The interaction between the Marek's Disease Virus (MDV) neurovirulence factor pp14 and the host transcription factor, CREB3

Keshalini Sabaratnam

A thesis submitted for the degree of

Doctor of Philosophy

St Edmund Hall
Division of Structural Biology
University of Oxford

Trinity term 2017

The interaction between the Marek's Disease Virus (MDV) neurovirulence factor pp14 and the host transcription factor, CREB3

Keshalini Sabaratnam, St Edmund Hall, University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Trinity term 2017

Abstract

Marek's Disease Virus (MDV) induces a wide range of neurological syndromes in susceptible hosts; however, the mechanisms behind the MDV-induced neuropathology are still poorly understood. The immediate-early 14kDa phosphoprotein, pp14, is associated with the neurovirulence phenotype of the virus. Yeast-two-hybrid screening identified the ER-bound transcription regulator, human CREB3 (cAMP Response Element-Binding protein), as an interacting partner of pp14, and fluorescence colocalisation between pp14 and chicken CREB3 (chCREB3) in MDV infected cells suggested an interaction between these proteins. The primary focus of this DPhil project was to further investigate this putative interaction using *in vitro* studies, with a view to determining if the interaction is linked to the neurovirulence of MDV. This investigation, which employed a combination of biochemical, cellular, and functional assays, found no conclusive evidence in support of the predicted interaction.

In addition, this project aimed to gain structural and functional insights into the MDV neurovirulence factor pp14 and the host transcription factor, chCREB3. Biophysical characterisation of recombinant pp14B identifies pp14 as a molten globule. The results reveal the protein, while possessing substantial secondary structure, is largely disordered lacking a stable tertiary structure. Multiple lines of evidence from this study also indicate pp14 is a putative zinc-binding protein. Moreover, phosphorylation analysis of recombinant pp14B, extracted from DF1 cells, by mass spectrometry provides conclusive evidence for the presence of two phosphorylation sites in the shared C-terminal region of pp14—serines 72 and 76 of pp14B. Structural flexibility, through a lack of a definite ordered tertiary structure, and functional features that can induce structural modifications indicate pp14 might interact with a number of binding partners and therefore could play multiple roles during MDV infection—a strong possibility due to the expression of the protein in all the different stages of virus infection.

Furthermore, this thesis presents the crystal structure of the homodimeric chCREB3 bZIP. The chCREB3 bZIP possesses a structured DNA binding region even in the absence of DNA, a feature that could potentially enhance both the DNA-binding specificity and affinity of chCREB3. Significantly, chCREB3 has a covalent intermolecular disulphide bond in the hydrophobic core of the bZIP, which may play a role in promoting stability. Moreover, sequence alignment of bZIP sequences from chicken, human and mouse reveals only members of the CREB3 subfamily possess this cysteine residue, indicating it could act as a redox-sensor. These results indicate members of the CREB3 subfamily, by possessing a putative redox-sensitive cysteine with the capacity to form an intermolecular disulphide bond, may be activated in response to oxidative stress.

This thesis is dedicated to my parents,

Thevaki and Ponnuthamby Sabaratnam

Table of contents

Abstract ii

Acknowledgements vii

Declaration of work viii

List of figures ix

List of tables xii

List of abbreviations xiii

1 Introduction 1

1.1 Marek's Disease 1

1.2 Marek's Disease Virus 3

1.2.1 Classification, structure and genome structure 3

1.2.2 Herpesvirus replication and latency: Gene expression regulation 5

1.2.3 Pathogenesis of MDV 10

1.3 Neurovirulence of MDV 13

1.3.1 Vaccination and emergence of neurovirulent strains of MDV 13

1.3.2 Neurologic manifestations of MD 15

1.3.3 Mechanisms underlying MDV neurovirulence 16

1.4 MDV neurovirulence factor, pp14 18

1.4.1 14kDa phosphoprotein, pp14 18

1.4.2 Interaction of pp14 with CREB3 21

1.4.3 Interaction of pp14 with chCREB3 in MDV infected cells 24

1.5 cAMP response element binding protein, CREB3 26

1.5.1 CREB3 subfamily of bZIP transcription factors 26

1.5.2 Biological functions of CREB3 30

1.6 bZIP transcription factors 34

1.7 Aim and objectives 37

2 Materials and Methods 40

2.1 Amino acid sequences 40

2.2 Plasmids 41

2.3 Vector maps in Appendix A 41

2.4 Antibodies 41

2.5 Antibiotics 42

2.6 Viruses 43

2.7 MDV gene nomenclature 44

2.8 Cell lines and primary cells 45

2.8.1 DF1 cells 45

2.8.2 HEK293T (human embryonic kidney cell line) 45

2.8.3 Insect cells 45

2.8.4 Chicken embryo fibroblast (CEF) cells 46

2.9 Molecular biology and microbiology techniques 46

2.9.1 Polymerase Chain Reaction 46

2.9.2 Agarose gel electrophoresis 48

2.9.3 Gel purification of DNA 48

2.9.4 In-Fusion ligation independent cloning 48

2.9.5 Classical cloning 51

2.9.6 Restriction endonuclease digestion of DNA 54

2.9.7 DNA ligation 54

2.9.8 Transformation of chemically competent *E. coli* 55

2.9.9 Glycerol stocks of *E. coli* 55

2.9.10 Preparation of LB and TB media 55
2.9.11 Preparation of agar plates 56
2.9.12 Purification of plasmid and BAC DNA 56
2.10 Transfection 56
2.10.1 Transfection of DF1 and HEK293T cells 56
2.10.2 Transfection of CEF cells with MDV-BAC for virus production 57
2.11 Protein analysis 58
2.11.1 SDS polyacrylamide gel electrophoresis (PAGE) 58
2.11.2 Instant Blue staining of proteins 58
2.11.3 Western blot 58
2.12 Co-immunoprecipitation 59
2.13 Bioimaging techniques 61
2.13.1 Immunofluorescence 61
2.13.2 Staining of live cells 62
2.13.3 Bioimaging and image processing 63
2.14 Förster Resonance Energy Transfer by Fluorescence lifetime imaging (FRET-FLIM) 63
2.15 Protein expression and purification 67
2.15.1 Size exclusion chromatography 67
2.15.2 Protein expression in *E. coli* 68
2.15.3 Protein expression in insect cells 72
2.16 Mass spectrometry analysis 76
2.17 Circular dichroism (CD) spectroscopy 76
2.18 ThermoFluor assay 78
2.19 Structural biology methods 79
2.19.1 Crystallisation screens 79
2.19.2 Duplex DNA formation 80
2.19.3 Crystallisation trials 80
2.19.4 Crystallisation optimisation screens 81
2.19.5 Cryoprotection 82
2.19.6 Data collection and processing 82
2.20 X-ray fluorescence (XRF) spectroscopy 83
2.21 Bioinformatic analyses 83

3 Interaction between pp14 and chCREB3 84
3.1 Introduction 84
3.2 Interaction of pp14 with chCREB3 in MDV infected cells 84
3.2.1 Biochemical characterisation of the interaction 84
3.2.2 Functional role of the interaction 86
3.3 Interaction studies with recombinant proteins 89
3.3.1 Determination of the domain of chCREB3 that interacts with pp14 89
3.3.2 Reverse co-immunoprecipitation 91
3.4 Expression and purification of the complex 93
3.5 Antibody cross-reactivity 97
3.6 Interaction between pp14 and chCREB3 in living cells 102
3.6.1 Introduction to FRET-FLIM 103
3.6.2 Construction of fluorescent fusion proteins 106
3.6.3 FRET-FLIM analysis 111
3.6.4 FRET-FLIM analysis of the interaction 115
3.7 Discussion 116

4 Structural and biophysical characterisation of pp14 122
4.1 Introduction 122
4.2 Protein sequence analysis 122
4.3 Expression and purification of pp14 124
4.3.1 Expression screening of pp14B 124

4.3.2 Scaling-up and purification of pp14B 126
4.3.3 Purification of pp14B from insect cells 127
4.3.4 Extraction of monomeric pp14B 128
4.4 ThermoFluor analysis 134
4.5 Crystallisation of pp14B 138
4.6 X-ray and X-ray fluorescence analyses of pp14B crystals 139
4.7 Secondary structure analysis using Circular Dichroism (CD) 140
4.8 Discussion 146

5 Structural characterisation of chicken CREB3 151
5.1 Introduction 151
5.2 Bioinformatic analyses of chCREB3 151
5.3 Characterisation of the N-terminal domain of chCREB3 152
5.3.1 Purification of the N-terminal domain of chCREB3 152
5.3.2 Purification of the bZIP domain of chCREB3 154
5.4 Purification of the C-terminal domain of chCREB3 171
5.4.1 Expression screening in insect cells 171
5.4.2 Scale-up and purification 172
5.5 Discussion 173

6 Conclusions and future work 175
6.1 Conclusions and future work 175
6.1.1 Interaction between pp14 and chCREB3 177
6.1.2 Insights into the structure of MDV pp14 180
6.1.3 Insights into the structure of chCREB3 186
6.2 Summary and final remarks 195

References 199

Appendix A – Vector maps 214

Acknowledgements

My DPhil was funded by the Biotechnology and Biological Science Research Council (BBSRC), the Pirbright Institute and the University of Oxford. This thesis would not have been possible without the above-mentioned financial support, and the help, support and guidance of many people.

First and foremost, I would like to thank my supervisors, Prof Jonathan Grimes, Prof Ray Owens, and Prof Venugopal Nair for their support and guidance throughout my DPhil. I am especially grateful to Prof Jonathan Grimes for his patience and constant support and guidance.

I also thank Dr Abdessamad Tahiri-Alaoui and Dr Ashley Roberts for their supervision, and Dr Pippa Hawes, and all past and present members of the Viral Oncogenesis Group for their help and support when I was working at the Pirbright Institute. I thank Prof Benedikt Kaufer of Free University of Berlin for providing me with the engineered BAC clones of MDV.

I am grateful to many people in the Division of Structural Biology (STRUBI) for providing such a supportive and friendly environment to work in. I am deeply indebted to Dr Guido Paesen, Dr Max Renner and Dr Dimple Karia for their continued support and encouragement. Many thanks to Max Renner for his help with the structural studies particularly. I also thank all past and present members of the Grimes Group, and all my friends in STRUBI for helpful discussions and their friendship. I am grateful to Dr Karl Harlos and Dr Tom Walter for their help with crystallisation experiments. Many thanks to Dr Sergi Padilla-Parra and Dr Luis Alvarez for their guidance with the FRET-FLIM experiments. Special thanks to Margaret Jones for her support in the lab and for her uplifting conversations, and to Weixian Lu for helping with cell culture.

I also thank all past and present members of the Oxford Protein Production Facility (OPPF), in particular Dr Louise Bird and Heather Rada, for their kind help and support during my time there. I thank Dr Rohanah Hussain of Diamond light source for B23 time (SM15935-1) and help with CD experiments.

Finally, thank you to my lovely Mum and Dad and my brother, whose love and support made this thesis possible.

Declaration of work

The work contained herein is my own except where explicitly stated otherwise in the text.

Keshalini Sabaratnam

St Edmund Hall

List of figures

Figure 1.1 Herpesvirus structure and the genomic structure of MDV. 4

Figure 1.2 Schematic showing the different stages of MD pathogenesis. 10

Figure 1.3 Vaccine-driven evolution of MDV and the changes in clinical features of MD. 14

Figure 1.4 MDV1 pp14 mediates neurovirulence. 18

Figure 1.5 pp14A and pp14B are expressed from alternatively spliced transcripts 20

Figure 1.6 Colocalisation of pp14 and endogenous huCREB3 in HeLA cells 22

Figure 1.7 Alignment of human and chicken CREB3 sequences 23

Figure 1.8 Subcellular localisation of chicken CREB3 constructs. 24

Figure 1.9 Colocalisation of pp14 and chCREB3 in MDV infected cells. 25

Figure 1.10 CREB3 subfamily of bZIP transcription factors 28

Figure 1.11 Schematic diagram showing the putative domains of huCREB3. 31

Figure 1.12 Crystal structures of bZIP domains bound to DNA. 34

Figure 1.13 Parallel coiled-coil structure of GCN4 ZIP homodimer 35

Figure 1.14 Interchain polar interactions found in the coiled-coil structure of the Fos–Jun ZIP heterodimer. 36

Figure 2.1 Schematic diagram of the BAC of MDV. 44

Figure 3.1 pp14 interacts with chCREB3 in MDV infected cells. 86

Figure 3.2 Localisation of chCREB3 was unaffected by pp14 in MDV infected CEF cells 88

Figure 3.3 pp14 interacts with the N-terminal domain of chCREB3. 91

Figure 3.4 Reverse co-immunoprecipitation failed to detect an interaction between pp14 and the N-terminal domain of chCREB3. 92

Figure 3.5 Purification using an anti-FLAG antibody. 95

Figure 3.6 GST-pull down assays. 96

Figure 3.7 Both pp14 and the NT domain of chCREB3 are detected only when anti-pp14 is used. 97

Figure 3.8 Anti-pp14 cross-reacts with the N-terminal domain of chCREB3. 98

Figure 3.9 Anti-pp14 epitope interacts with the N-terminal domain of chCREB3 in solution. 99

Figure 3.10 Sequence comparison between the pp14 peptide and chCREB3. 100

Figure 3.11 Cross-reactivity was not detected between anti-pp14 and the N-terminal domain of chCREB3 on a western blot or by immunofluorescence. 101
Figure 3.12 Förster resonance energy transfer. 103
Figure 3.13 Principles of FRET. 104
Figure 3.14 The principle of FLIM. 106
Figure 3.15 Schematic representations of the fluorescent protein fusion constructs. 107
Figure 3.16 Localisation of the fluorescent fusion proteins in DF1 cells. 109
Figure 3.17 ER-localisation of mRuby tagged NT-chCREB3-TM in DF1 cells 109
Figure 3.18 Donor fluorescence lifetime affected by the fluorescence intensity. 114
Figure 3.19 Interaction between pp14B and the NT-chCREB3 measured by FRET-FLIM. 116
Figure 4.1 Disorder prediction for pp14B. 123
Figure 4.2 3D structure prediction for pp14B. 124
Figure 4.3 High-throughput expression screening of pp14 fusion constructs in *E. coli*. 125
Figure 4.4 SEC analysis of pp14B extracted from *E. coli*. 126
Figure 4.5 SEC analysis of pp14B extracted from insect cells. 127
Figure 4.6 SEC analysis of pp14 treated with reducing agents. 129
Figure 4.7 Effect of additives/pH on the aggregation of pp14. 130
Figure 4.8 Extraction of monomeric pp14B using arginine. 130
Figure 4.9 Extraction of monomeric Cys-null pp14B. 132
Figure 4.10 Extraction of monomeric wild type pp14B using NP40. 133
Figure 4.11 ThermoFluor profile obtained for the monomeric pp14B extracted using arginine. 135
Figure 4.12 ThermoFluor profiles obtained for the Cys-null pp14B. 136
Figure 4.13 pH screen with the monomeric Cys-null pp14B. 136
Figure 4.14 ThermoFluor profiles obtained for the monomeric wild type pp14B extracted with NP40 (preparation 1). 137
Figure 4.15 ThermoFluor profiles obtained for the monomeric wild type pp14B extracted with NP40 (preparation 2). 137
Figure 4.16 Crystallisation experiments with the aggregated form of pp14B. 138
Figure 4.17 X–ray and X-ray fluorescence analysis of pp14B crystals. 139
Figure 4.18 Secondary structure analysis using far-UV CD spectroscopy. 144

Figure 4.19 Far-UV CD thermal denaturation experiment with monomeric wild type pp14B... 145

Figure 5.1 Disorder and structure prediction for chCREB3..... 152

Figure 5.2 Purification of the chCREB3 bZIP..... 155

Figure 5.3 Crystals of the chCREB3 bZIP. 156

Figure 5.4 Crystals obtained after optimisation. 157

Figure 5.5 The bZIP domain of chCREB3 forms a stable homodimer..... 159

Figure 5.6 The homodimeric chCREB3 bZIP has a structured basic region even in the absence of DNA..... 160

Figure 5.7 chCREB3 has a disulphide bond in the hydrophobic core of the bZIP..... 162

Figure 5.8 Asparagine residues at position 'a' of heptads promote homodimerisation and stability. 164

Figure 5.9 Sequence alignment of bZIP sequences..... 167

Figure 5.10 Crystals obtained from crystallisation trials with the chCREB3 bZIP in the presence of DNA..... 170

Figure 5.11 High-throughput expression screening of the CT-chCREB3 constructs in insect cells using the baculovirus system..... 172

Figure 5.12 Purification of the CT-domain of chCREB3 from insect cells..... 173

Figure 6.1 Crystal structures of bZIP domains. 188

Figure 6.2 Comparison of the basic regions of bZIP domains..... 190

Figure A 1 Vector maps (pcDNA3 vectors). 214

Figure A 2 Vector maps (pOPIN vectors). 216

List of tables

Table 1.1 Serotypes and pathotypes of Marek's Disease Virus. 3

Table 1.2 Known functions of members of the CREB3 subfamily. 30

Table 2.1 Amino acid sequences of proteins used in the study. 41

Table 2.2 Name and source of vectors used in this study 41

Table 2.3 Details and dilutions of primary and secondary antibodies used in this study 42

Table 2.4 Details of antibiotics used in this study. 42

Table 2.5 Typical temperature cycle used for PCR. 47

Table 2.6 pOPIN vectors used for cloning. 50

Table 2.7 Primers used for cloning into pOPIN vectors. 50

Table 2.8 Primers and restriction sites used for cloning into pcDNA3 vectors 52

Table 2.9 Transfection of fluorescent constructs. 64

Table 2.10 List of additives screened. 69

Table 2.11 Protein samples subjected to far-UV CD analysis. 77

Table 2.12 Protein samples subjected to ThermoFluor analysis. 78

Table 2.13 In-house pH screen buffers. 79

Table 2.14 Crystallisation trials. 81

Table 3.1 Experimental or control conditions subjected to FRET-FLIM analysis. 112

Table 4.1 List of additives screened 128

Table 5.1 Data collection and refinement statistics for the chCREB3 bZIP. 158

List of abbreviations

TF	Transferrin
% v/v	millilitre of solute/100 millilitre of solution
% w/v	the weight in grams of a solute per 100 millilitre of a solution
°C	Degrees Celsius
3D	Three dimensional
Å	Angström
a.u	Arbitrary units
aa	Amino acid
AD	Activation domain
AIDS	Acquired immunodeficiency syndrome
AP1	Activating protein 1
ARF4	Adenosine di phosphate (ADP) Ribosylation Factor 4
ATF	Activating transcription factor
BAC	Bacterial artificial chromosome
BATF	Basic activating transcription factor
BFA	Brefeldin A
BME	β-mercaptoethanol
bp	Base pair
BSA	Bovine serum albumin
bZIP	Basic leucine zipper
C-ter	Carboxyl terminal
C-terminal	Carboxyl terminal

C/EBP	CCAAT/Enhancer Binding Protein
cAMP	Cyclic AMP response element
CCCH zinc finger	(Cysteine)3(Histidine)1 zinc finger
CCL2	Chemokine (C-C motif) ligand 2
CCR	CC chemokine receptor
CD	Circular Dichroism
$CD4^+$	Cluster of differentiation 4 positive
Cdk	Cylin-dependent kinase
cDNA	Complementary DNA
CEF	Chicken embryo fibroblast
chCREB3	Chicken CREB3
cm	Centimetre
CNS	Central nervous system
Co-IP	Co-immunoprecipitation
CRE	cAMP response elements
CREB	Cyclic AMP response element binding protein
CREB3L1	CAMP-Responsive Element-Binding Protein 3-Like Protein 1
CREB3L2	CAMP-Responsive Element-Binding Protein 3-Like Protein 2
CREB3L3	CAMP-Responsive Element-Binding Protein 3-Like Protein 3
CREB3L4	CAMP-Responsive Element-Binding Protein 3-Like Protein 4
CT	Carboxyl terminal

CT-chCREB3	C-terminal domain of chicken CREB3
CTLs	Cytotoxic T lymphocytes
Cys	Cysteine
Da	Dalton
DAPI	4',6-diamidino-2-phenylindole
DBD	DNA binding domain
DC	Dendritic cells
DC-STAMP	Dendritic cell specific transmembrane protein
dH_20	Distilled water
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
DNA BD	DNA binding domain
dpi	Days post infection
dsDNA	Double stranded DNA
DTT	Dithiothreitol
E	Early
E. coli	Escherichia coli
EBV	Epstein Barr-virus
ECD	Electronic circular dichroism
ECL	Enhanced chemiluminescence
EDEM	ER degradation-enhancing alpha-mannosidase-like protein
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmatic reticulum

ERAD	Endoplasmic reticulum-associated degradation
ERSE	Endoplasmatic reticulum stress response element
FCS	Fetal calf serum
FFE	Feather follicle epithelium
FLIM	Fluorescence lifetime imaging
FRET	Förster resonance energy transfer
g	Constant of gravity
GFP	Green fluorescent protein
gK	Glycoprotein K
Gly	Glycine
GST	Glutathione-S-transferase
HA	Haemagglutinin
HALO	Modified haloalkane dehalogenase
HCF1	Host cell factor 1
HDAC	Histone deacetylase 3
HeLa	Henrietta Lacks' cells
Hep	Bam HI-H-encoded protein
Herp	Homocysteine-induced endoplasmic reticulum protein
his	Histidine
his6	Histidine x 6
HRP	Horseradish peroxidase
HSV1	Herpes simplex virus type 1
huCREB3	Human CREB3
HVT	Herpesvirus of turkeys
HVT	Herpes virus of turkeys

ICP	Infected cell peptide
ICP4	Infected cell peptide 4
IE	Immediate early
IF	Immunofluorescence
IFN-γ	Interferon gamma
IgG	Immunoglobulin
IP	Immunoprecipitation
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRES	Internal ribosome entry site
IRS	Internal repeat short
K	Kelvin
kDa	Kilo Dalton
KID	Kinase inhibitory domain
L	Late or litre
LAT	Latency-associated transcript
LB	Luria Bertani Broth
LDS	Lithium dodecyl sulfate
Leu	Leucine
Lkn-1	Leukotactin-1
LP	Late paralysis
MBP	Maltose-binding protein
mClover	Monomeric Clover
MD	Marek's Disease
mdeg	Millidegrees
MDV	Marek's Disease Virus

MDV1	Serotype 1 strains of Marek's disease virus
Meq	Marek’s disease virus EcoRI-Q fragment
MES	2-ethanesulfonic acid
Mg^{2+}	Magnesium ions
min	Minute
mini-F plasmid	Derivative of the fertility (F) factor of Escherichia coli
miRNA	Micro RNA
ml	Millilitre
mm	Millimetre
mM	Millimolar
mMDV	Mildly virulent MDV
MOPS	3-(N-morpholino) propanesulfonic acid
MPD	Methyl-2 4-pentanediol
MQ	Milli-Q ultra pure water
MR	Molecular replacement
mRNA	Messenger ribonucleic acid
mRuby	Monomeric Ruby
MS	Mass spectrometry
MSYB	bacterial protein MysB
MW	Molecular Weight
MWCO	Molecular Weight Cut-Off
Mys	Mystery protein
N-ter	Amino terminal
N-terminal	Amino terminal
NA	Numerical aperture

NaCl	Sodium chloride
NDSB	Non-detergent sulfobetaine
NFκB	Nuclear factor of kappa light polypeptide gene enhancer in B-cells
ng	Nanogram
Ni-NTA	Nickel-nitrilotriacetic acid
nm	Nanometre
NMR	Nuclear Magnetic Resonance
NP40	Nonidet P-40
ns	Nanosecond
NT	Amino terminal
NT-chCREB3	N-terminal domain of chicken CREB3
NT-chCREB3-TM	N-terminal domain plus the transmembrane region of chicken CREB3
OD_{600}	Optical density at 600 nm
ONEX	Overnight express auto-induction media
ORF	Open reading frame
PAGE	Polyacrylamide gel-electrophoresis
PAR4	Prostate apoptosis response 4
PBL	Peripheral blood lymphocytes
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PND	Permanent neurologic disease
PNS	Peripheral nervous system

POI	Protein of interest
pp14	14kDa phosphoprotein
pp38	Phosphoprotein 38
pRB-1B	BAC clone of the RB-1B strain MDV
psi	Pound force per square inch
Ref1	Redox effector factor 1
RFP	Red fluorescent protein
RIP	Regulated intramembrane proteolysis
RNA	Ribonucleic acid
ROI	Region of interest
RONN	Regional order neural network
ROS	Reactive oxidative species
rpm	Revolutions per minute
RT	Room temperature
S1P	Site 1 protease
S2P	Site 2 protease
SAXS	Small Angle X-ray Scattering
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
sec	Second
SOC	Super optimal broth
SUMO	Small ubiquitin-like modifier
TA	transcriptional activation
TAE	Tris-acetate-EDTA
TB	Terrific broth

TCEP	Tris (2-carboxyethyl) phosphine
TE	Tris-EDTA
TM	Transmembrane
Tm	Melting temperature
TP	Transient paralysis
Tris-HCl	Tris hydrochloride
TRL	Terminal repeat long
TRX	Thioredoxin
Tween	Polyoxyethylenesorbitanmonolaurate
U/ml	Units per millilitre
UB	Unbound
UL	Unique long
UPR	Unfolded protein response
UPRE	Unfolded protein response element
US	Unique short
UV	Ultraviolet
vIL	Viral interleukin
vTR	Viral telomerase RNA
vv+MDV	Very virulent plus MDV
vvMDV	Very virulent MDV
VZV	Varicella–zoster virus
WT	Wild type
XRF	X-ray fluorescence
Y2H	Yeast two-hybrid
$ZnCl_2$	Zinc chloride

α	Alpha
β	Beta
Δ	Delta
θ	Ellipticity
μg	Microgram
μl	Microlitre
μM	Micromolar
τ	Tau

1 Introduction

1.1 Marek's Disease

Marek's disease (MD) is an economically important disease of poultry (Davison and Nair, 2005) caused by serotype 1 strains of Marek's disease virus (MDV1) (Osterrieder et al., 2006). MDV has a worldwide distribution (Ahmed et al., 2016) and is an acute pathogen that induces a highly contagious disease, which can cause >90% morbidity and mortality in susceptible, unvaccinated hosts (Osterrieder et al., 2006). MD is controlled by vaccination with monovalent or multivalent live virus vaccines (Gupta et al., 2016). In spite of extensive vaccination, MD is estimated to cost the worldwide poultry industry $1–2 billion per year (Atkins et al., 2013; Hildebrandt et al., 2014).

MD pathology can present with a wide range of clinical symptoms depending on the strain of virus, vaccination status and host genetics of disease resistance (Witter and Schat, 2003). Four clinical forms have been described: classical, acute, transient paralysis (TP) and acute mortality syndrome. The classical form refers to the neural form of the disease. The main symptom of the classical form is flaccid paralysis, which is caused by lymphoproliferative lesions in the peripheral nerves (Witter and Schat, 2003). Small lymphomas can be present sometimes (Payne and Venugopal, 2000). The acute form of MD is a severe form of the disease, which causes severe depression, paralysis and sometimes death without clinical symptoms (Witter and Schat, 2003). It presents with lesions in the peripheral nerves with widespread lymphomatous infiltration affecting a variety of organs, including the viscera (liver, gonads, spleen, kidneys, lungs, proventriculus and

heart), skeletal muscle and skin (Witter and Schat, 2003). TP is an uncommon condition which occurs in young birds infected with highly virulent strains of MDV. It is characterised by a sudden onset of paralytic symptoms that often only last for 2–3 days after which the birds recover fully (Wexler, 2005; Witter and Schat, 2003; Witter et al., 1999). However, a more severe form of TP, termed acute TP, results in death of infected birds 1–3 days after the onset of paralytic signs (Wexler, 2005; Witter and Schat, 2003; Witter et al., 1999). Unlike the classical form of paralysis, TP is caused by inflammatory lesions of the central nervous system (CNS) and peripheral nervous system (PNS) (Mahy and Van Regenmortel, 2010). Another form of MD known as 'acute mortality syndrome' is induced by highly virulent pathotypes. This form of the disease is non-neoplastic, where the affected birds die with an early acute cytolytic disease (Payne and Venugopal, 2000).

