidinium ions, 25 μl samples were diluted to 100 μl in ddH<sub>2</sub>O and 100 μl 20% (v/v) TCA was added to samples on ice for 20 minutes. Samples were centrifuged at 13,000 rpm at 4°C for 15 minutes in a bench-top centrifuge and the supernatant was discarded. The protein pellet was washed with 100 μl ice-cold EtOH, air-dried for 30 minutes at room temperature and resuspended in 6x Laemmli sample buffer supplemented with 5% (w/v) SDS. A yellow color signified acidification by TCA and therefore samples were titrated with 1 N NaOH until the blue color was restored. Samples were then boiled at 95°C on a heat block for 5 minutes prior to analysis by SDS-PAGE.

### **6.2.16 Polyacrylamide Gel Electrophoresis, protein transfer and Western Blotting**

For all experiments, proteins were resolved using the method of Laemmli (1970). Briefly, 10-40 µg protein were diluted into 6x loading Dye supplemented with 150 mM Dithiothreitol (DTT) and heated to 95°C for 3 minutes in a Biorad PCR block prior to loading (except in the case of whole cell lysate samples, which were destined for blotting with an anti-PTP1B antibody). Samples were then loaded on either a 12% or 15% polyacrylamide gels with 4% stacking gels with an EZ-Run Pre-stained marker (Fisher Scientific). Gels were run for 60 minutes at 120-160V at 25°C (RT) in a BioRad Mini Electrophoresis tank in 1x gel running buffer (from 10x Stock; 250 mM Glycine pH 8.3, 2.5 M Tris-Base, 1% (w/v) SDS, pH 8.3) and were either stained with 0.5% (w/v) Coomassie blue total protein stain, or transferred to a PVDF membrane for Western blotting.

For electroblotting (protein transfer), the assembly cassette was setup in chilled 1x Transfer buffer while the membrane was activated in 100% (v/v) methanol for 5 minutes at RT followed by a brief wash with ddH<sub>2</sub>O and a 5 minute incubation in 1x Transfer buffer. Proteins were transferred overnight at 30V and 4°C in a Criterion Blotter tank (BioRad, Herts, UK). Once the proteins had transferred, the cassette was disassembled and the membranes were washed with ddH<sub>2</sub>O and blocked for 1 hour at 25°C in 5% (w/v) Milk/TBS (diluted to 10% (v/v) from 10x Stock TBS) with gentle shaking or 4% (v/v) BSA/TBS for 45 minutes at RT on a shaking platform, and incubated overnight at 4°C with primary antibody at a 1:1,000 dilution in 5% (w/v) Milk/TBS or 5% (v/v) BSA/TBS. The primary antibody was then removed, and the membrane was washed 2x with TBS buffer for 5 minutes each, before incubation with an HRP-conjugated secondary antibody in 5% (w/v) milk/TBS for 60 minutes at RT on a shaking platform. The secondary

antibody was then washed away with 3x-6x washes with TBST buffer (TBS buffer supplemented with 0.05% (v/v) Tween-20) for 5 minutes each. Finally, membranes were visualized by chemilluminescence using Labworks 4.1 software (Perkin-Elmer, Cambridge, UK) after a 3 minute incubation with 0.8 ml ECL reagent (SuperSignal West Femto Chemiluminescent Substrate, Thermo Scientific, DE, USA) prepared at a 1:1 (v/v) ratio of luminol:peroxidase. For PTP1B, an  $\alpha$ -PTP1B IgG2a primary antibody from mouse (Cat. 610140, BD Biosciences, Oxford, UK) was used in 5% (w/v) milk/TBS at a 1:1,000 dilution with an anti-mouse HRP conjugated IgG whole molecule peroxidase at a 1:5,000 dilution (Cat. A4416, Sigma, Dorset, UK). For the myc-epitope tag, a mouse anti-myc antibody (Clone 4A6, Millipore, MA, USA) was used in 5% (w/v) Milk/TBS at a 1:1,000 dilution. For the His<sub>6</sub> tag, an  $\alpha$ -His IgG1 antibody from mouse was used (Cat. 34660, Qiagen, Crawley, UK) at a 1:1,000 dilution in 5% (w/v) Milk/TBS and an anti-mouse HRP-conjugated (IgG whole molecule peroxidase) secondary antibody (Cat. A4416, Sigma, Dorset, UK), at a 1:1,500 dilution in 5% (w/v) Milk/TBS. For the actin loading control, an  $\alpha$ -Actin affinity purified primary antibody from rabbit (Cat. 2066, Sigma, Dorset, UK) was used in 5% BSA-TBS and developed with an HRP-conjugated whole IgG peroxidase secondary antibody at a 1:10,000 dilution in 5% BSA-TBS. Detection by chemilluminescence was done as before.

#### **6.2.17 Electromobility shift assay (EMSA) with C5.1, C12.2 and C13.2**

The potential of fusion proteins to bind dsRNA was assessed by electrophoretic mobility shift assay on a 6% Native PAGE gels (1.5 ml 40% (w/w) Bis-Acrylamide, 0.5 ml 10x Tris-Borate EDTA buffer, 7.89 ml ddH<sub>2</sub>O, 100.0  $\mu$ l 10% (w/v) APS, 10  $\mu$ l TEMED). Glassware was thoroughly cleaned with 70% (v/v) Ethanol followed by RNase-free water. The gel was prepared, poured, and allowed to solidify for 3 hours at 4°C. For the protein-

dsRNA coupling reaction in a 10  $\mu$ l final volume, 10 pmol of 10  $\mu$ M dsRNA (Stealth RNAi GFP reporter control, Life Technologies, CA, US) was diluted in complexation buffer (1xPBS for C5.1 and 25 mM HEPES (pH 7.4) for C12.2 and C13.2. Various amounts of protein were added to the dsRNA mix at molar concentrations calculated from values obtained by the BCA assay. The final dsRNA concentration was therefore 1  $\mu$ M. Protein-siRNA complex formation was allowed to proceed for 20 minutes on ice. 2 $\mu$ l sterilized 40% (w/v) sucrose was added. Gels were run at 4°C in 0.5x TBE buffer for 60-90 minutes at 90V. Gels were then stained with 1% (w/v) EtBr for 10 minutes, washed in ddH<sub>2</sub>O for 5 minutes and visualized under UV light.

#### **6.2.18 Cell culture**

Human Embryonic Kidney (HEK) 293 (ATCC #CRL-1573) were a kind gift from Professor Tinker (QMUL, London, UK). Human hepatocellular carcinoma (HepG2) cells (ATCC # HB-8065) were obtained from the American Type Culture Collection (LGC Standards, Middlesex, UK). HEK293 cells were routinely grown in Dulbecco's Modified Eagle Medium (DMEM) GlutaMax (Gibco, Paisley, UK), supplemented with 10% (w/v) fetal bovine serum (FBS) (Sigma, Dorset, UK). For the generation of the HEK293 reporter cell line stably expressing dEGFP, HEK293 cells were cultured in DMEM-AQ media (4,500 mg/L glucose, L-alanyl-glutamine, and sodium bicarbonate) (D0819, Sigma, Dorset, UK) supplemented with 1 mM Sodium Pyruvate, 10% FBS and penicillin-streptomycin at a final concentration of 100 U/ml and 100  $\mu$ g/ml, respectively (1x) (P4333, Sigma, Dorset, UK) in a humidified atmosphere with 5% CO<sub>2</sub>, to 80% confluency. Selection was done with 0.4-2 mg/ml G418 (Sigma, Dorset, UK). Stably expressing HEK293-dEGFP clones were maintained with 1.8 mg/ml G418. HepG2 cells and grown in DMEM (1.5 mg/L glucose, Lonza, Slough, UK), supplemented with 10% (w/v) FBS and

2mM L-glutamine. For passaging, cells were seeded in 10 ml complete growth media (GM) in 75 cm<sup>3</sup> Nunc plates (Nunc, NY, USA) with 100 U/ml penicillin and 100 µg/ml streptomycin (Sigma, Dorset, UK), incubated at 37°C and 5% CO<sub>2</sub> in a humidified atmosphere and passaged bi-weekly or when 80-90% confluence was reached. For experimental purposes, only cells of passage ≤ 20 were used.

For passaging, HEK293 cells were briefly trypsinised in 0.25x Trypsin/EDTA from 10x stock (Sigma, Dorset, UK) at room temperature, whereas HepG2 cells were trypsinised for 5 minutes at 37°C. Trypsin was neutralised with 5 ml complete growth media. Cells were centrifuged at 1,000 rpm for 3 minutes and media was removed. The cell pellet was resuspended in fresh media and plated at a 1:10 dilution with antibiotics.

#### **6.2.19 Trypan Blue exclusion assay and hemocytometry**

For experimental purposes, it was important that only adherent (live) cells were seeded. The trypan blue dye exclusion assay is based on the premise that healthy cells, with an impervious cell membrane, cannot take up the dye. Live cells therefore remain unstained, and can be counted by hemocytometry in an improved Neubauer haemocytometer. Cells were trypsinised, pelleted for 3 minutes at 1000 rpm and resuspended into 10 ml complete growth media without antibiotics. 100 µl cell culture was diluted into 150 µl sterile DPBS and 250 µl Trypan Blue dye (Sigma, Dorset, UK), were added at RT for 5 minutes. 10 µl of the cell suspension in PBS/Trypan Blue were added to the chambers and counted under phase-contrast microscopy. The number of live cells/ml were established using the formula:

$$\text{Average cell number} \times n \times 10^4 = \text{number of cells/ml}$$

where  $n$  = dilution factor (5) and the average cell number is derived from the number of cells within each counting chamber square (4), and where  $10^4$  is the multiplication factor to convert the chamber depth ( $0.1 \text{ mm}^3$ ) to  $\text{cm}^3$ .

## **6.2.20 Generation of a polyclonal HEK293 Reporter cell line**

### **A. Antibiotic Titration Curve**

HEK293 cells were seeded at a density of  $3 \times 10^4$  cells/ml in a 6-well plate, in 2 ml complete growth media without antibiotics. 24 hours after seeding, cells in each well were forward-transfected with 2.5  $\mu\text{g}$  pd1-EGFP-N1 vector DNA (Clontech, Saint-Germain-en-Laye, France). Plasmid DNA was diluted in 250  $\mu\text{l}$  Optimem with 2.5  $\mu\text{l}$  PLUS reagent. 5  $\mu\text{l}$  Lipofectamine LTX (Life Technologies, Paisley, UK) were diluted in 250  $\mu\text{l}$  Optimem. Plasmid DNA was then added to the Lipofectamine solution and Lipofectamine-DNA complexes were allowed to form at RT for 20 minutes according to the manufacturer's recommendations. 16 hours post-transfection, media were replaced with 2 ml fresh media, supplemented with G418 to the following final concentrations; 0 mg/ml, 0.4 mg/ml, 0.8 mg/ml, 1.2 mg/ml, 1.6 mg/ml and 2 mg/ml. Antibiotic-containing media were replaced every 24-48 hours for a period of 10 days, and cell viability was monitored daily under a microscope. To establish a titration curve, media were transferred to a 15 ml centrifugation tube. Adherent cells were washed 1x with 1 ml Dulbecco's Phosphate buffered saline (DPBS) (D8537, Sigma, Dorset, UK), trypsinised with 1 ml of 0.25x Trypsin-EDTA (T4174, Sigma, Dorset, UK), neutralised with 1 ml complete growth medium and centrifuged at 1,000 rpm for 3 minutes in a bench-top centrifuge. Live cells were then counted by hemocytometry and counts were standardized to the G418-untreated control.

## **B. Generation of stable HEK293-dEGFP cells**

For the generation of the pd1EGFP-N1- expressing HEK293 cell line, cells were seeded at a density of  $3 \times 10^4$  cells/ml in a 6-well plate and transfected with 2.5  $\mu$ g pd1EGFP-N1 DNA by Lipofectamine LTX (Invitrogen, Paisley, UK). The next day, G418 was added at a final concentration of 1 mg/ml, and replaced with fresh complete growth media every 48 hours for 12 days until cells became ~80% confluent. Cell viability was monitored daily by phase-contrast microscopy. Once confluent, cells were trypsinised and seeded in a T75 Nunc plate in 10 ml complete growth medium supplemented with 1 mg/ml G418 from 50 mg/ml stock solution prepared in DPBS. HEK293-dEGFP cells in a T75 Nunc plate were prepared for FACS by removing media, washing 1x with 5 ml DPBS, trypsinised with 4 ml 0.25x Trypsin-EDTA at room temperature for 2 minutes, followed by trypsin neutralisation with 4 ml complete growth media. Cells were then transferred to a 15 ml Eppendorf tube and centrifuged for 3 minutes at 1,000 rpm at room temperature. Media were removed without disturbing the cell pellet, and cells were resuspended in 10 ml FACS Analysis buffer (1% (v/v) FBS/DPBS supplemented with 2 mM EDTA). Live cells were counted by hemocytometry following the Trypan Blue Exclusion assay and diluted in FACS cell buffer to a final density of  $1 \times 10^6$  cells/ml into 2ml aliquots. Cells were kept on ice throughout. For fluorescence-activated cell sorting, confluent cells were washed 1x with PBS, trypsinised and viable cells were counted.  $1 \times 10^7$  cells were resuspended in 10 ml 1% (v/v) FBS/PBS to encourage monodispersion and kept on ice.  $3 \times 10^5$  untransfected cells in 1% (v/v) FBS/PBS were used as controls to 'normalize' auto-fluorescence levels. Cells were filtered through a 0.7  $\mu$ m filter, resuspended thoroughly by pipetting and analysed for fluorescence (dEGFP expression) on a MoFlo XDP Cell Sorter (Dako, Cambridgeshire UK) and viewed as histograms with Summit software (Beckman Coulter, High Wycombe,

UK) at the Wolfson Institute for Biomedical Research (UCL, London) at a rate of 1,300 events/sec into either medium or highly expressing populations.  $1 \times 10^5$  cells from each population were collected into tubes containing 3 ml 100% FBS supplemented with 100  $\mu\text{g/ml}$  penicillin/streptomycin on ice, and replated into  $75 \text{ cm}^3$  Nunc plates in complete growth media supplemented with the appropriate G418 concentration as a maintenance dose and 100  $\mu\text{g/ml}$  Penicillin/Streptomycin.

### **6.2.21 Cell viability assay**

The MTT assay (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) was used in order to assess potential cytotoxicity elicited on HEK293 and HepG2 by protein treatments at various concentrations for 24 hours.  $5 \times 10^3$  HEK293 cells were seeded in a 96-well plate in 100  $\mu\text{l}$  complete growth media without antibiotics and incubated at  $37^\circ\text{C}$  and 5%  $\text{CO}_2$  in a humidified atmosphere 16 hours before treatment. C11.2-C13.2 and C5.1 were thawed on ice and forward-transfect cells at a final molar concentration of 0.1, 0.3, 1 and 3.3  $\mu\text{M}$  in triplicate, in 100  $\mu\text{l}$  Optimem (Gibco, Paisley, UK). Optimem was used as a mock control. All treatments were done in triplicate ( $n=3$ ). For HepG2 cells, complexes were prepared within wells at final siRNA concentrations of 0.3, 1 and 3.3  $\mu\text{M}$  in triplicate, and overlaid with  $7.5 \times 10^3$  cells in Optimem for 5 hours, at a final volume of 100  $\mu\text{l}$ . Media were then aspirated and replaced with complete growth media supplemented with penicillin-streptomycin for 24 hours at  $37^\circ\text{C}$  and 5%  $\text{CO}_2$  in a humidified atmosphere. Results for HepG2 cells represent 3 independent experiments done in triplicate with biological replicates ( $n=9$ ). MTT (M5655, Sigma Dorset, UK) was dissolved in DPBS at a concentration of 5 mg/ml and 10  $\mu\text{l}$  were added to the cells in Optimem, to a final concentration of 0.45 mg/ml for 60 minutes at  $37^\circ\text{C}$  in a light-sealed environment. Media was carefully aspirated and 100  $\mu\text{l}$  dimethyl sulfoxide (DMSO) (Sigma, Dorset, UK) were

added to dissolve the formazan crystals. Cell viability was assessed at  $A_{570}$  in a Versamax microplate spectrophotometer (Molecular Devices, Oslo, Norway) and  $A_{640}$  to adjust for background fluorescence. Data acquisition was done using the SoftMax Pro Data Acquisition and Analysis software. For analysis,  $A_{640}$  values were subtracted from  $A_{570}$  values and values were normalized against the average, normalized mock-treated controls and mean cell viability was expressed as a % of mock-treated control.  $SD\pm$  was calculated based on each sample's deviation from the mean. Statistical analysis was done by two-way Anova followed by post-hoc analysis with Dunnett's 2-sided test (SPSS, IBM UK).

### **6.2.22 Protein-siRNA complex formation**

For GFP knockdown assays in the HEK reporter line, Stealth GFP RNAi Reporter control siRNA (Cat. # 12935-145, Ambion, Life Technologies, CA, US) in 1x RNA Annealing buffer (10 mM Tris-HCl, 20mM NaCl, 1 mM EDTA, pH 8) was used. For 100 nM final concentration, 60 pmoles were diluted into 50  $\mu$ l Optimem. For 10 nM final concentration, 6 pmoles were diluted as before. Previous studies have shown that 98% encapsulation of siRNA can be achieved with a molar excess of CPP at a ratio of 50:1 (Von Asbeck *et al*, 2013). CPPs were thus diluted at the appropriate molar concentration in 50  $\mu$ l Optimem (Life Technologies, Paisley, UK) on ice. Complex formation was allowed to proceed for 30 minutes on ice once the protein had been added to the siRNA.

### **6.2.23 Transient knockdown assays in a HEK293 reporter cell line**

For knockdown assays, highly expressing HEK293-dEGFP cells ( $6 \times 10^4$  cells/well) were seeded in a 24-well plate in a final volume of 1 ml complete growth media supplemented with 1.8 mg/ml G418 only. HEK293 cells were also seeded to correct for auto-fluorescence during analysis (in triplicate). At 80% confluency, protein-siRNA complexes with various proteins and Stealth GFP RNAi Reporter control siRNA were prepared as

described previously. Cells were briefly incubated in 500  $\mu$ l Optimem (Life Technologies, Paisley, UK) and 100  $\mu$ l of protein-siRNA complexes at a 50:1 molar ratio were directly added to the cells, to a final volume of 600  $\mu$ l and siRNA concentration of 10 nM or 100 nM. All treatments were done in triplicate (n=3). Complexes were incubated with the cells for 5 hours at 37°C and 5% CO<sub>2</sub> in a humidified atmosphere in Optimem (Life Technologies, Paisley, UK). Then, media were gently aspirated and replaced with complete growth media supplemented with 1.8 mg/ml G418, 50 mM chloroquine (Sigma, Dorset, UK) and 1x penicillin/streptomycin for 24 or 48 hours. Lipofectamine-transfected siRNA was used as a positive control. Scrambled siRNA (Stealth RNAi siRNA Negative control (LoGC) (Cat. #12935-200), Ambion, Life Technologies, Paisley, UK) and siRNA only treatments were used as a negative control. Following knockdown, plates were centrifuged at 1,000 rpm for 3 minutes at 4°C. Media were aspirated to minimize cell loss and cells prepared for flow cytometry analysis.