In both the classical and acute forms of the disease, the peripheral nerves are affected by three types of infiltrative changes, which are termed proliferative, inflammatory and minor changes, and denoted A-, B- and C-type lesions, respectively (Payne and Biggs, 1967). The A-type lesions consist mainly of proliferating lymphoblasts and macrophages (Payne and Biggs, 1967). The B-type lesion, characterised by interneuritic oedema, consists mainly of small lymphocytes and plasma cells, and Schwann cell proliferation (Payne and Biggs, 1967). The A- and B- type lesions cause demyelination of nerves, which is thought to responsible for the paralytic symptoms (Payne and Biggs, 1967). The C-type lesion, often observed in birds that show no gross lesions or clinical signs, is characterised by light infiltration of small lymphocytes and plasma cells (Payne

and Biggs, 1967). Lymphomas in the visceral organs and other tissues are cytologically similar to the A-type lymphoproliferative lesions that affect nerves (Payne and Venugopal, 2000).

1.2 Marek's Disease Virus

1.2.1 Classification, structure and genome structure

The causative agent, MDV1, is a member of the family *Herpesviridae*, subfamily *Alphaherpesvirinae*, genus *Mardivirus* (Osterrieder et al., 2006). Viruses of the genus *Mardivirus* have been classified into three serotypes based on antigenic differences, Serotypes 1, 2 and 3. Serotype 1 contains all the oncogenic strains and their attenuated forms. These strains, based on pathogenicity, have been further classified into different pathotypes: very virulent plus (vv+ MDV), very virulent (vvMDV), virulent (vMDV), mildly virulent (mMDV), and weakly virulent pathotypes (Table 1.1) (Kato and Hirai, 1985). Serotype 2 contains the non-oncogenic isolates of MDV from chickens and Serotype 3 contains the strains of naturally avirulent herpesvirus of turkeys (HVT) (Table 1.1) (Payne and Venugopal, 2000).

Serotype	Pathotype	Representative strains
1	Very virulent plus (vv+MDV)	648A, 610A
1	Very virulent (vvMDV)	Md/5, Md/11, Ala -8, RB-1B
1	Virulent (vMDV)	HPRS-16, JM, GA
1	Mild (mMDV)	HPRS-B14, Conn A
1	Weakly virulent	CU-2, CVI-988
2	Non-oncogenic	SB-1, HPRS-24, 301B/1, HN-1
3	Non-oncogenic	HVT-FC126, HVT-PB1

Table 1.1 Serotypes and pathotypes of Marek's Disease Virus.
(Adapted from Payne and Venugopal, 2000)

Like other herpesviruses, all MDV serotypes contain a linear double-stranded deoxyribonucleic acid (DNA) genome (Osterrieder et al., 2006) encased within an icosahedral protein capsid (Joy et al., 2006). The capsid is surrounded by a protein-filled region called the tegument, which is then enclosed by a lipid bilayer envelope embedded with glycoproteins (Joy et al., 2006) (Figure 1.1A). The genomes of MDV1 strains, which are approximately 175 to 180 kb in length (McPherson and Delany, 2016), have the capacity to encode more than 90 genes (Schumacher et al., 2008). The genomes consist of a unique-long (U_L) and a unique-short (U_S) segment, each flanked by inverted terminal and internal repeats [terminal repeat long (TR_L), terminal repeat short (TR_S), internal repeat long (IR_L) and internal repeat short (IR_S); Figure 1.1B] (Osterrieder et al., 2006).

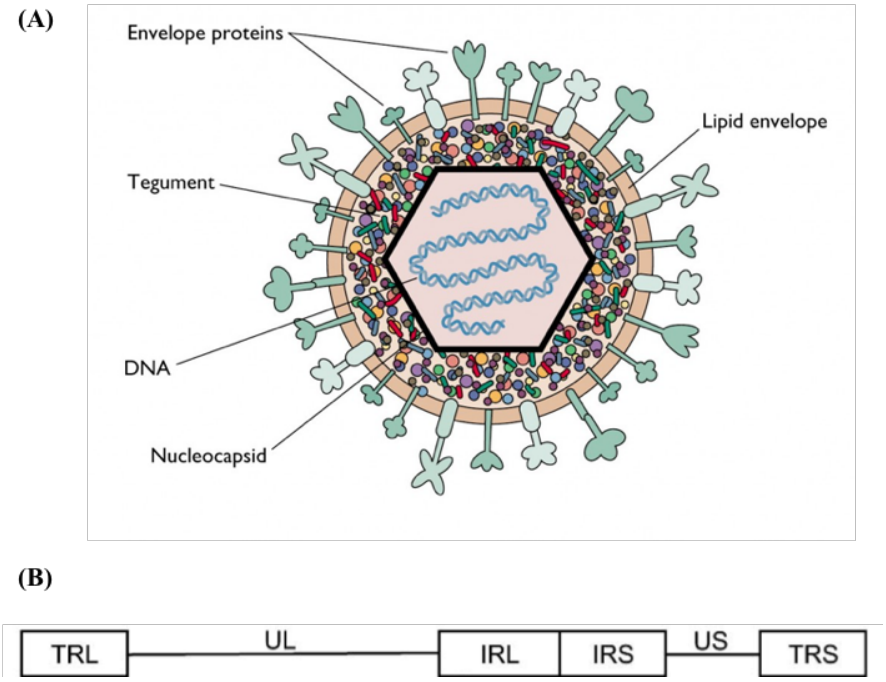


Figure 1.1 Herpesvirus structure and the genomic structure of MDV.
(A) Schematic diagram illustrating the multilayer organisation of herpesviruses. (Taken from: http://www.twiv.tv/virus-structure/) **(B)** Schematic representation of the MDV genome. The MDV genome consists of a unique long region (U_L) that is flanked by a terminal repeat long region (TR_L) and an internal repeat long region (IR_L), and a unique short region (US) that is flanked by a terminal repeat short region (TR_S) and an internal repeat short region (IR_S). (Picture taken from Tahiri-Alaoui et al., 2012)

MDV is genetically closely related to human herpesvirus 1 (herpes simplex virus type 1, HSV1) and human herpesvirus 3 (varicella–zoster virus, VZV). Genes contained in the U_L and U_S segments are largely homologous to those of HSV1 and VZV. The genus- and virus-specific genes are located in the inverted repeat regions, predominantly TR_L and IR_L. The genes that define the oncogenic serotype 1 include the Meq (Marek's disease virus EcoRI-Q fragment) oncogene and other unique genes encoding a viral interleukin (vIL)-8 homolog, phosphoprotein (pp)38, and viral telomerase RNA (vTR) (McPherson and Delany, 2016). MDV, especially due to its lymphotropism, is biologically similar to members of the *Gammaherpesvirinae*, such as Epstein–Barr virus (Osterrieder et al., 2006).

1.2.2 Herpesvirus replication and latency: Gene expression regulation

Like those of other herpesviruses, the life cycle of MDV consists of productive (lytic) and non-productive (latent) phases. Herpesvirus infection begins with the enveloped virus particle binding to the cell surface of a susceptible cell (Jenkins and Hoffman, 2000). MDV entry into cells may involve binding of glycoproteins B, C and D (gB, gC and gD) to host cell receptors to initiate fusion of the viral envelope to the cellular membrane leading to the release of the nucleocapsid into the cell cytoplasm (Davison and Nair, 2004). The released virion is uncoated by cellular enzymes, and the viral DNA circularises and enters the nucleus (Davison and Nair, 2004). During lytic infection, the genes are expressed in a sequential manner and are temporally subdivided into immediate-early (IE), early (E), and late (L) genes (Jenkins and Hoffman, 2000). MDV lytic infection involves expression of approximately 50 per cent of the viral genome, and about 50 viral polypeptides have been linked with this phase (Davison and Nair, 2004). The

initiation and temporal regulation of MDV gene expression is not as well understood as the study of MDV gene expression in cell culture is difficult due to the cell-associated nature of the virus (Davison and Nair, 2004). Thus, many of the MDV genes have been attributed function by homology to genes of more closely studied members of the alphaherpesvirus family (Schumacher et al., 2008).

The IE genes are transcribed immediately after infection. These gene products prepare the cell for viral gene expression, and initiate transcription of early genes (Arvin et al., 2007). In members of the alphaherpesvirus family that are closely related to the Mardiviruses, the binding of viral protein 16 (VP16) or α-transinducing factor (α-TIF), the U_L48 homologous protein, to a TAATGARAT motif found upstream of the ICP4 ORF initiates the expression of lytic genes (Davison and Nair, 2004). This interaction recruits cellular transcription factors which leads to the expression of the infected cell polypeptide 4 (ICP4) and other IE proteins (Davison and Nair, 2004). α-TIF is not absolutely required to initiate transcription since herpesvirus DNA itself has been found to be infectious and the U_L48 homologues of both varicella zoster virus (VZV) and MDV are not required for virus growth in cultured cells (Davison and Nair, 2004; Dorange et al., 2002). Herpesvirus ICP4 homologues are known as key IE transactivators of E gene expression. As infection develops, ICP4 homologues have the ability to self-regulate their expression in a temporal manner through recruitment of various cellular and viral transcription factors (Davison and Nair, 2004; Kato et al., 2002; Pratt et al., 1994). At least five sense transcripts that are expressed from the MDV ICP4 genomic region have been detected (Davison and Nair, 2004). The only other IE gene that has been identified in MDV is the ICP27 homologue which as has

been reported for other herpesvirus ICP27 homologues plays a role in post-transcriptional gene regulation by inhibiting both viral and host cell splicing (Amor et al., 2011; Moriuchi et al., 1994). Interestingly in MDV, the gK gene (U_L53) is expressed as an IE product (Ren et al., 1994), whereas in all the other herpesviruses analysed so far gK is expressed as a true-late protein. Another gene that is expressed with IE kinetics is the serotype 1 MDV specific L-ORF10 gene encoding the 14kDa phosphoprotein (pp14) (Hong and Coussens, 1994), which is associated with the neurovirulence phenotype of the virus (Tahiri-Alaoui et al., 2012) (discussed in Sections 1.3.3 and 1.4).

The E genes encode the proteins necessary for the viral replication (Arvin et al., 2007). Although as many as six E MDV genes have been reported, the only E MDV genes identified so far are R-LORF-14 and R-LORF14a encoding the phosphoproteins pp24 and pp38, respectively (Zhu et al., 1994). These genes are highly expressed during lytic infection (Zhu et al., 1994). It has been reported that these genes may be activated by ICP4 (Pratt et al., 1994). Notably, pp38 is one of the hallmarks of lytic infection (Davison and Nair, 2004) as deletion of the gene from MDV impairs lytic replication *in vivo* (Reddy et al., 2002). The circular herpesvirus DNA in infected cells are replicated through a rolling-circle mechanism, which generates concatemers that are cleaved and packaged into viral particles (Ben-Porat and Rixon, 1979). For genome replication, herpesviruses typically encode a DNA polymerase and a number of viral proteins required for DNA synthesis (Minarovits et al., 2007). These genes which are needed for viral DNA synthesis and the transcription of late structural genes are assumed to be regulated with early kinetics as the temporal regulation of the expression of MDV

enzymes and nucleic acid-binding proteins that take part in DNA synthesis is not known (Davison and Nair, 2004). The genes associated with MDV DNA synthesis include thymidine kinase (U_L23), dUTPase (U_L50), ribosyl reductase subunits (U_L39 and U_L40), origin-binding protein (U_L9), DNA polymerase subunits (U_L30 and U_L42), proteins forming the helicase-primase complex (U_L5, U_L8, and U_L52), and single-stranded DNA binding protein (U_L29) (Davison and Nair, 2004; Tulman et al., 2000).

After the onset of viral DNA replication, the late genes are expressed. These consist primarily of viral structural proteins, which include the nucleocapsid proteins, the tegument proteins, and the glycoproteins (Kato and Hirai, 1985), although some IE and E proteins have been shown to be structural components of the mature virion, e.g. HSV-l transactivating factors, ICP0 and ICP4 (Davison and Nair, 2004). During this stage, the viral DNA assembles with the preformed capsid structure and the resulting nucleocapsid obtains its envelope when it buds through the nuclear membrane (Jenkins and Hoffman, 2000). The identification of structural proteins of MDV is not easy due to the highly cell-associated nature of the virus, both *in vivo* and *in vitro* (Davison and Nair, 2004). In MDV, membrane glycoproteins (with the exception of gK), the major capsid protein VP5 and the major tegument proteins expressed from MDV059-MDV062 that are homologous to the HSV1 UL46-UL49 gene encoding VP11/12, VP13/14, VP16 and VP22 have been identified as L genes (Davison and Nair, 2004; Dorange et al., 2002; Lupiani et al., 2001). While VP11/12, VP13/14, VP16 are dispensable for MDV growth in cultured cells, VP22 is absolutely required for MDV replication (Dorange et al., 2002).

A hallmark of all herpesviruses is the ability to establish and maintain a latent infection (Jenkins and Hoffman, 2000). Latency is characterised by the absence of any viral replication, with limited amount of viral transcription, and the persistence of the viral genome in the infected cell (Jenkins and Hoffman, 2000). The overexpression of a limited subset of specific viral genes has been linked to MDV latency (Cantello et al., 1994), including the non-protein-coding latency-associated transcripts (LATs) that are antisense to the IE gene ICP4 (Cantello et al., 1994) and a cluster of viral latency-associated microRNAs (cluster mdv1-miR-M8-M10) in the LAT region (Burnside et al., 2006). These non-coding RNAs control the latent/lytic cycle balance by regulating viral or host cell gene expression, especially by inhibiting expression of viral IE genes (Strassheim et al., 2012). In addition, Meq is expressed during MDV latency and in the transformed cell. It has been shown to block apoptosis and transactivate gene expression (Parcells et al., 2003).

Latent herpesvirus infections can be reactivated resulting in the production of a lytic cycle of replication (Jenkins and Hoffman, 2000). The trigger mechanisms of MDV reactivation are not known, but it has been suggested it may be induced by a drop in cytokine levels in peripheral tissue (Parcells et al., 2003). During MDV reactivation, splice variants of Meq lacking several of the domains important to its transactivation and transrepression function are highly expressed (Parcells et al., 2003). In addition, genes from three open reading frames (ORFs) rightward from the origin of replication [Bam HI-H-encoded protein (Hep), Mystery protein (Mys), and pp38] are expressed during early stages of reactivation (Parcells et al., 2003).

1.2.3 Pathogenesis of MDV

MDV has a complex lifecycle. Figure 1.2 illustrates an overview model of MDV infection, pathogenesis, and transformation. The natural route of MDV infection is via the respiratory tract following inhalation of infectious cell-free MDV present in dander shed from feather follicles (Abdul-Careem et al., 2009). The early stages of infection in the lung are not well understood (Abdul-Careem et al., 2009). It is presumed macrophages in the lung pick up the virus and transfer it to lymphoid organs such as bursa of Fabricius, thymus and spleen (Payne and Venugopal, 2000). With the exception of the feather follicle epithelium, MDV infection is strictly cell-associated *in vivo* (Payne, 2012) and virus movement occurs through cell-to-cell spread (Nazerian et al., 1968).

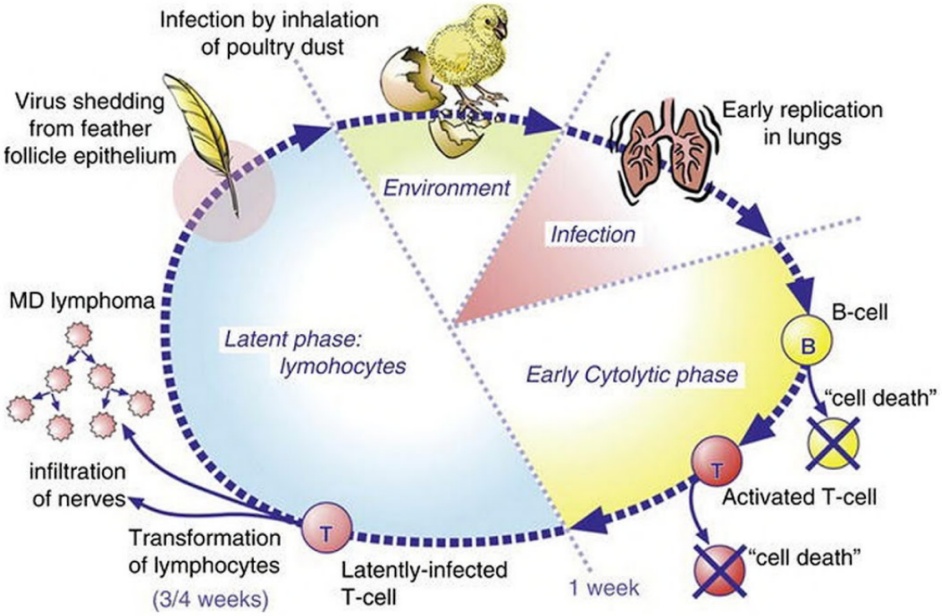


Figure 1.2 Schematic showing the different stages of MD pathogenesis.
The chickens become infected from dander in the environment by inhalation. The virus particles are taken up by macrophages in the lungs, which bring the virus to lymphoid organs. A cytolytic infection then occurs in B and activated T cells, leading to latency and transformation of the T cells. The virus is transmitted throughout the chicken by the circulating latently infected T cells. Then, the infection of the feather follicle epithelium results in the shedding of the virus. (Picture taken from Witter and Schat, 2003)

The pathogenesis of virulent MDV occurs through four overlapping infection stages:

Early cytolytic

MDV is transported to lymphoid organs within 2–3 days post infection (dpi) and a productive infection of the lymphoid cells can be detected between 5 and 6 days (Gimeno, 2008). The initial target cells are B lymphocytes, in which the virus induces a productive cytolytic infection resulting in the death of the infected cells. This process can elicit an immune response triggering the recruitment of several inflammatory cells including activated T cells (Gimeno, 2008). Subsequently, MDV infects activated CD4+ T cells (resting T cells are refractory to MDV infection), which can also support productive infection but in most cases infection becomes latent (Witter and Schat, 2003). These early cytolytic events targeting cells of the immune system results in a transient immunosuppression (Payne and Venugopal, 2000). MDV can also be detected in the brain, peripheral nerves and eye as early as 6 dpi (Gimeno, 2008). Depending on the virulence of the virus, reactivation of MDV can occur in some of the lymphocytes infiltrating the brain tissue, resulting in the induction of an acute inflammatory response characterised by vasculitis and vasogenic oedema that are responsible for the clinical signs of TP (Gimeno, 2008).

Latent infection

MDV establishes latent infection in activated $CD4^+$ T cells (Gimeno, 2008). Latent infection can be detected by 7–8 dpi not only in lymphoid organs but also in peripheral blood lymphocytes (PBL) (Gimeno, 2008), which disseminate the virus to other tissues of the chickens (Gimeno, 2008). In MDV, latency is associated

with genomic integration into chicken telomeres, which is facilitated by MDV telomeric repeats (McPherson et al., 2016).

Late cytolytic infection with immunosuppression

The reactivation phase of MDV is exhibited as a second cycle of lytic infection—detected by the end of the second week. This phase leads to permanent immunosuppression and can result in the development of lesions in all lymphatic tissues (Heidari et al., 2016). A fully productive infection also occurs in the skin and feather follicle epithelia, which results in the continuous shedding of cell-free virus, and virus transmission (Payne and Venugopal, 2000).

Transformation

Latently infected T cells can become transformed and develop into lymphomas (Gimeno, 2008). MDV-induced lymphomas can be detected in peripheral nerves, brain, skin, and any visceral organ as transformation can occur in lymphocytes infiltrating any tissue (Cho et al., 1998). Genomic integration and an efficient induction of the latent state of infection is considered a requirement for transformation (Kaufer et al., 2011), but the precise relationships of the latency and tumorigenesis stages are currently unknown (Davison and Nair, 2004). Several factors that contribute to MDV-induced lymphomagenesis have been identified. These include the major oncogene Meq, encoding a basic leucine zipper (bZIP) transcription factor which alters expression of several cellular and viral genes contributing to transformation, and the vIL-8 and vTR genes (Osterrieder et al., 2006).

1.3 Neurovirulence of MDV

1.3.1 Vaccination and emergence of neurovirulent strains of MDV

MD has been a major concern in the poultry industry due to the continuous emergence of increasingly virulent strains of MDV in the face of vaccination (Figure 1.3A) (Osterrieder et al., 2006). MD has been successfully controlled using vaccination since the early 1970s (Churchill et al., 1969). All three serotypes, the antigenically related, non-pathogenic serotype 2 and 3 strains and attenuated strains of serotype 1 MDV, are used as live vaccines in monovalent or multivalent forms (Payne and Venugopal, 2000). While the vaccines provide protection against tumours and clinical disease, by unknown mechanisms, they do not induce sterile immunity (Osterrieder et al., 2006). This means, despite the vaccination status, the birds carry and shed the vaccine viruses and any acquired pathogenic strains of MDV throughout their life (Atkins et al., 2013; Read et al., 2015), and thus allowing the virus to evolve. This type of vaccine is called a leaky vaccine (Read et al., 2015). Although the factors influencing the evolution of virus are unclear, it has been suggested it could be driven by the use of leaky anti-disease vaccines (Atkins et al., 2013; Read et al., 2015).

MD was first described as a polyneuritis by Jozsef Marek in 1907 (Davison and Nair, 2005; Gong et al., 2013). In the 1950s, intensification in the methods of poultry production led to the evolution of MDV towards greater virulence, causing the appearance of the more acute form of MD characterised by visceral lymphomas (Davison and Nair, 2005). In the early 1970s, the first vaccines against MD, based on HVT, were developed to control the disease (Baigent et al., 2006; Gong et al.,

2013). More virulent strains emerged in the mid-1980s leading to the introduction of serotype 2 vaccine such as strain SB-1, which was used in combination with HVT (Gong et al., 2013). In the 1990s, serotype 1 vaccine CVI988/Rispens was introduced to combat more virulent field viruses that emerged (Gong et al., 2013). Subsequently, strains with enhanced virulence may have evolved as the failure of CVI988/Rispens vaccination, used either alone or in combination with serotype 2 and/or serotype 3 vaccines, has been reported (Gong et al., 2013).

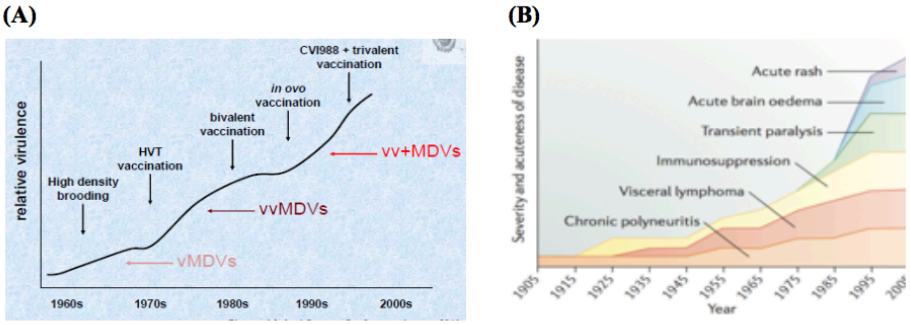


Figure 1.3 Vaccine-driven evolution of MDV and the changes in clinical features of MD.
(A) Graph displaying the vaccine-driven evolution of MDV towards greater virulence (Picture taken from Mark S. Parcells, Plant and Animal Genome Conference, January 2014) **(B)** Graph illustrating the changing clinical presentation of MD over time (Picture taken from Osterrieder et al., 2006)

With the evolution of MDV towards greater virulence, the clinical presentations of MD have also changed dramatically (Figure 1.3B) (Osterrieder et al., 2006). Chronic polyneuritis was prevalent until 1925; from 1925–1950, visceral lymphomas were observed; and from 1950 onwards more aggressive and faster-developing tumours appeared (Osterrieder et al., 2006). Significantly, in the past 20 years, the character of MD has changed into a syndrome that afflicts birds with severe neuropathology but which differs from the polyneuritis originally observed in the early 1900s (Osterrieder et al., 2006). Highly virulent strains have been

reported to cause a non-neoplastic disease, involving acute TP and acute brain oedema, capable of inducing higher rates of mortality (Osterrieder et al., 2006). These changes in MD highlight the importance of identifying and characterising both viral and host factors that contribute to MDV neuropathology, as it will pave the way towards developing improved intervention strategies against MD (Nair, 2005).

1.3.2 Neurologic manifestations of MD

MDV causes a wide range of neurological syndromes. The severity ranges from mild to severe depending on the strain of the virus. The neurologic manifestations of the virus include:

Neurolymphomatosis

The most common clinical sign observed in the classical form of MD is partial or complete paralysis of the legs and wings. The clinical paralysis is believed to be caused by enlargement of one or more of the peripheral nerves, resulting from lymphoid infiltration and demyelination (Powell, 1986; Witter and Schat, 2003).

Transient paralysis

TP, a rare condition that affects young birds, is induced by virulent pathotypes of MDV. Two forms of TP have been described, classical and acute. The classical form of TP is usually induced by virulent or very virulent strains of MDV (Abdul-Careem et al., 2006). In this form, affected birds display varying degrees of ataxia and flaccid paralysis of the neck or limbs 8–12 dpi but recover fully and rapidly within 2–3 days after the onset of paralytic signs (Witter and Schat, 2003; Xu et

al., 2012). Acute TP is a more severe form of the classical TP caused by highly virulent pathotypes (vvMDV and vv+MDV). This form is characterised by an irreversible flaccid paralysis of the neck and limbs, which usually results in death 1–3 days after the onset of paralytic signs (Xu et al., 2012). The lesions in birds affected by TP consist of proliferation of capillary endothelium leading to vasculitis, perivascular cuffing of mononuclear cells, vacuolation, and oedema (Abdul-Careem et al., 2006). These lesions may be detected in the brain, spinal cord and peripheral nerves (Abdul-Careem et al., 2006).

Permanent neurologic disease

Permanent neurologic disease (PND) is a non-paralytic syndrome which has been observed only in chickens that recovered from paralysis. However, it is not known whether it is a distinct syndrome or a by-product of TP (Gimeno et al., 2001; Witter and Schat, 2003). It is characterised by the onset of ataxia, torticollis and random spastic movements of the head, neck or eyelids, which do not resolve (Gimeno et al., 2001; Witter and Schat, 2003).

Late paralysis (LP)

LP, which is rarely observed, is characterised by the late onset of the paralytic stage (Witter and Schat, 2003).

1.3.3 Mechanisms underlying MDV neurovirulence

Despite the development of MDV strains towards greater neurovirulence, and the strong association of these strains with vaccine failures (Gimeno, 2008), the mechanisms underlying MDV-mediated neuropathology are still poorly

understood (Tahiri-Alaoui et al., 2012). The presence of B cells and antibody response, and enhanced expression of pro-inflammatory cytokines and interferon (IFN)-gamma in the brain have been associated with TP and the neurological disorders of MD (Abdul-Careem et al., 2006).

The current knowledge of the viral determinants of neurovirulence is limited. The 14kDa phosphoprotein, pp14, is the only identified, MDV1 specific neurovirulence factor (Tahiri-Alaoui et al., 2012). pp14 was identified as a neurovirulence factor by Tahiri-Alaoui et al. (2012) using reverse genetics (Tahiri-Alaoui et al., 2012). This study showed that the absence of pp14 expression correlated with increased survival of MDV-infected chickens (Tahiri-Alaoui et al., 2012). Within the experimental period, 100% of the birds inoculated with the double copy pp14 deletion virus, RB-1BΔpp14, survived compared with only 50% and 60% from the groups inoculated with the parent virus or the pp14 single copy revertant virus, RB-1B-SCR-pp14, respectively (Figure 1.4A) (Tahiri-Alaoui et al., 2012). Significantly, the study revealed birds inoculated with the double copy pp14 deletion virus showed no gross lesions in the peripheral nerves (Figure 1.4B) although typical MD pathology such as the presence of tumours in the spleen, kidney, and gonad was observed. Further, only 16.6% of the group infected with the double copy pp14 deletion virus showed microscopic lesions in the nerves, compared with 62.8% and 33.3% from groups inoculated with the parent wild type and the single copy pp14 revertant virus, RB-1B-SCR-pp14, respectively (Figure 1.4C). While the microscopic lesions in peripheral nerves of birds infected with the wild type RB-1B consisted of marked cellular infiltration, numerous proliferating lymphoblastic cells, and occasional oedema, the birds infected with

the double copy pp14 deletion virus exhibited only occasional cellular infiltrations. These results indicated pp14 may be required for efficient infiltration of mononuclear cells into the peripheral nerves of chickens during MDV pathogenesis. The DPhil project focussed on gaining further insights into the structure and function of pp14, and its mechanisms of action in neurovirulence.

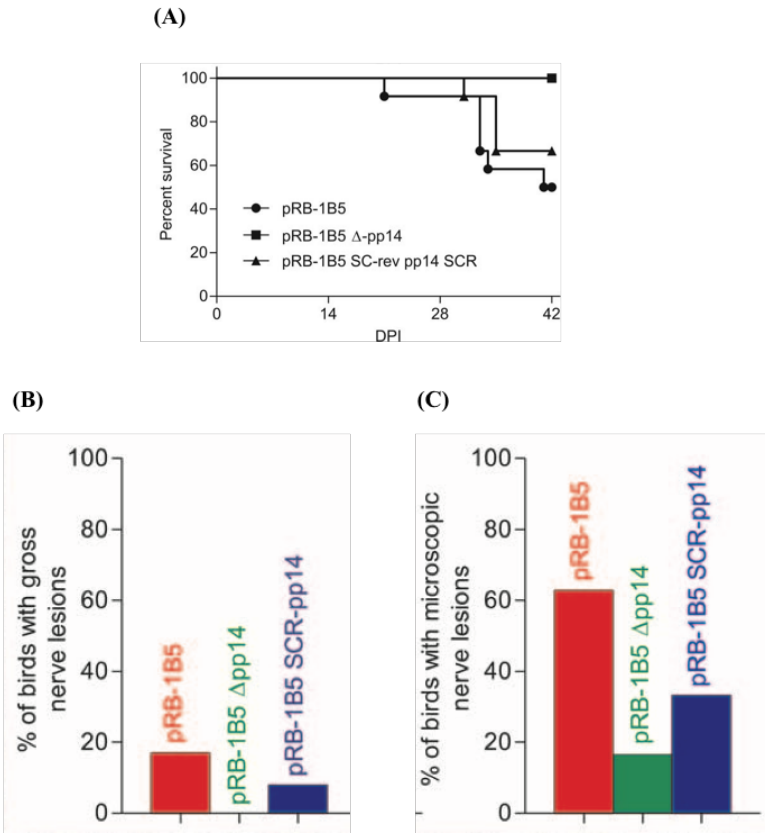


Figure 1.4 MDV1 pp14 mediates neurovirulence.
Chickens were infected with wild type RB-1B, double copy pp14 deletion virus, RB-1BΔpp14, and single copy revertant virus, RB-1B-SCR-pp14. **(A)** Survival curves of chickens infected with the indicated viruses. **(B)** Graph showing the percentage of birds with gross nerve lesions. **(C)** Graph showing the percentage of birds with microscopic nerve lesions. (Taken from Tahiri-Alaoui et al., 2012)

1.4 MDV neurovirulence factor, pp14

1.4.1 14kDa phosphoprotein, pp14

The 14kDa phosphoprotein, pp14, lacking homology to any know protein, represents a MDV serotype 1 specific antigen (Tahiri-Alaoui et al., 2009). It is

associated with the neurovirulence phenotype of MDV (Tahiri-Alaoui et al., 2012) and is dispensable for viral replication and oncogenesis (Tahiri-Alaoui et al., 2012).

The phosphoprotein pp14 is an immediate-early gene expressed from a bicistronic transcript that belongs to the 1.8-kb family of transcripts (Tahiri-Alaoui et al., 2009), which are associated with oncogenicity and maintenance of latency in lymphoblastoid cells (Parcells et al., 2003). The first open reading frame (ORF) of the bicistronic transcript encodes pp14 and, the second ORF encodes a 12kDa protein, RLORF9 (Tahiri-Alaoui et al., 2009). Two variants of pp14 (pp14A and pp14B) have been identified, which are coded by two alternatively spliced messages (Hong et al., 1995; Hong and Coussens, 1994). pp14A is produced via cap-dependent translation from mRNA variant 1, whereas pp14B is produced via an internal ribosome entry site (IRES)-dependent mechanism from mRNA variant 2 that possesses an IRES (Tahiri-Alaoui et al., 2009). pp14A and pp14B differ in the amino acid sequences at the amino-termini (N-termini), which are either 14 or 24 amino acids (aa) long, respectively, but have a common carboxyl terminal (C-terminal) domain, which is 71aa (Figure 1.5) (Hong et al., 1995; Hong and Coussens, 1994).

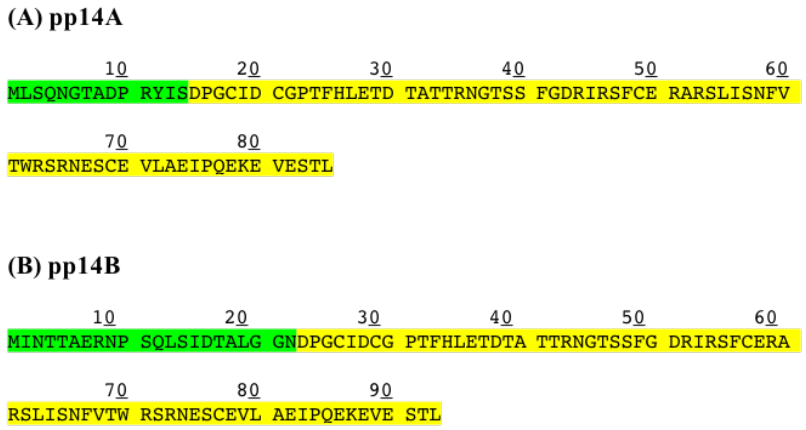


Figure 1.5 pp14A and pp14B are expressed from alternatively spliced transcripts.
(A) and **(B)** Amino acid sequences of the two forms of pp14 expressed from two alternatively spliced transcripts. The two isoforms of pp14 differ in the amino acid sequences at the N-termini (green), which are either 14aa (pp14A) or 22aa (pp14B) long, but have a common C-terminal domain, which is 71aa (yellow).

The mRNA variant that possesses an IRES (encoding pp14B) has been found to be more abundant in MDV-infected and transformed cells than the mRNA variant lacking the IRES (Tahiri-Alaoui et al., 2009). It has been proposed that the IRES-mediated translation initiation might be the preferred mode of translation for a subset of immediate-early transcripts of MDV1 during the early stages of viral infection when cap-dependent translation initiation might be compromised (Tahiri-Alaoui et al., 2009). The ability of this transcript to mediate cap-independent translation initiation may enable it to circumvent translation inhibition by the host double-stranded-RNA-dependent protein kinase PKR, as has been reported for the hepatitis C virus HCV IRES (Tahiri-Alaoui et al., 2009).

The difference in amino acid composition of the N-termini of the two forms of pp14 does not affect the cellular localisation of the proteins. Both forms of protein have been found to localise in the cytoplasm (Tahiri-Alaoui et al., 2012) and have been detected in cells lytically infected with oncogenic serotype 1 MDV strains

and their attenuated derivatives as well as latently MDV-infected and transformed cells (Tahiri-Alaoui et al., 2012) and in the chicken lymphoblastoid cell line, MSB-1, derived from a MD tumour (Hong et al., 1995; Hong and Coussens, 1994).

The estimated molecular weights of these two protein forms are 10.3kDa or 9.3kDa, but by western blotting the proteins are detected as 14kDa proteins (Hong et al., 1995; Hong and Coussens, 1994). Both forms of pp14 possess potential casein kinase and histone kinase consensus phosphorylation sites (Kato and Hirai, 1985) and have been found to be highly phosphorylated (Hong et al., 1995). Phosphorylation may be one factor contributing to the discrepancy between calculated and apparent sizes of pp14, but it is not the only factor contributing to the apparent size discrepancy (Hong et al., 1995). This was demonstrated when cell lysates from MDV-infected cells were treated with calf intestinal phosphatase and analysed by western blotting, only a slight shift in size (1–2kDa smaller) was observed relative to the untreated sample (Hong et al., 1995). Computer analysis predicts potential N-glycosylation sites within both forms; however, it was found that treatment of MDV-infected cells with tunicamycin (inhibitor of N-linked glycosylation) had little or no effect on the apparent size of pp14 (Hong et al., 1995).

1.4.2 Interaction of pp14 with CREB3

Dr Tahiri Alaoui et al. identified pp14 as a neurovirulence factor of MDV1. The availability of yeast-two-hybrid (Y2H) screen expertise in their laboratory prompted them to look for a cellular ligand for pp14. Using the shared C-terminal domain of pp14 as bait, a Y2H screen of the HeLa cell line cDNA library was

carried out and the screen identified the human cAMP response element binding protein (huCREB3) as an interacting partner (unpublished data). Subsequently, the colocalisation of recombinant pp14 and endogenous CREB3 was shown in HeLa cells by immunocytochemistry (Figure 1.6; unpublished data).

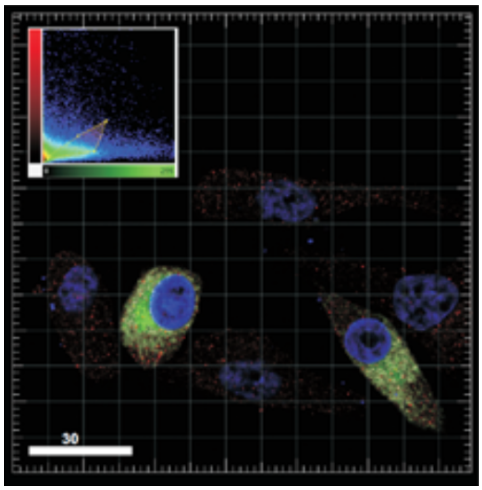


Figure 1.6 Colocalisation of pp14 and endogenous huCREB3 in HeLA cells. HeLa cells expressing recombinant pp14 were stained with an anti-pp14 antibody (green). Endogenous huCREB3 was stained in red. The colocalisation is shown as grey signal in cells expressing pp14. The inset is a 2D histogram of the colocalisation analysis using IMARIS software. The Pearson's coefficient in colocalised volume = 0.5585. (Unpublished data obtained from Dr Tahiri Alaoui)

Bioinformatics analyses revealed chicken CREB3 (chCREB3) and human CREB3 share a conserved domain organisation (Figure 1.7) and that the sequences share 45% amino acid sequence identity, with a high degree of similarity (73%) in the conserved bZIP domain. The predicted domain organisation of chCREB3 (Figure 1.8A) based on the sequence of huCREB3 was found to be consistent with the results of immunocytochemistry using DF1 cells transiently transfected with the full-length (FL) or truncations of chCREB3 (Figure 1.8, unpublished data). The full-length wild type chCREB3 showed ER/cytoplasmic localisation (Figure 1.8B). The deletion of the predicted transmembrane (TM) domain resulted in predominant nuclear localisation of chCREB3ΔTM (Figure 1.8C). The N-terminal domain of chCREB3 without the transmembrane domain (NT-chCREB3) showed

predominant nuclear localisation (Figure 1.8D). The C-terminal domain of chCREB3 without the transmembrane domain (CT-chCREB3) showed cytoplasmic and nuclear localisation (Figure 1.8E). These results led to the hypothesis that pp14 may mediate its neurovirulence through interaction with chCREB3.

Figure 1.7 Alignment of human and chicken CREB3 sequences.
Human CREB3 (Ensembl ID: ENST00000353704) and chicken CREB3 (Ensembl ID: ENSGALT00000003976) sequences were aligned using Clustal ["---" indicates gap. An asterisk "*" indicates all the sequences have the same nucleotide (fully conserved). Dots = no conservation. 1 dot "." = some conservation. 2 dots ":" = all pyrimidines or all purines at position.] Based on the sequence of human CREB3, the predicted N-terminal, transmembrane, and C-terminal domains of chicken CREB3 are highlighted in green, cyan and yellow, respectively.

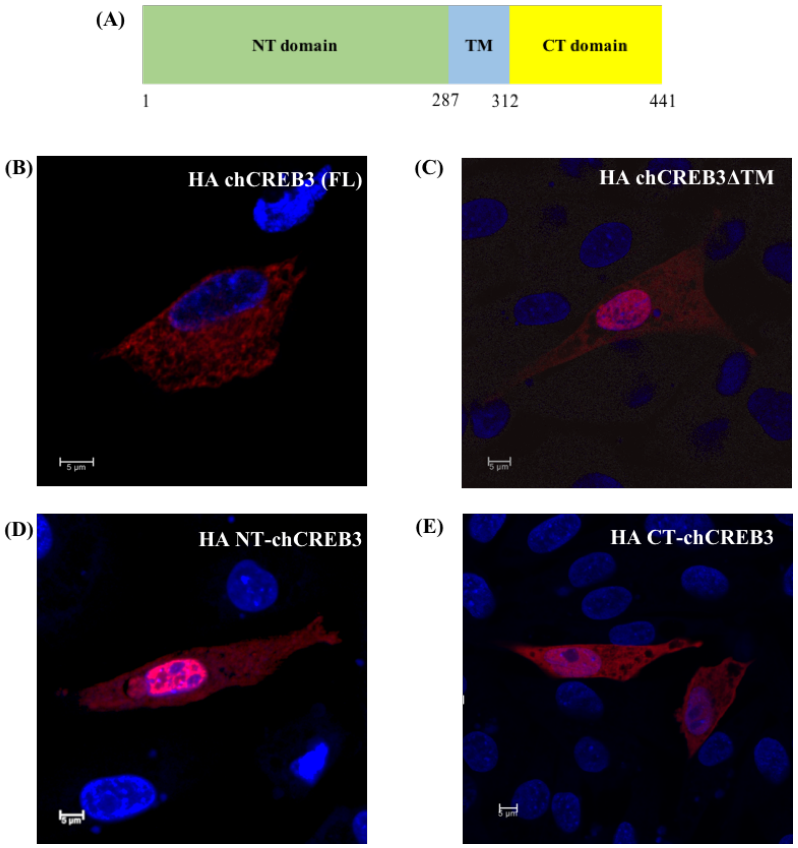


Figure 1.8 Subcellular localisation of chicken CREB3 constructs.
(A) Schematic diagram showing the predicted N-terminal (NT), transmembrane (TM) and C-terminal (CT) domains of chicken CREB3. The cDNA sequence of chicken CREB3 contains a coding region of 1,323bp, encoding a protein of 441 amino acids with a predicted molecular mass of 48.7kDa. Immunocytochemistry using an anti-HA antibody shows the cellular localisation of **(B)** HA tagged chCREB3 (FL); **(C)** HA tagged chCREB3ΔTM; **(D)** HA tagged NT-chCREB3 and **(E)** HA tagged CT-chCREB3 in DF1 cells. (Scale bars = 5μm) (Unpublished data obtained from Dr Tahiri-Alaoui)

1.4.3 Interaction of pp14 with chCREB3 in MDV infected cells

Dr Tahiri-Alaoui used a cellular infection model to investigate whether an interaction could be detected between pp14 and chCREB3 in the context of MDV infection. Primary chicken embryo fibroblast (CEF) cells that are permissive for MDV replication were used (Deng et al., 2010*)* because an efficient *in vitro* infection system for the primary target cells of MDV (B or T cells) was unavailable at the start of this project (Schermuly et al., 2015). Endogenous chCREB3 was not detected in CEF cells (Figure 1.9), and therefore as a strategy to introduce

chCREB3 in infected cells, the bacterial artificial chromosome (BAC) clone of the highly oncogenic RB-1B strain of MDV (pRB-1B) was engineered and the gene encoding chCREB3 with a N-terminal hemagglutinin (HA) epitope tag [denoted HA (N-ter) chCREB3] was inserted into the BAC, generating a BAC clone denoted pRB-1B/HA-chCREB3 (explained in detail in Section 2.6 of Chapter 2). By infecting cells with this engineered MDV virus encoding both the virally encoded pp14 and introduced chCREB3, the interaction was investigated in the context of MDV infection using immunofluorescence. Colocalisation of pp14 and chCREB3 was observed in infected cells (Figure 1.9), indicating a potential interaction between pp14 and chCREB3. These results formed the basis of this DPhil project aimed at investigating this putative interaction between pp14 and chCREB3 with a view to determining whether this interaction could be linked to the neurovirulence of MDV.

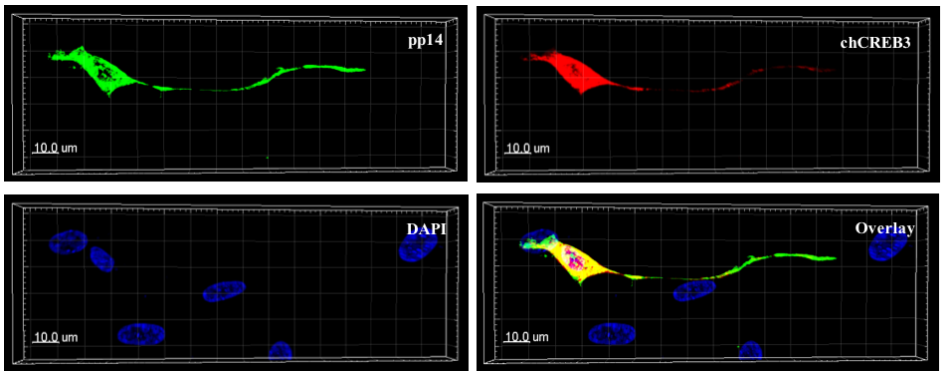


Figure 1.9 Colocalisation of pp14 and chCREB3 in MDV infected cells. Immunofluorescence was performed with RB-1B/HA-chCREB3 infected CEF cells. pp14 was stained with an anti-pp14 antibody (green). The overexpressed recombinant chCREB3 was stained with an anti-chCREB3 antibody (red) raised against the C-terminal domain of chCREB3 and the nuclei were stained with DAPI (blue). The colocalisation is shown as yellow in the merged picture. Endogenous chCREB3 was not detected in uninfected CEF cells. (Unpublished data obtained from Dr Tahiri-Alaoui).

1.5 cAMP response element binding protein, CREB3

1.5.1 CREB3 subfamily of bZIP transcription factors

CREB3 belongs to CREB3 subfamily of transcription factors (Chan et al., 2011), which is a branch of the CREB/ATF family within the bZIP superfamily (Lonze and Ginty, 2002). The CREB3 subfamily is highly conserved with orthologues identified in species ranging from sponges to humans (Fox and Andrew, 2015). Our current knowledge on CREB3 subfamily transcription factors comes from studies on members of the mammalian CREB3 subfamily. In mammals, the CREB3 subfamily of bZIP transcription factors consists of 5 members (Figure 1.9A): CREB3 (also known as LZIP or Luman), CREB3-Like Protein 1 (CREB3L1 or OASIS), CREB3-Like Protein 2 (CREB3L2 or BBF2H7), CREB3-Like Protein 3 (CREB3L3 or CREB-H) and CREB3-Like Protein 4 (CREB3L4 or AIbZIP) (Chan et al., 2011). The CREB3 subfamily members are closely related to *Drosophila* dCREB-A/BBF2 (Chan et al., 2011).

Members of the CREB3 subfamily exhibit a high degree of similarity, especially within the bZIP domain (Chan et al., 2011), and possess unique features conserved within the family (Barbosa et al., 2013). They are distinguished from all other bZIP transcription factors by a conserved domain of ~30 amino acids, termed the ATB (Adjacent to bZIP) domain adjacent to the bZIP DNA binding domain (Fox and Andrew, 2015). CREB3 proteins are structurally similar to ATF6 (Chan et al., 2011), but unlike ATF6 the ER localisation of all CREB3 members is determined by a cytosolic domain termed the ER retention motif (ERM) (Barbosa et al., 2013). CREB3 subfamily transcription factors are thought to recognise c-AMP response

element, CRE (TGACGTCA), box-B element (TACACGTAATC), endoplasmic reticulum (ER) stress responsive element II (ERSE-II; ATTGG-N-CCACG) and unfolded protein response (UPR) element (UPRE; TGACGTGG) (Chan et al., 2011).

CREB3 transcription factors exhibit quite different expression profiles in different tissues. CREB3 is ubiquitous; CREB3L1 is more abundantly expressed in some tissues including pancreas, prostate and bone; CREB3L2 can be detected in various tissues and the strongest expression occurs in placenta, lung, spleen, intestine and cartilage; CREB3L3 has a liver specific expression; and CREB3L4 is expressed in different organs such as pancreas, liver, and gonads, yet it is most abundant in prostate epithelial cells (Chan et al., 2011).

Similar to ATF6, CREB3 transcription factors are type II membrane-associated proteins with the N-terminal domain in the cytoplasm, the transmembrane segment spanning the ER membrane and the C-terminal domain in the ER lumen (Chan et al., 2011). The ER bound CREB3 subfamily transcription factors exist in an inactive form and are thought to be activated through regulated intramembrane proteolysis (RIP) (Chan et al., 2011). Upon activation, they are transported from the ER to the Golgi apparatus whether they are proteolytically cleaved by site 1 protease (S1P) and site 2 protease (S2P) sequentially to release the transcriptionally active N-terminal fragment, which is translocated into the nucleus to activate transcription of genes (Figure 1.9B) (Chan et al., 2011).

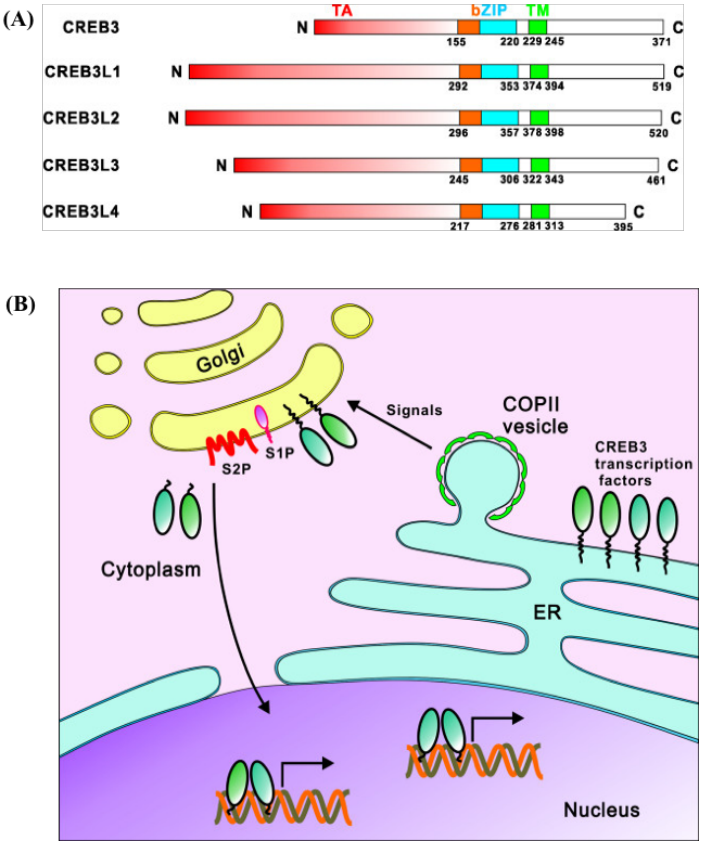


Figure 1.10 CREB3 subfamily of bZIP transcription factors.
(A) Diagrammatic representation of transcription factors of the human CREB3 subfamily showing the transcriptional activation (TA) domain, bZIP domain, the transmembrane (TM) domain**. (B)** CREB3 transcription factors are anchored to the ER membrane in an inactive form. Upon activation, CREB3 is translocated from the ER to the Golgi, where it is cleaved sequentially by S1P (Site-1-protease) and S2P (Site-2-protease), which liberates the N-terminal domain of CREB3 from the membrane allowing it to enter the nucleus where it activates transcription of target genes. (Picture taken from Chan et al., 2011)

CREB3 subfamily factors are thought to be activated in response to ER stress, and play important roles in unfolded protein response (UPR) (Chan et al., 2011). The UPR is a cellular stress response to accumulation of unfolded or misfolded protein in the lumen of the ER (Oslowski and Urano, 2011). The UPR aims to resolve ER stress by implementing one or more of the following: halting protein translation, degrading misfolded proteins through ER-associated degradation (ERAD), or activating pathways that promote production of molecular chaperones involved in

protein folding (Oslowski and Urano, 2011). When stress conditions cannot be resolved, the UPR promotes apoptosis (Oslowski and Urano, 2011).

Increasing data suggest viruses use the host UPR in the establishment of acute, chronic and latent infections (Carpenter and Grose, 2014), in particular viruses such as herpesviruses that cause latent infections. Herpesviruses express molecules that mimic UPR molecules to modify the UPR to help set up lytic infection and to break latency (Chan, 2014). VZV has been shown to manipulate the UPR to expand the ER to facilitate viral glycoprotein synthesis (Carpenter and Grose, 2014). MDV pathotypes, the vvMDV (RB1B and MD5) and vv+ MDV (TK), have also been shown to upregulate UPR gene expression during lytic replication (Neerukonda, 2015). In MDV TK strain-induced tumours, anti-apoptotic mechanisms triggered by UPR were found to be induced (Neerukonda, 2015). Tumour cells are subjected to severe conditions that are capable of inducing ER stress and UPR activation.

Not limited to ER stress, CREB3 subfamily transcription factors are activated by diverse physiological stimuli (Chan et al., 2011) although direct inducers of proteolytic activation of discrete CREB3 subfamily transcription factors remain to be identified. Whether ER stress induces proteolytic activation of CREB3 subfamily transcription factors remains controversial as contradictory reports have been published. While studies have reported proteolytic activation of CREB3 transcription factors by UPR induced by pharmaceutical agents such as tunicamycin or thapsigargin, several other studies found no activation of CREB3, CREB3L3 and CREB3L4 upon ER stress (Chan et al., 2011).

CREB3 transcription factors exhibit a diverse spectrum of biological functions beyond ER stress and UPR. CREB3L1, CREB3L2, CREB3L3 and CREB3L4 have been found to activate distinct subsets of target genes by microarray analyses (Chan et al., 2011). The factors have also been found to exhibit cell type-specific patterns in the activation of their target genes and have been reported to play critical roles in regulating development, metabolism, protein secretion, survival and oncogenesis (Chan et al., 2011). A summary of functions related to CREB3L1-4 is given in Table 1.2 and the known functions of CREB3 are reviewed in the next section.