#### **6.2.24 Analysis of dEGFP knockdown by flow cytometry**

Following dEGFP knockdown, HEK293-dEGFP cells were gently resuspended by pipetting and transferred to pre-chilled sterile 1.5 ml Eppendorf tubes. They were then centrifuged at 1000 rpm for 3 minutes at RT and the supernatant was discarded. The cell pellet was resuspended in chilled 1 ml FACS sample buffer (1% (v/v) FBS/DPBS, 2 mM EDTA), transferred to flow cytometry tubes and kept on ice. 1 mg/ml of propidium iodide (PI) were added to cells as a live/dead counterstain and analysed by flow cytometry on a Dako Cyan ADP cytometer.

#### **6.2.25 PTP1B knockdown assays in HepG2 cells.**

In order to optimise the minimum amount of cells needed for PTP1B detection by immunoblotting, a titration was done with varying cell numbers followed by the BCA

assay on whole cell lysates. In short, either  $5 \times 10^6$ ,  $2.5 \times 10^6$ ,  $5 \times 10^5$ ,  $2.5 \times 10^5$  and  $5 \times 10^4$  cells were seeded in either a T75 Nunc flask, a T25 Nunc flask or a 6-well plate in complete growth media and allowed to adhere overnight. The next day, media were removed, and cells were washed 1x with DPBS before treatment with each of 2 ml, 1 ml or 0.25 ml Lysis buffer, respectively. The total amount of protein present in the sample was determined by the BCA assay, and 10  $\mu$ g protein were analysed by WB with an  $\alpha$ -Actin primary antibody at a 1:200 dilution in 5% non-fat dry Milk-TBS and an anti-rabbit HRP conjugated secondary antibody and visualized by chemilluminescence.

Then, in order to optimise PTP1B transfection efficiency by Lipofectamine RNAiMax, two siRNA duplexes corresponding to different regions of the PTPN1 gene (VHS41290- 5'-CCAUAGUCGGAUUAACUACAU-3' and VHS41291- 5'-GGAAAGACCCUUCU-UCCGUUGAUAU-3') were assessed. 6, 15 or 30 pmol were complexes with 5  $\mu$ l Lipofectamine RNAiMAX in 500  $\mu$ l Optimem (Life Technologies, Paisley, UK). Cells were reverse-transfected inside wells by allowing complex formation inside wells and overlaying with  $2.5 \times 10^5$  HepG2 cells/well in 2 ml Optimem (Life Technologies, Paisley, UK) to yield final concentrations of siRNA of 10 nM, 25 nM and 50 nM, respectively. Cells were transfected for 4 hours to allow adherence at 37°C and 5% CO<sub>2</sub> in a humidified atmosphere. Adherence was monitored hourly by phase-contrast microscopy. Once adherence reached ~ 100%, media were removed and replaced with complete growth media and antibiotics. The cells were incubated for 72 hours at 37°C to allow efficient time for protein expression knockdown. Then, whole cell lysates were prepared and knockdown was analysed by western blotting 15  $\mu$ g crude protein with an anti-PTP1B antibody at a 1:1,000 dilution in 5% (w/v) Milk/TBS and a secondary HRP-conjugated anti-mouse secondary antibody at a 1:10,000 dilution in 5% (w/v) Milk-TBS.

### **6.2.26 Whole cell lysate preparation**

Whole cell lysates were prepared in Lysis buffer (150mM NaCl, 20mM Tris-HCl pH 7.5, 1mM EGTA, 1% Triton-X, 1% Nonidet-P40 (Igepal) supplemented with Protease Inhibitors at a 1:1,000 dilution, 100 mM NaF, 10 mM Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub> (sodium pyrophosphate tetrabasic) and 2 mM Na<sub>3</sub>VO<sub>4</sub> (sodium orthovanadate)). First, plates were centrifuged at 4°C and 1,000 rpm for 3 minutes, and media were carefully removed. Cells were lysed with 250 µl chilled lysis buffer for 15 minutes at RT. Lysed cells were then carefully scraped and transferred to chilled 1.5 ml eppendorf tubes and centrifuged at 13,000 rpm at 4°C for 15 minutes. The supernatant was transferred to clean, chilled 1.5 ml eppendorf tubes and stored at -80°C for future use.

### **6.2.27 Localisation studies in fixed HepG2 and HEK293 cells**

Confluent (~80%) HEK293 and HepG2 cells were trypsinised, counted by the Trypan Blue exclusion test on a haemocytometer and seeded at a density of  $6 \times 10^5$  cells in 2ml 10% (v/v) FBS/DMEM supplemented with 2 mM L-Glutamine on uncoated glass coverslips in 6-well plates. Cells were allowed to adhere overnight in a humidified atmosphere and 5% CO<sub>2</sub> until confluency was reached (~60-80%; approximately 18 hours). Media were replaced following a wash with 2ml filter-sterilized 1xPBS (Ca<sup>2+</sup>/Mg<sup>2+</sup>). Cells were treated with either a low (0.1µM), medium (1 µM) or high (10 µM) dosage of C5.1-myc (from 290 µM stock in 1xPBS) in 500 µl PBS or Optimem for 2 or 24 hours at 37°C in a humidified atmosphere and 5% CO<sub>2</sub>. 500µl sterile Dulbecco's PBS (Ca<sup>2+</sup>/Mg<sup>2+</sup>-free) was also used as a control.

Protein concentration was measured by Nanodrop at  $A_{280}$  using 1xPBS as a blank. Protein concentrations and MW were used with following conversion formula to calculate its molar concentration:

*Molar concentration of protein (mmol/L)*

$= \text{Concentration at } A_{280} (\text{mg/ml}) \times \text{MW (mMol/mg)} \times \text{Volume (1,000ml/L)}$

$\cdot \cdot [( \text{mg/ml} ) / \text{MW (Da)}] \times 1,000$

C5.1-myc to be used for either 10  $\mu\text{M}$ , 1  $\mu\text{M}$  or 0.1  $\mu\text{M}$  final molarities were calculated using the following dilution formula:

$$C_i (\mu\text{M}) \times V_i (\mu\text{l}) = C_f (\mu\text{M}) \times V_f (\mu\text{l})$$

Cell viability and confluence were assessed by phase-contrast microscopy prior to treatment whilst the protein was thawed on ice. Cells were incubated in a humidified atmosphere and 5%  $\text{CO}_2$  for either 2 or 24 hours. Following treatment, cells were aspirated, gently washed with 2 ml filter-sterilized 1xPBS ( $\text{Ca}^{2+}/\text{Mg}^{2+}$ ). Cells were fixed and permeabilized with fresh filter-sterilized 0.2% (v/v) Triton-X / 4% (v/v) paraformaldehyde (from 36% (v/v) formalin stock, in 1xPBS ( $\text{Ca}^{2+}/\text{Mg}^{2+}$ )) for 15 minutes at 25°C. Cells were then gently washed again 3x with 2.5 ml/well filter-sterilized PBS and coverslips (VWR, Leicestershire, UK) were transferred to a 12-well plate to minimize antibody volumes. Fixed cells were then blocked with 300 $\mu\text{l}$  10% (v/v) goat serum/PBS (from 10x stock) supplemented with 1.5 $\mu\text{l}$  RNase A and incubated with an anti-myc tag antibody (Clone 4A6, Millipore, MA, USA) in 10% (v/v) goat serum/PBS at a 1:1,000 dilution for 1.5 hours at 25°C. Cells were then washed 3x with 300  $\mu\text{l}$  1xPBS ( $\text{Ca}^{2+}/\text{Mg}^{2+}$ ) for 15 minutes each and incubated with a secondary anti-mouse FITC-labeled antibody (Sigma, Dorset, UK) at 1:500 dilution in 300  $\mu\text{l}$  10% (v/v) goat serum/PBS for another

hour. Nuclei were stained with 300  $\mu$ l ToPro3 (Life Technologies, Paisley, UK) in Dulbecco's PBS at a 1:750 dilution, for 30 minutes at 25°C in a light-sealed container. Coverslips were mounted on glass slides using Vectashield H1000 and visualized with a Leica TCS-SP2 spectral confocal microscope and Leica LCS imaging software (described in Section 6.2.30).

### **6.2.28 Localisation studies in live HepG2 cells**

HepG2 cells were seeded in a 6-well plate at a density of  $2.5 \times 10^5$  cells/well in 2.5ml complete growth media without antibiotics, 16 hours before transfection with protein-siRNA complexes. Protein-siRNA-FITC (Block-iT fluorescent oligo, Life Technologies, Paisley, UK) complexes were allowed to form at 4°C at 50:1 molar ratios and applied to cells in Optimem for 60 minutes at 37°C at a final volume of 2.5 ml. Complexes were removed and replaced with fresh complete growth media without antibiotics. Wheat germ agglutinin-dsRed (WGA-dsRed) was added at a final concentration of 1 mg/ml for 10 or 30 minutes at RT. Media were removed and cells were gently washed 2x with PBS. Visualisation in 2.5ml PBS was done by confocal microscopy with a Leica SP2 laser scanning microscope.

### **6.2.29 Confocal microscopy**

Fixed cells were mounted on slides and visualized on a Leica SP2 DM-RXE confocal microscope under a 63x oil immersion lens, whereas live cells were analysed in 1x PBS. Images were taken using the proprietary Leica confocal software (Leica Microsystems, Heidelberg, Germany). To-Pro3 nuclear staining was acquired with Ar/HeNe lasers at an excitation wavelength ( $Ex_{max}$ ) of 633nm and an emission wavelength ( $Em_{max}$ ) of 650-720 nm. FITC was detected at 488 nm ( $Ex_{max} = 490$  nm,  $Em_{max} = 525$  nm). WGA-dsRed

( $E_{x_{\max}} = 557$  nm and  $E_{m_{\max}} = 572$  nm) was detected at 563 nm. Laser intensities were 35-60%. To minimize noise-to-signal ratios, the pinhole diameter was set to 1 Airy unit and the speed of image lines/second when taking horizontal xy-sections in scan mode was set to 30. Cells were focused in continuous scan mode.

# Appendix I

## I.1 Bioinformatic analysis on constructs C1.1-C6.1

**Table I.1 Amino acid sequences of constructs C1.1-C6.1<sup>3</sup>.**  $\epsilon$ , extinction coefficient at 280 nm measured in water. Important sequences are highlighted; AntpHD, green; Penetratin, yellow; DRBM1, light blue; histidine tags, dark grey (His<sub>10</sub> or His<sub>6</sub>); GST tag, light grey; Factor X $\alpha$  cleavage site, red (C1.1, C2.1); TEV protease cleavage site, purple (C3.1, C4.1); Prescission protease cleavage site, dark purple (C5.1, C6.1). Underlined amino acids correspond to restriction enzyme cleavage sites following translation.

Name	Protein	Amino acid Sequence	No. of amino acids	Mw (kDa)	pI	$\epsilon$ (M <sup>-1</sup> cm <sup>-1</sup> ),
C1.1	His <sub>10</sub> -AntpHD-DRBM1	MGHHHHHHHHSSGH <u>IEGR</u> AMGRKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKENGGGGSGGGGSFFMEELNTYRQKQGVVLKYQELPNSGPPHRRRFTFQVIIDGREFPEGEGRSKKEAKNAAAKLAVEILNKKKA-	164	19.25	10.3	18.45
C2.1	His <sub>10</sub> -Pen-DRBM1	MGHHHHHHHHSSGH <u>IEGR</u> AMGRQIKIWFQNRRMKWKKENGGGGSGGGGSFFMEELNTYRQKQGVVLKYQELPNSGPPHRRRFTFQVIIDGREFPEGEGRSKKEAKNAAAKLAVEILNKKKA-	122	13.9	10.1	14
C3.1	AntpHD-DRBM1-His <sub>6</sub>	MGRKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKENGGGGSGGGGSFFMEELNTYRQKQGVVLKYQELPNSGPPHRRRFTFQVIIDGREFPEGEGRSKKEAKNAAAKLAVEILNKKATENLYFQ <u>ALE</u> HHHHHH-	159	18.8	10.1	19.9
C4.1	Pen-DRBM1-His <sub>6</sub>	MRQIKIWFQNRRMKWKKENGGGGSGGGGSFFMEELNTYRQKQGVVLKYQELPNSGPPHRRRFTFQVIIDGREFPEGEGRSKKEAKNAAAKLAVEILNKKATENLYFQ <u>SLE</u> HHHHHH-	116	13.4	9.8	15.5

<sup>3</sup> Continued overleaf

<b>C5.1</b>	GST-AntpHD-DRBM1	MSPILGYWKIKGLVQPTRLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMAIRYIADKHNMLGGCPKERAIEISMLEGAVLDIRYGVSRISYKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAI PQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLEVLFGGPLGKRRKRGRQTYTRYQTLEI EKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRMMKWKKENGGGGSGGGGSFFMEKLNT YRQKQGVVLKYLELPNSGPPHRRRFTFLVIIDGRQFPEGEGRKKKKAKNAAAKLAVEILKEK KAFI-	376	44.1	9.5	61.56
<b>C6.1</b>	GST-Pen:DRBM1	MSPILGYWKIKGLVQPTRLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMAIRYIADKHNMLGGCPKERAIEISMLEGAVLDIRYGVSRISYKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAI PQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLEVLFGGPLGSRQIKIWFQTRRMKWKKE NGGGGSGGGGSFFMEELNTYRQKQGVVLKYQELPNSGPPHRRRFTFQVIIDGREFPEGERSK KEAKNAAAKLAVEILKEKKGVS-	333	38.4	8.6	57.1

## I.2 Bioinformatic analysis on myc-tagged constructs

**Table I.2** Amino acid sequences for constructs C5.1-myc and C6.1-myc<sup>4</sup>. Important sequences are highlighted as in Table I.1. The myc epitope tag is in bold.

Name	Protein	Amino acid Sequence	No.of Amino acids	Mw (kDa)	pI	$\epsilon$ ( $M^{-1} \text{ cm}^{-1}$ ),
C5.1-myc	GST-AntpHD-DRBM1-myc	MSPILGYWKIKGLVQPTRLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAIEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSD <b>LEVLFGPLGSGRKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFONRRMKWKKEN</b> GGGGSGGGGSFFMEELNTYRQKQGVLKYQELPNSGPPHRRRFTFQVIIDGREFPEGEGRSCKKAKNAAAKLAVEILK <b>EKKA</b> VSE <b>QKLISEEDLN</b> -	387	45.1	9.0	61.56
C6.1-myc	<b>GST-Pen: DRBM1-myc</b>	MSPILGYWKIKGLVQPTRLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAIEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSD <b>LEVLFGPLGSRQIKIWFONRRMKWKKEN</b> GGGGSGGGG <b>IFFMEELNTYRQKQGVLKYQELPNSGPPHRRRFTFLV</b> IIDGRQFPEGDGRNKKKTKNAAAKLAVEILK <b>EKKA</b> VSE <b>QKLISEEDLN</b> -	344	39.8	8.6	57.1

<sup>4</sup> Continued overleaf

**Table I.3 Amino acid sequences of GST-tagged constructs following purification and tag cleavage by Precision protease. Important features are highlighted as in Table I.1. The myc epitope tag is in bold.**

Name	Protein	Amino acid Sequence	No. of Amino acids	Mw (kDa)	pI	$\epsilon$ ( $M^{-1} \text{ cm}^{-1}$ ),
C5.1	AntpHD-DRBM1	<b>GPLGKRRKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNR</b> <b>RMKWKKEN</b> GGGGSGGGGSFFMEKLN <b>TYRQKQG</b> VVLKYLELPNSGPPH <b>DRRFTFLVI</b> <b>IDGRQFPEGEGR</b> KKKKAKNAAAKLAVEIL <b>KEKKAFI</b> -	150	17.4	10.7	18.45
C6.1	Pen-DRBM1	<b>GPLGSRQIKIWFQTRRMKWKKEN</b> GGGGSGGGGSFFMEELN <b>TYRQKQG</b> VVLKYQELP NSGPPH <b>DRRFTFQVI</b> IDGREFPEGEGRSK <b>KEAKNAAAKLAVEILKEK</b> KGVS-	107	12.0	10.0	14
C5.1-myc	AntpHD-DRBM1-myc	<b>GPLGSRKRRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNR</b> <b>RMKWKKEN</b> GGGGSGGGGSFFMEELN <b>TYRQKQG</b> VVLKYQELPNSGPPH <b>DRRFTFQVI</b> <b>IDGREFPEGEGRS</b> KKKAKNAAAKLAVEIL <b>KEKKA</b> VSE <b>QKLISEEDLN</b> -	161	18.7	10.0	18.45
C6.1-myc	Pen-DRBM1-myc	<b>GPLGSRQIKIWFNRRMKWKKEN</b> GGGGSGGGG <b>IF</b> MEELN <b>TYRQKQG</b> VVLKYQELP NSGPPH <b>DRRFTFLVI</b> IDGRQFPEGDGRN <b>KKKTKNAAAKLAVEILKEKKA</b> VSE <b>QKLISE</b> <b>EDLN</b> -	118	13.3	9.9	14

**Table I.4 Amino acid sequences of constructs C7-C10.2.** Important features are highlighted as in Table I.1. Amino acid sequences for constructs C7.2-C10.2. These are comprised of the Penetratin analogue EB1 fused to the DRBMx2, or the inverted sequence, DRBMx2-EB1. A His<sub>6</sub> tag (dark grey) was conferred by cloning into a modified pET32-a vector, and a GST tag (light grey), from the pGEX-6P-2 vector. EB1 is highlighted in yellow; DRBMx2 is shown in light blue.  $\epsilon$ , extinction coefficient