Transcription factor	Functions
CREB3L1	▪ Plays a critical role in bone formation through the transcription of COL1A1, and possibly COL1A2, and the secretion of bone matrix proteins ▪ Required to protect astrocytes from ER stress-induced cell death ▪ Upon infection with diverse DNA and RNA viruses, inhibits cell-cycle progression by binding to promoters and activating transcription of genes encoding cell-cycle inhibitors
CREB3L2	▪ Activates the Sec23a-dependent pathway required for secretion of collagens and other extracellular matrix proteins during normal ▪ Protects cells from ER stress-induced death in a neuroblastoma cell line
CREB3L3	▪ May activate the expression of acute phase response (APR) gene in acute inflammatory response ▪ May be involved in growth suppression
CREB3L4	▪ Highly tissue-specific expression profile, androgen regulation, chromosomal localisation, and expression profile of CREB3L4 in prostate tumours suggest that CREB3L4 may play an important role in prostate cancer and in androgen receptor signalling in prostate cells

Table 1.2 Known functions of members of the CREB3 subfamily.
(Denard et al., 2011; Ishikura-Kinoshita et al., 2012; Kondo et al., 2005, 2007; Murakami et al., 2009; Nakagawa et al., 2014; Qi et al., 2002; Zhang et al., 2006)6)

1.5.2 Biological functions of CREB3

The current knowledge on CREB3 comes from studies on the mammalian CREB3 transcription factor. CREB3 is the primary member of the mammalian CREB3 subfamily of CRE-binding transcription factors (Figure 1.11) (Chan et al., 2011). Although CREB3 transcripts are detected quite broadly, the protein has been

observed only in the trigeminal ganglion neurons, monocytes and bone marrow dendritic cells (Fox and Andrew, 2015). In THP-1 monocytic cells, an isoform of human CREB3 lacking the putative transmembrane region has been reported (Kang et al., 2009), which was shown to suppress the transcription activation of target genes of glucocorticoid receptor (Kang et al., 2009). Two isoforms of mouse CREB3, LZIP-1 and LZIP-2, have been reported, with the N-terminal region of LZIP-2 having an additional 25 amino acids (Burbelo et al., 1994). As described below, a range of stimuli can activate CREB3 showing that its role is not limited to ER stress (Chan et al., 2011).

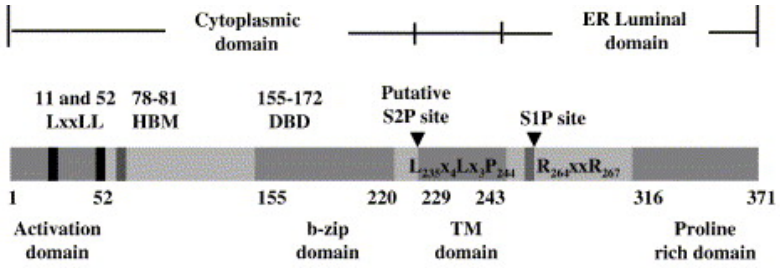


Figure 1.11 Schematic diagram showing the putative domains of huCREB3. The positions of the transcriptional activation domain (residues 1 to 52), the bZIP domain (residues 155 to 220), the DNA-binding domain (DBD, residues 155 to 172), the transmembrane domain (residues 229 to 243), and the proline-rich domain (residues 316 to 371) are shown. HCF-binding motif (HBM), S1P recognition motif ($R_{264}XXR_{267}$), putative S2P site ($L_{235}X_4LX_3P_{244}$), putative nuclear receptor (NR) box (LxxLL) are also shown. (Taken from Blot et al., 2006)

A few studies support a role for CREB3 in the ER stress response, especially in the ERAD process. In one study, CREB3 was reported to induce the ERAD-related Herp (homocysteine-induced endoplasmic reticulum protein or Mif1) via transactivation of an ER stress response element (ERSE)-II in the Herp promoter (Liang et al., 2006). In another study, CREB3 was shown to induce another ERAD protein, EDEM (ER degradation-enhancing alpha-mannosidase-like protein), by binding to and activating transcription from the unfolded protein response element (UPRE) (DenBoer et al., 2005). In addition, Brefeldin A (BFA), a very common

inducer of the UPR or ER stress response, has also been shown to induce CREB3 expression and its proteolytic cleavage (Reiling et al., 2013). In this study, CREB3 was shown to transcriptionally upregulate ARF4 (ADP ribosylation factor 4) in response to BFA. CREB3-ARF4 signalling has recently been implicated in the Golgi stress response (Chan et al., 2011).

CREB3 is a cellular target of several viruses. CREB3 was first identified as an interacting partner of the transcriptional co-activator host cell factor (HCF)-1 (Lu et al., 1997). CREB3 shares a consensus HCF1 binding motif with VP16, the HSV IE gene expression activator protein, and interacts with the same region of HCF1. This viral mimicry led to the hypothesis that CREB3 may have a role in cell proliferation or certain cellular stress response, which might be exploited by HSV to establish latency and/or reactivation (Lu and Misra, 2000). It has been suggested CREB3 plays a role in the establishment of HSV latency by sequestering HCF1 in the ER membrane of sensory neurons (Lu and Misra, 2000). CREB3 may also play a role in HSV reactivation, as it could activate the promoters of IE110 and LAT, two genes that facilitate reactivation of HSV1 from latency (Lu and Misra, 2000). In HCV infection, the HCV core protein inhibits CREB3 transcriptional activity, by sequestering it in the cytoplasm, to facilitate cellular transformation (Jin et al., 2000). This study, in which the HCV core targets CREB3 to induce cellular transformation, suggested a cellular tumour suppressor role for CREB3 (Jin et al., 2000). CREB3 is also implicated in human immunodeficiency virus (HIV) gene regulation. The activated nuclear form of CREB3 interacts with Tat and inhibits Tat-mediated HIV1 transactivation, which acts negatively on Tat-mediated HIV1 transactivation of the HIV1 long terminal repeat (Blot et al., 2006). Another major

HIV1 protein, the TMgp41 subunit of HIV1 envelope (Env) protein, interacts with full-length ER-bound CREB3 to destabilise it.

CREB3 may play a role in cell migration. CREB3 has been shown to interact with CC chemokine receptor (CCR) 1 and induce leukotactin-1 (Lkn-1)-dependent leukocyte cell migration, through activation of the nuclear factor (NF)-κB pathway (Jang et al., 2007). CREB3 has also been shown to enhance chemokine (C-C motif) ligand 2 (CCL2)-induced cell migration by inducing monocyte CC chemokine receptor 2 (Sung et al., 2008).

CREB3 is expressed in dendritic cells (DC), the professional antigen-presenting cells of the immune system. CREB3 is proteolytically activated upon DC maturation induced by different DC maturation stimuli (Eleveld-Trancikova et al., 2010). This study implicated a role for CREB3 in early immune responses. The activation of CREB3 in dendritic cells is believed to involve DC specific transmembrane protein (DC-STAMP) and OS-9 (Endoplasmic Reticulum Lectin), both of which interact with CREB3 at the cytosolic side of the ER (Eleveld-Trancikova et al., 2010). In another study, CREB3 was identified as a key factor that is involved in osteoclastogenesis because of its effect on DC-STAMP expression, stability and localisation (Kanemoto et al., 2015).

CREB3 has been implicated in cancer development. Histone deacetylase 3 (HDAC3) is another interacting partner of CREB3 (Kim et al., 2010), which through its interaction with CREB3 inhibits CREB3-mediated transcriptional activation and migration of metastatic breast cancer cells (Kim et al., 2010).

1.6 bZIP transcription factors

bZIP transcription factors, which have been structurally classified by their DNA-binding domain, bZIP, represent one of the largest families of transcription factors in eukaryotic cells (Miotto and Struhl, 2006). The conserved bZIP domain, which is 60 to 80 amino acids in length, consists of a DNA binding basic region and a leucine zipper dimerisation region (Figure 1.12A) (Zhang et al., 2013). bZIP transcription factors have been subdivided into classes of proteins according to their DNA-binding specificity or dimerisation properties (Miotto and Struhl, 2006).

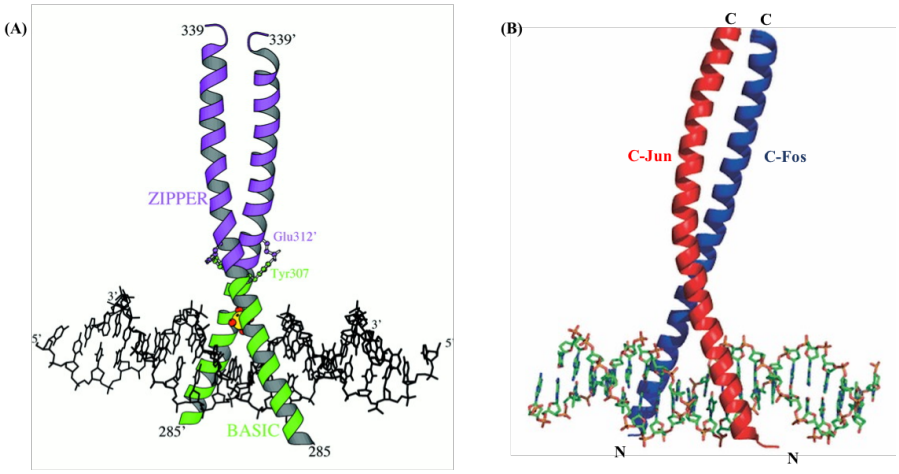


Figure 1.12 Crystal structures of bZIP domains bound to DNA.
(A) Crystal structure of the homodimeric bZIP of CREB bound to a DNA oligomer containing the CRE site (TGACGTCA). The basic region residues are coloured green and leucine zipper residues are coloured purple. **(B)** Crystal structure of the heterodimeric bZIP of c-Fos:c-Jun bound to a DNA oligomer containing the AP1 site (TGACTCA). (PDB codes: 1DH3 and 1FOS; References: Schumacher et al., 2000, König and Richmond, 1993 and Worrall and Mason, 2011)

bZIP family members form dimers via their characteristic leucine zipper helices, which consist of four to five heptad (seven-residue) repeats of amino acids (Vinson et al., 2002), that form a parallel coiled-coil structure (Figure 1.13) (Hakoshima, 2014). The seven positions of a heptad are labelled (**a-b-c-d-e-f-g**) in one helix and

(**a'-b'-c'-d'-e'-f '-g'**) in the other. Amino acids in the **a** and **d** positions project from the same face of the α-helix (Figure 1.13) (Vinson et al., 2002). Hydrophobic residues occupy the **a** and **d** positions preferentially, with the **a** positions occupied by a variety of amino acids and the **d** positions typically occupied by leucine (Vinson et al., 2002). The occurrence of leucine at every **d** or 7th position is characteristic of the ZIP motif (Figure 1.13) (Vinson et al., 2002). The amino acids in the **a** and **d** positions of one monomer and the amino acids at the complementary **a′** and **d′** positions of the opposite monomer interact with each other (Figure 1.13) creating a hydrophobic core, which is essential for dimer stability (Thompson et al., 1993; Vinson et al., 2002).

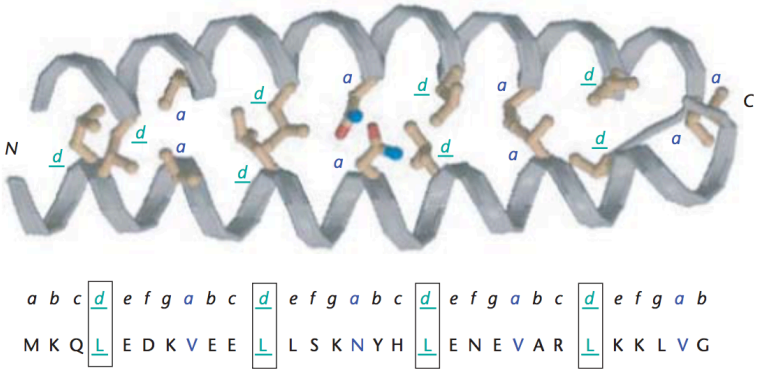


Figure 1.13 Parallel coiled-coil structure of GCN4 ZIP homodimer.
The main chains of the two peptide chains are represented as ribbons in grey. The side chains participating in the dimer association are represented as stick models. The positions of the heptad repeat are labelled a,b,c,d,e,f,g. The leucines at position **d** are boxed and highlighted in green and underlined. The residues at position **a** are highlighted in blue. A polar asparagine residue located in the middle of the ZIP motif at position **a** promotes homodimerisation. (PDB code 1gd2; Hakoshima, T., et al, 2014)

bZIP proteins can form both homodimers and heterodimers (Figure 1.12) (Hakoshima, 2014). However, only specific bZIP protein pairs can form heterodimers, and some bZIP proteins do not form homodimers, but only form heterodimers (Hakoshima, 2014). For example, Fos preferentially forms a stable heterodimer with Jun, while the homodimer of Fos is extremely unstable

(Hakoshima, 2014). The dimerisation specificity of the ZIP motifs is primarily determined by interactions between polar residues at the **g** and **e'** (**g↔e′**) and **e** and **g'** (**e↔g'**) positions of the interacting helices (Figure 1.14) (Cohen and Parry, 1990; Vinson et al., 1993, 2002). These positions are typically occupied by charged amino acids. Crystal structures of bZIP proteins reveal **g↔e′** interhelical interactions between charged amino acids in the **g** and **e′** positions can be either attractive or repulsive and thus can regulate both homodimerisation and heterodimerisation. In addition to **g↔e′** pairs, exceptionally polar residues at the **a** and **d** positions regulate dimerisation specificity (Hakoshima, 2014; Vinson et al., 2002). For example, asparagine in the **a** position has been found to dimerise preferentially with another asparagine in **a** position, promoting homodimerisation between bZIP proteins, while basic amino acids in the **a** position are repulsive and have thus been found to drive heterodimerisation (Vinson et al., 2002; Zeng et al., 1997). Similar to residues in the **a** position, a few residues in the **d** position are exceptionally charged or polar and contribute to dimerisation specificity (Hakoshima, 2014; Vinson et al., 2002).

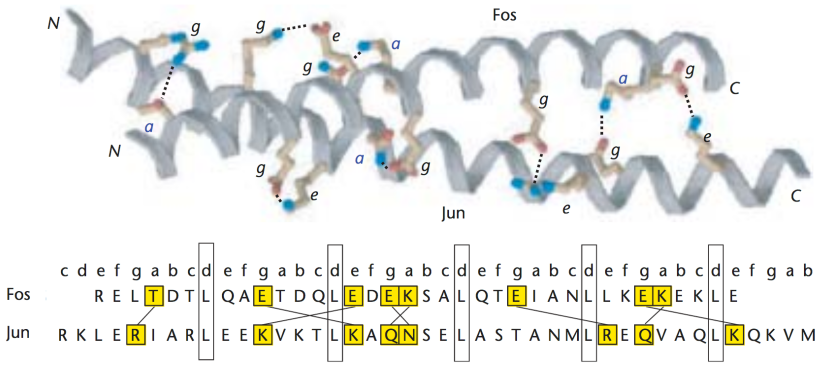


Figure 1.14 Interchain polar interactions found in the coiled-coil structure of the Fos–Jun ZIP heterodimer.
Hydrogen bonding interactions are indicated by dotted lines. The dimerisation specificity of the ZIP motifs is primarily determined by interactions between polar residues at the **g** and **e** positions. A few polar residues at the **a** positions also contribute to the specificity. (PDB code: 1FOS; Hakoshima, T., et al, 2014.)

Transcription factors of the bZIP superfamily bind to short palindromic or pseudo palindromic target DNA sequences in a sequence specific manner (Hakoshima, 2014). Many bZIP proteins share partially overlapping sites, and are capable of binding two distinct sites (Jolma et al., 2013). The DNA-binding region, N-terminus to the leucine zipper, is rich in basic residues (hence called the basic region). It is composed of ~16 amino acid residues consisting of an invariant Asn-x7-Arg/Lys motif (Zhang et al., 2013). Generally, the basic region of bZIP proteins is largely disordered, which upon binding with the DNA ligand adopts a helical conformation with a surface that is complementary to the target DNA sequence (Miller et al., 2003; O'Neil et al., 1991; Patel et al., 1990; Shuman et al., 1990). bZIP proteins bind to DNA like a pair of forceps with the basic regions interacting with the major groove of the DNA (Figure 1.12) (Glover and Harrison, 1995; König and Richmond, 1993; Montclare et al., 2001; Schumacher et al., 2000).

1.7 Aim and objectives

The current knowledge on viral and host cell factors and the underlying mechanisms that contribute to MDV neurovirulence is limited. While pp14 is the only viral factor that has been associated with MDV neurovirulence to date, its mechanisms of action are not known. In this regard, identifying host cell factors that interact with pp14, and understanding the role and implications of these virus–host interactions on MD will help us gain a better understanding of the mechanisms involved in MDV neuropathology. Using reverse genetics, it was shown pp14 is necessary for successful lymphocytic infiltration and manifestation of lesions in the nerves of infected chickens (Tahiri-Alaoui et al., 2012). These peripheral lesions are associated with the clinical paralysis observed in the classical form of

the disease (Payne and Biggs, 1967); however, the precise cellular and molecular mechanisms that mediate cellular infiltration into nerves are still not understood. Hence, unravelling the mechanisms by which pp14 mediates neurovirulence in chickens will shed light on our understanding of MDV neuropathogenesis.

Yeast-two-hybrid screening identified the ER-bound transcription regulator, human CREB3, as an interacting partner of pp14. Bioinformatic analyses revealed chicken CREB3 and human CREB3 share a conserved domain organisation, and that the sequences share 45% amino acid sequence identity, with a high degree of similarity (73%) in the conserved bZIP domain. Subsequently, pp14 was shown to colocalise with recombinant chicken CREB3 in MDV infected CEF cells indicating a potential interaction between pp14 and chicken CREB3. These results led to the hypothesis that pp14 might be interacting with chicken CREB3, and the interaction could be linked to the neurovirulence of MDV. The primary goal of this DPhil research project was to further investigate this putative interaction between pp14 and chicken CREB3 using *in vitro* studies, with a view to determining whether the interaction could be linked to the neurovirulence of MDV.

In addition, other objectives of the project were to gain structural and functional insights into the MDV neurovirulence factor pp14 and the bZIP transcription factor chCREB3. No crystallographic structures were available for CREB3 although it is an important, multifunctional protein, which has been shown to play a role in the ER-stress induced unfolded protein response (UPR) (Chan et al., 2011), cell proliferation and migration (Jang et al., 2007; Kim et al., 2010), tumour suppression (Jin et al., 2000), inflammatory gene expression (Eleveld-Trancikova

et al., 2010; Sung et al., 2008), and has been shown to be a target of viruses such as HSV, HCV and HIV (Blot et al., 2006; Jin et al., 2000; Lu and Misra, 2000). The knowledge gained from the structural studies of these proteins may one day be exploited to develop new interventions against disease processes.

2 Materials and Methods

2.1 Amino acid sequences

Protein/MW	Amino acid sequence
pp14B, 10,254Da	10 20 30 40 50 60 MINTTAERNP SQLSIDTALG GNDPGCIDCG PTFHLETDTA TTRNGTSSFG DRIRSFCERA 70 80 90 RSLISNFVTW RSRNESCEVL AEIPQEKEVE STL
Cys-null pp14B, 10,254Da	10 20 30 40 50 60 MINTTAERNP SQLSIDTALG GNDPG**A**ID**A**G PTFHLETDTA TTRNGTSSFG DRIRSF**A**ERA 70 80 90 RSLISNFVTW RSRNES**A**EVL AEIPQEKEVE STL
Chicken CREB3 (full-length), 48,745Da	10 20 30 40 50 60 MAGPEPLAAL PDGDLLDFLL NDDAPAAGSP AGDGPLLGDW VLPELELLDK EMDSFISSLL 70 80 90 100 110 120 SPSEDEPCLL QGCLPADSDS STSEDQNPFH SSGSDVASSP QSSDIVQFDH NYSLHQDWPA 130 140 150 160 170 180 LESVRSDMAE GDVSIDLETW MCLEGTNETL QESCSFPVTV AVDAGPQLMP GAPMQSNFPE 190 200 210 220 230 240 LVLTEEERQL LEKDGVSLPT CLPLTKAEER LLKKVRRKIR NKQSAQDSRR RKKIYVDGLE 250 260 270 280 290 300 NRVAACTAQN HELQKKVQLL QKQNMSLLEQ LRKLQALVRQ STTKTTTAST CIMVLVLSFC 310 320 330 340 350 360 LILSPSLYSF GSRGPQPEFR VLSRQIREFP NKVWQAAPAV QEEVVLERLS PEPEDTSLGS 370 380 390 400 410 420 LNQSREEGQS PPKPDPRSAF NGNSSSDPPA GSELGHPQPQ EQQSRRDALH SEVLAPWKDE 430 440 NQEWVERTTS VVIQQHRTDE M
NT-chCREB3 (1-288), 31,600Da	10 20 30 40 50 60 MAGPEPLAAL PDGDLLDFLL NDDAPAAGSP AGDGPLLGDW VLPELELLDK EMDSFISSLL 70 80 90 100 110 120 SPSEDEPCLL QGCLPADSDS STSEDQNPFH SSGSDVASSP QSSDIVQFDH NYSLHQDWPA 130 140 150 160 170 180 LESVRSDMAE GDVSIDLETW MCLEGTNETL QESCSFPVTV AVDAGPQLMP GAPMQSNFPE 190 200 210 220 230 240 LVLTEEERQL LEKDGVSLPT CLPLTKAEER LLKKVRRKIR NKQSAQDSRR RKKIYVDGLE 250 260 270 280 NRVAACTAQN HELQKKVQLL QKQNMSLLEQ LRKLQALVRQ STTKTTTA
NT-chCREB3-TM (1-314), 34,376Da	10 20 30 40 50 60 MAGPEPLAAL PDGDLLDFLL NDDAPAAGSP AGDGPLLGDW VLPELELLDK EMDSFISSLL 70 80 90 100 110 120 SPSEDEPCLL QGCLPADSDS STSEDQNPFH SSGSDVASSP QSSDIVQFDH NYSLHQDWPA 130 140 150 160 170 180 LESVRSDMAE GDVSIDLETW MCLEGTNETL QESCSFPVTV AVDAGPQLMP GAPMQSNFPE 190 200 210 220 230 240 LVLTEEERQL LEKDGVSLPT CLPLTKAEER LLKKVRRKIR NKQSAQDSRR RKKIYVDGLE 250 260 270 280 290 300 NRVAACTAQN HELQKKVQLL QKQNMSLLEQ LRKLQALVRQ STTKTTTAST CIMVLVLSFC 310 LILSPSLYSF GSRG

CT-chCREB3 (313-441), 14,599Da	322 332 342 352 362 372 RGPQPEFRVL SRQIREFPNK VWQAAPAVQE EVVLERLSPE PEDTSLGSLN QSREEGQSPP 382 392 402 412 422 432 KPDPRSAFNG NSSSDPPAGS ELGHPQPQEQ QSRRDALHSE VLAPWKDENQ EWVERTTSVV IQQHRTDEM
chCREB3 bZIP, (206-280), 8950Da	215 225 235 245 255 265 KAEERLLKKV RRKIRNKQSA QDSRRRKKIY VDGLENRVAA CTAQNHELQK KVQLLQKQNM 275 SLLEQLRKLQ ALVRQ

Table 2.1 Amino acid sequences of proteins used in the study.

2.2 Plasmids

Vector	Source
pOPIN vectors	OPPF
pcDNA3.1(+)	Invitrogen
pcDNA3-mClover	Dr Sergi Padilla-Parra (Oxford University)
pcDNA3-mRuby	Dr Sergi Padilla-Parra (Oxford University)
pcDNA3.1-Untagged pp14B	Dr Tahiri-Alaoui (Pirbright Institute)
pcDNA3.1-HA (n-terminal) NT-chCREB3	Dr Tahiri-Alaoui (Pirbright Institute)
pcDNA3.1-HA (n-terminal) CT-chCREB3	Dr Tahiri-Alaoui (Pirbright Institute)
pp14B cloned into various pOPIN vectors	OPPF
NT-chCREB3 cloned into various pOPIN vectors	OPPF
CT-chCREB3 cloned into various pOPIN vectors	OPPF

Table 2.2 Name and source of vectors used in this study

2.3 Vector maps in Appendix A

2.4 Antibodies

The polyclonal antibodies against MDV1 pp14, chCREB3 and MDV1 RLORF9 were raised in rabbit against synthetic peptides EVLAEIPQEKEVESTL, CEVLAPWKDENQEWVERT, and VLAPWKDENQEWVERT, respectively, and were purified by affinity chromatography with immobilised peptides (PickCell Laboratories BV, The Netherlands). Details and dilutions of all antibodies used in this study are given in Table 2.3.

Antibody	Antigen	Species	Conjugate	Source	Isotype	Dilution
α-pp14	pp14	Rabbit	-	PickCell Laboratories BV	-	WB 1:2000 IF1: 2000
α -chCREB3	chCREB3	Rabbit	-	PickCell Laboratories BV	-	WB 1:2000
α-RLORF9	RLORF9	Rabbit	-	PickCell Laboratories BV	-	-
BD1	MDV-pp38	Rabbit	-	Pirbright	IgG2A	IF 1:1000
α-HA	Influenza-hemagglutinin	Rabbit	-	Covance	IgG	IF 1:1000
		Mouse	-	Covance	IgG	IF 1:1000
		Mouse	HRP	Cell signalling	IgG1	WB 1:2000
α-flag	FLAG	Mouse	-	Sigma	IgG1	-
		Rat	HRP	Sigma	IgG1	WB 1:2000
α-His	C-terminal His	Mouse	-	Sigma	IgG1	WB 1:2000
	C-terminal His	Mouse	HRP	Invitrogen	IgG2b	WB 1:5000
	N-terminal His	Mouse	-	Sigma	IgG2a	WB 1:2000
α -GST	GST	Mouse	HRP	Sigma	IgG1	WB 1:2000
Normal mouse IgG	-	Mouse	-	Sigma	IgG	-
α- rabbit IgG (secondary)	-	Goat	HRP	Sigma	IgG (H+L)	WB 1:2000
α- mouse IgG (secondary)	-	Rabbit	HRP	Sigma	IgG (H+L)	WB 1:2000
α- rabbit IgG (secondary)	-	Goat	Alexa488	Molecular Probes	IgG (H+L)	IF 1:500
		Goat	Alexa568	Molecular Probes	IgG (H+L)	IF 1:500
α- mouse IgG (secondary)	-	Goat	Alexa488	Molecular Probes	IgG (H+L)	IF 1:500
		Goat	Alexa568	Molecular Probes	IgG (H+L)	IF 1:500

Table 2.3 Details and dilutions of primary and secondary antibodies used in this study.

2.5 Antibiotics

Antibiotic	Antibiotic concentration
Ampicillin	100μg/mL
Kanamycin	50μg/mL
Chloramphenicol	34μg/mL
Penicillin	100Units/ml
Streptomycin	100μg/mL

Table 2.4 Details of antibiotics used in this study.

2.6 Viruses

The genetically engineered RB-1B BAC clones, pRB-1B/HA-chCREB3 and pRB-1BΔpp14/HA-chCREB3, were obtained from Benedikt Kaufer (Free University of Berlin). The deletion of the two copies of pp14 gene from the pRB-1B BAC, and the insertion of foreign gene encoding HA tagged chCREB3 into the mini-F region of the pRB-1B BAC (Figure 2.1) were carried by recombination-mediated genetic engineering (Tischer BK et al., 2006; Yu et al., 2000).

BAC is a prokaryotic cloning vector derived from a single copy or low-copy-number minimal fertility factor replicon (mini-F), derivative of the fertility factor of *Escherichia coli* (*E. coli)* (Petherbridge et al., 2003), that can stably maintain a large DNA insert (average size 150-300 kb) and be propagated in *E. coli* (Osterrieder et al., 2006). Maintenance and faithful replication of the BAC construct in *E. coli* are facilitated by components encoded in the mini-F, the backbone of the BAC vector (Tischer and Kaufer, 2012). BACs allow straightforward genetic manipulation of the genomes in *E. coli* (Osterrieder et al., 2006) and reconstitution of recombinant viruses can be accomplished by transfection of purified BAC DNA into susceptible cells (Tischer and Kaufer, 2012). Viruses derived from BAC clones are stable after *in vitro* and *in vivo* passages and show characteristics and growth kinetics similar to those of the parental virus (Petherbridge et al., 2003, 2004). The genomes of four MDV strains (three avirulent vaccine strains and one vv strain, RB-1B) have been cloned as BACs (Osterrieder et al., 2006).