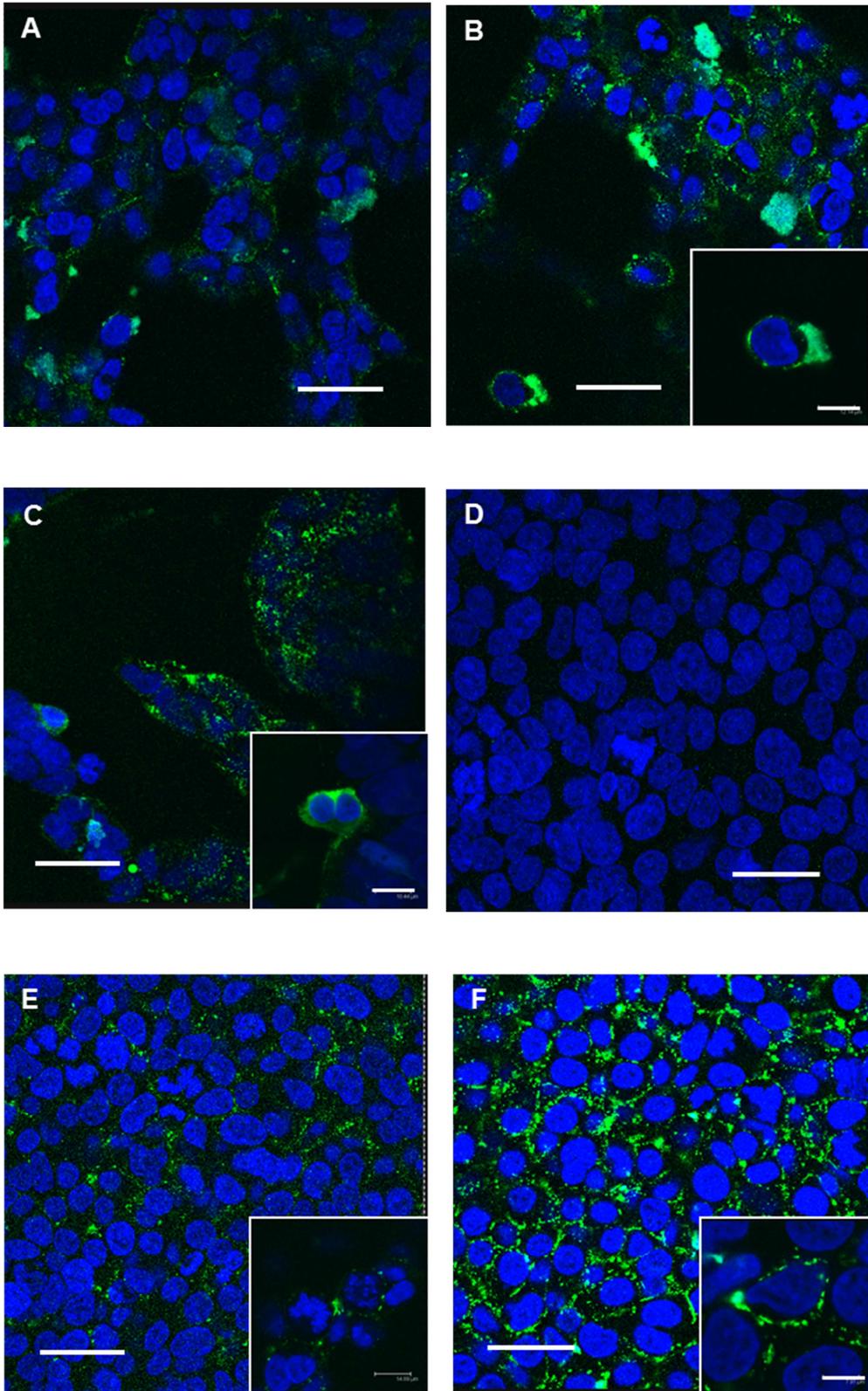
Name	Protein	Amino acid Sequence	No. of Amino acids	Mw (kDa)	pI	$\epsilon$ x1000 (M <sup>-1</sup> cm <sup>-1</sup> ),
C7.2	EB1: DRBMx2- His <sub>6</sub>	MLIRLXSHLIHIWFQNRRLKWKKKGGGGSGGGGSAGDLSAGFFMEELNTYRQKQGVLKYQELPNSGPPHRRFTFQVIIDGREFPEGEGRSKKEAKNAAAKLAVEILNKEKKA VSPLLLTTTNSSEGLSMGNYIGLINRIAQKKRLTVNYEQCASGVHGPEGFHYKCKMGQKEYSIGTGSTKQEAQLAAKLAYLQILSEETSVKSDLEHHHHHH-	215	24.05	9.6	21.5
C8.2	DRBMx2:E B1-His <sub>6</sub>	MAGDLTAGFFMEELNTYRQKQGVLKYQELPNSGPPHRRFTFQVIIDGREFPEGEGRSKKEAKNAAAKLAVEILNKEKKA VSPLLLTTTNSSEGLSMGNYIGLINRIAQKKRLTVNYEQCASGVHGPEGFHYKCKMGQKEYSIGTGSTKQEAQLAAKLAYLQILSEETSVKSDGGGGSGGGGSLIRLWSHLIHIWFQNRRLKWKKKLEHHHHHH-	216	24.14	9.6	27.0
C9.2	GST-EB1: DRBMx2	MSPILGYWKIKGLVQPTRLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDV KLTQSMAIRYIADKHNMLGGCPKERA EISMLEGAVLDIRYGVSR IAYSKDFETLKVDFLSKLP EMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAI PQI DKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLEVLFGPLGSLIRLWSHLIHIWFQNRRLKWKKKGGGGSGGGGSAGDLSAGFFMEELNTYRQKQGVLKYQELPNSGPPHRRFTFQVIIDGR EFPEGEGRSKKEAKNAAAKLAVEILNKEKKA VSPLLLTTTNSSEGLSMGNYIGLINRIAQKKRL TVNYEQCASGVHGPEGFHYKCKMGQKEYSIGTGSTKQEAQLAAKLAYLQILSEETSVKSD	438	49.8	8.8	70.2
C10.2	GST- DRBMx2:E B1	MSPILGYWKIKGLVQPTRLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDV KLTQSMAIRYIADKHNMLGGCPKERA EISMLEGAVLDIRYGVSR IAYSKDFETLKVDFLSKLP EMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAI PQI DKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLEVLFGPLGSLIRLWSHLIHIWFQNRRLKWKKKGGGGSGGGGSAGDLSAGFFMEELNTYRQK QGVVLKYQELPNSGPPHRRFTFQVIIDGREFPEGEGRSKKEAKNAAAKLAVEILNKEKKA VSPLLLTTTNSSEGLSMGNYIGLINRIAQKKRLTVNYEQCASGVHGPEGFHYKCKMGQKEYSIGT GSTKQEAQLAAKLAYLQILSEETSVKSDGGGGSGGGGSLIRLWSHLIHIWFQNRRLKWKKK-	438	49.8	8.8	70.2

### I.3 Bioinformatic analysis on constructs C11.2- C13.

**Table I.5 Amino acid sequences for constructs C11.2-C13.2.** These are comprised of Tat-DRBMx2, Penetratin-DRBMx2 or DRBMx2 only, respectively. A His<sub>6</sub> tag (dark grey) was conferred by cloning into a pET11-d vector. Important features are highlighted; Tat, dark blue; Penetratin, yellow; DRBMx2, light blue.

Name	Protein	Amino acid Sequence	No. of Amino acids	Mw (kDa)	pI	$\epsilon$ (M <sup>-1</sup> cm <sup>-1</sup> ),
C11.2	Tat-DRBMx2	MGRKKRXQRRRGAPAGDLSAGFFMEELNTYRQKQGVLKYQELPNSGPPHRRFTFQVII DGREFPEGEGRSKKEAKNAAKLA VEILNKEKKA VSPLLLTTNSSEGLSMGNYIGLINRIA QKKRLTVNYEQCASGVHGPEGFHYKCKMGQKEYSIGTGSTKQEAQLAAKLAYLQILSEE TSVKSDHHHHHH	194	21.7	9.7	10.56
C12.2	Pen-DRBMx2	MGRQIKXWFQNRMMKWKKGSGGSGGSGGSGGSGGSGGSGGAPAGDLSAGFFMEEL NTYRQKQGVLKYQELPNSGPPHRRFTFQVIIDGREFPEGEGRSKKEAKNAAKLA VEIL NKEKKA VSPLLLTTNSSEGLSMGNYIGLINRIAQKKRLTVNYEQCASGVHGPEGFHYKCK MGQKEYSIGTGSTKQEAQLAAKLAYLQILSEETS VKSDHHHHHH	225	24.3	9.7	21.56
C13.2	DRBMx2	MGAPAGDLSAGFFMEELNTYRQKQGVLKYQELPNSGPPHRRFTFQVIIDGREFPEGEGR SKKEAKNAAKLA VEILNKEKKA VSPLLLTTNSSEGLSMGNYIGLINRIAQKKRLTVNYE QCASGVHGPEGFHYKCKMGQKEYSIGTGSTKQEAQLAAKLAYLQILSEETS VKSDHHHH HH	184	20.4	9.1	10.6

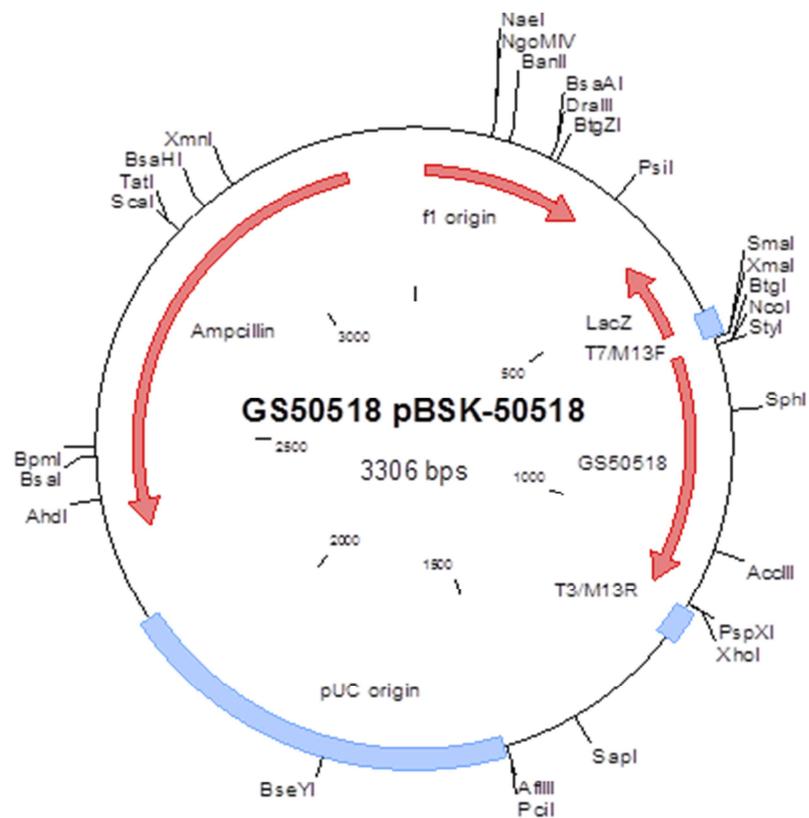
I.4 Figure 2.19, Enlarged



# Appendix II

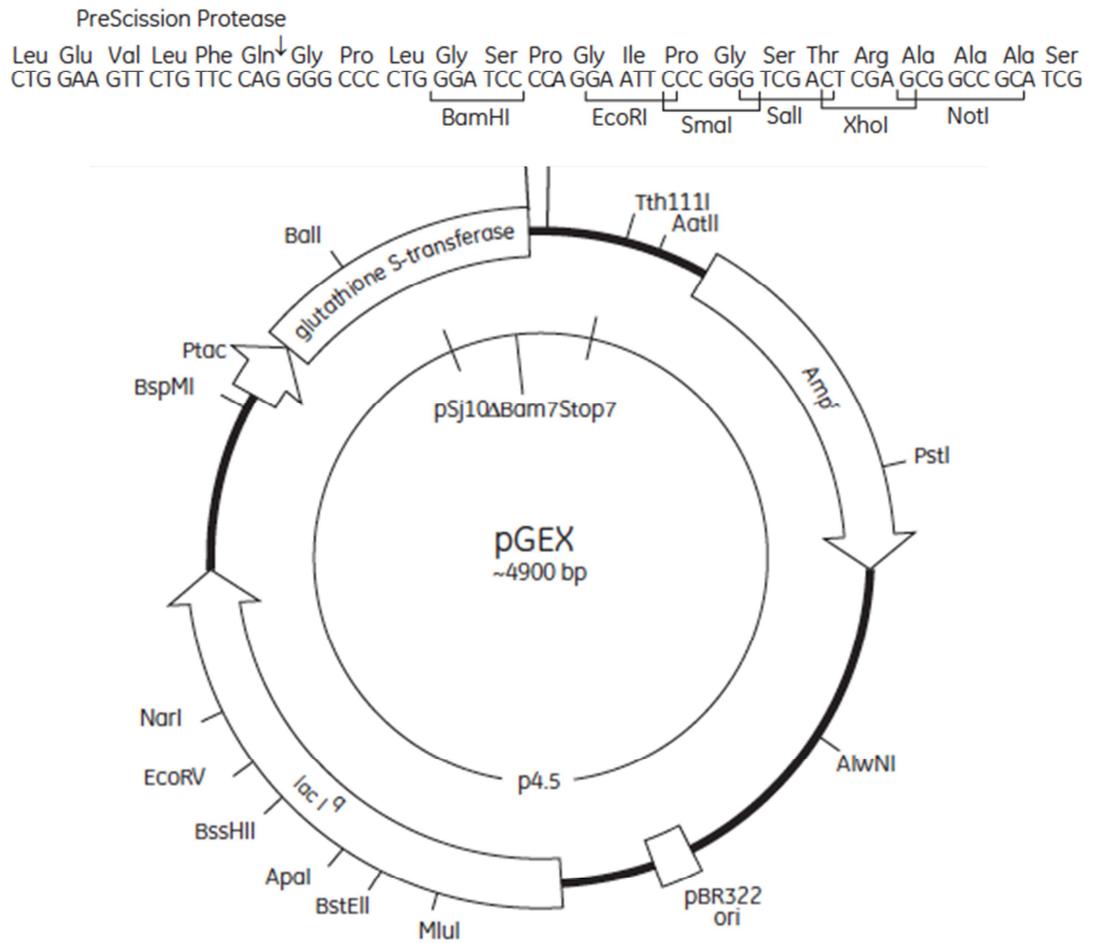
## II.1 Plasmid Vector Maps

Figure II.1.1 Restriction map and multiple cloning site (MCS) of the pBSK-50518 plasmid vector.



The AntpHD-DRBM1 synthetic gene (GS50518) was cloned into the *EcoRV*-digested pBluescript II SK derivative lacking its multiple cloning site.

**Figure II.1.2 Restriction map and multiple cloning site (MCS) of the pGEX-6P-2 plasmid vector**

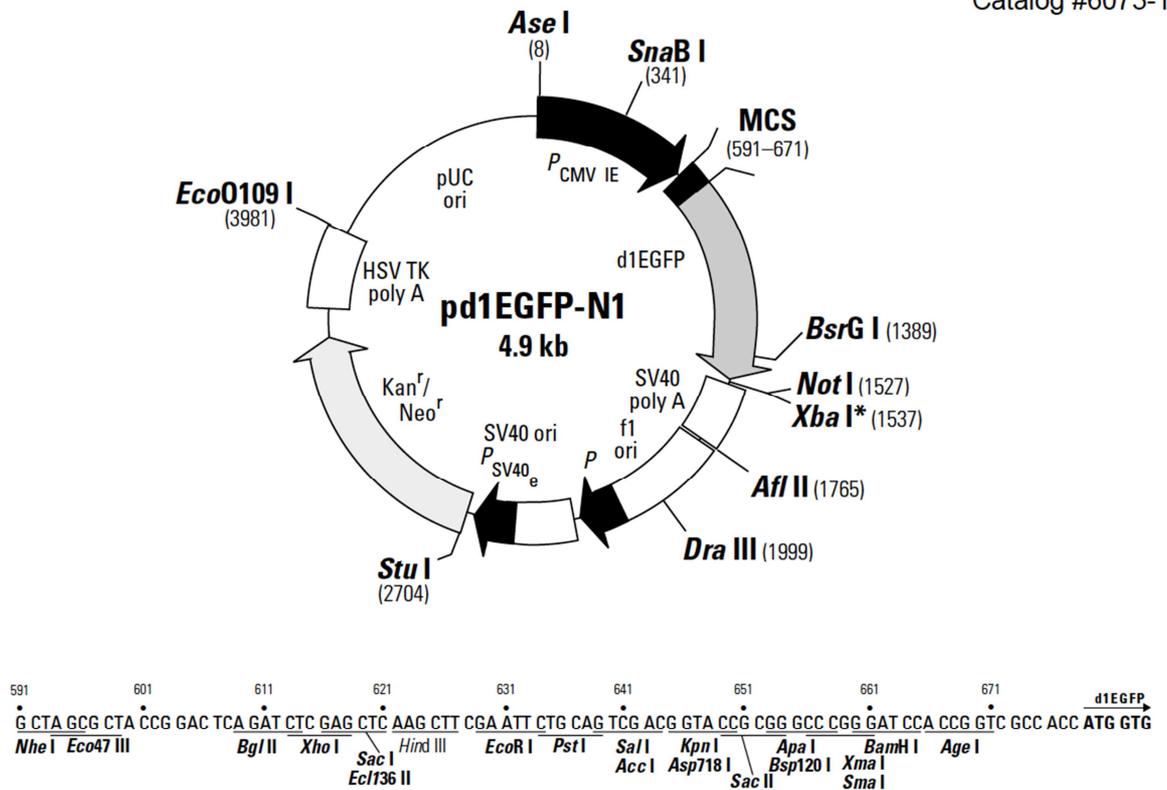


Inserts cloned into pGEX-6P-2 downstream of the glutathione-S-transferase (GST) tag were purified by affinity chromatography. A PreScission protease cleavage site allows for the proteolytic cleavage of the tag following protein purification.



**Figure II.1.4 Restriction map and multiple cloning site (MCS) of pd1EGFP-N1 plasmid**

Catalog #6073-1



Unique restriction sites are in bold. The *Not* I site follows the d1EGFP stop codon. The *Xba* I site is methylated in the DNA provided by BD Biosciences Clontech. This plasmid was used to generate the dEGFP expressing HEK293-dEGFP reporter cell line.

## II.2 Statistical Analysis of MTT assays in HEK293 and HepG2 cells by 2-way Anova (SPSS).

### II.2.1 MTT assay in HEK293 cells (n=3). P-value limits were set at $\leq 0.05$ .

**Table II.2.1.1 Tests of Between-Subjects Effects.** Important entries are in bold.

Source	Type III Sum of Squares	df	Mean Square	F	Significance	Partial R Squared
Corrected Model	11.747	19	.618	2.094	.024	.499
Intercept	236.013	1	236.013	799.476	.000	.952
<b>Protein Level</b>	3.085	4	.771	2.612	<b>.050</b>	.207
<b>Concentration Level</b>	5.327	3	1.776	6.015	<b>.002</b>	.311
<b>Protein-Concentration combined</b>	3.335	12	.278	.941	<b>.517</b>	.220
Error	11.808	40	.295			
Total	259.568	60				
Corrected Total	23.555	59				

R Squared = 0.499 (Adjusted R<sup>2</sup> = 0.261)

**Table II.2.1.2 Pairwise comparisons between proteins by 2-way Anova.** After mean viability at each concentration is normalised against the mean mock-treated value, Anova compares each concentration for each protein with the rest<sup>5</sup>.

Protein	Concentration Level I	Concentration Level II	Mean Difference (I-II)	Std. Error	Significance. <sup>b</sup>	95% Confidence Interval for Difference <sup>b</sup>	
						Lower Bound	Upper Bound
C5.1	.10	.30	.079	.279	.778	-.484	.642
		1.00	.237	.279	.401	-.327	.800
		3.30	.432	.279	.128	-.131	.996
	.30	.10	-.079	.279	.778	-.642	.484
		1.00	.157	.279	.575	-.406	.720
		3.30	.353	.279	.212	-.210	.916
	1.00	.10	-.237	.279	.401	-.800	.327
		.30	-.157	.279	.575	-.720	.406
		3.30	.196	.279	.486	-.367	.759
	3.30	.10	-.432	.279	.128	-.996	.131
		.30	-.353	.279	.212	-.916	.210
		1.00	-.196	.279	.486	-.759	.367
C13.2	.10	.30	.033	.279	.906	-.530	.596
		1.00	.146	.279	.602	-.417	.709
		3.30	.701 <sup>*</sup>	.279	.016 <sup>*</sup>	.138	1.264
	.30	.10	-.033	.279	.906	-.596	.530
		1.00	.113	.279	.687	-.450	.676
		3.30	.668 <sup>*</sup>	.279	.021 <sup>*</sup>	.105	1.231
	1.00	.10	-.146	.279	.602	-.709	.417
		.30	-.113	.279	.687	-.676	.450
		3.30	.555	.279	.053	-.008	1.118

<sup>5</sup> Continued overleaf

		3.30	.10 .30 1.00	-.701* -.668* -.555	.279 .279 .279	.016* .021* .053	-1.264 -1.231 -1.118	-.138 -.105 .008	
Mock control	.10		.30 1.00 3.30	.000 1.110E-016 1.110E-016	.279 .279 .279	1.000 1.000 1.000	-.563 -.563 -.563	.563 .563 .563	
		.30		.10 1.00 3.30	.000 1.110E-016 1.110E-016	.279 .279 .279	1.000 1.000 1.000	-.563 -.563 -.563	.563 .563 .563
			1.00		.10 .30 3.30	-1.110E-016 -1.110E-016 .000	.279 .279 .279	1.000 1.000 1.000	-.563 -.563 -.563
	3.30			.10 .30 1.00	-1.110E-016 -1.110E-016 .000	.279 .279 .279	1.000 1.000 1.000	-.563 -.563 -.563	.563 .563 .563
		C12.2	.10		.30 1.00 3.30	-.267 -.168 .098	.279 .279 .279	.344 .550 .726	-.830 -.731 -.465
	.30				.10 1.00 3.30	.267 .099 .365	.279 .279 .279	.344 .725 .198	-.296 -.464 -.198
1.00					.10 .30 3.30	.168 -.099 .266	.279 .279 .279	.550 .725 .345	-.395 -.662 -.297
	3.30			.10 .30 1.00	-.098 -.365 -.266	.279 .279 .279	.726 .198 .345	-.661 -.928 -.829	.465 .198 .297
C11.2			.10		.30 1.00 3.30	-.001 .275 .969	.279 .279 .279	.996 .330 .001*	-.565 -.288 .406

	.30	.10	.001	.279	.996	-.562	.565
		1.00	.276	.279	.328	-.287	.839
		3.30	.970*	.279	.001*	.407	1.533
	1.00	.10	-.275	.279	.330	-.838	.288
		.30	-.276	.279	.328	-.839	.287
		3.30	.694*	.279	.017*	.131	1.257
	3.30	.10	-.969*	.279	.001*	-1.532	-.406
		.30	-.970*	.279	.001*	-1.533	-.407
		1.00	-.694*	.279	.017*	-1.257	-.131

Based on estimated marginal means.

\*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

**Table II.2.1.3 Dunnett's post-hoc test based on observed means following MTT assays in HEK293 cells.** This test compares each concentration against the highest concentration tested, regardless of which protein is being tested to indicate dose-dependent effects. Significance values are highlighted.

Concentration Level I	Concentration Level II	Mean Difference (I-II)	Std. Error	Significance	95% Confidence Interval	
					Lower Bound	Upper Bound
.10	3.30	.44019*	.124595	.003	.13600	.74439
.30	3.30	.47134*	.124595	.001	.16715	.77554
1.00	3.30	.34227*	.124595	.024	.03807	.64646

Based on observed means.

The error term is Mean Square (Error) = .116.

\*. The mean difference is significant at the .05 level.

a. Dunnett t-tests treat one group as a control, and compare all other groups against it.

## II.2.2 MTT assay in HepG2 cells (n=9)

**Table II.2.1.4 Tests of Between-Subjects Effects.** Important entries are in bold.

Source	Type III Sum of Squares	df	Mean Square	F	Significance	Partial R Squared
Corrected Model	.158 <sup>a</sup>	14	.011	.657	.811	.071
Intercept	6.338	1	6.338	368.519	.000	.754
<b>Protein Level</b>	.117	4	.029	1.704	<b>.153</b>	.054
<b>Concentration Level</b>	.001	2	.000	.020	<b>.980</b>	.000
<b>Protein Level * Concentration Level</b>	.040	8	.005	.293	<b>.967</b>	.019
Error	2.064	120	.017			
Total	8.560	135				
Corrected Total	2.222	134				

a. R Squared = .071 (Adjusted R Squared = -.037)

**Table II.2.1.5 Pairwise comparisons between proteins by 2-way Anova.** After mean viability at each concentration is normalised against the mean mock-treated value, Anova was used to compare each concentration for each protein with the rest<sup>6</sup>.

Protein	Concentration Level 1	Concentration Level 2	Mean Difference (1-2)	Std. Error	Significance. <sup>a</sup>	95% Confidence Interval for Difference <sup>a</sup>	
						Lower Bound	Upper Bound
C12.2	.33	1.00	-.022	.062	.728	-.144	.101
		3.33	.013	.062	.840	-.110	.135
	1.00	.33	.022	.062	.728	-.101	.144
		3.33	.034	.062	.582	-.088	.157
	3.33	.33	-.013	.062	.840	-.135	.110
		1.00	-.034	.062	.582	-.157	.088
C11.2	.33	1.00	-.019	.062	.763	-.141	.104
		3.33	.025	.062	.681	-.097	.148
	1.00	.33	.019	.062	.763	-.104	.141
		3.33	.044	.062	.476	-.078	.167
	3.33	.33	-.025	.062	.681	-.148	.097
		1.00	-.044	.062	.476	-.167	.078
C5.1	.33	1.00	.048	.062	.437	-.074	.171
		3.33	-.019	.062	.761	-.141	.104
	1.00	.33	-.048	.062	.437	-.171	.074
		3.33	-.067	.062	.280	-.189	.055
	3.33	.33	.019	.062	.761	-.104	.141
		1.00	.067	.062	.280	-.055	.189
C13.2	.33	1.00	-.002	.062	.974	-.124	.120

<sup>6</sup> Continued overleaf

		<b>3.33</b>		-.012	.062	.849	-.134	.111
	<b>1.00</b>	<b>.33</b>		.002	.062	.974	-.120	.124
		<b>3.33</b>		-.010	.062	.875	-.132	.113
	<b>3.33</b>	<b>.33</b>		.012	.062	.849	-.111	.134
		1.00		.010	.062	.875	-.113	.132
Mock control	.33	1.00		.003	.062	.961	-.119	.125
		3.33		-.026	.062	.677	-.148	.097
	1.00	.33		-.003	.062	.961	-.125	.119
		3.33		-.029	.062	.641	-.151	.094
3.33	.33		.026	.062	.677	-.097	.148	
	1.00		.029	.062	.641	-.094	.151	

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference

**Table II.2.1.6 Dunett's post-hoc analysis for each protein against C11.2 to indicate protein-specific effects in HepG2 cells .** This test compares each protein against C11.2, regardless of which protein concentration is being tested. Significance values are highlighted.

(I) Drug_type	(J) Drug_type	Mean Difference (I-J)	Std. Error	Significance.	95% Confidence Interval	
					Lower Bound	Upper Bound
C5.1	C11.2	.02247	.035693	.923	-.06586	.11080
C13.2	C11.2	.05210	.035693	.396	-.03623	.14043
Mock control	C11.2	.02810	.035693	.847	-.06023	.11643
C12.2	C11.2	.08659	.035693	.050	-.00174	.17492

**Table II.2.1.7 Dunett's post-hoc test based for each concentration against the highest concentration used to indicate dose-dependent effects in HepG2 cells.** This test compares each concentration against the highest concentration tested, regardless of which protein is being tested.

Concentration Level 1	Concentration Level 2	Mean Difference (1-2)	Std. Error	Significance	95% Confidence Interval	
					Lower Bound	Upper Bound
.33	3.33	-.00368	.027647	.987	-.06557	.05820
1.00	3.33	-.00548	.027647	.972	-.06737	.05641

# Bibliography

- AAGAARD, L. & ROSSI, J. J. 2007. RNAi therapeutics: principles, prospects and challenges. *Adv Drug Deliv Rev*, 59, 75-86.
- ABES, R., MOULTON, H. M., CLAIR, P., YANG, S. T., ABES, S., MELIKOV, K., PREVOT, P., YOUNGBLOOD, D. S., IVERSEN, P. L., CHERNOMORDIK, L. V. & LEBLEU, B. 2008. Delivery of steric block morpholino oligomers by (R-X-R)<sub>4</sub> peptides: structure-activity studies. *Nucleic Acids Res*, 36, 6343-54.
- ABES, S., TURNER, J. J., IVANOVA, G. D., OWEN, D., WILLIAMS, D., ARZUMANOV, A., CLAIR, P., GAIT, M. J. & LEBLEU, B. 2007. Efficient splicing correction by PNA conjugation to an R6-Penetratin delivery peptide. *Nucleic Acids Res*, 35, 4495-502.
- AGOONI, A., MODY, N., OWEN, C., CZOPEK, A., ZIMMER, D., BENTIREN-ALJ, M., BENCE, K. K. & DELIBEGOVIC, M. 2011. Liver-specific deletion of protein tyrosine phosphatase (PTP) 1B improves obesity- and pharmacologically induced endoplasmic reticulum stress. *Biochem J*, 438, 369-78.
- AHMAD, F., LI, P. M., MEYEROVITCH, J. & GOLDSTEIN, B. J. 1995. Osmotic loading of neutralizing antibodies demonstrates a role for protein-tyrosine phosphatase 1B in negative regulation of the insulin action pathway. *J Biol Chem*, 270, 20503-8.
- AKASHI, S., OSAWA, R. & NISHIMURA, Y. 2005. Evaluation of protein-DNA binding affinity by electrospray ionization mass spectrometry. *J Am Soc Mass Spectrom*, 16, 116-25.
- ALVES, I. D., BECHARA, C., WALRANT, A., ZALTSMAN, Y., JIAO, C. Y. & SAGAN, S. 2011. Relationships between membrane binding, affinity and cell internalization efficacy of a cell-penetrating peptide: penetratin as a case study. *PLoS One*, 6, e24096.
- AMARZGUIOUI, M., LUNDBERG, P., CANTIN, E., HAGSTROM, J., BEHLKE, M. A. & ROSSI, J. J. 2006. Rational design and in vitro and in vivo delivery of Dicer substrate siRNA. *Nat Protoc*, 1, 508-17.
- AMBROS, V. & LEE, R. C. 2004. Identification of microRNAs and other tiny noncoding RNAs by cDNA cloning. *Methods Mol Biol*, 265, 131-58.
- AMERSHAM 2013. GST Gene Fusion System Handbook.
- AOKI, K., MORIGUCHI, H., YOSHIOKA, T., OKAWA, K. & TABARA, H. 2007. In vitro analyses of the production and activity of secondary small interfering RNAs in *C. elegans*. *EMBO J*, 26, 5007-19.

- ARAI, R., UEDA, H., KITAYAMA, A., KAMIYA, N. & NAGAMUNE, T. 2001. Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng*, 14, 529-32.
- ARAKAWA, T. & TIMASHEFF, S. N. 1983. Preferential interactions of proteins with solvent components in aqueous amino acid solutions. *Arch Biochem Biophys*, 224, 169-77.
- ARAKAWA, T. & TIMASHEFF, S. N. 1984a. Mechanism of protein salting in and salting out by divalent cation salts: balance between hydration and salt binding. *Biochemistry*, 23, 5912-23.
- ARAKAWA, T. & TIMASHEFF, S. N. 1984b. Protein stabilization and destabilization by guanidinium salts. *Biochemistry*, 23, 5924-9.
- ARAKAWA, T., TSUMOTO, K., KITA, Y., CHANG, B. & EJIMA, D. 2007. Biotechnology applications of amino acids in protein purification and formulations. *Amino Acids*, 33, 587-605.
- ARNAU, J., LAURITZEN, C. & PEDERSEN, J. 2006a. Cloning strategy, production and purification of proteins with exopeptidase-cleavable His-tags. *Nat Protoc*, 1, 2326-33.
- ARNAU, J., LAURITZEN, C., PETERSEN, G. E. & PEDERSEN, J. 2006b. Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expr Purif*, 48, 1-13.
- ARYA, S. K., GUO, C., JOSEPHS, S. F. & WONG-STAAAL, F. 1985. Trans-activator gene of human T-lymphotropic virus type III (HTLV-III). *Science*, 229, 69-73.
- ASANTE-APPIAH, E. & KENNEDY, B. P. 2003. Protein tyrosine phosphatases: the quest for negative regulators of insulin action. *Am J Physiol Endocrinol Metab*, 284, E663-70.
- BABIARZ, J. E. & BLELLOCH, R. 2008. Small RNAs - their biogenesis, regulation and function in embryonic stem cells. *StemBook*. Cambridge (MA).
- BABIARZ, J. E., RUBY, J. G., WANG, Y., BARTEL, D. P. & BLELLOCH, R. 2008. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev*, 22, 2773-85.
- BARFORD, D., FLINT, A. J. & TONKS, N. K. 1994. Crystal structure of human protein tyrosine phosphatase 1B. *Science*, 263, 1397-404.
- BARR, A. J. 2010. Protein tyrosine phosphatases as drug targets: strategies and challenges of inhibitor development. *Future Med Chem*, 2, 1563-76.
- BARR, A. J., UGOCHUKWU, E., LEE, W. H., KING, O. N., FILIPPAKOPOULOS, P., ALFANO, I., SAVITSKY, P., BURGESS-BROWN, N. A., MULLER, S. & KNAPP, S. 2009. Large-scale structural analysis of the classical human protein tyrosine phosphatome. *Cell*, 136, 352-63.

- BERNAS, T. & DOBRUCKI, J. 2002. Mitochondrial and nonmitochondrial reduction of MTT: interaction of MTT with TMRE, JC-1, and NAO mitochondrial fluorescent probes. *Cytometry*, 47, 236-42.
- BERNAS, T. & DOBRUCKI, J. W. 2000. The role of plasma membrane in bio-reduction of two tetrazolium salts, MTT, and CTC. *Arch Biochem Biophys*, 380, 108-16.
- BERNSTEIN, E., CAUDY, A. A., HAMMOND, S. M. & HANNON, G. J. 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, 409, 363-6.
- BERRIDGE, M. V., HERST, P. M. & TAN, A. S. 2005. Tetrazolium dyes as tools in cell biology: new insights into their cellular reduction. *Biotechnol Annu Rev*, 11, 127-52.
- BEVILACQUA, P. C. & CECH, T. R. 1996. Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR. *Biochemistry*, 35, 9983-94.
- BEVILACQUA, P. C., GEORGE, C. X., SAMUEL, C. E. & CECH, T. R. 1998. Binding of the protein kinase PKR to RNAs with secondary structure defects: role of the tandem A-G mismatch and noncontiguous helices. *Biochemistry*, 37, 6303-16.
- BIDDINGER, S. B., HERNANDEZ-ONO, A., RASK-MADSEN, C., HAAS, J. T., ALEMAN, J. O., SUZUKI, R., SCAPA, E. F., AGARWAL, C., CAREY, M. C., STEPHANOPOULOS, G., COHEN, D. E., KING, G. L., GINSBERG, H. N. & KAHN, C. R. 2008. Hepatic insulin resistance is sufficient to produce dyslipidemia and susceptibility to atherosclerosis. *Cell Metab*, 7, 125-34.
- BIGGIN, P. C. & SANSOM, M. S. 1999. Interactions of alpha-helices with lipid bilayers: a review of simulation studies. *Biophys Chem*, 76, 161-83.
- BLOCH-GALLEGO, E., LE ROUX, I., JOLIOT, A. H., VOLOVITCH, M., HENDERSON, C. E. & PROCHIANTZ, A. 1993. Antennapedia homeobox peptide enhances growth and branching of embryonic chicken motoneurons in vitro. *Journal of Cell Biology*, 120, 485-92.
- BONFANTI, M., TAVERNA, S., SALMONA, M., D'INCALCI, M. & BROGGINI, M. 1997. p21WAF1-derived peptides linked to an internalization peptide inhibit human cancer cell growth. *Cancer Research*, 57, 1442-6.
- BORNHORST, J. A. & FALKE, J. J. 2000. Purification of proteins using polyhistidine affinity tags. *Methods Enzymol*, 326, 245-54.
- BOUSSIF, O., LEZOUALC'H, F., ZANTA, M. A., MERGNY, M. D., SCHERMAN, D., DEMENEIX, B. & BEHR, J. P. 1995. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc Natl Acad Sci U S A*, 92, 7297-301.
- BRAMSEN, J. B. & KJEMS, J. 2012. Development of Therapeutic-Grade Small Interfering RNAs by Chemical Engineering. *Front Genet*, 3, 154.

- BRAMSEN, J. B., LAURSEN, M. B., NIELSEN, A. F., HANSEN, T. B., BUS, C., LANGKJAER, N., BABU, B. R., HOJLAND, T., ABRAMOV, M., VAN AERSCHOT, A., ODADZIC, D., SMICIUS, R., HAAS, J., ANDREE, C., BARMAN, J., WENSKA, M., SRIVASTAVA, P., ZHOU, C., HONCHARENKO, D., HESS, S., MULLER, E., BOBKOV, G. V., MIKHAILOV, S. N., FAVA, E., MEYER, T. F., CHATTOPADHYAYA, J., ZERIAL, M., ENGELS, J. W., HERDEWIJN, P., WENGEL, J. & KJEMS, J. 2009. A large-scale chemical modification screen identifies design rules to generate siRNAs with high activity, high stability and low toxicity. *Nucleic Acids Res*, 37, 2867-81.
- BULLOUGH, P. A., HUGHSON, F. M., SKEHEL, J. J. & WILEY, D. C. 1994. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature*, 371, 37-43.
- BURROUGHS, A. M., ANDO, Y., DE HOON, M. J., TOMARU, Y., SUZUKI, H., HAYASHIZAKI, Y. & DAUB, C. O. 2011. Deep-sequencing of human Argonaute-associated small RNAs provides insight into miRNA sorting and reveals Argonaute association with RNA fragments of diverse origin. *RNA Biol*, 8, 158-77.
- BUSSO, D., DELAGOUTTE-BUSSO, B. & MORAS, D. 2005. Construction of a set Gateway-based destination vectors for high-throughput cloning and expression screening in *Escherichia coli*. *Analytical Biochemistry*, 343, 313-21.
- BYCROFT, M., GRUNERT, S., MURZIN, A. G., PROCTOR, M. & ST JOHNSTON, D. 1995. NMR solution structure of a dsRNA binding domain from *Drosophila* staufen protein reveals homology to the N-terminal domain of ribosomal protein S5. *EMBO J*, 14, 3563-71.
- CALNAN, B. J., BIANCALANA, S., HUDSON, D. & FRANKEL, A. D. 1991. Analysis of arginine-rich peptides from the HIV Tat protein reveals unusual features of RNA-protein recognition. *Genes Dev*, 5, 201-10.
- CARPINO, P. A. & GOODWIN, B. 2010. Diabetes area participation analysis: a review of companies and targets described in the 2008 - 2010 patent literature. *Expert Opin Ther Pat*, 20, 1627-51.
- CASTANOTTO, D., SAKURAI, K., LINGEMAN, R., LI, H., SHIVELY, L., AAGAARD, L., SOIFER, H., GATIGNOL, A., RIGGS, A. & ROSSI, J. J. 2007. Combinatorial delivery of small interfering RNAs reduces RNAi efficacy by selective incorporation into RISC. *Nucleic Acids Res*, 35, 5154-64.
- CAZES, J. 2001. *Encyclopedia of chromatography*, New York, Marcel Dekker.
- CHALFIE, M., TU, Y., EUSKIRCHEN, G., WARD, W. W. & PRASHER, D. C. 1994. Green fluorescent protein as a marker for gene expression. *Science*, 263, 802-5.
- CHEN, C. X., CHO, D. S., WANG, Q., LAI, F., CARTER, K. C. & NISHIKURA, K. 2000. A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains. *RNA*, 6, 755-67.