The BAC clone of the highly oncogenic RB-1B strain (Figure 2.1) was generated by replacing MDV094 (US2) of RB-1B, which is dispensable for virus growth *in vivo* and tumorigenesis, with the mini-F plasmid pHA1 (Osterrieder et al., 2006; Petherbridge et al., 2004).

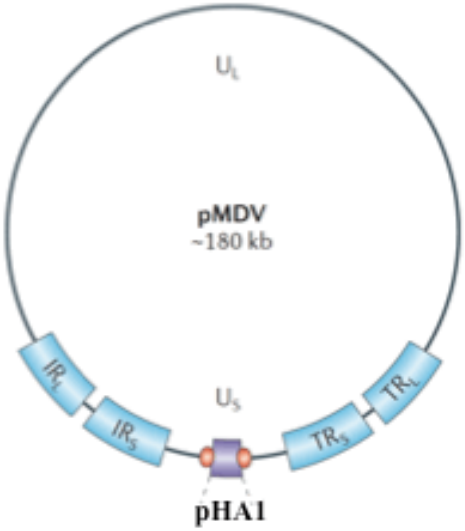


Figure 2.1 Schematic diagram of the BAC of MDV.
The pRB-1B BAC contains the minimal fertility factor replicon pHA1 instead of MDV094 (US2), which is dispensable for virus growth *in vivo* and tumorigenesis. The maintenance and faithful replication of the BAC construct in *E. coli* are facilitated by components encoded in the mini-F. (Picture taken from Osterrieder et al., 2006)

2.7 MDV gene nomenclature

MDV genes that are homologous to all of the HSV1 genes are named after the UL genes of HSV. MDV-specific ORFs in the UL region are referred to as LORFs, those in the long repeat as R-LORFs and those in the US region are referred to as SORFs (Brunovskis and Velicer, 1995; Lee et al., 2000).

2.8 Cell lines and primary cells

2.8.1 DF1 cells

DF1 is a continuous cell line of chicken embryo fibroblasts (Foster, 1998). DF1 cells were cultured in Dulbecco's modified Eagles' medium (DMEM) (Invitrogen) supplemented with 10% fetal calf serum (FCS) (PAA), 1% L-glutamine (Gibco, Life Technologies) and 1% non-essential amino acids (Gibco, Life Technologies). Cells were grown and maintained at 37°C in 5% CO_2.

2.8.2 HEK293T (human embryonic kidney cell line)

The 293T cell line is a permanent line derived from primary embryonic human kidney (HEK) (Graham et al., 1977). HEK293T cells were maintained in DMEM supplemented with 10% FCS, 1% L-glutamine and 1% non-essential amino acids. Cells were grown and maintained at 37°C in 5% CO_2.

2.8.3 Insect cells

The insect lines Sf9 or Sf21 (*Spodoptera frugiperda*) were used for expression of recombinant proteins using the baculovirus system. Sf21 cells (IPLB-Sf21-AE) are derived from the ovarian tissue of the Fall Armyworm, *Spodoptera frugiperda* (Vaughn et al., 1977). The Sf9 insect cell line is a clonal isolate derived from the parental *Spodoptera frugiperda* cell line IPLB-Sf-21-AE (Pasumarthy and Murhammer, 1994). Sf21 and Sf9 cells were grown and maintained in Sf-900 II SFM medium (Thermofisher) supplemented with 3.5% heat inactivated FCS, and 1% penicillin and streptomycin (Gibco, Life Technologies). Cells were grown and maintained in suspension at 28°C with shaking at 150rpm (revolutions per minute).

2.8.4 Chicken embryo fibroblast (CEF) cells

Primary chicken embryo fibroblasts (CEFs) were prepared from embryos of 10-day-old specific-pathogen-free (SPF) line 0 chickens (Payne et al., 1991). Primary cultures were grown in CEF growth medium for 24 hours. A volume of 104.9ml of CEF growth medium was prepared by mixing 76ml Milli-Q (ultrapure water), 10ml 10x 199 medium (Sigma), 10ml tryptose phosphate broth (Pirbright Institute), 5ml heat inactivated FCS (Sigma), 2.7ml of 7.5% sodium bicarbonate (Pirbright Institute), 1ml of 200mM L-glutamine (Sigma), 0.1ml 1000x penicillin and streptomycin (Sigma) and 0.1ml of 1mg/ml Fungizone (Sigma). After 24 hours, the cells were maintained in CEF maintenance medium. A volume of 101.9ml of CEF maintenance medium was prepared by mixing 76ml Milli-Q, 10ml 10x 199 medium, 10ml tryptose phosphate broth, 2ml heat inactivated newborn calf serum (Pirbright Institute), 2.7ml of 7.5% sodium bicarbonate, 1ml of 200mM L-glutamine, 0.1ml 1000x penicillin and streptomycin and 0.1ml of 1mg/ml Fungizone. Cells were grown and maintained at 37.5°C in 5% CO_2.

2.9 Molecular biology and microbiology techniques

2.9.1 Polymerase Chain Reaction

All PCR reactions were carried out using the Phusion® High-Fidelity DNA polymerase (New England Biolabs) provided as 2x master mix containing Phusion DNA Polymerase, nucleotides and optimised reaction buffer. The primers for PCR were obtained from Eurofins MWG Operon (Ebersberg, Germany). For In-Fusion molecular cloning into pOPIN expression vectors, primers were designed with the appropriate 5'- and 3'- extensions for the specific plasmid in question (Section

2.9.4). For cloning by restriction enzyme digestion and ligation, primers were designed with the appropriate restriction sites. The pp14 and chCREB3 constructs obtained from Dr Tahiri-Alaoui (Table 2.2) were used as template for PCR amplification of these two genes. The pcDNA3/mRuby vector obtained from Dr Sergi Padilla-Parra was used as template to PCR amplify the mRuby gene. PCR reactions to amplify DNA fragments were performed in a total volume of 50µL. Each reaction composed of 2x master mix, 10µM of each primer, and 10ng-50ng of template DNA made up to 50µl with water. To screen bacterial colonies after ligation, PCR reactions were performed in a total volume of 20µL. Each reaction composed of 2x master mix, and 10µM of each primer (one vector specific and one insert specific) made up to 20µl with water. Colonies were picked with a pipette tip and transferred to the PCR solution. All PCR reactions were carried out in PTC-225 Peltier Thermal Cycler (MJ Research) using Expand Long Template PCR System (Roche). A typical temperature cycle for PCR is shown in Table 2.5. The PCR products were analysed on a 1% agarose gel. For cloning, the amplified fragments were purified using the QIAquick Gel Extraction System (Qiagen).

Steps	Temperature	Duration	
Initial denaturation	95°C	1:30min	
Denaturation	95°C	30 sec	
Annealing	°C*	30 sec	20-30 cycles
Elongation	72°C	30 sec per kb	
Final Elongation	72°C	5 min	
Final hold	10°C	∞	

Table 2.5 Typical temperature cycle used for PCR.
*Optimal annealing temperature for primer pair

2.9.2 Agarose gel electrophoresis

1% (w/v) agarose gels were prepared in 1x TAE buffer (0.04M Tris-Acetate and 0.001M EDTA). SYBR Safe DNA gel stain (Invitrogen) was added for visualisation of DNA in agarose gels. The samples were mixed with 6x loading buffer (NEB) prior to loading onto gels alongside appropriate molecular weight markers (GeneRuler 1kb DNA ladder, Thermo Scientific). The gels were run at 100 volts in 1x TAE buffer. DNA bands were visualised using the Bio-Rad Gel-Chemi doc imaging system.

2.9.3 Gel purification of DNA

In order to purify DNA fragments after restriction digestion or PCR amplification, the DNA samples were electrophoresed on a 1% agarose gel stained with SYBR Safe. The bands were visualised by UV illumination, and excised using a scalpel blade, and purified using the QIAquick Gel Extraction System following the manufacturer's instructions.

2.9.4 In-Fusion ligation independent cloning

The protein constructs were cloned into pOPIN vectors, developed by the OPPF (Bird et al., 2014), for expression and purification. The advantage of using pOPIN expression vectors is that the majority of vectors of the pOPIN family allow expression of proteins from multiple hosts, i.e. a single vector capable of expression in *E. coli*, mammalian cell lines and insect cell lines, which enables parallel screening of constructs in multiple hosts (Berrow et al., 2009). A cleavable

histidine (His) tag has been incorporated into most of the pOPIN vectors (Berrow et al., 2009) to allow simple affinity purification of gene products.

Various constructs of pp14B, NT-chCREB3 and CT-chCREB3 cloned into different pOPIN vectors were obtained from the OPPF. However, in order to attach new epitope tags to pp14B and the NT-chCREB3, additional constructs were generated (Table 2.7). These were cloned into pOPIN vectors using the ligation independent In-Fusion® HD cloning system (Clontech). The vector maps of pOPIN vectors used for cloning are shown in Appendix A, Figure A 1. Information about each vector, and the forward and reverse primer extensions used for a given vector is given in Table 2.6. The sequences of forward and reverse primers used for PCR amplification are listed in Table 2.7. The pOPINE vector (encoding a C-terminal His tag) was used to clone pp14B or NT-chCREB3 with a C-terminal HA or FLAG tag (Table 2.7). To introduce these tags, the reverse PCR primer was designed to contain the tag of interest together with a stop codon to prevent inclusion of the C-terminal His tag encoded by the vector. To generate a cysteine-null (Cys-null) mutant of pp14B (Table 2.1) with a C-terminal His tag, a gBlock gene fragment of full-length pp14B in which all the cysteine codons were substituted with alanine codons (C26A, C29A, C57A and C77A) was ordered from Integrated DNA Technology (IDT) and cloned into pOPINE. This fragment was designed to also contain appropriate extensions for pOPINE for cloning.

Vector	Fusion Tag	Forward Primer Extension	Reverse Primer Extension
pOPIN E	C-terminal His tag --KHHHHHH	AGGAGATATACCATG	GTGATGGTGATGTTT
pOPIN F	N-terminal His tag MAHHHHHHSSGLEVLFQ*GP	AAGTTCTGTTTCAGGGCCCG	ATGGTCTAGAAAGCTTTA
pOPIN J	N-terminal His-GST MAHHHHHHSSG-GST- LEVLFQ*GP	AAGTTCTGTTTCAGGGCCCG	ATGGTCTAGAAAGCTTTA

Table 2.6 pOPIN vectors used for cloning.
List of pOPIN vectors used for cloning with details about the tag encoded by the vector and the forward and reverse primer extensions required for In-Fusion cloning. * represents 3C cleavage site

Gene	Primer sequence	Vector
FLAG (C-ter) pp14B	5'-AGG AGA TAT ACC ATG GTA AAT ACT ACG GCA GAA AGA AAC CC -3' (Fwd) 5'-GTG ATG GTG ATG TTT TCA TCA CTT GTC GTC ATC GTC TTT GTA GTC GAG GGT CGA TTC CAC TTC TTT CTC-3' (Rev)	pOPINE
His- GST(N-ter) pp14B	5'-AAG TTC TGT TTC AGG GCC CGG TAA ATA CTA CGG CAG AAA GAA ACC CTT CTC AG -3' (Fwd) 5'-ATG GTC TAG AAA GCT TTA TCA GAG GGT CGA TTC CAC TTC TTT C-3'(Rev)	pOPINJ
HA (C-ter) NT-chCREB3	5'-AGG AGA TAT ACC ATG GCA GGT CCG GAA CCG CTG G-3' (Fwd) 5'-GTG ATG GTG ATG TTT TCA AGC GTA ATC TGG AAC ATC GTA TGG GTA TGC GGT GGT GGT TTT GGT GG-3' (Rev)	pOPINE
FLAG (C-ter) NT-chCREB3	5'-AGG AGA TAT ACC ATG GCA GGT CCG GAA CCG CTG G -3' (Fwd) 5'-GTG ATG GTG ATG TTT TCA CTT GTC GTC ATC GTC TTT GTA GTC TGC GGT GGT GGT TTT GGT GG-3' (Rev)	pOPINE
His (C-ter) NT-chCREB3	5'-AGG AGA TAT ACC ATG GCA GGT CCG GAA CCG CTG G -3' (Fwd) 5'-GTG ATG GTG ATG TTT TGC GGT GGT GGT TTT GGT GG-3' (Rev)	pOPINE
His (N-ter) chCREB3 bZIP	5'-AAG TTC TGT TTC AGG GCC CGA AAG CAG AAG AAC GCC TGC TG-3' (Fwd) 5'- ATG GTC TAG AAA GCT TTA CTG ACG AAC CAG TGC CTG CAG-3' (Rev)	pOPINF

Table 2.7 Primers used for cloning into pOPIN vectors.

In-Fusion enzyme can join any two pieces of DNA that have a ~15bp overlap at their ends by a mechanism equivalent to homologous recombination (Bird et al., 2014). First, the target DNA was amplified by PCR using primers with appropriate

extensions (Berrow et al., 2009), and the amplified products were purified using the PCR purification kit (Qiagen). The target vectors were linearised by restriction digestion (using HindIII and KpnI for pOPINF and pOPINJ and NcoI and PmeI for pOPINE). The In-Fusion reaction was carried out by mixing 200ng of the purified PCR product, 100ng of the linearised vector and water to a total volume of 10µl, which was then mixed with lyophilized enzyme from the In-Fusion® Dry-Down PCR Cloning Kit (Clontech). The reaction mixture was incubated for 15 minutes at 37°C followed by 15 minutes at 50°C in a PCR thermocycler (Bio-Rad). Subsequently, the mixture was diluted with 40µl of TE (Tris-EDTA) buffer and 2.5µl of this dilution was transformed into library efficiency *E. coli* DH5α competent cells (Life technologies). Colonies containing DNA inserts were identified by colony PCR screening (Section 2.9.1). From positive clones, the plasmids were extracted, and the integrity of the inserts was confirmed by Sanger sequencing using Source Bioscience sequencing services.

2.9.5 Classical cloning

The protein constructs were cloned into pcDNA3.1(+) or pcDNA3 for expression in DF1 cells. The vector maps of these vectors are shown in Appendix A, Figure A 2. Untagged pp14B and HA tagged chCREB3 constructs cloned into pcDNA3.1 were obtained from Dr Tahiri-Alaoui (Table 2.2). To perform interaction studies using FRET-FLIM in DF1 cells, pp14B and the C-and N-terminal domains of chCREB3 were labelled with a fluorescent tag (mClover or mRuby), with a short linker (Serine-Glycine-Leucine-Arginine-Serine-Arginine-Glycine-Aspartate-Proline-Proline-Valine-Alanine-Threonine) between the fluorescent protein and the protein of interest. The genes encoding the florescent fusion proteins were

cloned into either pcDNA3 or pcDNA3.1 using classical cloning by restriction enzyme digestion and ligation. The primers used for cloning are given in Table 2.8.

Gene	Primer sequence	Vector	Restriction sites
mClover (N-ter) pp14B	5'-GCC GCT TCT AGA TCC GGA CTC AGA TCT CGA GGG GAC CCA CCG GTC GCC ACC GTA AAT ACT ACG GCA GAA A-3' (Fwd) 5'-GGT AAC GGG CCC TCA GAG GGT CGA TTC CAC TTC-3' (Rev)	pcDNA3	XbaI/ApaI
mClover (C-ter) pp14B	5'-ACC GCT GGA TCC ATG GTA AAT ACT ACG GCA GAA-3' (Fwd) 5'-GGC AAT GAA TTC GGT GGC GAC CGG TGG GTC CCC TCG AGA TCT GAG TCC GGA GAG GGT CGA TTC CAC TTC-3' (Rev)	pcDNA3	BamH1/EcoR1
mRuby (N-ter) PCR	5'-GAT CTA GGA TCC ATG GTG TCT AAG GGC GAA GAG-3' (Fwd) 5'-TAG ATC AAG CTT GGT GGC GAC CGG TGG GTC CCC TCG AGA TCT GAG TCC GGA CTT GTA CAG CTC GTC CAT CCC-3' (Rev)	pcDNA3.1(+)	BamHI/HindIII
mRuby (N-ter) NT-chCREB3-TM	5'-GAC TCA AAG CTT GCG GGC CCC GAG CCT CTG GCC-3' (Fwd) 5'-TGA GTC GAA TTC TCA ACC TCT ACT GCC AAA CGA GTA GAG-3' (Rev)	pcDNA3.1(+)	HindIII/EcoR1
mRuby (N-ter) NT-chCREB3	5'-GAC TCA AAG CTT GCG GGC CCC GAG CCT CTG GCC-3' (Fwd) 5'-TGA GTC GAA TTC TCA CGC TGT GGT AGT TTT GGT GGT GGA CTG-3' (Rev)	pcDNA3.1(+)	HindIII/EcoR1
mRuby (C-ter) NT-chCREB3	5'-GAC CTG AAG CTT ATG GCG GGC CCC GAG CCT CTG GCC-3' (Fwd) 5'-ACG GTC GGA TCC GTG GCG ACC GGT GGG TCC CCT CGA GAT CTG AGT CCG GAC GCT GTG GTA GTT TTG GTG GTG GAC TG-3' (Rev)	pcDNA3	HindIII/BamHI
mRuby (N-ter) CT-chCREB3	5'-GAC TCA AAG CTT AGA GGG CCG CAG CCA GAG TTC-3' (Fwd) 5'-TGA GTC GAA TTC TCA CAT CTC ATC GGT GCG GTG CTG-3' (Rev)	pcDNA3.1(+)	HindIII/EcoR1
mRuby (C-ter) CT-chCREB3	5'-GAC CTG AAG CTT ATG AGA GGG CCG CAG CCA GAG TTC-3' (Fwd) 5'-ACG GTC GGA TCC GTG GCG ACC GGT GGG TCC CCT CGA GAT CTG AGT CCG GAC ATC TCA TCG GTG CGG TGC TG-3' (Rev)	pcDNA3	HindIII/BamHI

Table 2.8 Primers and restriction sites used for cloning into pcDNA3 vectors.

To create pp14B with a C-terminal mClover tag, the sequence encoding pp14B was generated by PCR with primers (Table 2.8) that contained the linker sequence and the required restriction sites. The PCR product was digested with XbaI/ApaI,

and then inserted in-frame into pcDNA3/mClover vector digested with the same enzymes.

To generate pp14B with a N-terminal mClover tag, the sequence encoding pp14B was generated by PCR with primers (Table 2.8) that contained the linker sequence and the required restriction sites. The PCR product was digested with BamH1/EcoR1, and then inserted in-frame into pcDNA3/mClover vector digested with the same enzymes.

To attach a N-terminal mRuby tag to NT-chCREB3, NT-chCREB3-TM and CT-chCREB3, two PCR fragments, one encoding mRuby and the linker, and the other encoding the protein of interest, were generated with primers (Table 2.8) that included the required restriction sites and the linker sequence. The PCR fragment encoding mRuby with the linker sequence was digested with BamHI/HindIII, and the PCR fragment encoding NT-chCREB3, NT-chCREB3-TM or CT-chCREB3 was digested with HindIII/EcoRI, such that the two PCR fragments had stick ends to form a fragment composed of mRuby-linker-POI, and this fragment in turn had sticky ends to ligate with pcDNA3.1 digested with BamHI/EcoRI. The two PCR fragments, encoding mRuby-linker and the protein of interest, were ligated together into the digested pcDNA3.1 vector.

To attach a C-terminal mRuby tag to NT-chCREB3, and CT-chCREB3, the gene encoding NT-chCREB3 or CT-chCREB3 was generated by PCR with primers (Table 2.8) that included the linker sequence and the required restriction sites. The

PCR products were digested with HindIII/BamHI, and then inserted in-frame into pcDNA3/mRuby vector digested with the same enzymes.

PCR amplifications, restriction enzyme digestions, and DNA ligations were undertaken as described in appropriate sections. Colonies containing DNA inserts were identified using colony PCR screening. From positive clones, the plasmids were extracted and the integrity of the inserts was confirmed by Sanger sequencing.

2.9.6 Restriction endonuclease digestion of DNA

Restriction enzymes were obtained from New England Biolabs (NEB). Restriction digests were performed as per the instructions on the manufacturer's website (https://www.neb.com/). DNA was visualised by agarose gel electrophoresis, and the fragments were purified using the QIAquick Gel Extraction System.

2.9.7 DNA ligation

Ligations were carried out using T4 DNA Ligase (NEB). The insert DNA digested with appropriate enzymes was mixed with linearised vector DNA digested with the same enzymes in the molar ratio 1 vector: 3 insert in a 20µl reaction and ligated using 1µl T4 DNA Ligase (NEB) by incubation at room temperature for 10 minutes following the manufacturer's instructions. 2µl of the ligation mixture was transformed into library efficiency *E. coli* DH5α competent cells and the cells were plated onto Luria Bertani (LB) agar plates containing appropriate selection antibiotic and left overnight at 37℃. Colonies containing DNA inserts were identified using colony PCR screening.

2.9.8 Transformation of chemically competent *E. coli*

To transform chemically competent *E. coli* cells such as DH5α, and Rosetta2 pLysS, 50µL aliquots of chemically competent cells were thawed on ice. 100ng DNA was added per tube and left on ice for 30 minutes. The cells were then placed in a 42°C water bath for 30 seconds and then immediately placed on ice for a further 2 minutes. Super optimal broth (SOC; Invitrogen) was added to the cells to make the total volume up to 300µl and incubated at 37°C while shaking at 180rpm for 1 hour. They were then plated onto LB agar containing appropriate selection antibiotic and left overnight at 37°C.

2.9.9 Glycerol stocks of *E. coli*

Individual colonies were grown overnight in 5ml of LB with appropriate antibiotic selection, and 750µl of the culture was transferred to a cryovial. Glycerol was added to a final concentration of 25% before storing at -80°C.

2.9.10 Preparation of LB and TB media

To prepare LB medium, 25g of the LB broth (Miller) was resuspended in 1litre (L) distilled water (dH_2O) and autoclaved at 121°C for 15–20 minutes at 15psi (pound force per square inch) pressure.

To prepare terrific broth (TB), 48g of TB powder (Melford) was dissolved in 1L dH_2O, and 4ml of 100% glycerol was added. The mixture was then autoclaved at 121°C for 15–20 minutes at 15psi pressure.

2.9.11 Preparation of agar plates

LB agar was prepared as LB medium followed by addition of 15g/L of agar before autoclaving. To prepare agar plates, the agar was melted in a microwave, and when the temperature of the melted agar reached 40°C appropriate antibiotic(s) was added before pouring into required plates.

2.9.12 Purification of plasmid and BAC DNA

Individual colonies were grown overnight in 5ml (small scale) or 50ml (medium scale) of LB supplemented with appropriate antibiotics, and cells were harvested by centrifugation at 4,000xg for 20 minutes. Isolation of plasmid DNA was carried out using either the QIAprep® Spin Miniprep Kit (small scale) or QIAGEN Plasmid Midi Kit (Qiagen) (medium-scale) following the manufacturer's instructions. The low-copy number BAC vectors were isolated using the QIAGEN Plasmid Maxi Kit (Qiagen) following the manufacturer's instructions from 250ml bacterial culture. DNA samples were quantified using a Nanodrop 1000 Spectrophotometer (Thermo Scientific).

2.10 Transfection

2.10.1 Transfection of DF1 and HEK293T cells

Transfection of plasmids into DF1 and HEK293T cells was carried out using Lipofectamine 2000 (Invitrogen). Cells were sub-cultured the previous day and seeded to be at about 80–90% confluence prior to transfection. For a 6 well plate, DNA and Lipofectamine 2000 were both diluted in reduced serum medium Opti-MEM (Gibco), to a total volume of 200µl each. They were then mixed to make a

volume of 400µl, left for 20 minutes at room temperature to allow DNA/Lipofectamine 2000 complexes to form and added directly to the cell media. Transfections were scaled up and down depending on the size of the culture vessel used.

2.10.2 Transfection of CEF cells with MDV-BAC for virus production

The transfection of purified MDV BAC DNA into CEF cells leads to reconstitution of infectious virus. Transfection of BAC DNA plasmids was carried out using Lipofectamine (Invitrogen). CEF cells were sub-cultured the previous day and seeded to be at about 80–90% confluence prior to transfection. To transfect a 6 well plate, 10µL of Lipofectamine was diluted to a volume of 100µL with Opti-MEM, and 1µg of BAC DNA was also diluted to a total volume of 100µL with Opti-MEM and the mixtures were left for 5 minutes at room temperature. They were then mixed to make a total volume of 200µL and left for 20 minutes at room temperature to allow DNA/Lipofectamine complexes to form. The media from the cells was removed and they were washed once with Opti-MEM. The Lipofectamine/DNA mix was made up to 1ml with Opti-MEM and added to the cells and left for 8 hours at 37°C. The mixture was then removed and CEF growth medium was added and the cells were incubated at 37°C. After 24 hours, the growth medium was replaced with CEF maintenance medium and left until the appearance of MDV plaques (which typically occurs 4–5 days after transfection).

2.11 Protein analysis

2.11.1 SDS polyacrylamide gel electrophoresis (PAGE)

Protein samples were mixed with of 4x SDS (Sodium dodecyl sulphate) sample loading buffer (250mM Tris pH6.8, 8% w/v SDS, 0.02% w/v Bromophenol blue, 40% v/v Glycerol, and 8% β-mercaptoethanol (BME)] or 4x LDS [Lithium dodecyl sulfate) sample buffer (Invitrogen) containing 8% BME, heated to 99°C for 10 minutes, and spun down before loading. Samples were resolved on a 4–12% or 10% Bis-Tris gels (Invitrogen), which were run in 1x MES buffer (Invitrogen) using the NuPage Novex gel system (Invitrogen). 10µl samples were loaded per well alongside appropriate molecular weight markers [SeeBlue® Plus2 Prestained Protein Standard (3–198kDa), Invitrogen or PageRuler Prestained protein ladder (10–250kDa), Invitrogen]. The gels were run at 200 volts until the blue dye reached the bottom of the gels.

2.11.2 Instant Blue staining of proteins

To visualise proteins separated by SDS-PAGE, the gels were stained with Instant Blue (ISB01L Expedon) for 30–60 minutes. After removing the stain, the gels were distained by washing with water and photographed using the Bio-Rad Gel-Chemi doc imaging system.

2.11.3 Western blot

For western blot analysis, proteins from SDS-PAGE gels were transferred either onto PVDF membrane using the iBlot western detection stack/iBlot® Dry Blotting System (Invitrogen) or onto nitrocellulose membrane using the mini blot wet

transfer module (Invitrogen) in transfer buffer [25mM Tris base, 192 mM Glycine, 10% (v/v) Ethanol] as per manufacturers' instructions. After transfer, the membranes were blocked in blocking buffer [PBS containing 5% (w/v) dried skimmed milk (Marvel) and 0.1% (w/v) Tween-20] for 1 hour at room temperature with shaking. They were then incubated with appropriate primary antibodies diluted in blocking buffer at room temperature for 1 hour or overnight at 4°C with shaking, and then washed three times in wash buffer [PBS containing 0.1% (w/v) Tween-20] for 5 minutes each time. The membranes were then incubated with the appropriate horseradish peroxidase (HRP) conjugated secondary antibody diluted in blocking buffer for 1 hour at room temperature, then washed again three times for 10 minutes each time. The blots were then developed using an enhanced chemiluminescence (ECL) reagent (Bio-Rad laboratories) and visualised using the Bio-Rad Gel-Chemi doc imaging system.