- CHEN, X., BAI, Y., ZARO, J. L. & SHEN, W. C. 2010. Design of an in vivo cleavable disulfide linker in recombinant fusion proteins. *Biotechniques*, 49, 513-8.
- CHENG, A., UETANI, N., SIMONCIC, P. D., CHAUBEY, V. P., LEE-LOY, A., MCGLADE, C. J., KENNEDY, B. P. & TREMBLAY, M. L. 2002. Attenuation of leptin action and regulation of obesity by protein tyrosine phosphatase 1B. *Dev Cell*, 2, 497-503.
- CHENG, C. J. & SALTZMAN, W. M. 2011. Enhanced siRNA delivery into cells by exploiting the synergy between targeting ligands and cell-penetrating peptides. *Biomaterials*, 32, 6194-203.
- CHESSHIRE, J. A. & HIPKISS, A. R. 1989. Low-Temperatures Stabilize Interferon-Alpha-2 against Proteolysis in *Methylophilus-Methylotrophus* and *Escherichia-Coli*. *Appl Microbiol Biotechnol*, 31, 158-162.
- CHIU, Y. L., ALI, A., CHU, C. Y., CAO, H. & RANA, T. M. 2004. Visualizing a correlation between siRNA localization, cellular uptake, and RNAi in living cells. *Chem Biol*, 11, 1165-75.
- CHIU, Y. L. & RANA, T. M. 2003. siRNA function in RNAi: a chemical modification analysis. *RNA*, 9, 1034-48.
- CHRISTIAENS, B., SYMOENS, S., VERHEYDEN, S., ENGELBORGHES, Y., JOLIOT, A., PROCHIANTZ, A., VANDEKERCKHOVE, J., ROSSENEU, M. & VANLOO, B. 2002. Tryptophan fluorescence study of the interaction of penetratin peptides with model membranes. *Eur J Biochem*, 269, 2918-26.
- CONSOLE, S., MARTY, C., GARCIA-EICHEVERRIA, C., SCHWENDENER, R. & BALLMER-HOFER, K. 2003. Antennapedia and HIV transactivator of transcription (TAT) "protein transduction domains" promote endocytosis of high molecular weight cargo upon binding to cell surface glycosaminoglycans. *J Biol Chem*, 278, 35109-14.
- COOLIDGE, C. J. & PATTON, J. G. 2000. A new double-stranded RNA-binding protein that interacts with PKR. *Nucleic Acids Res*, 28, 1407-17.
- COSGAYA, J. M., ARANDA, A., CRUCES, J. & MARTIN-BLANCO, E. 1998. Neuronal differentiation of PC12 cells induced by engrailed homeodomain is DNA-binding specific and independent of MAP kinases. *J Cell Sci*, 111, 2377-2384.
- COURSINDEL, T., JARVER, P. & GAIT, M. J. 2012. Peptide-based in vivo delivery agents for oligonucleotides and siRNA. *Nucleic Acid Ther*, 22, 71-6.
- CRICK, F. 1970. Central dogma of molecular biology. *Nature*, 227, 561-3.
- CROMBEZ, L., ALDRIAN-HERRADA, G., KONATE, K., NGUYEN, Q. N., MCMASTER, G. K., BRASSEUR, R., HEITZ, F. & DIVITA, G. 2009. A new potent secondary amphipathic cell-penetrating peptide for siRNA delivery into mammalian cells. *Mol Ther*, 17, 95-103.

- CROMBEZ, L., CHARNET, A., MORRIS, M. C., ALDRIAN-HERRADA, G., HEITZ, F. & DIVITA, G. 2007. A non-covalent peptide-based strategy for siRNA delivery. *Biochem Soc Trans*, 35, 44-6.
- CULLEN, B. R. 1986. Trans-activation of human immunodeficiency virus occurs via a bimodal mechanism. *Cell*, 46, 973-82.
- CZECH, B. & HANNON, G. J. 2011. Small RNA sorting: matchmaking for Argonautes. *Nat Rev Genet*, 12, 19-31.
- DABO, S. & MEURS, E. F. 2012. dsRNA-dependent protein kinase PKR and its role in stress, signaling and HCV infection. *Viruses*, 4, 2598-635.
- DAVIDSON, T. J., HAREL, S., ARBOLEDA, V. A., PRUNELL, G. F., SHELANSKI, M. L., GREENE, L. A. & TROY, C. M. 2004. Highly efficient small interfering RNA delivery to primary mammalian neurons induces MicroRNA-like effects before mRNA degradation. *J Neurosci*, 24, 10040-6.
- DE FOUGEROLLES, A., VORNLOCHER, H. P., MARAGANORE, J. & LIEBERMAN, J. 2007. Interfering with disease: a progress report on siRNA-based therapeutics. *Nat Rev Drug Discov*, 6, 443-53.
- DEFRONZO, R. A. 2009. Banting Lecture. From the triumvirate to the ominous octet: a new paradigm for the treatment of type 2 diabetes mellitus. *Diabetes*, 58, 773-95.
- DELIBEGOVIC, M., BENCE, K. K., MODY, N., HONG, E. G., KO, H. J., KIM, J. K., KAHN, B. B. & NEEL, B. G. 2007. Improved glucose homeostasis in mice with muscle-specific deletion of protein-tyrosine phosphatase 1B. *Mol Cell Biol*, 27, 7727-34.
- DELIBEGOVIC, M., ZIMMER, D., KAUFFMAN, C., RAK, K., HONG, E. G., CHO, Y. R., KIM, J. K., KAHN, B. B., NEEL, B. G. & BENCE, K. K. 2009. Liver-specific deletion of protein-tyrosine phosphatase 1B (PTP1B) improves metabolic syndrome and attenuates diet-induced endoplasmic reticulum stress. *Diabetes*, 58, 590-9.
- DEROSSI, D., CALVET, S., TREMBLEAU, A., BRUNISSEN, A., CHASSAING, G. & PROCHIANTZ, A. 1996. Cell internalization of the third helix of the Antennapedia homeodomain is receptor-independent. *J Biol Chem*, 271, 18188-93.
- DEROSSI, D., JOLIOT, A. H., CHASSAING, G. & PROCHIANTZ, A. 1994. The third helix of the Antennapedia homeodomain translocates through biological membranes. *J Biol Chem*, 269, 10444-50.
- DESHAYES, S., MORRIS, M., HEITZ, F. & DIVITA, G. 2008. Delivery of proteins and nucleic acids using a non-covalent peptide-based strategy. *Adv Drug Deliv Rev*, 60, 537-47.
- DESPLAN, C., THEIS, J. & O'FARRELL, P. H. 1985. The Drosophila developmental gene, engrailed, encodes a sequence-specific DNA binding activity. *Nature*, 318, 630-5.

- DETZER, A., OVERHOFF, M., MESCALCHIN, A., ROMPF, M. & SCZAKIEL, G. 2008. Phosphorothioate-stimulated cellular uptake of siRNA: a cell culture model for mechanistic studies. *Curr Pharm Des*, 14, 3666-73.
- DETZER, A., OVERHOFF, M., WUNSCH, W., ROMPF, M., TURNER, J. J., IVANOVA, G. D., GAIT, M. J. & SCZAKIEL, G. 2009. Increased RNAi is related to intracellular release of siRNA via a covalently attached signal peptide. *RNA*, 15, 627-36.
- DIABETESUK. 2013. *Diabetes Prevalence 2013* [Online]. Available: [http://www.diabetes.org.uk/About\\_us/What-we-say/Statistics/Diabetes-prevalence-2013/](http://www.diabetes.org.uk/About_us/What-we-say/Statistics/Diabetes-prevalence-2013/).
- DIAZ-PENDON, J. A., LI, F., LI, W. X. & DING, S. W. 2007. Suppression of antiviral silencing by cucumber mosaic virus 2b protein in Arabidopsis is associated with drastically reduced accumulation of three classes of viral small interfering RNAs. *Plant Cell*, 19, 2053-63.
- DIEBOLD, S. S., MASSACRIER, C., AKIRA, S., PATUREL, C., MOREL, Y. & REIS E SOUSA, C. 2006. Nucleic acid agonists for Toll-like receptor 7 are defined by the presence of uridine ribonucleotides. *European Journal of Immunology*, 36, 3256-67.
- DIETZ, G. P. & BAHR, M. 2004. Delivery of bioactive molecules into the cell: the Trojan horse approach. *Mol Cell Neurosci*, 27, 85-131.
- DINGWALL, C., ERNBERG, I., GAIT, M. J., GREEN, S. M., HEAPHY, S., KARN, J., LOWE, A. D., SINGH, M. & SKINNER, M. A. 1990. HIV-1 tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure. *EMBO J*, 9, 4145-53.
- DOENCH, J. G., PETERSEN, C. P. & SHARP, P. A. 2003. siRNAs can function as miRNAs. *Genes Dev*, 17, 438-42.
- DOM, G., SHAW-JACKSON, C., MATIS, C., BOUFFIOUX, O., PICARD, J. J., PROCHIANTZ, A., MINGEOT-LECLERCQ, M. P., BRASSEUR, R. & REZSOHAZY, R. 2003. Cellular uptake of Antennapedia Penetratin peptides is a two-step process in which phase transfer precedes a tryptophan-dependent translocation. *Nucleic Acids Res*, 31, 556-61.
- DONZE, O., JAGUS, R., KOROMILAS, A. E., HERSHEY, J. W. & SONENBERG, N. 1995. Abrogation of translation initiation factor eIF-2 phosphorylation causes malignant transformation of NIH 3T3 cells. *EMBO J*, 14, 3828-34.
- DRAKE, P. G. & POSNER, B. I. 1998. Insulin receptor-associated protein tyrosine phosphatase(s): role in insulin action. *Mol Cell Biochem*, 182, 79-89.
- DRIN, G., COTTIN, S., BLANC, E., REES, A. R. & TEMSAMANI, J. 2003. Studies on the internalization mechanism of cationic cell-penetrating peptides. *J Biol Chem*, 278, 31192-201.

- DUCHARDT, F., FOTIN-MLECZEK, M., SCHWARZ, H., FISCHER, R. & BROCK, R. 2007. A comprehensive model for the cellular uptake of cationic cell-penetrating peptides. *Traffic*, 8, 848-66.
- DUCHARDT, F., RUTTEKOLK, I. R., VERDURMEN, W. P., LORTAT-JACOB, H., BURCK, J., HUFNAGEL, H., FISCHER, R., VAN DEN HEUVEL, M., LOWIK, D. W., VUISTER, G. W., ULRICH, A., DE WAARD, M. & BROCK, R. 2009. A cell-penetrating peptide derived from human lactoferrin with conformation-dependent uptake efficiency. *J Biol Chem*, 284, 36099-108.
- DYSON, M. R., SHADBOLT, S. P., VINCENT, K. J., PERERA, R. L. & MCCAFFERTY, J. 2004. Production of soluble mammalian proteins in *Escherichia coli*: identification of protein features that correlate with successful expression. *BMC Biotechnol*, 4, 32.
- DZANANOVIC, E., PATEL, T. R., DEO, S., MCELENEY, K., STETEFELD, J. & MCKENNA, S. A. 2013. Recognition of viral RNA stem-loops by the tandem double-stranded RNA binding domains of PKR. *RNA*, 19, 333-44.
- EGUCHI, A., MEADE, B. R., CHANG, Y. C., FREDRICKSON, C. T., WILLERT, K., PURI, N. & DOWDY, S. F. 2009. Efficient siRNA delivery into primary cells by a peptide transduction domain-dsRNA binding domain fusion protein. *Nat Biotechnol*, 27, 567-71.
- EIRIKSDOTTIR, E., KONATE, K., LANGEL, U., DIVITA, G. & DESHAYES, S. 2010. Secondary structure of cell-penetrating peptides controls membrane interaction and insertion. *Biochim Biophys Acta*, 1798, 1119-28.
- EJIMA, D., WATANABE, M., SATO, Y., DATE, M., YAMADA, N. & TAKAHARA, Y. 1999. High yield refolding and purification process for recombinant human interleukin-6 expressed in *Escherichia coli*. *Biotechnol Bioeng*, 62, 301-10.
- EL-ANDALOUSSI, S., JARVER, P., JOHANSSON, H. J. & LANGEL, U. 2007a. Cargo-dependent cytotoxicity and delivery efficacy of cell-penetrating peptides: a comparative study. *Biochem J*, 407, 285-92.
- EL-ANDALOUSSI, S., JOHANSSON, H., MAGNUSDOTTIR, A., JARVER, P., LUNDBERG, P. & LANGEL, U. 2005. TP10, a delivery vector for decoy oligonucleotides targeting the Myc protein. *J Control Release*, 110, 189-201.
- EL-ANDALOUSSI, S., JOHANSSON, H. J., HOLM, T. & LANGEL, U. 2007b. A novel cell-penetrating peptide, M918, for efficient delivery of proteins and peptide nucleic acids. *Mol Ther*, 15, 1820-6.
- EL-ANDALOUSSI, S., JOHANSSON, H. J., LUNDBERG, P. & LANGEL, U. 2006. Induction of splice correction by cell-penetrating peptide nucleic acids. *J Gene Med*, 8, 1262-73.
- EL-ANDALOUSSI, S. E., LEHTO, T., MAGER, I., ROSENTHAL-AIZMAN, K., OPREA, II, SIMONSON, O. E., SORK, H., EZZAT, K., COPOLOVICI, D. M., KURRIKOFF, K., VIOLA, J. R., ZAGHLOUL, E. M., SILLARD, R.,

- JOHANSSON, H. J., SAID HASSANE, F., GUTERSTAM, P., SUHORUTSENKO, J., MORENO, P. M., OSKOLKOV, N., HALLDIN, J., TEDEBARK, U., METSPALU, A., LEBLEU, B., LEHTIO, J., SMITH, C. I. & LANGEL, U. 2011. Design of a peptide-based vector, PepFect6, for efficient delivery of siRNA in cell culture and systemically in vivo. *Nucleic Acids Res*, 39, 3972-87.
- EL-SAYED, A., FUTAKI, S. & HARASHIMA, H. 2009. Delivery of macromolecules using arginine-rich cell-penetrating peptides: ways to overcome endosomal entrapment. *AAPS J*, 11, 13-22.
- ELBASHIR, S. M., HARBORTH, J., LENDECKEL, W., YALCIN, A., WEBER, K. & TUSCHL, T. 2001a. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, 411, 494-8.
- ELBASHIR, S. M., MARTINEZ, J., PATKANIOWSKA, A., LENDECKEL, W. & TUSCHL, T. 2001b. Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO J*, 20, 6877-88.
- ELCHEBLY, M., PAYETTE, P., MICHALISZYN, E., CROMLISH, W., COLLINS, S., LOY, A. L., NORMANDIN, D., CHENG, A., HIMMS-HAGEN, J., CHAN, C. C., RAMACHANDRAN, C., GRESSER, M. J., TREMBLAY, M. L. & KENNEDY, B. P. 1999. Increased insulin sensitivity and obesity resistance in mice lacking the protein tyrosine phosphatase-1B gene. *Science*, 283, 1544-8.
- ELLIOTT, G. & O'HARE, P. 1997. Intercellular trafficking and protein delivery by a herpesvirus structural protein. *Cell*, 88, 223-33.
- ELMEN, J., ZHANG, H. Y., ZUBER, B., LJUNGBERG, K., WAHREN, B., WAHLESTEDT, C. & LIANG, Z. 2004. Locked nucleic acid containing antisense oligonucleotides enhance inhibition of HIV-1 genome dimerization and inhibit virus replication. *FEBS Lett*, 578, 285-90.
- FABANI, M. M. & GAIT, M. J. 2008. miR-122 targeting with LNA/2'-O-methyl oligonucleotide mixmers, peptide nucleic acids (PNA), and PNA-peptide conjugates. *RNA*, 14, 336-46.
- FAHRAEUS, R., LAIN, S., BALL, K. L. & LANE, D. P. 1998. Characterization of the cyclin-dependent kinase inhibitory domain of the INK4 family as a model for a synthetic tumour suppressor molecule. *Oncogene*, 16, 587-96.
- FERRANDON, D., ELPHICK, L., NUSSLEIN-VOLHARD, C. & ST JOHNSTON, D. 1994. Stufen protein associates with the 3'UTR of bicoid mRNA to form particles that move in a microtubule-dependent manner. *Cell*, 79, 1221-32.
- FIERCEBIOTECH 2014. Novartis slams the brakes on once-ambitious RNAi development efforts.
- FIERRO-MONTI, I. & MATHEWS, M. B. 2000. Proteins binding to duplexed RNA: one motif, multiple functions. *Trends Biochem Sci*, 25, 241-6.

- FILIPPOV, V., SOLOVYEV, V., FILIPPOVA, M. & GILL, S. S. 2000. A novel type of RNase III family proteins in eukaryotes. *Gene*, 245, 213-21.
- FIRE, A., XU, S., MONTGOMERY, M. K., KOSTAS, S. A., DRIVER, S. E. & MELLO, C. C. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, 391, 806-11.
- FONSECA, S. B., PEREIRA, M. P. & KELLEY, S. O. 2009. Recent advances in the use of cell-penetrating peptides for medical and biological applications. *Adv Drug Deliv Rev*, 61, 953-64.
- FRANCIS, D. M. & PAGE, R. 2001. Strategies to Optimize Protein Expression in *E. coli*. *Current Protocols in Protein Science*. John Wiley & Sons, Inc.
- FRANKEL, A. D. & PABO, C. O. 1988. Cellular uptake of the tat protein from human immunodeficiency virus. *Cell*, 55, 1189-93.
- FROKJAER, S. & OTZEN, D. E. 2005. Protein drug stability: a formulation challenge. *Nat Rev Drug Discov*, 4, 298-306.
- FUCHS, S. M. & RAINES, R. T. 2004. Pathway for polyarginine entry into mammalian cells. *Biochemistry*, 43, 2438-44.
- FUJIMOTO, K., HOSOTANI, R., MIYAMOTO, Y., DOI, R., KOSHIBA, T., OTAKA, A., FUJII, N., BEAUCHAMP, R. D. & IMAMURA, M. 2000. Inhibition of pRb phosphorylation and cell cycle progression by an antenapedia-p16(INK4A) fusion peptide in pancreatic cancer cells. *Cancer Letters*, 159, 151-8.
- FUTAKI, S., SUZUKI, T., OHASHI, W., YAGAMI, T., TANAKA, S., UEDA, K. & SUGIURA, Y. 2001. Arginine-rich peptides. An abundant source of membrane-permeable peptides having potential as carriers for intracellular protein delivery. *J Biol Chem*, 276, 5836-40.
- GALABRU, J. & HOVANESSIAN, A. 1987. Autophosphorylation of the protein kinase dependent on double-stranded RNA. *J Biol Chem*, 262, 15538-44.
- GASTEIGER E., H. C., GATTIKER A., DUVAUD S., WILKINS M.R., APPEL R.D., BAIROCH A. 2005. *Protein Identification and Analysis Tools on the ExPASy Server*, Swiss Institute of Bioinformatics, Switzerland, Humana Press (2005). .
- GATIGNOL, A., BUCKLER-WHITE, A., BERKHOUT, B. & JEANG, K. T. 1991. Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR. *Science*, 251, 1597-600.
- GEOGHEGAN, J. C., GILMORE, B. L. & DAVIDSON, B. L. 2012. Gene Silencing Mediated by siRNA-binding Fusion Proteins Is Attenuated by Double-stranded RNA-binding Domain Structure. *Mol Ther Nucleic Acids*, 1, e53.
- GERLSMA, S. Y. 1968. Reversible denaturation of ribonuclease in aqueous solutions as influenced by polyhydric alcohols and some other additives. *J Biol Chem*, 243, 957-61.

- GOLDSTEIN, B. J., AHMAD, F., DING, W., LI, P. M. & ZHANG, W. R. 1998. Regulation of the insulin signalling pathway by cellular protein-tyrosine phosphatases. *Mol Cell Biochem*, 182, 91-9.
- GOODING, M., BROWNE, L. P., QUINTEIRO, F. M. & SELWOOD, D. L. 2012. siRNA delivery: from lipids to cell-penetrating peptides and their mimics. *Chem Biol Drug Des*, 80, 787-809.
- GRASLUND, S., NORDLUND, P., WEIGELT, J., HALLBERG, B. M., BRAY, J., GILEADI, O., KNAPP, S., OPPERMANN, U., ARROWSMITH, C., HUI, R., MING, J., DHE-PAGANON, S., PARK, H. W., SAVCHENKO, A., YEE, A., EDWARDS, A., VINCENTELLI, R., CABBILLAU, C., KIM, R., KIM, S. H., RAO, Z., SHI, Y., TERWILLIGER, T. C., KIM, C. Y., HUNG, L. W., WALDO, G. S., PELEG, Y., ALBECK, S., UNGER, T., DYM, O., PRILUSKY, J., SUSSMAN, J. L., STEVENS, R. C., LESLEY, S. A., WILSON, I. A., JOACHIMIAK, A., COLLART, F., DEMENTIEVA, I., DONNELLY, M. I., ESCHENFELDT, W. H., KIM, Y., STOLS, L., WU, R., ZHOU, M., BURLEY, S. K., EMTAGE, J. S., SAUDER, J. M., THOMPSON, D., BAIN, K., LUZ, J., GHEYI, T., ZHANG, F., ATWELL, S., ALMO, S. C., BONANNO, J. B., FISER, A., SWAMINATHAN, S., STUDIER, F. W., CHANCE, M. R., SALI, A., ACTON, T. B., XIAO, R., ZHAO, L., MA, L. C., HUNT, J. F., TONG, L., CUNNINGHAM, K., INOUE, M., ANDERSON, S., JANJUA, H., SHASTRY, R., HO, C. K., WANG, D., WANG, H., JIANG, M., MONTELIONE, G. T., STUART, D. I., OWENS, R. J., DAENKE, S., SCHUTZ, A., HEINEMANN, U., YOKOYAMA, S., BUSSOW, K. & GUNSALUS, K. C. 2008a. Protein production and purification. *Nat Methods*, 5, 135-46.
- GRASLUND, S., SAGEMARK, J., BERGLUND, H., DAHLGREN, L. G., FLORES, A., HAMMARSTROM, M., JOHANSSON, I., KOTENYOVA, T., NILSSON, M., NORDLUND, P. & WEIGELT, J. 2008b. The use of systematic N- and C-terminal deletions to promote production and structural studies of recombinant proteins. *Protein Expr Purif*, 58, 210-21.
- GREEN, I., CHRISTISON, R., VOYCE, C. J., BUNDELL, K. R. & LINDSAY, M. A. 2003. Protein transduction domains: are they delivering? *Trends Pharmacol Sci*, 24, 213-5.
- GREGORY, R. I., CHENDRIMADA, T. P., COOCH, N. & SHIEKHATTAR, R. 2005. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell*, 123, 631-40.
- GRIMM, D., STREETZ, K. L., JOPLING, C. L., STORM, T. A., PANDEY, K., DAVIS, C. R., MARION, P., SALAZAR, F. & KAY, M. A. 2006. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature*, 441, 537-41.
- GRISSHAMMER, R. & TUCKER, J. 1997. Quantitative evaluation of neurotensin receptor purification by immobilized metal affinity chromatography. *Protein Expr Purif*, 11, 53-60.

- GU, W., CLAYCOMB, J. M., BATISTA, P. J., MELLO, C. C. & CONTE, D. 2011. Cloning Argonaute-associated small RNAs from *Caenorhabditis elegans*. *Methods Mol Biol*, 725, 251-80.
- GUO, S. 2014. Insulin signaling, resistance, and the metabolic syndrome: insights from mouse models into disease mechanisms. *J Endocrinol*, 220, T1-T23.
- GUO, S. & KEMPHUES, K. J. 1995. par-1, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell*, 81, 611-20.
- GUTERSTAM, P., MADANI, F., HIROSE, H., TAKEUCHI, T., FUTAKI, S., EL ANDALOUSSI, S., GRASLUND, A. & LANGEL, U. 2009. Elucidating cell-penetrating peptide mechanisms of action for membrane interaction, cellular uptake, and translocation utilizing the hydrophobic counter-anion pyrenebutyrate. *Biochim Biophys Acta*, 1788, 2509-17.
- HA, M. & KIM, V. N. 2014. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol*, 15, 509-24.
- HAI, F. G., MARKOVA, B., KLAMAN, L. D., BOHMER, F. D. & NEEL, B. G. 2003. Regulation of receptor tyrosine kinase signaling by protein tyrosine phosphatase-1B. *J Biol Chem*, 278, 739-44.
- HALL, A. H., WAN, J., SPESOCK, A., SERGUEEVA, Z., SHAW, B. R. & ALEXANDER, K. A. 2006. High potency silencing by single-stranded boranophosphate siRNA. *Nucleic Acids Res*, 34, 2773-81.
- HAMMARSTROM, M., HELLGREN, N., VAN DEN BERG, S., BERGLUND, H. & HARD, T. 2002. Rapid screening for improved solubility of small human proteins produced as fusion proteins in *Escherichia coli*. *Protein Science*, 11, 313-21.
- HANSEN, A., SCHAFFER, I., KNAPPE, D., SEIBEL, P. & HOFFMANN, R. 2012. Intracellular toxicity of proline-rich antimicrobial peptides shuttled into mammalian cells by the cell-penetrating peptide penetratin. *Antimicrob Agents Chemother*, 56, 5194-201.
- HARBORTH, J., ELBASHIR, S. M., VANDENBURGH, K., MANNINGA, H., SCARINGE, S. A., WEBER, K. & TUSCHL, T. 2003. Sequence, chemical, and structural variation of small interfering RNAs and short hairpin RNAs and the effect on mammalian gene silencing. *Antisense Nucleic Acid Drug Dev*, 13, 83-105.
- HASSANE, F. S., ABES, R., EL ANDALOUSSI, S., LEHTO, T., SILLARD, R., LANGEL, U. & LEBLEU, B. 2011. Insights into the cellular trafficking of splice redirecting oligonucleotides complexed with chemically modified cell-penetrating peptides. *J Control Release*, 153, 163-72.
- HAUBER, J., MALIM, M. H. & CULLEN, B. R. 1989. Mutational analysis of the conserved basic domain of human immunodeficiency virus tat protein. *Journal of Virology*, 63, 1181-7.

- HEINICKE, L. A., WONG, C. J., LARY, J., NALLAGATLA, S. R., DIEGELMAN-PARENTE, A., ZHENG, X., COLE, J. L. & BEVILACQUA, P. C. 2009. RNA dimerization promotes PKR dimerization and activation. *J Mol Biol*, 390, 319-38.
- HEITZ, F., MORRIS, M. C. & DIVITA, G. 2009. Twenty years of cell-penetrating peptides: from molecular mechanisms to therapeutics. *Br J Pharmacol*, 157, 195-206.
- HELLMAN, L. M. & FRIED, M. G. 2007. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat Protoc*, 2, 1849-61.
- HENCHEY, L. K., JOCHIM, A. L. & ARORA, P. S. 2008. Contemporary strategies for the stabilization of peptides in the alpha-helical conformation. *Curr Opin Chem Biol*, 12, 692-7.
- HIGASHIJIMA, T., UZU, S., NAKAJIMA, T. & ROSS, E. M. 1988. Mastoparan, a peptide toxin from wasp venom, mimics receptors by activating GTP-binding regulatory proteins (G proteins). *J Biol Chem*, 263, 6491-4.
- HOLM, T., RAAGEL, H., ANDALOUSSI, S. E., HEIN, M., MAE, M., POOGA, M. & LANGEL, U. 2011. Retro-inversion of certain cell-penetrating peptides causes severe cellular toxicity. *Biochim Biophys Acta*, 1808, 1544-51.
- HORNUNG, V., GUENTHNER-BILLER, M., BOURQUIN, C., ABLASSER, A., SCHLEE, M., UEMATSU, S., NORONHA, A., MANOHARAN, M., AKIRA, S., DE FOUGEROLLES, A., ENDRES, S. & HARTMANN, G. 2005. Sequence-specific potent induction of IFN-alpha by short interfering RNA in plasmacytoid dendritic cells through TLR7. *Nature Medicine*, 11, 263-70.
- HU, G. 2009. Peptide Synthesis in Pharmaceutical Manufacturing. . *BioProcess J*, 8, 51-53.
- HUNTZINGER, E. & IZAURRALDE, E. 2011. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet*, 12, 99-110.
- HUOTARI, J. & HELENIUS, A. 2011. Endosome maturation. *EMBO J*, 30, 3481-500.
- IKI, T., YOSHIKAWA, M., NISHIKIORI, M., JAUDAL, M. C., MATSUMOTO-YOKOYAMA, E., MITSUHARA, I., MESHI, T. & ISHIKAWA, M. 2010. In vitro assembly of plant RNA-induced silencing complexes facilitated by molecular chaperone HSP90. *Mol Cell*, 39, 282-91.
- INVITROGEN 2007. Stealth RNAi Reporter Control Duplexes Manual 1-4.
- IRONS, B. K. & MINZE, M. G. 2014. Drug treatment of type 2 diabetes mellitus in patients for whom metformin is contraindicated. *Diabetes Metab Syndr Obes*, 7, 15-24.
- IVANOVA, G. D., ARZUMANOV, A., ABES, R., YIN, H., WOOD, M. J., LEBLEU, B. & GAIT, M. J. 2008. Improved cell-penetrating peptide-PNA conjugates for

- splicing redirection in HeLa cells and exon skipping in mdx mouse muscle. *Nucleic Acids Res*, 36, 6418-28.
- JACKSON, A. L., BARTZ, S. R., SCHELTER, J., KOBAYASHI, S. V., BURCHARD, J., MAO, M., LI, B., CAVET, G. & LINSLEY, P. S. 2003. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol*, 21, 635-7.
- JACOB, F. & MONOD, J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol*, 3, 318-56.
- JARVER, P., COURSIINDEL, T., ANDALOUSSI, S. E., GODFREY, C., WOOD, M. J. & GAIT, M. J. 2012. Peptide-mediated Cell and In Vivo Delivery of Antisense Oligonucleotides and siRNA. *Mol Ther Nucleic Acids*, 1, e27.
- JENSSEN, H. & ASPMO, S. I. 2008. Serum stability of peptides. *Methods Mol Biol*, 494, 177-86.
- JIAO, C. Y., DELAROCHE, D., BURLINA, F., ALVES, I. D., CHASSAING, G. & SAGAN, S. 2009. Translocation and endocytosis for cell-penetrating peptide internalization. *J Biol Chem*, 284, 33957-65.
- JINEK, M. & DOUDNA, J. A. 2009. A three-dimensional view of the molecular machinery of RNA interference. *Nature*, 457, 405-12.
- JOHNSON, K. J., PECK, A. R., LIU, C., TRAN, T. H., UTAMA, F. E., SJOLUND, A. B., SCHABER, J. D., WITKIEWICZ, A. K. & RUI, H. 2010. PTP1B suppresses prolactin activation of Stat5 in breast cancer cells. *Am J Pathol*, 177, 2971-83.
- JOHNSTON, M., GEOFFROY, M. C., SOBALA, A., HAY, R. & HUTVAGNER, G. 2010. HSP90 protein stabilizes unloaded argonaute complexes and microscopic P-bodies in human cells. *Mol Biol Cell*, 21, 1462-9.
- JOLIOT, A., PERNELLE, C., DEAGOSTINI-BAZIN, H. & PROCHIANTZ, A. 1991a. Antennapedia homeobox peptide regulates neural morphogenesis. *Proc Natl Acad Sci U S A*, 88, 1864-8.
- JOLIOT, A. H., TRILLER, A., VOLOVITCH, M., PERNELLE, C. & PROCHIANTZ, A. 1991b. alpha-2,8-Polysialic acid is the neuronal surface receptor of antennapedia homeobox peptide. *New Biol*, 3, 1121-34.
- JONES, S. W., CHRISTISON, R., BUNDELL, K., VOYCE, C. J., BROCKBANK, S. M., NEWHAM, P. & LINDSAY, M. A. 2005. Characterisation of cell-penetrating peptide-mediated peptide delivery. *Br J Pharmacol*, 145, 1093-102.
- JORGENSEN, L., HOSTRUP, S., MOELLER, E. H. & GROHGANZ, H. 2009. Recent trends in stabilising peptides and proteins in pharmaceutical formulation - considerations in the choice of excipients. *Expert Opin Drug Deliv*, 6, 1219-30.
- JULIANO, R., ALAM, M. R., DIXIT, V. & KANG, H. 2008. Mechanisms and strategies for effective delivery of antisense and siRNA oligonucleotides. *Nucleic Acids Res*, 36, 4158-71.

- JUNE, R. K., GOGOI, K., EGUCHI, A., CUI, X. S. & DOWDY, S. F. 2010. Synthesis of a pH-sensitive nitrilotriacetic linker to peptide transduction domains to enable intracellular delivery of histidine imidazole ring-containing macromolecules. *J Am Chem Soc*, 132, 10680-2.
- KAPLAN, I. M., WADIA, J. S. & DOWDY, S. F. 2005. Cationic TAT peptide transduction domain enters cells by macropinocytosis. *J Control Release*, 102, 247-53.
- KARIKO, K., BHUYAN, P., CAPODICCI, J. & WEISSMAN, D. 2004. Small interfering RNAs mediate sequence-independent gene suppression and induce immune activation by signaling through toll-like receptor 3. *Journal of Immunology*, 172, 6545-9.
- KATO, D., MIYAZAWA, K., RUAS, M., STARBORG, M., WADA, I., OKA, T., SAKAI, T., PETERS, G. & HARA, E. 1998. Features of replicative senescence induced by direct addition of antenapedia-p16INK4A fusion protein to human diploid fibroblasts. *FEBS Lett*, 427, 203-8.
- KELEMEN, B. R., HSIAO, K. & GOUELI, S. A. 2002. Selective in vivo inhibition of mitogen-activated protein kinase activation using cell-permeable peptides. *J Biol Chem*, 277, 8741-8.
- KERKIS, A., HAYASHI, M. A., YAMANE, T. & KERKIS, I. 2006. Properties of cell penetrating peptides (CPPs). *IUBMB Life*, 58, 7-13.
- KHARRAT, A., MACIAS, M. J., GIBSON, T. J., NILGES, M. & PASTORE, A. 1995. Structure of the dsRNA binding domain of E. coli RNase III. *EMBO J*, 14, 3572-84.
- KIERZEK, E., PASTERNAK, A., PASTERNAK, K., GDANIEC, Z., YILDIRIM, I., TURNER, D. H. & KIERZEK, R. 2009. Contributions of stacking, preorganization, and hydrogen bonding to the thermodynamic stability of duplexes between RNA and 2'-O-methyl RNA with locked nucleic acids. *Biochemistry*, 48, 4377-87.
- KIM, D. H. & ROSSI, J. J. 2007. Strategies for silencing human disease using RNA interference. *Nature Reviews Genetics*, 8, 173-84.
- KIM, J., LEE, S. H., CHOE, J. & PARK, T. G. 2009a. Intracellular small interfering RNA delivery using genetically engineered double-stranded RNA binding protein domain. *J Gene Med*, 11, 804-12.
- KIM, V. N., HAN, J. & SIOMI, M. C. 2009b. Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol*, 10, 126-39.
- KIM, W. J., CHRISTENSEN, L. V., JO, S., YOCKMAN, J. W., JEONG, J. H., KIM, Y. H. & KIM, S. W. 2006. Cholesteryl oligoarginine delivering vascular endothelial growth factor siRNA effectively inhibits tumor growth in colon adenocarcinoma. *Mol Ther*, 14, 343-50.

- KLAMAN, L. D., BOSS, O., PERONI, O. D., KIM, J. K., MARTINO, J. L., ZABOLOTNY, J. M., MOGHAL, N., LUBKIN, M., KIM, Y. B., SHARPE, A. H., STRICKER-KRONGRAD, A., SHULMAN, G. I., NEEL, B. G. & KAHN, B. B. 2000. Increased energy expenditure, decreased adiposity, and tissue-specific insulin sensitivity in protein-tyrosine phosphatase 1B-deficient mice. *Mol Cell Biol*, 20, 5479-89.
- KOKKINOPOULOS, I., JORDAN, W. J. & RITTER, M. A. 2005. Toll-like receptor mRNA expression patterns in human dendritic cells and monocytes. *Mol Immunol*, 42, 957-68.
- KOREN, S. & FANTUS, I. G. 2007. Inhibition of the protein tyrosine phosphatase PTP1B: potential therapy for obesity, insulin resistance and type-2 diabetes mellitus. *Best Pract Res Clin Endocrinol Metab*, 21, 621-40.
- KOSHKIN, A. A. & WENGEL, J. 1998. Synthesis of Novel 2',3'-Linked Bicyclic Thymine Ribonucleosides. *J Org Chem*, 63, 2778-2781.
- KOSUGE, M., TAKEUCHI, T., NAKASE, I., JONES, A. T. & FUTAKI, S. 2008. Cellular internalization and distribution of arginine-rich peptides as a function of extracellular peptide concentration, serum, and plasma membrane associated proteoglycans. *Bioconjug Chem*, 19, 656-64.
- KOUSPAROU, C. A., YIACOUMI, E., DEONARAIN, M. P. & EPEMETOS, A. A. 2012. Generation of a selectively cytotoxic fusion protein against p53 mutated cancers. *BMC Cancer*, 12, 338.
- KROVAT, B. C. & JANTSCH, M. F. 1996. Comparative mutational analysis of the double-stranded RNA binding domains of *Xenopus laevis* RNA-binding protein A. *J Biol Chem*, 271, 28112-9.
- KUMAR, P., WU, H., MCBRIDE, J. L., JUNG, K. E., KIM, M. H., DAVIDSON, B. L., LEE, S. K., SHANKAR, P. & MANJUNATH, N. 2007. Transvascular delivery of small interfering RNA to the central nervous system. *Nature*, 448, 39-43.
- KWAK, P. B. & TOMARI, Y. 2012. The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nat Struct Mol Biol*, 19, 145-51.
- LAEMMLI, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227, 680-5.
- LAUFER, S. D. & RESTLE, T. 2008. Peptide-mediated cellular delivery of oligonucleotide-based therapeutics in vitro: quantitative evaluation of overall efficacy employing easy to handle reporter systems. *Curr Pharm Des*, 14, 3637-55.
- LEDER, P. & NIRENBERG, M. W. 1964. Rna Codewords and Protein Synthesis, 3. On the Nucleotide Sequence of a Cysteine and a Leucine Rna Codeword. *Proc Natl Acad Sci U S A*, 52, 1521-9.
- LEE, M. H. & SCHEDL, T. 2006. RNA-binding proteins. *WormBook*, 1-13.