2.12 Co-immunoprecipitation

Cell culture: Co-immunoprecipitation assays were carried out with cell lysates prepared from different cell lines (DF1 cells, CEF cells and insect cells). DF1 and CEF cells were sub-cultured in a 6-well plate the previous day and seeded to be at about 80–90% confluence prior to transfection. DF1 cells were transfected with 2.5µg DNA of each construct per well (total 5µg DNA) and processed 48 hours post transfection. CEF cells were transfected with 1µg BAC pRB-1B/HA-chCREB3, and infected cells were processed 4 days post transfection. Approximately 1×10^6 Sf9 cells were seeded with 2ml insect cell culture media in a 6-well plate the day before infection. Cells were co-infected with 10µl of each

P1 virus (Section 2.15.3.1) per well. Infected Sf9 cells were processed 48 hours post infection.

Preparation of cell lysate: The cells were washed three times with ice cold PBS after removing the growth medium. They were then lysed on ice for 20 minutes using 1ml lysis buffer [50mM Tris-HCl pH 7.5, 150mM NaCl, 0.5% Nonidet P-40 (NP40), 5% glycerol, 3x Halt™ Protease and Phosphatase Inhibitor Single-Use Cocktail, Ethylenediaminetetraacetic acid (EDTA)-Free (100x), 2x Roche Complete EDTA free protease inhibitors] per sample. Lysed cells were centrifuged at 14000rpm at 4ºC for 15minutes, and the supernatant (lysate) was transferred to a clean tube.

Co-immunoprecipitation: 400µl of clarified lysate was used per co-immunoprecipitation reaction. For each sample, 50µl protein A Dynabeads (Invitrogen) was washed once with wash buffer (50mM Tris-HCl pH 7.5, 150mM NaCl, 0.5% NP40, 5% glycerol) and incubated with 5µg of appropriate IP antibody or control IgG (rabbit anti-RLORF9 or normal mouse IgG) in wash buffer for 1 hour at room temperature. The beads were then incubated with the cell lysate overnight at 4°C or for 3 hours at room temperature on a rotary shaker. The supernatant (unbound, UB) was then removed and the beads were washed 3 times with 1ml wash buffer each time. The bound materials or immunopreciptates (IP) were then eluted by boiling the beads with 20µl 1x LDS sample buffer containing 2% BME for 8 minutes. The samples were separated by SDS-PAGE and analysed by western blotting with appropriate antibodies.

2.13 Bioimaging techniques

2.13.1 Immunofluorescence

The standard procedure was followed. Cells were washed once with PBS after removing the culture medium and fixed in 4% paraformaldehyde for 10 minutes at room temperature and then washed with PBS. The cells were permeabilised with 0.1% Triton X-100 for 15 minutes and washed again in PBS, and then blocked in PBS with 0.5% BSA and 0.1% Tween-20 for 1 hour at room temperature. The primary antibody was then added for 1 hour diluted in blocking buffer, afterwards the cells were washed three times for five minutes each time in PBS and then incubated with the secondary species-specific fluorescent conjugated antibody for 1 hour in blocking buffer. Cells were washed a further three times for five minutes in PBS and then stained with DAPI (1:10,000) made up in water before dipping in water and mounting on slides using Vectashield (Vector Labs). Coverslips were then sealed with nail varnish and visualised by confocal microscopy.

To assess colocalisation between pp14 and chCREB3, and the localisation of chCREB3 in infected CEF cells, the following procedure was followed. CEF cells were sub-cultured in a 6-well plate the previous day and seeded to be at about 80–90% confluence prior to transfection. The cells were transfected with 1μg BAC pRB-1B/HA-chCREB3 or 1μg BAC pRB-1BΔpp14/HA-chCREB3 per well. At day 4, the cells were transferred to coverslips in a 24-well plate and were fixed on the fifth day after transfection. The cells infected with RB-1B/HA-chCREB3 were immunostained with rabbit anti-pp14 and mouse anti-HA, and then probed with Alexa 568-conjugated goat anti-rabbit and Alexa 488-conjugated goat anti-mouse

secondary antibodies. The cells infected with RB-1BΔpp14/HA-chCREB3 were immunostained with mouse anti-pp38 (to indicate infection) and rabbit anti-HA and then probed with Alexa 488-conjugated goat anti-rabbit and Alexa 568 conjugated goat anti-mouse secondary antibodies.

To test potential cross reactivity between anti-pp14 and chCREB3 by immunofluorescence, DF1 cells were sub-cultured on coverslips in a 24-well plate and seeded to be at about 80–90% confluence prior to transfection. The cells were transiently transfected with 0.5μg of mRuby (N-ter) NT-chCREB3, mRuby (N-ter) NT-chCREB3-TM, or untagged pp14 (positive control) and were fixed 24 hours after transfection. The cells were then stained with rabbit anti-pp14, followed by Alexa 488 conjugated goat anti-rabbit secondary antibody.

2.13.2 Staining of live cells

DF1 cells were sub-cultured in 35-mm glass-bottomed culture dishes (MatTek, Ashland, MA) the previous day and seeded to be at about 70–80% confluence prior to transfection. Cells were transiently transfected with mRuby NT-chCREB3-TM and 48 hours after transfection, living DF1 cells were incubated with 500mM ER-Tracker™ Blue-White DPX (Invitrogen) diluted in Fluorobrite (Gibco) supplemented with 3% FCS for 30 minutes at 37°C. After removing the staining solution, the cells were imaged in Fluorobrite supplemented with 3% FCS using confocal microscopy.

2.13.3 Bioimaging and image processing

Confocal images were taken using a Leica TCS SP5 microscope (Leica Microsystems, Manheim, Germany) at the Pirbright Institute or Leica SP8 SMD X confocal microscope (Leica Microsystems, Manheim, Germany) in STRUBI. The parameters used for samples imaged on the Leica SP8 SMD X confocal microscope were as follows. Images were taken with a 63x oil-immersion objective with a numerical aperture of 1.40. The ER-Tracker Blue-White DPX dye and DAPI were excited with a 405nm laser. Alexa 488 was excited with a white light laser tuned to 488nm. Alexa 568 and mRuby were excited with a white light laser tuned to 561nm. Fluorescence emission was detected with HyD detectors capable of photon counting and a photomultiplier tube. The emission windows for the fluorescent proteins used were selected as follows: DAPI and ER-Tracker Blue-White DPX (418nm–491nm), Alexa-488 (501nm–552nm), mRuby and Alexa-568 (570nm–621nm). The pinhole was set at one Airy unit. Acquired images were analysed and processed using the LAS AF software (Leica Microsystems).

2.14 Förster Resonance Energy Transfer by Fluorescence lifetime imaging (FRET-FLIM)

DF1 cells were sub-cultured in 35-mm glass-bottomed culture dishes (MatTek, Ashland, MA) the previous day and seeded to be at about 70–80% confluence prior to transfection. As given in Table 2.9, cells were transfected or co-transfected with appropriate plasmid(s) using Lipofectamine 2000. To determine the amount of DNA and Lipofectamine 2000 to use per construct (to achieve appropriate levels

of expression), multiple rounds of transfection varying the amounts of plasmid DNA and Lipofectamine 2000 were carried out.

Construct (s)/ Amount of DNA	Lipofectamine 2000
Unlinked mClover alone (10ng)	1µl
mClover-mRuby tandem (25 ng)	1µl
mClover pp14B alone (10ng)	1µl
mClover pp14B (5ng) and mRuby NT-chCREB3 (800ng)	2µl
mClover pp14B (10ng) and mRuby NT chCREB3-TM (200ng)	2µl
mClover pp14B (15ng) and mRuby CT-chCREB3 (400ng)	2µl

Table 2.9 Transfection of fluorescent constructs.
Amounts of plasmid DNA and Lipofectamine 2000 used for transfections.

FRET was detected and quantified by fluorescence lifetime imaging (FLIM) using a time correlated single photon counting approach (TCSPC). The TCSPC FLIM technique uses a pulsed laser light source to excite the sample at a high frequency, which is coupled to a detector that measures the time taken for each photon to be emitted after the pulse (Baird et al., 2014; Broussard et al., 2013). The total photon count, displayed in different time bins, creates a decay curve (Figure 2.2) (Baird et al., 2014; Broussard et al., 2013). In the absence of FRET, a single exponential model is used to fit the donor fluorescence decay and in FRET situations, a double exponential model can approximate the resulting donor fluorescence decay (Baird et al., 2014; Broussard et al., 2013). The fluorescence lifetime is calculated from the slope of the decay curves by applying appropriate equations as described below (Berezin and Achilefu, 2010).

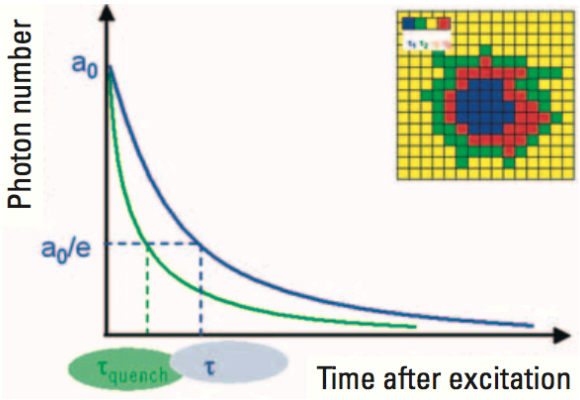


Figure 2.2 Fluorescence decay curve.
Following the excitation pulse, a_0, the initial number of emitted photons decays exponentially. The time take for the fluorescence to decay to a_0/e (~ 37%) is the fluorescence lifetime. In FRET conditions, the lifetime (τ) shifts to shorter times (τ_{quench}). FLIM generates a spatial map of fluorescence lifetimes by measuring at each position of the scanning region (see inset showing a pseudocoloured lifetime image). (Picture obtained from: http://www.leica-microsystems.com)

Living DF1 cells were imaged 36 hours post transfection, in FluoroBrite supplemented with 3% FCS, in a humidified chamber maintained at 37°C using a SP8-SMD Leica microscope. Areas of interest were chosen under a 63x/1.4 NA oil immersion objective. The cells were selected to have similar initial mClover fluorescence intensity (approximately 10–20AU). For each FLIM image, mClover was excited with a white light laser tuned to 488nm and emitted single photons from mClover were subsequently detected by a HyD detector (Leica Microsystems).

Data processing for lifetime analysis was performed using SymPhoTime (Picoquant GmbH). Fifteen cells were analysed per condition. To achieve high signal to noise and minimise artefacts, only cells with 600 photons per pixel and negligible amount of bleaching were included in the analysis. Suitable regions of interest (ROIs) were selected for each sample based on the localisation pattern of proteins of interest. The expression of mClover pp14B and mRuby NT-chCREB3-

TM was predominant around the ER region immediately near the nucleus. Therefore, the ROI in cells expressing mClover pp14B alone, mClover pp14B and mRuby NT-chCREB3-TM together, and mClover pp14B and mRuby CT-chCREB3 together comprised the ER region immediately near the nucleus. The ROI in cells co-expressing mClover pp14B and mRuby NT-chCREB3 included the nucleus, where NT-chCREB3 expression was predominant, as well as the ER region immediately near the nucleus. The distribution of mClover was uniform in the cytoplasm and nucleus, and therefore any region of the cell was selected in cells expressing mClover alone or mClover-mRuby tandem. The acquired fluorescence decays of each pixel within ROIs were tail-fitted with one or two–exponential theoretical models (Kong et al., 2016; Padilla-Parra et al., 2008). To ensure precise fitting, selected ROIs contained at least 500 000 total photon counts.

The mean fluorescence lifetimes (Tau, τ) were calculated using SymPhoTime as previously described (Kong et al., 2016; Padilla-Parra et al., 2008). In non-FRET conditions, a single exponential model was used to fit the donor fluorescence decay, and the following equation, which defines the mean fluorescence lifetime, τ, of a fluorescence decay, i(t), was applied:

$$\langle \tau \rangle = \int t \times i(t) dt \, / \int i(t) dt$$

In FRET conditions, a two-species model was applied due to the presence of an interacting fraction corresponding to a population that relaxes through FRET (f_D) and a non-interacting fraction in which the donor lifetime remains undisturbed

($1-f_D$) and the following equation was applied to determine the fluorescence lifetime of the donor:

$$i(t) = (1 - f_D)e^{-t/\tau_D} + f_D e^{-t/\tau_F}$$

where f_D stands for the fraction of interacting donor, τ_D is the fixed donor lifetime and τ_F is the discrete FRET lifetime.

To determine the fluorescence lifetime of mClover in cells expressing mClover-mRuby tandem (predicted to give maximum FRET), the mean fluorescence lifetime from cells expressing mClover alone was first assigned. This value, fixed as the lifetime of the non-interacting fraction, was used to perform two-exponential fitting.

2.15 Protein expression and purification

2.15.1 Size exclusion chromatography

Affinity purified protein samples were subjected to size exclusion chromatography (SEC), which separates proteins on the basis of their size and shape, to increase the purity and homogeneity of the protein samples. After affinity purification, samples of pp14B, CT-chCREB3 and bZIP were analysed using a Superdex 75 10/300 GL (GE Healthcare) column equilibrated with appropriate running buffer. SEC was run at a flow rate of 1ml/min collecting 0.5ml fractions over the whole run. Following purification, the protein-containing fractions were analysed by SDS-PAGE.

2.15.2 Protein expression in *E. coli*

2.15.2.1 Large-scale expression of constructs

The appropriate pOPIN vector containing the protein of interest was transformed into Rosetta2 (DE3) pLysS cells. Cultures were started from overnight precultures in terrific broth (TB) supplemented with ampicillin and chloramphenicol and were grown at 37°C while shaking at 160rpm to an OD_{600} of 0.6. Expression was induced with 0.5mM IPTG (isopropyl β-D-1- thiogalactopyranoside) for 18 hours at 20°C with shaking at 160rpm. The cells were harvested by centrifugation at 5000xg for 20 minutes at 18°C.The pellets were either processed immediately or flash-frozen and stored at -20°C for later use.

2.15.2.2 Additive/pH screening

The full length pp14B with a C-terminal His tag was expressed using the pOPINE vector in Rosetta2 (DE3) pLysS cells as described in Section 2.15.2.1. To test the effect of different additives and pH on the aggregation of pp14B, protein was extracted from half a litre of *E. coli* culture per condition and subjected to SEC.

For additive screening, the pellets were resuspended in 50mM Tris-HCl, 500mM NaCl, 0.2% Tween-20, 1x Roche Complete EDTA free protease inhibitors supplemented with an additive at a particular concentration (Table 2.10) in the presence or absence of 50mM L-arginine hydrochloride (L-arginine HCl). To screen buffers between pH 7 and pH 8, the pellets were resuspended in appropriate buffers (PBS pH 7, MOPS pH 7, Tris-HCl pH 7, or Tris-HCl pH 8) containing 500mM NaCl, 0.2% Tween-20, 1x Roche Complete EDTA free protease inhibitors

in the presence or absence of 50mM L-arginine HCl. To screen buffers below pH 7 (CH_3COONa pH 5, MES pH 6, and PBS pH 6), the pellets were resuspended in 50mM Tris-HCl, 500mM NaCl, 0.2% Tween-20, 50mM L-arginine HCl, and buffer exchanged after Ni-NTA (nickel-nitrilotriacetic acid) purification. The resuspended cells were lysed by sonification, spun down and the clarified lysate was loaded onto a column containing pre-equilibrated Ni-NTA agarose (Qiagen). After extensive washes, the recombinant protein was eluted using the corresponding binding buffer supplemented with 300mM Imidazole (without Tween-20). The samples were then subjected to SEC on a Superdex 75 10/300 GL column equilibrated with the corresponding binding buffer (without Tween-20).

Additive	**Concentration (s)**
Urea	0.5M, 1M, 2M
Arginine L-HCl	0.05M, 0.5M, 1M
BME/DTT/TCEP	1mM, 5mM, 10mM
NDSB	1M
Zinc	1mM, 10mM
Glycerol	10%
Arginine L-HCl + L-glutamic acid (L-Glu)	50mM/50mM

Table 2.10 List of additives screened.

2.15.2.3 Refolding of pp14 with zinc chloride

A rapid on-column refolding technique was employed during the purifications. The full-length pp14B with a C-terminal His tag was expressed using the pOPINE vector in Rosetta2 (DE3) pLysS cells as described in Section 2.14.2.1. The bacterial pellet was resuspended in 50mM Tris-HCl pH 7.0, 500mM NaCl, 0.2% Tween-20, 8M urea, and lysed by sonication. The lysate was clarified by centrifugation at 20,000xg for 1 hour at 4°C and the supernatant was loaded onto a column containing pre-equilibrated Ni-NTA agarose. After extensive washes

with wash buffer (binding buffer supplemented with 5mM $ZnCl_2$), the urea concentration was rapidly reduced by dilution to facilitate refolding of the protein while still bound to the nickel beads. The sample was then eluted in 50mM Tris-HCl pH7.0, 500mM NaCl, 5mM $ZnCl_2$, 300mM Imidazole and was subjected to SEC on a Superdex 75 10/300 GL column equilibrated with 50mM Tris-HCl pH7.0, 500mM NaCl, 5mM $ZnCl_2$. A control refolding experiment with buffers containing no $ZnCl_2$ was also carried out in parallel.

2.15.2.4 Extraction of monomeric pp14B

Using arginine: The full length pp14B with a C-terminal His tag was expressed in Rosetta2 (DE3) pLysS cells as described in Section 2.14.2.1. For protein extraction, the pellets were resuspended in 50mM Tris-HCl pH8, 500mM NaCl, 50mM L-arginine HCl, 0.1% Tween 20, 1x Roche Complete EDTA free protease inhibitors. The cells were lysed by sonication, spun down and the clarified lysate was loaded onto a column containing pre-equilibrated Ni-NTA agarose. After extensive washes, the recombinant protein was eluted with 50mM Tris-HCl pH8, 500mM NaCl, 1M L-arginine HCl, 300mM Imidazole. The sample was then subjected to SEC on a Superdex 75 10/300 GL column equilibrated with 50mM Tris-HCl pH8, 500mM NaCl, 50mM L-arginine HCl. Following purification, the fractions containing protein were analysed by SDS-PAGE.

Using NP40: The wild type or Cys-null pp14B with a C-terminal His tag was expressed using the pOPINE vector in Rosetta2 (DE3) pLysS cells as described in Section 2.14.2.1. For protein extraction, the pellets were resuspended in 50 mM Tris-HCl pH 7.5, 150mM NaCl, NP40 (0.5% or 2%), glycerol (5% or 10%), 1x

Roche Complete EDTA free protease inhibitors and lysed by sonication. The clarified lysate was then incubated with Ni-NTA agarose for 2 hours at 4°C with rotation and was loaded onto a column to perform washes with wash buffers (50mM Tris-HCl pH 7.5, 150mM NaCl, 1M NaCl and 50mM Tris-HCl pH 7.5, 150mM NaCl, 10mM imidazole). The proteins were eluted with 50mM Tris-HCl pH 7.5, 150 mM NaCl, 300mM Imidazole. The samples were then subjected to SEC on a Superdex 75 10/300 GL in 50mM Tris-HCl pH 7.5, 150mM NaCl. Following purification, the fractions containing protein were analysed by SDS-PAGE.

2.15.2.5 Expression and purification of the chCREB3 bZIP

The bZIP domain of chCREB3 with a N-terminal His tag was expressed using the pOPINF vector in Rosetta2 (DE3) pLysS cells. The expression was carried out as described in Section 2.14.2.1 except LB media was used for expression, and the expression was induced for 3 hours at 37°C. For protein extraction, the pellets were resuspended in 50mM Tris-HCl pH 8, 150mM NaCl, 0.1% Tween 20, 1x Roche Complete EDTA free protease inhibitors and lysed by sonication. The clarified lysate was incubated with Ni-NTA agarose for 2 hours at 4°C with rotation and was loaded onto a column to perform washes with wash buffers (50mM Tris-HCl pH 8, 150mM NaCl, 1M NaCl and 50mM Tris-HCl, pH 8, 150mM NaCl, 10mM imidazole). The protein was eluted with 50mM Tris-HCl pH 8, 150mM NaCl, 300mM Imidazole. The sample was then subjected to SEC on a Superdex 75 10/300GL column equilibrated with 50mM Tris-HCl, pH 8, 150mM NaCl. Following purification, the fractions containing protein were analysed by SDS-PAGE.

2.15.3 Protein expression in insect cells

2.15.3.1 Generation and amplification of recombinant baculovirus

Sf21 or Sf9 cells were transferred from a suspension culture to a 6-well plate (1×10^6 cells/well) and allowed to settle for 30 minutes. The insect cell culture medium was then replaced with plating medium (SF900II medium supplemented with 1.5% FCS). 0.5µg of the appropriate pOPIN vector containing the gene of interest and 0.15µg of linearised bacmid (donated by Prof Ian Jones, Reading University, Zhao et al. 2003) were diluted to a volume of 100µl in serum free Sf900II medium, and mixed with 8µl of transfection reagent Cellfectin II, (Life Technologies) diluted to 100µl in serum free Sf900II medium. The mixture was left for 30 minutes at room temperature to allow Cellfectin II/DNA complexes to form, before adding the mixture dropwise to the cells. The 6-well plate was then incubated at 28°C for 5 hours, after which the medium was replaced with insect cell culture medium, and the plate put back at 28°C. After 72 hours, 1ml of medium, now containing recombinant baculovirus (P0 virus), from each well was used to inoculate a 2L flask containing 40ml suspension culture of Sf9 or Sf21 cells at a density of 1×10^6 cells/ml, and the virus was allowed to amplify for 72 hours at 28°C with shaking at 135rpm. The supernatant containing amplified virus was harvested by centrifugation at 1000xg for 5 minutes at 20°C, and the virus stock (P1 virus) was stored at 4°C for later use. The cell pellets were analysed by western blotting to test for protein expression.

2.15.3.2 Small-scale expression screening in Sf9 cells

The expression screening of pp14 and NT-chCREB3 constructs in Sf9 cells was conducted in 6-well plates. The expression of the NT-chCREB3 was tested with the following fusion tags: His (N-ter), His-sumo (N-ter), His-TRX (N-ter), His-MBP (N-ter), His-halo (N-ter), His-GST (N-ter), FLAG (C-ter), HA (C-ter) and His (C-ter). The expression of pp14B was tested with the following fusion tags: His-GST (N-ter), FLAG (C-ter), and His (C-ter). Approximately 1×10^6 Sf9 cells were seeded with 2ml insect cell culture media the day before infection. Cells were infected with 10µl of each recombinant baculovirus per well. Infected Sf9 cells were harvested 48 hours post infection by centrifugation at 1,000xg for 10 minutes. The pellets were resuspended in 50µl 4x LDS sample buffer containing 2% BME and boiled at 95°C for 1 hour. 10µl of each sample was analysed by SDS-PAGE and western blotting with appropriate antibodies.

The expression screening of CT-chCREB3 in Sf9 cells was conducted in 24 deep well plates (round bottoms) at the OPPF, following OPPF protocols. Plates were prepared with 5ml of Sf9 cells at 1×10^6 cells/ml in each well. 5µl or 50µl P1 virus was added to each well including a GFP control and incubated for 2 and 3 days at 27°C, 250rpm. At 2 and 3 days, 1ml from each well was transferred to a 96-well deep well block and centrifuged for 15 minutes at 6,000g to remove the supernatant and the cells frozen at -80°C for 30 minutes. The cells were then processed using the standard OPPF Ni-NTA expression screen protocol on the Qiagen BioRobot 8000. The cells were resuspended in 210µl of lysis buffer [50mM NaH2PO4 pH8, 300mM NaCl, 10mM imidazole, 1% v/v Tween 20 supplemented with 1mg/ml lysozyme and 400 units/ml of DNAse Type I (Sigma D-4527)] and incubated for

30 minutes before clearing the lysates by centrifuging the deep-well block at 5000xg for 30 minutes at 4°C. The supernatants were then transferred to wells in a flat-bottomed 96-well microtitre plate (Greiner 655101) and incubated with 20µl of Ni-NTA magnetic beads per well for 30 minutes at room temperature. The 96-well microplate was then placed on the 96-well magnet for 1 minute and the supernatants removed carefully from the beads and the beads washed with 200µl of wash buffer (50mM NaH2PO4 pH8, 300mM NaCl, 10mM imidazole, 0.05% v/v Tween 20) two times. The samples were eluted with 50µl of elution buffer (50mM NaH2PO4 pH8, 300mM NaCl, 250mM imidazole, 0.05% v/v Tween 20). 10µl of each sample was mixed with 10µl 2x SDS loading buffer supplemented with BME and 10µl of each sample was analysed by SDS-PAGE.

2.15.3.3 GST pull down (small scale)

Approximately 1×10^6 Sf9 cells were seeded with 2ml insect cell culture media in 6-well plates the day before infection. Cells were co-infected with 10µl of each virus per well and incubated for 2 days at 27°C and harvested by centrifugation at 4,000xg for 10 minutes. Cell lysates were prepared as described in Section 2.12. 500µl of the cell lysate was incubated with 75µl of Glutathione Sepharose 4 Fast flow beads (GE Healthcare) for 3 hours at 4°C on a rotary shaker. The beads were spun down by centrifugation at 1000xg for 10 minutes. After removing the supernatant, the beads were washed 3 times with 1ml of ice cold lysis buffer. To elute the bound proteins, the beads were resuspended in 20µl 1xLDS sample buffer containing 2% BME and boiled for 10 minutes at 99°C.

2.15.3.4 Scaling-up expression and purification

Extraction of pp14B: Five 2L flasks were prepared with 200ml of Sf21 cells at 0.5×10^6 cells/ml the day before infection. 200µl P1 virus expressing pp14B with a C-terminal His tag was inoculated per 200ml. Two days after infection, the cells were harvested by centrifugation at 1000xg for 10 minutes. Protein was extracted from 1L of cell culture. Cells were lysed in 50mM Tris-HCl pH 8, 500mM NaCl, 1% Triton X-100, 1x Roche Complete EDTA free protease inhibitors. The lysate was cleared by centrifugation at 4000xg for 20 minutes at 4°C and the supernatant loaded onto a column containing pre-equilibrated Ni-NTA agarose. The beads were washed extensively with wash buffer (50mM Tris-HCl pH 8, 500mM NaCl, 1% Triton X-100 10mM imidazole), and the protein was eluted with elution buffer (50mM Tris-HCl pH 8, 500mM NaCl, 300mM Imidazole). The sample was then subjected to SEC on a Superdex 75 10/300 GL column equilibrated with 50mM Tris-HCl pH8, 500mM NaCl. Following purification, the fractions containing protein were analysed by SDS-PAGE.

Extraction of CT-chCREB3: Five 2L flasks were prepared with 200ml of Sf21 cells at 0.5×10^6 cells/ml the day before infection. 200µl P1 virus expressing CT-chCREB3 with a C-terminal His tag was inoculated per 200ml culture. Three days post infection, the cells were harvested by centrifugation at 1000xg for 10 minutes. Protein was extracted from 1L of cell culture as described for pp14B above. Following purification, the fractions containing protein were analysed by SDS-PAGE.

2.16 Mass spectrometry analysis

DF1 cells were transfected with untagged pp14B in T75 flasks. Immunoprecipitations were performed using 10µg anti-pp14 antibody with cell lysates prepared from these cells. The samples were resolved on a 4–12% Bis-Tris gel. The gel was stained with Instant Blue. After destaining with water, the protein bands were excised and sent to Dr Dong Xia of the University of Liverpool for phosphorylation analysis using mass spectrometry (MS).

2.17 Circular dichroism (CD) spectroscopy

CD analyses were performed using electronic circular dichroism (ECD) at beamline B23 of Diamond Light Source (Harwell, UK). ECD spectra were recorded using a Chirascan Plus instrument (Applied Photophysics Ltd) equipped with a Peltier temperature control system. The measurements on samples in Table 2.11 were performed at 5°C using a 0.02mm path length demountable cell filled with 3µl of protein solution at a concentration of 4.2mg/ml or 4.6mg/ml. The following parameters were used: 4 acquisitions per sample, 1nm step, 1s collection time per step, 1.0nm spectral bandwidth, 260nm–185nm range (far-UV region).