- LEMAIRE, P. A., ANDERSON, E., LARY, J. & COLE, J. L. 2008. Mechanism of PKR Activation by dsRNA. *J Mol Biol*, 381, 351-60.
- LEVIN, D. & LONDON, I. M. 1978. Regulation of protein synthesis: activation by double-stranded RNA of a protein kinase that phosphorylates eukaryotic initiation factor 2. *Proc Natl Acad Sci U S A*, 75, 1121-5.
- LEWIS, E. B. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature*, 276, 565-70.
- LEWIS, R. A., KAUFMAN, T. C., DENELL, R. E. & TALLERICO, P. 1980. Genetic Analysis of the Antennapedia Gene Complex (Ant-C) and Adjacent Chromosomal Regions of *DROSOPHILA MELANOGASTER*. I. Polytene Chromosome Segments 84b-D. *Genetics*, 95, 367-81.
- LIPINSKI, C. A., LOMBARDO, F., DOMINY, B. W. & FEENEY, P. J. 2001. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev*, 46, 3-26.
- LIU, J., CARMELL, M. A., RIVAS, F. V., MARSDEN, C. G., THOMSON, J. M., SONG, J. J., HAMMOND, S. M., JOSHUA-TOR, L. & HANNON, G. J. 2004. Argonaute2 is the catalytic engine of mammalian RNAi. *Science*, 305, 1437-41.
- LU, X. Q., MALUMBRES, R., SHIELDS, B., JIANG, X. Y., SAROSIEK, K. A., NATKUNAM, Y., TIGANIS, T. & LOSSOS, I. S. 2008. PTP1B is a negative regulator of interleukin 4-induced STAT6 signaling. *Blood*, 112, 4098-4108.
- LUKANOWSKA, M., HOWL, J. & JONES, S. 2013. Bioportides: bioactive cell-penetrating peptides that modulate cellular dynamics. *Biotechnol J*, 8, 918-30.
- LUNDBERG, M., WIKSTROM, S. & JOHANSSON, M. 2003. Cell surface adherence and endocytosis of protein transduction domains. *Mol Ther*, 8, 143-50.
- LUNDBERG, P., EL-ANDALOUSSI, S., SUTLU, T., JOHANSSON, H. & LANGEL, U. 2007. Delivery of short interfering RNA using endosomolytic cell-penetrating peptides. *FASEB J*, 21, 2664-71.
- LUNDIN, P., JOHANSSON, H., GUTERSTAM, P., HOLM, T., HANSEN, M., LANGEL, U. & S, E. L. A. 2008. Distinct uptake routes of cell-penetrating peptide conjugates. *Bioconjug Chem*, 19, 2535-42.
- MAACHUPALLI-REDDY, J., KELLEY, B. D. & DE BERNARDEZ CLARK, E. 1997. Effect of inclusion body contaminants on the oxidative renaturation of hen egg white lysozyme. *Biotechnol Prog*, 13, 144-50.
- MAGZOUB, M., KILK, K., ERIKSSON, L. E., LANGEL, U. & GRASLUND, A. 2001. Interaction and structure induction of cell-penetrating peptides in the presence of phospholipid vesicles. *Biochim Biophys Acta*, 1512, 77-89.
- MALHOTRA, P., DASARADHI, P. V., KUMAR, A., MOHMMED, A., AGRAWAL, N., BHATNAGAR, R. K. & CHAUHAN, V. S. 2002. Double-stranded RNA-mediated

- gene silencing of cysteine proteases (falcipain-1 and -2) of *Plasmodium falciparum*. *Mol Microbiol*, 45, 1245-54.
- MANCHE, L., GREEN, S. R., SCHMEDT, C. & MATHEWS, M. B. 1992. Interactions between double-stranded RNA regulators and the protein kinase DAI. *Mol Cell Biol*, 12, 5238-48.
- MANIATAKI, E. & MOURELATOS, Z. 2005. A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes Dev*, 19, 2979-90.
- MANN, D. A. & FRANKEL, A. D. 1991. Endocytosis and targeting of exogenous HIV-1 Tat protein. *EMBO J*, 10, 1733-9.
- MARCINIAK, R. A., CALNAN, B. J., FRANKEL, A. D. & SHARP, P. A. 1990. HIV-1 Tat protein trans-activates transcription in vitro. *Cell*, 63, 791-802.
- MARTY, C., MEYLAN, C., SCHOTT, H., BALLMER-HOFER, K. & SCHWENDENER, R. A. 2004. Enhanced heparan sulfate proteoglycan-mediated uptake of cell-penetrating peptide-modified liposomes. *Cell Mol Life Sci*, 61, 1785-94.
- MASLIAH, G., BARRAUD, P. & ALLAIN, F. H. 2013. RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. *Cell Mol Life Sci*, 70, 1875-95.
- MASQUIDA, B. & WESTHOF, E. 2000. On the wobble GoU and related pairs. *RNA*, 6, 9-15.
- MAXWELL, K. L., BONA, D., LIU, C. S., ARROWSMITH, C. H. & EDWARDS, A. M. 2003. Refolding out of guanidine hydrochloride is an effective approach for high-throughput structural studies of small proteins. *Protein Science*, 12, 2073-2080.
- MAZZOLA, N. 2012. Review of current and emerging therapies in type 2 diabetes mellitus. *Am J Manag Care*, 18, S17-26.
- MEADE, B. R. & DOWDY, S. F. 2007. Exogenous siRNA delivery using peptide transduction domains/cell penetrating peptides. *Adv Drug Deliv Rev*, 59, 134-40.
- MEADE, B. R. & DOWDY, S. F. 2009. The road to therapeutic RNA interference (RNAi): Tackling the 800 pound siRNA delivery gorilla. *Discov Med*, 8, 253-6.
- MEISTER, G. 2013. Argonaute proteins: functional insights and emerging roles. *Nat Rev Genet*, 14, 447-59.
- MEISTER, G., LANDTHALER, M., PATKANIOWSKA, A., DORSETT, Y., TENG, G. & TUSCHL, T. 2004. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol Cell*, 15, 185-97.
- MEISTER, G., LANDTHALER, M., PETERS, L., CHEN, P. Y., URLAUB, H., LUHRMANN, R. & TUSCHL, T. 2005. Identification of novel argonaute-associated proteins. *Curr Biol*, 15, 2149-55.

- MERRICK, W. C. 1992. Mechanism and regulation of eukaryotic protein synthesis. *Microbiol Rev*, 56, 291-315.
- MEURS, E., CHONG, K., GALABRU, J., THOMAS, N. S., KERR, I. M., WILLIAMS, B. R. & HOVANESSIAN, A. G. 1990. Molecular cloning and characterization of the human double-stranded RNA-activated protein kinase induced by interferon. *Cell*, 62, 379-90.
- MIDOUX, P. & MONSIGNY, M. 1999. Efficient gene transfer by histidylated polylysine/pDNA complexes. *Bioconjug Chem*, 10, 406-11.
- MIDOUX, P., PICHON, C., YAOUANC, J. J. & JAFFRES, P. A. 2009. Chemical vectors for gene delivery: a current review on polymers, peptides and lipids containing histidine or imidazole as nucleic acids carriers. *Br J Pharmacol*, 157, 166-78.
- MILLETTI, F. 2012. Cell-penetrating peptides: classes, origin, and current landscape. *Drug Discov Today*, 17, 850-60.
- MO, R. H., ZARO, J. L. & SHEN, W. C. 2012. Comparison of cationic and amphipathic cell penetrating peptides for siRNA delivery and efficacy. *Mol Pharm*, 9, 299-309.
- MORRIS, M. C., DEPOLLIER, J., MERY, J., HEITZ, F. & DIVITA, G. 2001. A peptide carrier for the delivery of biologically active proteins into mammalian cells. *Nat Biotechnol*, 19, 1173-6.
- MOSCHOS, S. A., JONES, S. W., PERRY, M. M., WILLIAMS, A. E., ERJEFALT, J. S., TURNER, J. J., BARNES, P. J., SPROAT, B. S., GAIT, M. J. & LINDSAY, M. A. 2007. Lung delivery studies using siRNA conjugated to TAT(48-60) and penetratin reveal peptide induced reduction in gene expression and induction of innate immunity. *Bioconjug Chem*, 18, 1450-9.
- MOSMANN, T. 1983. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J Immunol Methods*, 65, 55-63.
- MUELLER, J., KRETZSCHMAR, I., VOLKMER, R. & BOISGUERIN, P. 2008. Comparison of cellular uptake using 22 CPPs in 4 different cell lines. *Bioconjug Chem*, 19, 2363-74.
- MUJEEB, A., BISHOP, K., PETERLIN, B. M., TURCK, C., PARSLow, T. G. & JAMES, T. L. 1994. NMR structure of a biologically active peptide containing the RNA-binding domain of human immunodeficiency virus type 1 Tat. *Proc Natl Acad Sci U S A*, 91, 8248-52.
- MULLER, M., AFFOLTER, M., LEUPIN, W., OTTING, G., WUTHRICH, K. & GEHRING, W. J. 1988. Isolation and sequence-specific DNA binding of the Antennapedia homeodomain. *EMBO J*, 7, 4299-304.
- MURATOVSKA, A. & ECCLES, M. R. 2004. Conjugate for efficient delivery of short interfering RNA (siRNA) into mammalian cells. *FEBS Lett*, 558, 63-8.

- NALLAGATLA, S. R. & BEVILACQUA, P. C. 2008. Nucleoside modifications modulate activation of the protein kinase PKR in an RNA structure-specific manner. *RNA*, 14, 1201-13.
- NALLAGATLA, S. R., TORONEY, R. & BEVILACQUA, P. C. 2008. A brilliant disguise for self RNA: 5'-end and internal modifications of primary transcripts suppress elements of innate immunity. *RNA Biol*, 5, 140-4.
- NANDURI, S., CARPICK, B. W., YANG, Y., WILLIAMS, B. R. & QIN, J. 1998. Structure of the double-stranded RNA-binding domain of the protein kinase PKR reveals the molecular basis of its dsRNA-mediated activation. *EMBO J*, 17, 5458-65.
- NANDURI, S., RAHMAN, F., WILLIAMS, B. R. & QIN, J. 2000. A dynamically tuned double-stranded RNA binding mechanism for the activation of antiviral kinase PKR. *EMBO J*, 19, 5567-74.
- NAPOLI, C., LEMIEUX, C. & JORGENSEN, R. 1990. Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell*, 2, 279-289.
- NGUYEN, Q. N., CHAVLI, R. V., MARQUES, J. T., CONRAD, P. G., JR., WANG, D., HE, W., BELISLE, B. E., ZHANG, A., PASTOR, L. M., WITNEY, F. R., MORRIS, M., HEITZ, F., DIVITA, G., WILLIAMS, B. R. & MCMASTER, G. K. 2006. Light controllable siRNAs regulate gene suppression and phenotypes in cells. *Biochim Biophys Acta*, 1758, 394-403.
- NICHOLSON, A. W. 1996. Structure, reactivity, and biology of double-stranded RNA. *Prog Nucleic Acid Res Mol Biol*, 52, 1-65.
- NOLAND, C. L., MA, E. & DOUDNA, J. A. 2011. siRNA repositioning for guide strand selection by human Dicer complexes. *Mol Cell*, 43, 110-21.
- O'CONNELL, M. A., KRAUSE, S., HIGUCHI, M., HSUAN, J. J., TOTTY, N. F., JENNY, A. & KELLER, W. 1995. Cloning of cDNAs encoding mammalian double-stranded RNA-specific adenosine deaminase. *Mol Cell Biol*, 15, 1389-97.
- OEHLKE, J., BEYERMANN, M., WIESNER, B., MELZIG, M., BERGER, H., KRAUSE, E. & BIENERT, M. 1997. Evidence for extensive and non-specific translocation of oligopeptides across plasma membranes of mammalian cells. *Biochim Biophys Acta*, 1330, 50-60.
- OHLENDORF, D. H., ANDERSON, W. F., FISHER, R. G., TAKEDA, Y. & MATTHEWS, B. W. 1982. The molecular basis of DNA-protein recognition inferred from the structure of cro repressor. *Nature*, 298, 718-23.
- OTTING, G., QIAN, Y. Q., MULLER, M., AFFOLTER, M., GEHRING, W. & WUTHRICH, K. 1988. Secondary structure determination for the Antennapedia homeodomain by nuclear magnetic resonance and evidence for a helix-turn-helix motif. *EMBO J*, 7, 4305-9.

- PACK, D. W., HOFFMAN, A. S., PUN, S. & STAYTON, P. S. 2005. Design and development of polymers for gene delivery. *Nat Rev Drug Discov*, 4, 581-93.
- PAIN, V. M. 1996. Initiation of protein synthesis in eukaryotic cells. *Eur J Biochem*, 236, 747-71.
- PARK, H., DAVIES, M. V., LANGLAND, J. O., CHANG, H. W., NAM, Y. S., TARTAGLIA, J., PAOLETTI, E., JACOBS, B. L., KAUFMAN, R. J. & VENKATESAN, S. 1994. TAR RNA-binding protein is an inhibitor of the interferon-induced protein kinase PKR. *Proc Natl Acad Sci U S A*, 91, 4713-7.
- PARKER, J. S., ROE, S. M. & BARFORD, D. 2005. Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature*, 434, 663-6.
- PATEL, C. V., HANDY, I., GOLDSMITH, T. & PATEL, R. C. 2000. PACT, a stress-modulated cellular activator of interferon-induced double-stranded RNA-activated protein kinase, PKR. *J Biol Chem*, 275, 37993-8.
- PATEL, R. C. & SEN, G. C. 1992. Identification of the double-stranded RNA-binding domain of the human interferon-inducible protein kinase. *J Biol Chem*, 267, 7671-6.
- PATTERSON, J. B., THOMIS, D. C., HANS, S. L. & SAMUEL, C. E. 1995. Mechanism of interferon action: double-stranded RNA-specific adenosine deaminase from human cells is inducible by alpha and gamma interferons. *Virology*, 210, 508-11.
- PECOT, C. V., CALIN, G. A., COLEMAN, R. L., LOPEZ-BERESTEIN, G. & SOOD, A. K. 2011. RNA interference in the clinic: challenges and future directions. *Nat Rev Cancer*, 11, 59-67.
- PEREZ, F., JOLIOT, A., BLOCH-GALLEGO, E., ZAHRAOUI, A., TRILLER, A. & PROCHIANTZ, A. 1992. Antennapedia homeobox as a signal for the cellular internalization and nuclear addressing of a small exogenous peptide. *J Cell Sci*, 102 ( Pt 4), 717-22.
- PETERS, G. A., HARTMANN, R., QIN, J. & SEN, G. C. 2001. Modular structure of PACT: distinct domains for binding and activating PKR. *Mol Cell Biol*, 21, 1908-20.
- PIRES-DASILVA, A., NAYERNIA, K., ENGEL, W., TORRES, M., STOYKOVA, A., CHOWDHURY, K. & GRUSS, P. 2001. Mice deficient for spermatid perinuclear RNA-binding protein show neurologic, spermatogenic, and sperm morphological abnormalities. *Developmental Biology*, 233, 319-28.
- POOGA, M., SOOMETTS, U., HALLBRINK, M., VALKNA, A., SAAR, K., REZAEI, K., KAHL, U., HAO, J. X., XU, X. J., WIESENFELD-HALLIN, Z., HOKFELT, T., BARTFAI, T. & LANGEL, U. 1998. Cell penetrating PNA constructs regulate galanin receptor levels and modify pain transmission in vivo. *Nat Biotechnol*, 16, 857-61.

- POON, G. M. & GARIEPY, J. 2007. Cell-surface proteoglycans as molecular portals for cationic peptide and polymer entry into cells. *Biochem Soc Trans*, 35, 788-93.
- RAMOS, A., GRUNERT, S., ADAMS, J., MICKLEM, D. R., PROCTOR, M. R., FREUND, S., BYCROFT, M., ST JOHNSTON, D. & VARANI, G. 2000. RNA recognition by a Staufen double-stranded RNA-binding domain. *EMBO J*, 19, 997-1009.
- RAO, D. D., VORHIES, J. S., SENZER, N. & NEMUNAITIS, J. 2009. siRNA vs. shRNA: similarities and differences. *Adv Drug Deliv Rev*, 61, 746-59.
- RAUB, T. J., KOROLY, M. J. & ROBERTS, R. M. 1990. Endocytosis of wheat germ agglutinin binding sites from the cell surface into a tubular endosomal network. *J Cell Physiol*, 143, 1-12.
- REGBERG, J., SRIMANEE, A. & LANGEL, U. 2012. Applications of cell-penetrating peptides for tumor targeting and future cancer therapies. *Pharmaceuticals (Basel)*, 5, 991-1007.
- REYNOLDS, A., ANDERSON, E. M., VERMEULEN, A., FEDOROV, Y., ROBINSON, K., LEAKE, D., KARPILOW, J., MARSHALL, W. S. & KHVOROVA, A. 2006. Induction of the interferon response by siRNA is cell type- and duplex length-dependent. *RNA*, 12, 988-93.
- RICE, A. P., DUNCAN, R., HERSHEY, J. W. & KERR, I. M. 1985. Double-stranded RNA-dependent protein kinase and 2-5A system are both activated in interferon-treated, encephalomyocarditis virus-infected HeLa cells. *Journal of Virology*, 54, 894-8.
- RICHARD, J. P., MELIKOV, K., VIVES, E., RAMOS, C., VERBEURE, B., GAIT, M. J., CHERNOMORDIK, L. V. & LEBLEU, B. 2003. Cell-penetrating peptides. A reevaluation of the mechanism of cellular uptake. *J Biol Chem*, 278, 585-90.
- RISS TL, M. R., NILES AL, HELENE HA , BENINK A, WORZELLA TJ, MINOR L. 2013. Cell Viability Assays. Assay Guidance Manual [Internet].  
*Bethesda (MD): Eli Lilly & Company and the National Center for Advancing Translational Sciences.*
- ROBBINS, M., JUDGE, A. & MACLACHLAN, I. 2009. siRNA and innate immunity. *Oligonucleotides*, 19, 89-102.
- RONDINONE, C. M., TREVILLYAN, J. M., CLAMPIT, J., GUM, R. J., BERG, C., KROEGER, P., FROST, L., ZINKER, B. A., REILLY, R., ULRICH, R., BUTLER, M., MONIA, B. P., JIROUSEK, M. R. & WARING, J. F. 2002. Protein tyrosine phosphatase 1B reduction regulates adiposity and expression of genes involved in lipogenesis. *Diabetes*, 51, 2405-11.
- ROSS, J. P. & KASSIR, Z. 2014. The varied roles of nuclear argonaute-small RNA complexes and avenues for therapy. *Mol Ther Nucleic Acids*, 3, e203.

- ROUSSELLE, C., CLAIR, P., LEFAUCCONNIER, J. M., KACZOREK, M., SCHERRMANN, J. M. & TEMSAMANI, J. 2000. New advances in the transport of doxorubicin through the blood-brain barrier by a peptide vector-mediated strategy. *Mol Pharmacol*, 57, 679-86.
- ROUSSELLE, C., SMIRNOVA, M., CLAIR, P., LEFAUCCONNIER, J. M., CHAVANIEU, A., CALAS, B., SCHERRMANN, J. M. & TEMSAMANI, J. 2001. Enhanced delivery of doxorubicin into the brain via a peptide-vector-mediated strategy: saturation kinetics and specificity. *J Pharmacol Exp Ther*, 296, 124-31.
- RUDA, V. M., CHANDWANI, R., SEHGAL, A., BOGORAD, R. L., AKINC, A., CHARISSE, K., TARAKHOVSKY, A., NOVOBRANTSEVA, T. I. & KOTELIANSKY, V. 2014. The roles of individual mammalian argonautes in RNA interference in vivo. *PLoS One*, 9, e101749.
- RYDSTROM, A., DESHAYES, S., KONATE, K., CROMBEZ, L., PADARI, K., BOUKHADDAOUI, H., ALDRIAN, G., POOGA, M. & DIVITA, G. 2011. Direct translocation as major cellular uptake for CADY self-assembling peptide-based nanoparticles. *PLoS One*, 6, e25924.
- RYTER, J. M. & SCHULTZ, S. C. 1998. Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J*, 17, 7505-13.
- SAKAI, N. & MATILE, S. 2003. Anion-mediated transfer of polyarginine across liquid and bilayer membranes. *J Am Chem Soc*, 125, 14348-56.
- SAMUEL, C. E., DUNCAN, R., KNUTSON, G. S. & HERSHEY, J. W. 1984. Mechanism of interferon action. Increased phosphorylation of protein synthesis initiation factor eIF-2 alpha in interferon-treated, reovirus-infected mouse L929 fibroblasts in vitro and in vivo. *J Biol Chem*, 259, 13451-7.
- SANNES-LOWERY, K. A., HU, P., MACK, D. P., MEI, H. Y. & LOO, J. A. 1997. HIV-1 Tat peptide binding to TAR RNA by electrospray ionization mass spectrometry. *Analytical Chemistry*, 69, 5130-5.
- SAUNDERS, L. R. & BARBER, G. N. 2003. The dsRNA binding protein family: critical roles, diverse cellular functions. *FASEB J*, 17, 961-83.
- SAUNDERS, L. R., PERKINS, D. J., BALACHANDRAN, S., MICHAELS, R., FORD, R., MAYEDA, A. & BARBER, G. N. 2001. Characterization of two evolutionarily conserved, alternatively spliced nuclear phosphoproteins, NFAR-1 and -2, that function in mRNA processing and interact with the double-stranded RNA-dependent protein kinase, PKR. *J Biol Chem*, 276, 32300-12.
- SCHEIN, C. H. 1989. Production of Soluble Recombinant Proteins in Bacteria. *Bio-Technology*, 7, 1141-1147.
- SCHELLER, A., WIESNER, B., MELZIG, M., BIENERT, M. & OEHLKE, J. 2000. Evidence for an amphipathicity independent cellular uptake of amphipathic cell-penetrating peptides. *Eur J Biochem*, 267, 6043-50.

- SCHMEDT, C., GREEN, S. R., MANCHE, L., TAYLOR, D. R., MA, Y. & MATHEWS, M. B. 1995. Functional characterization of the RNA-binding domain and motif of the double-stranded RNA-dependent protein kinase DAI (PKR). *J Mol Biol*, 249, 29-44.
- SCHWARZ, D. S., HUTVAGNER, G., DU, T., XU, Z., ARONIN, N. & ZAMORE, P. D. 2003. Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, 115, 199-208.
- SCHWARZE, S. R., HO, A., VOCERO-AKBANI, A. & DOWDY, S. F. 1999. In vivo protein transduction: delivery of a biologically active protein into the mouse. *Science*, 285, 1569-72.
- SIMEONI, F., MORRIS, M. C., HEITZ, F. & DIVITA, G. 2003. Insight into the mechanism of the peptide-based gene delivery system MPG: implications for delivery of siRNA into mammalian cells. *Nucleic Acids Res*, 31, 2717-24.
- SMITH, D. B. & JOHNSON, K. S. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene*, 67, 31-40.
- SNYDER, E. L. & DOWDY, S. F. 2005. Recent advances in the use of protein transduction domains for the delivery of peptides, proteins and nucleic acids in vivo. *Expert Opin Drug Deliv*, 2, 43-51.
- SODROSKI, J., ROSEN, C., WONG-STAAAL, F., SALAHUDDIN, S. Z., POPOVIC, M., ARYA, S., GALLO, R. C. & HASELTINE, W. A. 1985. Trans-acting transcriptional regulation of human T-cell leukemia virus type III long terminal repeat. *Science*, 227, 171-3.
- SORENSEN, H. P. & MORTENSEN, K. K. 2005. Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microb Cell Fact*, 4, 1.
- SPAGNOU, S., MILLER, A. D. & KELLER, M. 2004. Lipidic carriers of siRNA: differences in the formulation, cellular uptake, and delivery with plasmid DNA. *Biochemistry*, 43, 13348-56.
- ST JOHNSTON, D., BROWN, N. H., GALL, J. G. & JANTSCH, M. 1992. A conserved double-stranded RNA-binding domain. *Proc Natl Acad Sci U S A*, 89, 10979-83.
- SUGITA, T., YOSHIKAWA, T., MUKAI, Y., YAMANADA, N., IMAI, S., NAGANO, K., YOSHIDA, Y., SHIBATA, H., YOSHIOKA, Y., NAKAGAWA, S., KAMADA, H., TSUNODA, S. I. & TSUTSUMI, Y. 2008. Comparative study on transduction and toxicity of protein transduction domains. *Br J Pharmacol*, 153, 1143-52.
- SUNDQVIST, G., BENKESTOCK, K. & ROERAADE, J. 2005. Investigation of multiple binding sites on ribonuclease A using nano-electrospray ionization mass spectrometry. *Rapid Commun Mass Spectrom*, 19, 1011-6.

- SURABHI, R. M. & GAYNOR, R. B. 2002. RNA interference directed against viral and cellular targets inhibits human immunodeficiency Virus Type 1 replication. *Journal of Virology*, 76, 12963-73.
- TAETZ, S., BOCHOT, A., SURACE, C., ARPICCO, S., RENOIR, J. M., SCHAEFER, U. F., MARSAUD, V., Kerdine-Roemer, S., LEHR, C. M. & FATTAL, E. 2009. Hyaluronic acid-modified DOTAP/DOPE liposomes for the targeted delivery of anti-telomerase siRNA to CD44-expressing lung cancer cells. *Oligonucleotides*, 19, 103-16.
- TAM, O. H., ARAVIN, A. A., STEIN, P., GIRARD, A., MURCHISON, E. P., CHELOUFI, S., HODGES, E., ANGER, M., SACHIDANANDAM, R., SCHULTZ, R. M. & HANNON, G. J. 2008. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature*, 453, 534-8.
- TANEJA, S. & AHMAD, F. 1994. Increased thermal stability of proteins in the presence of amino acids. *Biochem J*, 303 ( Pt 1), 147-53.
- TANIGUCHI, C. M., EMANUELLI, B. & KAHN, C. R. 2006. Critical nodes in signalling pathways: insights into insulin action. *Nat Rev Mol Cell Biol*, 7, 85-96.
- TAYLOR, D. R., TIAN, B., ROMANO, P. R., HINNEBUSCH, A. G., LAI, M. M. & MATHEWS, M. B. 2001. Hepatitis C virus envelope protein E2 does not inhibit PKR by simple competition with autophosphorylation sites in the RNA-binding domain. *Journal of Virology*, 75, 1265-73.
- THOMAS, P. & SMART, T. G. 2005. HEK293 cell line: a vehicle for the expression of recombinant proteins. *J Pharmacol Toxicol Methods*, 51, 187-200.
- TIAN, B. & MATHEWS, M. B. 2001. Functional characterization of and cooperation between the double-stranded RNA-binding motifs of the protein kinase PKR. *J Biol Chem*, 276, 9936-44.
- TIGANIS, T. 2013. PTP1B and TCPTP--nonredundant phosphatases in insulin signaling and glucose homeostasis. *FEBS J*, 280, 445-58.
- TOLIA, N. H. & JOSHUA-TOR, L. 2006. Strategies for protein coexpression in *Escherichia coli*. *Nat Methods*, 3, 55-64.
- TRINH, R., GURBAXANI, B., MORRISON, S. L. & SEYFZADEH, M. 2004. Optimization of codon pair use within the (GGGS)<sub>3</sub> linker sequence results in enhanced protein expression. *Mol Immunol*, 40, 717-22.
- TSOU, R. C. & BENCE, K. K. 2012. The Genetics of PTPN1 and Obesity: Insights from Mouse Models of Tissue-Specific PTP1B Deficiency. *J Obes*, 2012, 926857.
- TSUMOTO, K., EJIMA, D., KUMAGAI, I. & ARAKAWA, T. 2003. Practical considerations in refolding proteins from inclusion bodies. *Protein Expr Purif*, 28, 1-8.

- TSUMOTO, K., SHINOKI, K., KONDO, H., UCHIKAWA, M., JUJI, T. & KUMAGAI, I. 1998. Highly efficient recovery of functional single-chain Fv fragments from inclusion bodies overexpressed in *Escherichia coli* by controlled introduction of oxidizing reagent--application to a human single-chain Fv fragment. *J Immunol Methods*, 219, 119-29.
- TUNNEMANN, G., MARTIN, R. M., HAUPT, S., PATSCH, C., EDENHOFER, F. & CARDOSO, M. C. 2006. Cargo-dependent mode of uptake and bioavailability of TAT-containing proteins and peptides in living cells. *FASEB J*, 20, 1775-84.
- TUNNEMANN, G., TER-AVETISYAN, G., MARTIN, R. M., STOCKL, M., HERRMANN, A. & CARDOSO, M. C. 2008. Live-cell analysis of cell penetration ability and toxicity of oligo-arginines. *J Pept Sci*, 14, 469-76.
- TYAGI, M., RUSNATI, M., PRESTA, M. & GIACCA, M. 2001. Internalization of HIV-1 tat requires cell surface heparan sulfate proteoglycans. *J Biol Chem*, 276, 3254-61.
- UCCI, J. W., KOBAYASHI, Y., CHOI, G., ALEXANDRESCU, A. T. & COLE, J. L. 2007. Mechanism of interaction of the double-stranded RNA (dsRNA) binding domain of protein kinase R with short dsRNA sequences. *Biochemistry*, 46, 55-65.
- UMBACH, J. L. & CULLEN, B. R. 2009. The role of RNAi and microRNAs in animal virus replication and antiviral immunity. *Genes Dev*, 23, 1151-64.
- VAISHNAW, A. K., GOLLOB, J., GAMBA-VITALO, C., HUTABARAT, R., SAH, D., MEYERS, R., DE FOUGEROLLES, T. & MARAGANORE, J. 2010. A status report on RNAi therapeutics. *Silence*, 1, 14.
- VAN ASBECK, A. H., BEYERLE, A., MCNEILL, H., BOVEE-GEURTS, P. H., LINDBERG, S., VERDURMEN, W. P., HALLBRINK, M., LANGEL, U., HEIDENREICH, O. & BROCK, R. 2013. Molecular parameters of siRNA--cell penetrating peptide nanocomplexes for efficient cellular delivery. *ACS Nano*, 7, 3797-807.
- VAN DE WATER, F. M., BOERMAN, O. C., WOUTERSE, A. C., PETERS, J. G., RUSSEL, F. G. & MASEREEUW, R. 2006. Intravenously administered short interfering RNA accumulates in the kidney and selectively suppresses gene function in renal proximal tubules. *Drug Metab Dispos*, 34, 1393-7.
- VARKOUHI, A. K., SCHOLTE, M., STORM, G. & HAISMA, H. J. 2011. Endosomal escape pathways for delivery of biologicals. *J Control Release*, 151, 220-8.
- VERDINE, G. L. & HILINSKI, G. J. 2012. Stapled peptides for intracellular drug targets. *Methods Enzymol*, 503, 3-33.
- VICKERS, T. A., LIMA, W. F., NICHOLS, J. G. & CROOKE, S. T. 2007. Reduced levels of Ago2 expression result in increased siRNA competition in mammalian cells. *Nucleic Acids Res*, 35, 6598-610.

- WADIA, J. S., STAN, R. V. & DOWDY, S. F. 2004. Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. *Nature Medicine*, 10, 310-5.
- WAKIMOTO, B. T. & KAUFMAN, T. C. 1981. Analysis of larval segmentation in lethal genotypes associated with the antennapedia gene complex in *Drosophila melanogaster*. *Developmental Biology*, 81, 51-64.
- WANG, P., HENNING, S. M. & HEBER, D. 2010. Limitations of MTT and MTS-based assays for measurement of antiproliferative activity of green tea polyphenols. *PLoS One*, 5, e10202.
- WANG, Q., KHILLAN, J., GADUE, P. & NISHIKURA, K. 2000. Requirement of the RNA editing deaminase ADAR1 gene for embryonic erythropoiesis. *Science*, 290, 1765-8.
- WANG, W. 2005. Protein aggregation and its inhibition in biopharmaceutics. *Int J Pharm*, 289, 1-30.
- WARING, J. F., CIURLIONIS, R., CLAMPIT, J. E., MORGAN, S., GUM, R. J., JOLLY, R. A., KROEGER, P., FROST, L., TREVILLYAN, J., ZINKER, B. A., JIROUSEK, M., ULRICH, R. G. & RONDINONE, C. M. 2003. PTP1B antisense-treated mice show regulation of genes involved in lipogenesis in liver and fat. *Mol Cell Endocrinol*, 203, 155-68.
- WATANABE, T., TOTOKI, Y., TOYODA, A., KANEDA, M., KURAMOCHI-MIYAGAWA, S., OBATA, Y., CHIBA, H., KOHARA, Y., KONO, T., NAKANO, T., SURANI, M. A., SAKAKI, Y. & SASAKI, H. 2008. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature*, 453, 539-43.
- WAUGH, D. S. 2005. Making the most of affinity tags. *Trends Biotechnol*, 23, 316-20.
- WEEKS, K. M. & CROTHERS, D. M. 1993. Major groove accessibility of RNA. *Science*, 261, 1574-7.
- WEISS, W. A., TAYLOR, S. S. & SHOKAT, K. M. 2007. Recognizing and exploiting differences between RNAi and small-molecule inhibitors. *Nat Chem Biol*, 3, 739-44.
- WEN, Y. & MENG, W. S. 2014. Recent In Vivo Evidences of Particle-Based Delivery of Small-Interfering RNA (siRNA) into Solid Tumors. *J Pharm Innov*, 9, 158-173.
- WHITE, M. F. 2003. Insulin signaling in health and disease. *Science*, 302, 1710-1.
- WHITEHEAD, K. A., LANGER, R. & ANDERSON, D. G. 2009. Knocking down barriers: advances in siRNA delivery. *Nat Rev Drug Discov*, 8, 129-38.
- WRIGHT, C. S., GAVILANES, F. & PETERSON, D. L. 1984a. Primary structure of wheat germ agglutinin isolectin 2. Peptide order deduced from X-ray structure. *Biochemistry*, 23, 280-7.

- WRIGHT, H. T., BROOKS, D. M. & WRIGHT, C. S. 1984b. Evolution of the multidomain protein wheat germ agglutinin. *J Mol Evol*, 21, 133-8.
- WU, S. & KAUFMAN, R. J. 1996. Double-stranded (ds) RNA binding and not dimerization correlates with the activation of the dsRNA-dependent protein kinase (PKR). *J Biol Chem*, 271, 1756-63.
- WU, Z., MILLER, E., AGBANDJE-MCKENNA, M. & SAMULSKI, R. J. 2006. Alpha2,3 and alpha2,6 N-linked sialic acids facilitate efficient binding and transduction by adeno-associated virus types 1 and 6. *Journal of Virology*, 80, 9093-103.
- WU, Z. W., CHIEN, C. T., LIU, C. Y., YAN, J. Y. & LIN, S. Y. 2012. Recent progress in copolymer-mediated siRNA delivery. *J Drug Target*, 20, 551-60.
- XIAO, P. J. & SAMULSKI, R. J. 2012. Cytoplasmic trafficking, endosomal escape, and perinuclear accumulation of adeno-associated virus type 2 particles are facilitated by microtubule network. *Journal of Virology*, 86, 10462-73.
- YI, R., DOEHLE, B. P., QIN, Y., MACARA, I. G. & CULLEN, B. R. 2005. Overexpression of exportin 5 enhances RNA interference mediated by short hairpin RNAs and microRNAs. *RNA*, 11, 220-6.
- YIP, S. C., SAHA, S. & CHERNOFF, J. 2010. PTP1B: a double agent in metabolism and oncogenesis. *Trends Biochem Sci*, 35, 442-9.
- ZEINEDDINE, D., PAPADIMOU, E., CHEBLI, K., GINESTE, M., LIU, J., GREY, C., THURIG, S., BEHFAR, A., WALLACE, V. A., SKERJANC, I. S. & PUCEAT, M. 2006. Oct-3/4 dose dependently regulates specification of embryonic stem cells toward a cardiac lineage and early heart development. *Dev Cell*, 11, 535-46.
- ZHANG, F., ROMANO, P. R., NAGAMURA-INOUE, T., TIAN, B., DEVER, T. E., MATHEWS, M. B., OZATO, K. & HINNEBUSCH, A. G. 2001. Binding of double-stranded RNA to protein kinase PKR is required for dimerization and promotes critical autophosphorylation events in the activation loop. *J Biol Chem*, 276, 24946-58.
- ZHANG, H., KOLB, F. A., JASKIEWICZ, L., WESTHOF, E. & FILIPOWICZ, W. 2004. Single processing center models for human Dicer and bacterial RNase III. *Cell*, 118, 57-68.
- ZHOU, J., SHUM, K. T., BURNETT, J. C. & ROSSI, J. J. 2013. Nanoparticle-Based Delivery of RNAi Therapeutics: Progress and Challenges. *Pharmaceuticals (Basel)*, 6, 85-107.
- ZIMMER, S., STEINMETZ, M., ASDONK, T., MOTZ, I., COCH, C., HARTMANN, E., BARCET, W., WASSMANN, S., HARTMANN, G. & NICKENIG, G. 2011. Activation of endothelial toll-like receptor 3 impairs endothelial function. *Circ Res*, 108, 1358-66.
- ZINKER, B. A., RONDINONE, C. M., TREVILLYAN, J. M., GUM, R. J., CLAMPIT, J. E., WARING, J. F., XIE, N., WILCOX, D., JACOBSON, P., FROST, L.,

KROEGER, P. E., REILLY, R. M., KOTERSKI, S., OPGENORTH, T. J.,  
ULRICH, R. G., CROSBY, S., BUTLER, M., MURRAY, S. F., MCKAY, R. A.,  
BHANOT, S., MONIA, B. P. & JIROUSEK, M. R. 2002. PTP1B antisense  
oligonucleotide lowers PTP1B protein, normalizes blood glucose, and improves  
insulin sensitivity in diabetic mice. *Proc Natl Acad Sci U S A*, 99, 11357-62.