Thermal melt CD measurements with wild type pp14B (monomeric) in 50mM Tris-HCl pH 7.5, 150mM NaCl were performed using a 0.1mm path length cylindrical cell filled with 15µl of protein solution at a concentration of 1mg/ml. CD spectra were recorded every 10°C from 10–90°C and a final scan taken after cooling down to 10°C. The following parameters were used: 4 acquisitions per

sample, 1nm step, 1s collection time per step, 1.0nm spectral bandwidth, 260 nm–185nm range.

Data were processed and analysed using proprietary software, CDApps (Hussain et al., 2015), developed at the beamline B23 at Diamond Light Source. Quadruplicate measurements from each sample were averaged and buffer subtracted and the raw data were converted into the per residue molar absorption unit, delta epsilon ($\Delta\varepsilon$) expressed in M^{-1} cm^{-1}, by normalisation with respect to protein concentration and path length. The secondary structure estimation from CD spectra was calculated using the CONTIN/LL algorithm (Hussain et al., 2015). The CDApps software fits a melting curve (based on the Boltzmann algorithm) at any selected wavelength and gives an estimation for the melting temperature (T_m) (Hussain et al., 2015). For wild type pp14B (monomeric), the melting curve was obtained by plotting the ellipticity (θ) (measured in millidegrees, mdeg) at 205nm versus temperature.

Sample	Buffer	Protein concentration
Wild type pp14B (monomeric)	50mM Tris-HCl pH 7.5, 150mM NaCl	4.6mg/ml
Wild type pp14B (monomeric)	50mM Tris-HCl pH 7.5, 150mM NaCl, 10mM BME	4.2mg/ml
Wild type pp14B (aggregated form)	50mM Tris-HCl pH 7.5, 150mM NaCl	4.6mg/ml
Cys-null pp14B (monomeric)	50mM Tris-HCl pH 7.5, 150mM NaCl	4.6mg/ml
Cys-null pp14B (aggregated form)	50mM Tris-HCl pH 7.5, 150mM NaCl	4.6mg/ml

Table 2.11 Protein samples subjected to far-UV CD analysis.
The buffer composition and the concentration of the samples subjected to CD analysis.

2.18 ThermoFluor assay

5000x stock of SYPRO orange fluorescent dye was obtained from Molecular Probes (Life Technologies, Paisley, UK), supplied in dimethylsulphoxide (DMSO). Working stocks of 50x SYPRO orange was produced freshly for each experiment by dilution in sample buffer. 50µl volume reactions were set up per well. A standard reaction was set up by mixing appropriate amount of protein of interest (Table 2.12) with 3× SYPRO orange and made up to 50µl with sample buffer. For pH screening, 3µg of Cys-null pp14B (monomeric) in 50mM Tris-HCl pH 7.5, 150mM NaCl was mixed with 10µl of in-house (STRUBI) buffer prepared at different pH (Table 2.13), 3x SYPRO and made up to 50µl with sample buffer. Reactions were set up in 96-well PCR plates (Agilent), mixed by pipetting, and then centrifuged at 2000xg to remove any bubbles. The plates were then sealed using optically clear foils (BioRad Microseal B, MSB-1001). The experiments were carried out using a MX3005p qPCR machine (Agilent, Edinburgh, UK). The excitation and emission bandpass filters used for SYPRO orange were 492nm and 585nm, respectively. The experiments were ramped from 25°C to 99°C recording a fluorescence reading every 1°C.

Sample	Sample buffer	Amount of protein
Wild type pp14B (monomeric)	50mM Tris-HCl pH 8, 500mM NaCl, 50mM L-Arginine HCl	4µg
Wild type pp14B (monomeric)	50mM Tris-HCl pH 7.5, 150mM NaCl, +/- 10mM BME	5µg
Wild type pp14B (monomeric)	50mM Tris-HCl pH 7.5, 150mM NaCl, +/- 50mM BME	10µg
Cys-null pp14B (monomeric or aggregated)	50mM Tris-HCl pH 7.5, 150mM NaCl	5µg

Table 2.12 Protein samples subjected to ThermoFluor analysis.

pH	Buffer composition
pH 4	50mM Sodium acetate pH 4.0, 500mM NaCl
pH 5	50mM Citric acid pH 5.0, 500mM NaCl
pH 6	50mM Sodium cacodylate pH 6.0, 500mM NaCl
pH 7	50mM HEPES pH 7.0, 500mM NaCl
pH 7.4	50mM HEPES pH 7.4, 500mM NaCl
pH 8	50mM Tris-HCl pH 8.0, 500mM NaCl
pH 9	50mM CAPSO pH 9.0, 500mM NaCl

Table 2.13 In-house pH screen buffers.

2.19 Structural biology methods

2.19.1 Crystallisation screens

Initial crystallisation screening was based on a nanolitre (nl) scale sitting drop vapour-diffusion method in 96-well crystallisation plates. The Division of Structural Biology (STRUBI), Oxford, UK has a standard set of crystallisation screening blocks. Each block contains 96 different crystallisation conditions prepared from commercially available sparse matrix crystallisation screens obtained from Hampton Research, Emerald BioSystems, MiTeGen and Molecular Dimensions Limited, etc.

Crystallisation screens were set up by dispensing 100µl of the precipitant solutions into the Greiner plate reservoir chambers of CrystalQuick 96 well crystallisation plate (Greiner) using Robbins Hydra 96 micro-dispenser (Matrix Technologies Ltd, UK). 100nl of protein was then dispensed as sitting drops on to the central well of the raised platform by Cartesian robot (Cartesian Technologies Microsys MIC4000) followed by addition of 100nl of reservoir to each crystallisation platform from the corresponding reservoir wells. The crystallisation plate was then immediately sealed manually with transparent self-adhesive foils (Viewseal,

Greiner) and were then stored at 20°C in the Homebase automated storage vault (The Automation Partnership, Ltd, UK) or at 4°C in a Rock Imager1000 (Formulatrix). Plates at 20°C were imaged using an Oasis 1700 imaging system (Vecco Instruments) and plates at 4°C were imaged using the in-built imaging system of the Rock Imager 1000 (Formulatrix).

2.19.2 Duplex DNA formation

The 20-mer and 12-mer DNA oligomers containing CREB3 target sequences were obtained from Eurofins: 5′-CTTGGC**TGACGTCA**GCCAAG-3′ (CRE site in bold) (Schumacher et al., 2000) and 5′-G**ATGACGTCAT**C-3′ (ATF/CREB site in bold) (Fujii et al., 1998). These sequences were adapted from previously published sequences. The oligonucleotides were self-annealed by heating at 95°C for 5 minutes in 50mM Tris-HCl pH 8.0, 150mM NaCl, and allowed to cool slowly to room temperature. The oligomers were then stored at 4°C.

2.19.3 Crystallisation trials

Crystallisation trials were carried out with pp14B and chCREB3 fragments by mixing 100nl of protein and 100nl of the reservoir solution as described in Section 2.19.1. Details of the crystallisation trials conducted are given in Table 2.14. To determine appropriate starting protein concentration for crystallisation, pre-crystallisation tests were performed using the pre-crystallisation test (PCT™) kit (Hampton Research).

The purified chCREB3 bZIP was mixed with the duplex DNA oligomers just before setting up trials. The molar ratio of bZIP dimer (0.5mM) to DNA duplex (0.5mM) was 1. Mixing of the protein and DNA resulted in precipitation of the protein, which was re-solubilised by addition of $MgCl_2$ to a final concentration of 30–50mM.

Protein	Sample buffer	Concentration	Screens
His-pp14B (aggregated form)	50mM Tris-HCl pH 7.5, 500mM NaCl	6.67mg/ml	Block 1, Block 3, Index, Morpheus 1 (4°C and 18°C)
His-pp14B (monomeric)	50mM Tris-HCl pH 7.5, 150mM NaCl	4mg/ml	Block 1, Block 2, Block 3, Block 4N, Block 5, Index, Morpheus 1, Limited, Nuc-Pro, SaltRx and PEGRx (4°C and 18°C)
His-pp14B (monomeric)	50mM Tris-HCl pH 7.5, 150mM NaCl, 10mM BME	4mg/ml	Block 1, Block 2, Block 3, Block 4N, Block 5, Index, Morpheus 1, Limited, Nuc-Pro, SaltRx and PEGRx (4°C and 18°C)
His-Cys-null pp14B (monomeric)	50mM Tris-HCl pH 7.5, 150mM NaCl	10mg/ml	Block 1, Block 2, Block 3, Block 5, PACT premier, Index, PGA, PE, SaltRx, PEGRx, Morpheus 1, Morpheus 2 (4°C and 18°C)
His- CT-chCREB3	50mM Tris-HCl pH 8, 500mM NaCl	4mg/ml	Block 1, Block 3, Index, Morpheus (4°C and 18°C)
His-chCREB3 bZIP	50mM Tris-HCl pH 8, 150mM NaCl	44mg/ml	Block 1, Block 2, Block 4N, Index, PACT premier, NucPro (4°C and 18°C)
His-chCREB3 bZIP/20-mer DNA	50mM Tris-HCl pH 8.0, 150mM NaCl, 30–50mM $MgCl_2$	bZIP dimer (0.5mM) to DNA duplex (0.5mM), 1:1	Block 4N, NucPro (4°C and 18°C)
His-chCREB3 bZIP/12-mer DNA	50mM Tris, pH 8.0, 150mM NaCl, 30–50mM MgCl2	bZIP dimer (0.5 mM) to DNA duplex (0.5 mM), 1:1	Block 4N, NucPro (4°C and 18°C)

Table 2.14 Crystallisation trials.
A list of the sample subjected to crystallisation trials giving information. The buffer composition, protein concentration used for crystallisation, and the names of screens set up for each sample are given in the table.

2.19.4 Crystallisation optimisation screens

To obtain larger, diffraction quality crystals of the chCREB3 bZIP, a three-row optimisation screen and a pH screen were conducted at 4°C around the condition:

50% v/v 2-Methyl-2,4-pentanediol (MPD), 100mM Tri-sodium citrate; pH 5.6, 10mM Magnesium chloride. In three-row optimisation, only F, G and H rows (36 wells) of a 96-well plate were used. The original screen was dispensed on the crystallisation plate reservoir wells and simultaneously also diluted with Milli-Q water in 1%-3% steps by Robogo (MWG Biotech, UK) leading to the formation of mother liquor, which contains 100% to 67% of the screen components. Crystallisation droplets were generated by the Cartesian robot (Cartesian Technologies Microsys MIC4000) programmed to optimise protein: mother liquor ratio by dispensing protein such that the final protein: mother liquor ratios were 1:1, 1:2 and 2:1 in rows F, G and H of the crystallisation plate, respectively. The plates were then immediately sealed manually with transparent self-adhesive foils and stored in the storage vault at 4°C. In the pH screen, the concentration of MPD was varied against pH (50–100% MPD vs pH 3–8) in a 96-well plate.

2.19.5 Cryoprotection

The pp14 crystals were harvested from the drop, briefly soaked in the reservoir solution supplemented with 25% glycerol and flash frozen in liquid nitrogen. The chCREB3 bZIP crystals were frozen by immersing in the reservoir solution, which contained MPD, without any additional cryoprotection.

2.19.6 Data collection and processing

All diffraction data for the chCREB3 bZIP were collected on beamline I04 at Diamond Light Source. For structure solution, 180° of data collected on a single crystal of the chCREB3 bZIP were auto-processed using Xia2 (Winter et al., 2013),

integrated using DIALS (Waterman et al., 2013) and scaled and merged using Aimless (Evans, 2006). The structure was solved by molecular replacement with the program PHASER (McCoy et al., 2007) using a dimer extracted from the CREB bZIP-CRE complex structure (PDB code: 1DH3) after deletion of the bound duplex DNA. The structure was refined by iterative rounds of manual model building in COOT (Emsley and Cowtan, 2004) and using BUSTER (Smart et al., 2012). Data collection and refinement statistics for the chCREB3 bZIP are summarised in Table 5.1 of Chapter 5. The pp14 crystals were tested on the I02 beamline.

2.20 X-ray fluorescence (XRF) spectroscopy

Elemental analysis of the pp14 crystals was carried out using synchrotron radiation X-ray fluorescence (XRF) at the I02 microfocus beamline of Diamond Light Source (Didcot, UK). The data were treated with the XRF analysis package PyMCA.

2.21 Bioinformatic analyses

bZIP sequences were obtained from the Uniprot database (http://www.uniprot.org/). Multiple alignments were performed using Clustal Omega (Sievers et al., 2011).

3 Interaction between pp14 and chCREB3

3.1 Introduction

MDV induces a wide range of neurological syndromes in susceptible hosts. However, there is a lack of understanding of the underlying mechanisms and information on viral and host factors that contribute to neuropathology. The immediate-early 14kDa phosphoprotein, pp14, is associated with the neurovirulence phenotype of the virus (Tahiri-Alaoui et al., 2012). Yeast-two-hybrid screening identified the ER-bound transcription regulator, human CREB3, as an interacting partner of pp14, and fluorescence colocalisation between pp14 and chicken CREB3 in RB-1B/HA-chCREB3 infected cells indicated a potential interaction between these proteins, leading to the hypothesis that this interaction could be linked to the neurovirulence of MDV. To further investigate this putative interaction between pp14 and chicken CREB3, various *in vitro* experiments were undertaken. The results these experiments are presented in this chapter.

3.2 Interaction of pp14 with chCREB3 in MDV infected cells

3.2.1 Biochemical characterisation of the interaction

As described in Section 1.4.3 of Chapter 1, initially in order to investigate whether pp14 interacts with chCREB3, immunofluorescence was carried out with RB-1B/HA-chCREB3 infected CEF cells, and fluorescence colocalisation was observed between pp14 and chCREB3 indicating a potential interaction between these proteins. To confirm the interaction, co-immunoprecipitations (co-IPs) were carried out. Co-immunoprecipitation identifies protein-protein interactions using target protein-specific antibodies to indirectly capture proteins that are bound to

the target protein. The procedure involves using a target-specific antibody to separate the target protein and the associated proteins from all the other proteins in the cell lysate (lysate). The immune complex is then washed after removing the unbound (UB) material. The immunoprecipitated (IP) proteins are separated by gel electrophoresis, which are then analysed by immunoblotting with specific antibodies to identify binding partners.

For analysis, CEF cells were infected with RB-1B/HA-chCREB3, and the cell lysate prepared from the infected cells was subjected to immunoprecipitation with rabbit anti-pp14 (a polyclonal antibody raised against a peptide of pp14). The immunoprecipitated proteins were probed with anti-pp14 and anti-HA to detect pp14 and HA (N-ter) chCREB3, respectively. Co-immunoprecipitation of chCREB3 and pp14 was observed indicating an interaction between pp14 and chCREB3 (Figure 3.1). A modified form of chCREB3, corresponding to the size of the N-terminal domain of chCREB3 (NT-chCREB3), was also detected in the cell lysate, and the lane containing proteins immunoprecipitated with anti-pp14 (Figure 3.1), indicating the NT domain of chCREB3 could be interacting with pp14. The size of the NT-chCREB3 was determined by western blot analysis performed with cell lysate prepared from DF1 cells transiently transfected with HA (N-ter) NT-chCREB3 (Figure 3.3).

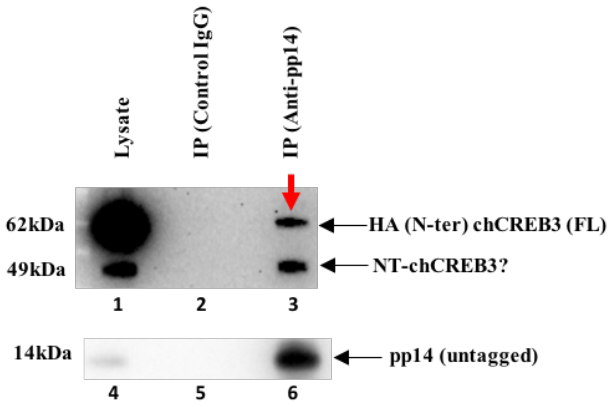


Figure 3.1 pp14 interacts with chCREB3 in MDV infected cells.
CEF cells were infected with RB-1B/HA-chCREB3. The lysate prepared from these cells was subjected to co-immunoprecipitations with anti-pp14 and control IgG. The immunoprecipitates were immunoblotted with anti-HA (lanes 1-3) and anti-pp14 (lanes 4-6) antibodies.

3.2.2 Functional role of the interaction

Next, in an effort to determine the functional role of the interaction of pp14 with chCREB3 during MDV infection, a pp14 deletion mutant of pRB-1B/HA-chCREB3 (denoted pRB-1BΔpp14/HA-chCREB3) was generated. The engineered BAC clones of RB-1B were kindly provided by Benedikt Kaufer of Free University of Berlin. By infecting CEF cells with RB-1B/HA-chCREB3 or RB-1BΔpp14/HA-chCREB3, the localisation of chCREB3 could be assessed in the presence and absence of pp14 and compared. CREB3, which is anchored to the ER membrane in an inactive form (Chan et al., 2011), is activated through regulated intramembrane proteolysis (RIP). RIP liberates the N-terminal fragment of CREB3 from the membrane allowing it to translocate to the nucleus (Chan et al., 2011). Since activation of CREB3 results in a change in its localisation pattern, i.e. both ER-bound/cytoplasmic and nuclear forms can be observed upon activation (Chan et al., 2011), a hypothesis was proposed that in MDV infected cells the association of pp14 with chCREB3 could modify the localisation of the ER-bound factor. The role of the interaction could be either to prevent or induce its activation.

To investigate this hypothesis, infection was induced by transfecting CEF cells with BAC DNA of either pRB-1B/HA-chCREB3 or pRB-1BΔpp14/HA-chCREB3. As described in Sections 2.6 and 2.10.2 of Chapter 2, the transfection of purified MDV BAC DNA into CEF cells leads to reconstitution of infectious virus. Five days after transfection, because MDV plaques typically appear 4–5 days after transfection, the cells were fixed, and immunostained with appropriate antibodies against proteins of interest (Section 2.13.1) and the localisation of chCREB3 was examined using confocal microscopy. Fifty infected cells were assessed per experimental condition (Figure 3.2).

The localisation of chCREB3 was found to be unaffected by pp14 in MDV infected CEF cells (Figure 3.2A). While the majority of RB-1B/HA-chCREB3 or RB-1BΔpp14/HA-chCREB3 infected cells exhibited ER-bound/cytoplasmic chCREB3 only (Figure 3.2A, C, D, F), the activated nuclear form of chCREB3 was detected in a small fraction of the infected cells (approximately 20% and 10% of the RB-1B/HA-chCREB3 and RB-1BΔpp14/HA-chCREB3 infected cells, respectively) (Figure 3.2 C, E, G).

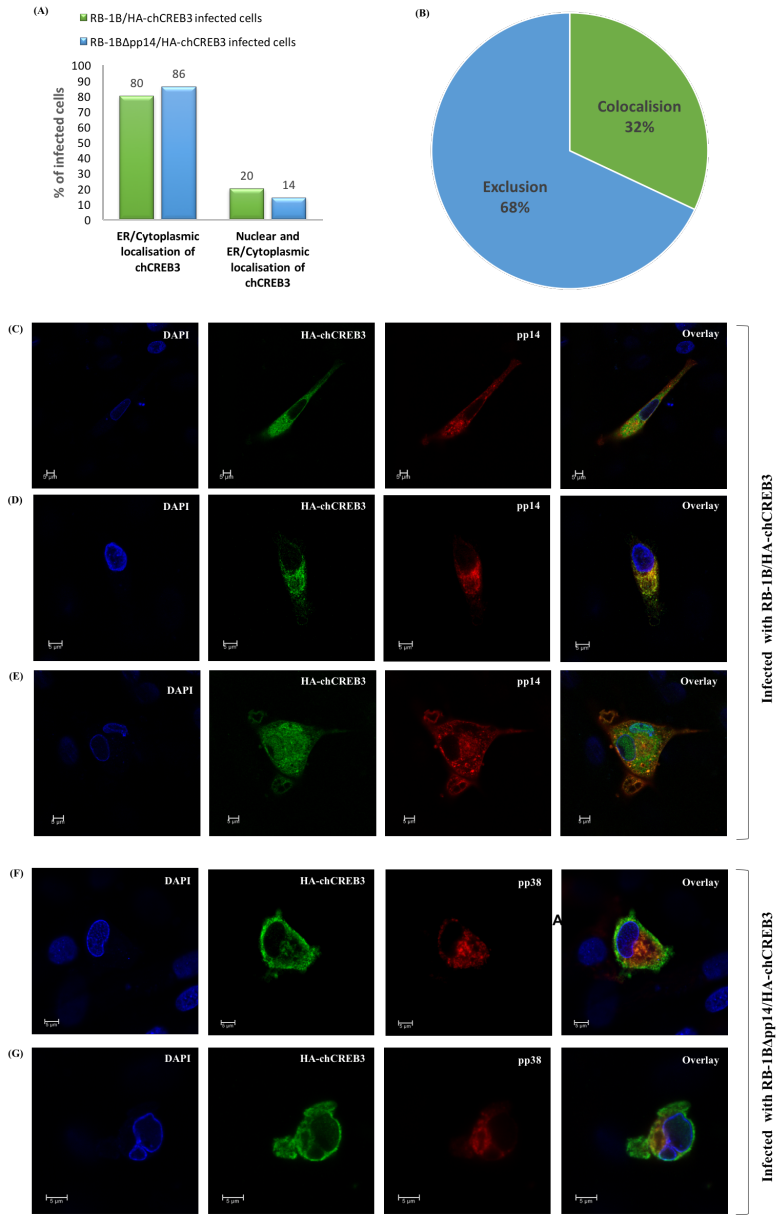


Figure 3.2 Localisation of chCREB3 was unaffected by pp14 in MDV infected CEF cells.
The localisation of chCREB3 was assessed in CEF cells infected with either RB-1B /HA-chCREB3 or RB-1BΔpp14/HA-chCREB3. Fifty cells were analysed per condition. The RB-1B/HA-chCREB3 infected CEF cells were stained with rabbit anti-pp14 and mouse anti-HA followed by secondary anti-rabbit Alexa 568 and anti-mouse Alexa 488. The RB-1BΔpp14/HA-chCREB3 infected CEF cells were stained with mouse anti-pp38 (to indicate infection) and rabbit anti-HA followed by secondary anti-mouse Alexa 568 and anti-rabbit Alexa 488. All cells were co-stained with DAPI. **(A)** The bar chart shows the results of the chCREB3 localisation analysis. The majority of infected cells (RB-1B/HA-chCREB3 or RB-1BΔpp14/HA-chCREB3) exhibited ER-bound/cytoplasmic chCREB3 only (**C, D and F** show representative cells), and a subpopulation exhibited both ER-bound/cytoplasmic and nuclear chCREB3 (**E and G** show representative cells). **(B)** The pie chart shows the results of the colocalisation analysis performed with fifty CEF cells infected with RB-1B/HA-chCREB3. Colocalisation of pp14 and chCREB3 was observed only in a subpopulation of the infected cells (D shows a representative cell exhibiting colocalisation). (Scale bars = 5µm)

Colocalisation of pp14 and chCREB3 was also assessed in RB-1B/HA-chCREB3 infected cells by assessing the colour overlay of the two fluorescent markers (green and red fluorescence resulting in a yellow colour). Surprisingly, colocalisation of pp14 and chCREB3 was not observed in the majority of infected cells. Only 32% of the RB-1B/HA-chCREB3 infected CEF cells exhibited colocalisation (Figure 3.2B,C,D,E).

It was difficult to draw clear conclusions from this investigation as it is not known whether chCREB3 is expressed in MDV infected cells *in vivo*, and if so whether MDV infection would suppress or promote activation of chCREB3. The role of the interaction could not be inferred from this experiment as the expression of pp14 was found to have no effect on the localisation of chCREB3. Because an interaction was detected between pp14 and chCREB3 by co-immunoprecipitation, it was predicted that colocalisation would be observed in all infected cells. However, colocalisation was observed only in a subpopulation of the infected cells. This observation raised the question whether the detected interaction could be a transient interaction, one that may be occurring at a particular time-point post infection. However, this could not be investigated further as time-course experiments were not easy due to the cell- associated nature of the virus.

3.3 Interaction studies with recombinant proteins

3.3.1 Determination of the domain of chCREB3 that interacts with pp14

While fluorescence colocalisation was observed in a subpopulation of the MDV infected CEF cells, and the co-immunoprecipitation data supported an interaction

between pp14 and chCREB3 in these cells, the results of the functional assay devised to identify the role of the interaction were inconclusive. However, to further examine the putative interaction, and with the aim to characterise the interacting domain of chCREB3, co-immunoprecipitation assays were carried out with recombinant proteins, which were co-expressed in DF1 cells (a continuous cell line of chicken embryo fibroblasts). For this investigation, the gene encoding pp14B was cloned into pcDNA3.1. Since the Y2H screen used the common C-terminal region of pp14 as bait, both forms were considered suitable for the purpose of interaction studies. pp14B was cloned without a tag because of its small size—a tag attached to a small protein can sometimes interfere with folding and function. The NT (residues 2–288) and CT (residues 312–441) domains of chCREB3 were cloned into pcDNA3.1 with a N-terminal HA tag (YPYDVPDYA) [denoted HA (N-ter) NT-chCREB3 and HA (N-ter) CT-chCREB3]. Intended to serve as an ER-bound cytoplasmic form, an additional construct of the NT-chCREB3 containing the transmembrane region (residues 2–312) was also produced [denoted HA (N-ter) NT-chCREB3-TM].

For co-immunoprecipitation analysis, cell lysates were prepared from DF1 cells co-transfected with pp14B (untagged) and HA (N-ter) NT-chCREB3, HA (N-ter) NT-chCREB3-TM or HA (N-ter) CT-chCREB3, and the lysates were subjected to co-immunoprecipitations with anti-pp14. The immunoprecipitates were immunoblotted with anti-pp14 and anti-HA to detect pp14B and the HA tagged chCREB3 constructs, respectively. Co-immunoprecipitation of pp14B with both constructs of the NT-chCREB3 (Figure 3.3A, C) was observed—indicating the viral protein targets the functionally important domain of the transcription factor.

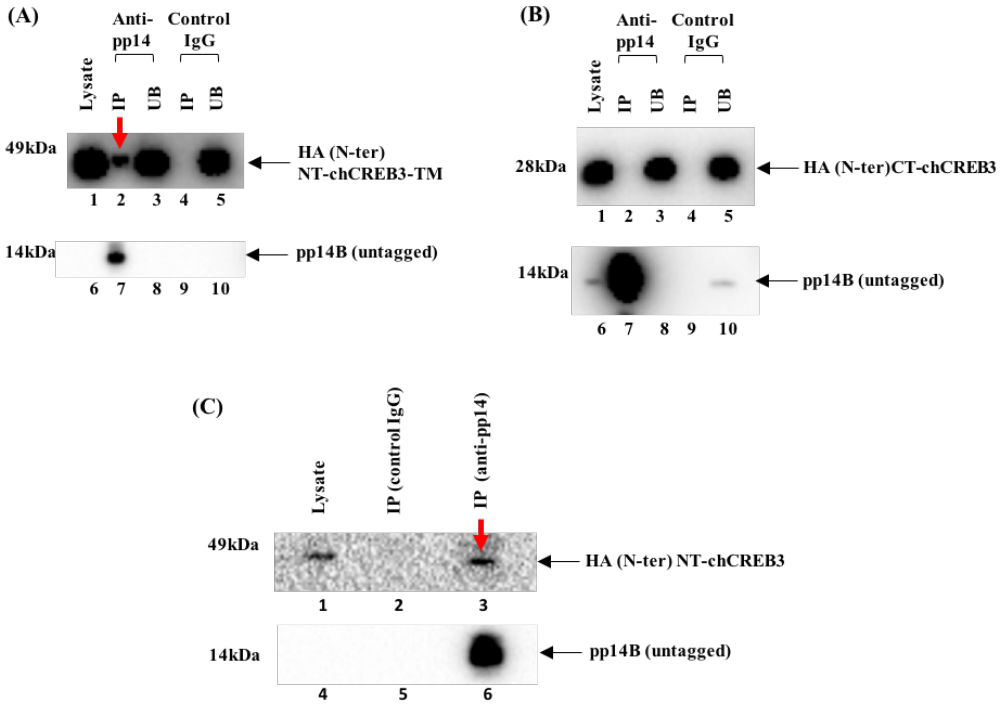


Figure 3.3 pp14 interacts with the N-terminal domain of chCREB3.
Co-immunoprecipitation assays were carried out with cell lysates prepared from DF1 cells co-transfected with **(A)** HA (N-ter) NT-chCREB3-TM and pp14B (untagged), **(B)** HA (N-ter) CT-chCREB3 and pp14B (untagged) and **(C)** HA (N-ter) NT-chCREB3 and pp14B (untagged) using anti-pp14 and control IgG. The immunoprecipitates were immunoblotted with anti-HA [lanes 1-5 of **(A)** and **(B)**, lanes 1-3 of **(C)**] and anti-pp14 [lanes 6-10 of **(A)** and **(B)**, lane 4-6 of **(C)**]. The red arrows in **(A)** and **(C)** indicate the co-immunoprecipitation of pp14 with the NT-chCREB3 constructs.

3.3.2 Reverse co-immunoprecipitation

To verify the interaction detected between pp14B and the NT domain of chCREB3, reverse co-immunoprecipitation targeting the NT-chCREB3 was performed. DF1 cells were co-transfected with pp14B (untagged) and HA (N-ter) NT-chCREB3-TM, and the cell lysate prepared from these cells was subjected to immunoprecipitations with both anti-HA, and anti-pp14 and the immunoprecipitates were immunoblotted with anti-pp14 and anti-HA. Co-immunoprecipitation of pp14 and the NT-chCREB3-TM construct was observed with co-immunoprecipitation using anti-pp14, but the reverse co-

immunoprecipitation with anti-HA failed to detect an interaction between pp14 and the NT-chCREB3-TM construct (Figure 3.4).

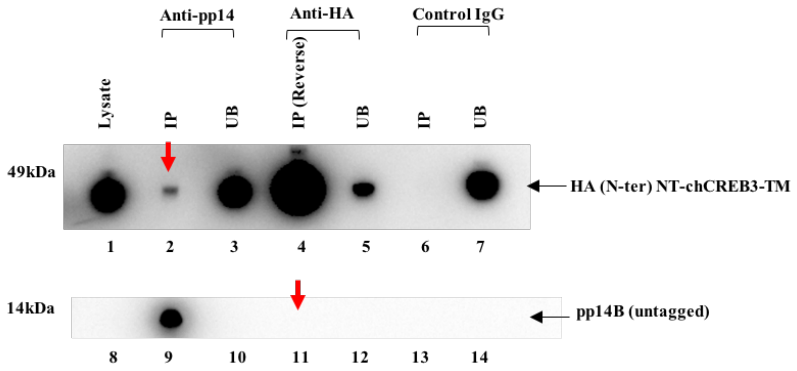


Figure 3.4 Reverse co-immunoprecipitation failed to detect an interaction between pp14 and the N-terminal domain of chCREB3.
DF1 cells were co-transfected with HA (N-ter) NT-chCREB3-TM and pp14B (untagged). The lysate prepared from these cells was subjected to co-immunoprecipitations with anti-pp14, anti-HA and control IgG. The immunoprecipitates were immunoblotted with anti-HA (lanes 1-7) and anti-pp14 (8-14). The red arrows indicate where a band was expected if the other protein co-immunoprecipitated with the target protein.

These contradictory results needed further investigation using alternative approaches. A negative reverse co-IP could result from steric hindrance that may prevent an antibody from binding to the target region within the protein complex during co-immunoprecipitation, and in such cases, it would be necessary to probe by using antibodies against different regions of the target protein. As this project also aimed to characterise the interaction at the structural level, a way to move forward, and to verify the putative interaction was to determine if a complex of pp14B bound to the NT-chCREB3 could be purified, and analysed by size exclusion chromatography (SEC), which separates biomolecules based on their size.

However, it was recognised that a number of limitations are associated with cell-free biochemical techniques used for co-purification. An important limitation is that detection of the bound proteins is dependent on the stability of protein-protein interactions. The ability to maintain stable physiological interactions throughout the mechanical and chemical stresses of the purification process is a critical factor. Due to this limitation, low-affinity interactions could go undetected. Therefore, in parallel, a cellular approach was also employed to investigate the putative interaction. The technique chosen for this investigation was Förster resonance energy transfer (FRET), a well-established technique that provides high spatial resolution for detection of protein-protein interactions in living cells (Bacskai et al., 2003; Elangovan et al., 2002; Krishnan et al., 2003). The results of this investigation in living DF1 cells is presented in Section 3.6.

3.4 Expression and purification of the complex

Two approaches can be employed for co-purification of proteins. One method is *in vitro* complex formation, which is attained by mixing separately purified recombinant proteins. Alternatively, preformed protein complexes can be isolated after co-expression of the binding partners. With either approach, to separate the bound (target protein bound to its ligand) from unbound species, the interacting proteins are required to be tagged with two different affinity/epitope tags.

For expression and purification, the genes encoding the NT-chCREB3 and pp14B were cloned into pOPIN expression vectors harbouring different fusion tags (Section 2.9.4). Although pp14 constructs expressed in *E. coli*, the NT-chCREB3 constructs did not express in *E. coli* (results obtained from Louise Bird of OPPF).

Therefore, recombinant baculoviruses expressing different fusion constructs of pp14B and the NT-chCREB3 were generated (Sections 2.15.3.1), and an expression screen was conducted in insect cells (Section 2.15.3.2). Although the pp14B fusion constructs expressed in insect cells, which could also be scaled-up and purified, the expression of the NT-chCREB3 constructs that were found to express in insect cells [HA (C-ter) NT-chCREB3, His (C-ter) NT-chCREB3 and FLAG (C-ter) NT-chCREB3] could not be scaled up for purification (results not shown). This expression issue of the NT-chCREB3 constructs made *in vitro* complex formation not possible.

Alternatively, to co-purify the proteins, a co-expression screen was conducted in insect cells. Those fusion constructs of pp14B and the NT-chCREB3 that were found to express in insect cells from the initial expression screen were used. Co-expression was tested with different combinations of fusion tags, and the following combinations were found suitable for co-expression: 1) FLAG (C-ter) pp14B and HA (C-ter) NT-chCREB3, 2) His-GST (N-ter) pp14B and FLAG (C-ter) NT-chCREB3, and 3) His-GST (N-ter) pp14B and His (C-ter) NT-chCREB3.

Due to the small size of HA (YPYDVPDYA, 1102Da) and FLAG (DYKDDDDK, 1012 Da) tags (not expected to interfere with structural and functional studies), the FLAG (C-ter) pp14B and HA (C-ter) NT-chCREB3 pair was chosen for co-expression and purification. A small-scale pulldown experiment was performed first. For this, FLAG (C-ter) pp14B and HA (C-ter) NT-chCREB3 were co-expressed in 2ml of Sf9 cells, and the cell lysate prepared from these cells was used to pull down the FLAG tagged pp14B using an anti-FLAG antibody. The eluates

were immunoblotted with anti-FLAG and anti-HA to detect FLAG (C-ter) pp14B and HA (C-ter) NT-chCREB3, respectively. A band corresponding to pp14 was detected, however as expected a band corresponding to the NT-chCREB3 was not observed, indicating there was no co-elution of the NT-chCREB3 and pp14B (Figure 3.5).

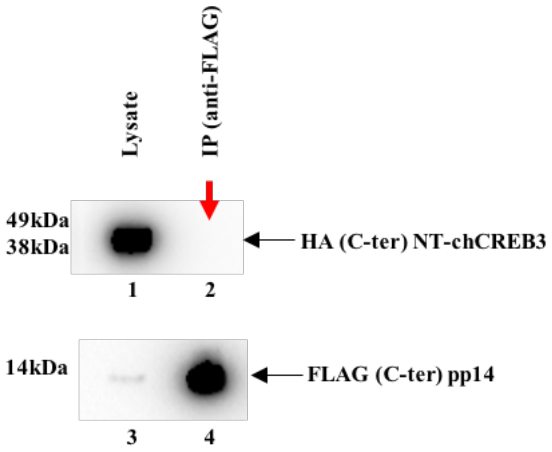


Figure 3.5 Purification using an anti-FLAG antibody.
Sf9 cells were co-infected with baculoviruses expressing FLAG (C-ter) pp14B and HA (C-ter) NT-chCREB3. The FLAG tagged pp14B was pulled down using an anti-FLAG antibody from the cell lysate prepared from these cells. The eluates were immunoblotted with anti-HA (lanes1,2) and anti-FLAG (3,4). The red arrow indicates where a band was expected if the proteins formed a complex

As an alternative, GST-pull down experiments were also carried out with cell lysates prepared from insect cells co-infected with baculoviruses expressing His-GST (N-ter) pp14B and FLAG (C-ter) NT-chCREB3, and His-GST (N-ter) pp14B and His (C-ter) NT-chCREB3. Glutathione-sepharose beads were used to pull down the GST tagged pp14B from the lysates. However, co-elution of the NT-chCREB3 and pp14B was not observed (Figure 3.6).

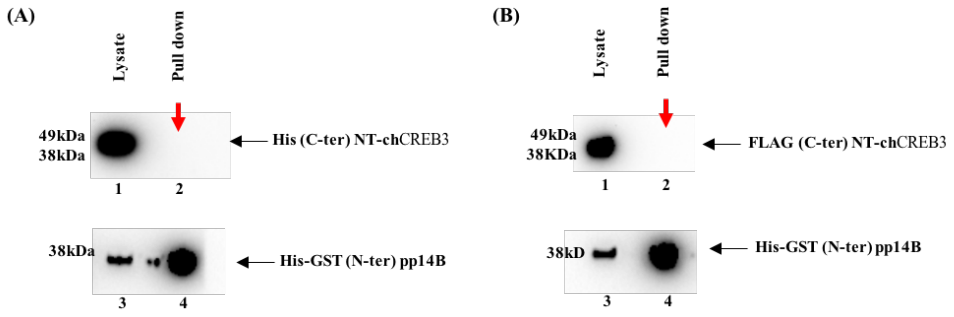


Figure 3.6 GST-pull down assays.
Sf9 cells were co-infected with baculoviruses expressing **(A)** His-GST (N-ter) pp14B and His (C-ter) NT-chCREB3 and **(B)** His-GST (N-ter) pp14B and FLAG (C-ter) NT-chCREB3. The GST tagged pp14B was pulled down using glutathione-agarose beads from cell lysates prepared from these cells. The eluates were immunoblotted with anti-His [lanes 1,2 of **(A)**], anti-FLAG [lanes 1,2 of **(B)**] and anti-GST [lanes 3,4 of **(A)** and **(B)**]. The red arrow indicates where a band was expected if the proteins formed a complex.

Many factors could have caused irreproducibility. As the previous pull-down experiments were carried out using avian cells, it may be that a host cell factor was a requirement for the interaction to occur. In previous experiments, untagged pp14 was used, and therefore it was possible the FLAG or GST tag added to pp14B could have sterically hindered the interaction. As previously anti-pp14 was used for pulling down pp14B, it was possible that the anti-FLAG antibody, targeting a different region of the target protein, could have been sterically hindered to interact with the epitope (which may be buried) within the complex.

To rule out the possibility that steric hindrance may have prevented anti-FLAG from interacting with the epitope region on the target protein within the complex, a pull down was attempted using anti-pp14—the same antibody used in previous co-immunoprecipitation experiments. FLAG (C-ter) pp14B and HA (C-ter) NT-chCREB3 were co-expressed in insect cells, and the cell lysate prepared from these cells was subjected to immunoprecipitations with both anti-pp14, and anti-FLAG (both antibodies targeting pp14). The immunoprecipitates were analysed using

anti-pp14 and anti-HA to detect FLAG (C-ter) pp14B and HA (C-ter) NT-chCREB3, respectively. While bands corresponding to pp14 and the NT-chCREB3 were observed in the lane containing proteins immunoprecipitated with anti-pp14, a band corresponding to the NT-chCREB3 was not observed in the lane containing proteins immunoprecipitated with anti-FLAG although a band corresponding to pp14 was present (Figure 3.7). These results indicated the FLAG tag attached to pp14 was not hindering the interaction, and a host cell factor was not required to mediate the interaction. However, because both pp14B and the NT-chCREB3 could be detected following immunoprecipitation with anti-pp14 only, possible cross-reactivity was suspected between anti-pp14 and the N-terminal domain of chCREB3. Further experiments were carried out to rule out the possibility of cross-reactivity.

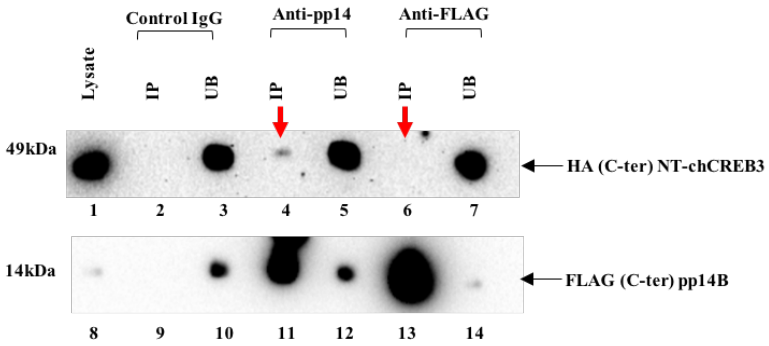


Figure 3.7 Both pp14 and the NT domain of chCREB3 are detected only when anti-pp14 is used.
Sf9 cells were co-infected with baculoviruses expressing FLAG (C-ter) pp14B and HA (C-ter) NT-chCREB3. The lysate prepared from these cells was subjected to co-immunoprecipitations with anti-pp14, anti-FLAG and control IgG. The immunoprecipitates were immunoblotted with anti-HA (lanes 1-7) and anti-pp14 (lanes 8-14). The red arrow indicates where a band was expected if the proteins formed a complex.

3.5 Antibody cross-reactivity

To investigate potential cross-reactivity between anti-pp14 and the N-terminal domain of chCREB3, DF1 cells were transfected with HA (N-ter) NT-chCREB3-

TM alone and the lysate prepared from these cells was subjected to immunoprecipitation with anti-pp14. The immunoprecipitated proteins were probed with anti-HA to determine if the HA tagged chCREB3 construct could be detected. A strong band corresponding to the size of the NT-chCREB3-TM construct was observed indicating cross-reactivity between anti-pp14 and the N-terminal domain of chCREB3 (Figure 3.8).

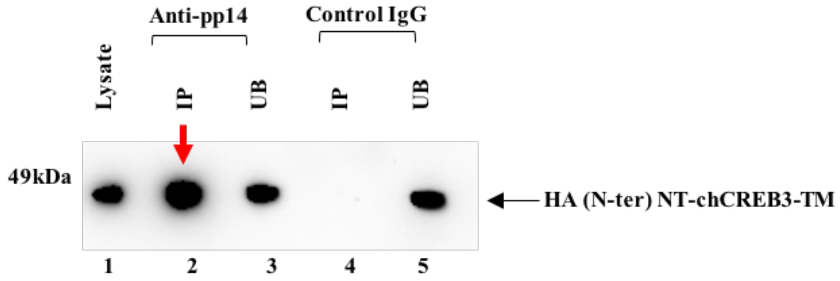


Figure 3.8 Anti-pp14 cross-reacts with the N-terminal domain of chCREB3. DF1 cells were transfected with HA (N-ter) NT-chCREB3-TM alone. The lysate prepared from these cells was subjected to immunoprecipitations with anti-pp14 and control IgG. The immunoprecipitates were analysed using an anti-HA antibody. The cross-reactivity observed between anti-pp14 and the NT domain of chCREB3 is indicated by the red arrow.

In order to confirm the cross-reactivity observed, DF1 cells were transiently transfected with HA (N-ter) NT-chCREB3-TM alone and the lysate prepared from these cells was subjected to immunoprecipitation with anti-pp14 in the presence and absence of the pp14 peptide (EVLAEIPQEKEVESTL) used for antibody production. In the presence of the pp14 peptide, a band corresponding to the NT-chCREB3 was not observed (Figure 3.9), which indicted the pp14 peptide competitively inhibited binding of the NT-chCREB3 to the antibody. These results indicate an interaction between the N-terminal domain of chCREB3 and the antibody epitope suggesting the residues that are common between the sequences of the NT-chCREB3 and the peptide of pp14 could be responsible for the cross-

reactivity (Figure 3.10). Importantly, cross-reactivity between the proteins was observed only in solution. The anti-pp14 was not found to be cross-reactive when it was probed against denatured NT-chCREB3 on a western blot (Figure 3.11A), suggesting the antibody could be interacting with a conformational epitope. Cross-reactivity was also not observed by immunofluorescence when fixed DF1 cells transiently transfected with mRuby tagged NT-chCREB3 construct alone were probed with anti- pp14 (Figure 3.11B,C,D).

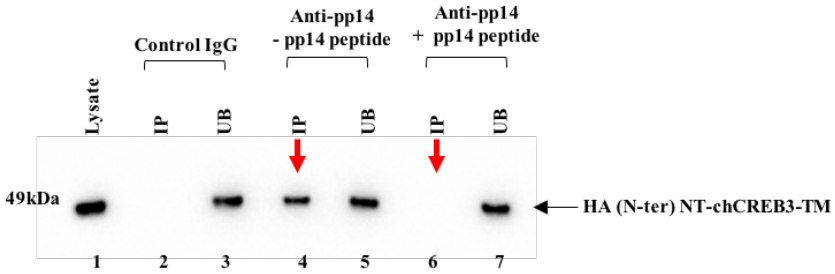


Figure 3.9 Anti-pp14 epitope interacts with the N-terminal domain of chCREB3 in solution.
DF1 cells were transfected with HA (N-ter) NT-chCREB3-TM alone. The lysate prepared from these cells was subjected to immunoprecipitations with anti-pp14 in the presence and absence of the pp14 peptide (EVLAEIPQEKEVESTL) and control IgG. The immunoprecipitates were analysed with an anti-HA antibody.

(A) Anti-pp14 peptide (EVLAEIPQEKEVESTL)

(B) Chicken CREB3 (Green = NT domain, Blue/underlined = TM domain, Yellow = CT domain)

```
        10         20         30         40         50         60
MAGPEPLAAL PDGDLLDFLL NDDAPAAGSP AGDGPLLGDW VLPELELLDK EMDSFISSLL

        70         80         90        100        110        120
SPSEDEPCLL QGCLPADSDS STSEDQNPFH SSGSDVASSP QSSDIVQFDH NYSLHQDWPA

       130        140        150        160        170        180
LESVRSDMAE GDVSIDLETW MCLEGTNETL QESCSFPVTV AVDAGPQLMP GAPMQSNFPE

       190        200        210        220        230        240
LVLTEEERQL LEKDGVSLPT CLPLTKAEER LLKKVRRKIR NKQSAQDSRR RKKIYVDGLE

       250        260        270        280        290        300
NRVAACTAQN HELQKKVQLL QKQNMSLLEQ LRKLQALVRQ STTKTTTAST CIMVLVLSFC

       310        320        330        340        350        360
LILSPSLYSF GSRGPQPEFR VLSRQIREFP NKVWQAAPAV QEEVVLERLS PEPEDTSLGS

       370        380        390        400        410        420
LNQSREEGQS PPKPDPRSAF NGNSSSDPPA GSELGHPQPQ EQQSRRDALH SEVLAPWKDE

       430        440
NQEWVERTTS VVIQQHRTDE M
```

Figure 3.10 Sequence comparison between the pp14 peptide and chCREB3. **(A)** Amino acid sequence of the pp14 peptide used for antibody production. **(B)** Amino acid sequence of chCREB3. The NT, TM and CT domains of chCREB3 are coloured as indicated in the figure. The residues that are common between the pp14 peptide and the NT-TM domain only are highlighted in red. The residues that are common between the pp14 peptide and the CT domain only are highlighted in red. The residues that are common between the pp14 peptide, and both NT- and CT-domains of chCREB3 are shown in bold face/highlighted in magenta.

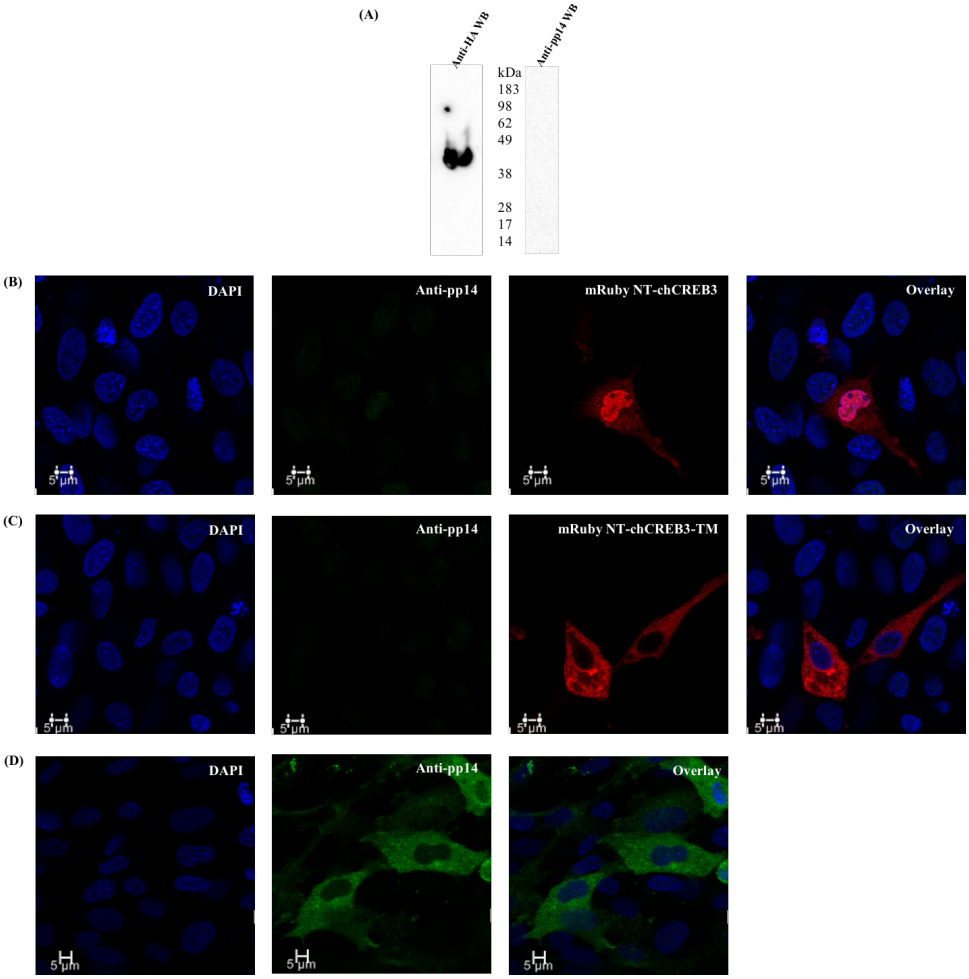


Figure 3.11 Cross-reactivity was not detected between anti-pp14 and the N-terminal domain of chCREB3 on a western blot or by immunofluorescence. **(A)** DF1 cells were transfected with HA (N-ter) NT-chCREB3-TM alone, and the whole cell lysate prepared from these cells was subjected to western blotting (WB) analysis with anti-HA and anti-pp14. DF1 cells transiently transfected with mRuby tagged NT-chCREB3 alone **(B)**, mRuby tagged NT-chCREB3-TM alone **(C)**, or untagged pp14 alone **(D)** were fixed 24 hours after transfection and stained with rabbit anti-pp14, followed by secondary anti-rabbit Alexa 488 and co-stained with DAPI. (Scale bars = 5µm)

3.6 Interaction between pp14 and chCREB3 in living cells

Conflicting results were observed with co-immunoprecipitation assays conducted with cell lysate prepared from DF1 cells co-transfected with pp14B (untagged) and the N-terminal domain of chCREB3 (HA tagged). While co-immunoprecipitation with anti-pp14 indicated an interaction between pp14 and the N-terminal domain of chCREB3, the reverse co-immunoprecipitation with anti-HA targeting the HA tagged NT-chCREB3 failed to detect an interaction. These contradictory results prompted further investigation using alternative techniques to verify the putative interaction.

An alternative biochemical approach, aimed at co-purifying the proteins for SEC analysis, was chosen to further investigate the putative interaction (Section 3.4); however, it was recognised that it was also important to address the limitations associated with biochemical experiments in detecting protein-protein interactions. An important limitation associated with techniques used for co-purification of proteins or isolation of protein complexes is that detection of the bound proteins is dependent on the stability of protein-protein interactions and their ability to withstand the mechanical and chemical stress of the procedures. Further, attaching epitope/affinity tags to the binding partners to facilitate isolation may present with other issues. For example, interaction between two proteins could be hindered by steric effects exerted by the tag moiety on either of the interacting proteins, and sometimes steric hindrance may prevent an antibody binding to the epitope region on the target protein. Another major limitation of biochemical experiments is that they are performed outside the cellular environment. If the proteins are not in their natural physiological environment, they may not fold properly, or may not have

optimal conditions (e.g. pH, post-translational modifications, or cofactors) to interact. In consideration of these limitations, in parallel, a cellular technique was also employed to investigate the putative interaction detected between pp14B and the N-terminal domain of chCREB3 in DF1 cells. The technique chosen for this investigation was Förster resonance energy transfer by fluorescence lifetime imaging (FRET-FLIM), which is a powerful cellular technique that provides high spatial resolution for detection of protein-protein interactions in living cells (Bacskai et al., 2003; Elangovan et al., 2002; Krishnan et al., 2003).

3.6.1 Introduction to FRET-FLIM

FRET is based on the ability of a donor fluorescent molecule in an excited electronic state to transfer some of its excitation energy to an acceptor fluorescent molecule (or chromophore) in its vicinity, in a non-radiative way (Figure 3.12) (Kappel, 2012; Vogel et al., 2006).

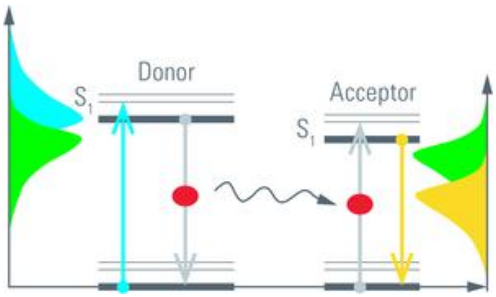


Figure 3.12 Förster resonance energy transfer.
The donor molecule is excited upon absorption of a photon of appropriate energy (blue arrow). The excited donor can relax either by fluorescence (grey arrow, left) or by resonance energy transfer to the acceptor molecule (black arrow). (Picture taken from Kappel, 2012)

Three conditions must be met for efficient FRET: 1) The donor and acceptor molecules must be within ~1–10 nanometre (nm) of each other (Figure 3.13A,B) (Kappel, 2012; Lleres et al., 2007). As directly interacting proteins are typically positioned within this distance, FRET measurements are a valuable tool for probing molecular interactions (Kappel, 2012; Lleres et al., 2007). For protein-protein interaction studies by FRET, the protein of interest is tagged with the donor molecule, and the test protein is tagged with the acceptor molecule; 2) The absorption spectrum of the acceptor must overlap the fluorescence emission spectrum of the donor (Figure 3.13C) (Kappel, 2012; Vogel et al., 2006); 3) The donor and acceptor molecules must have the appropriate relative orientation (Kappel, 2012; Vogel et al., 2006).

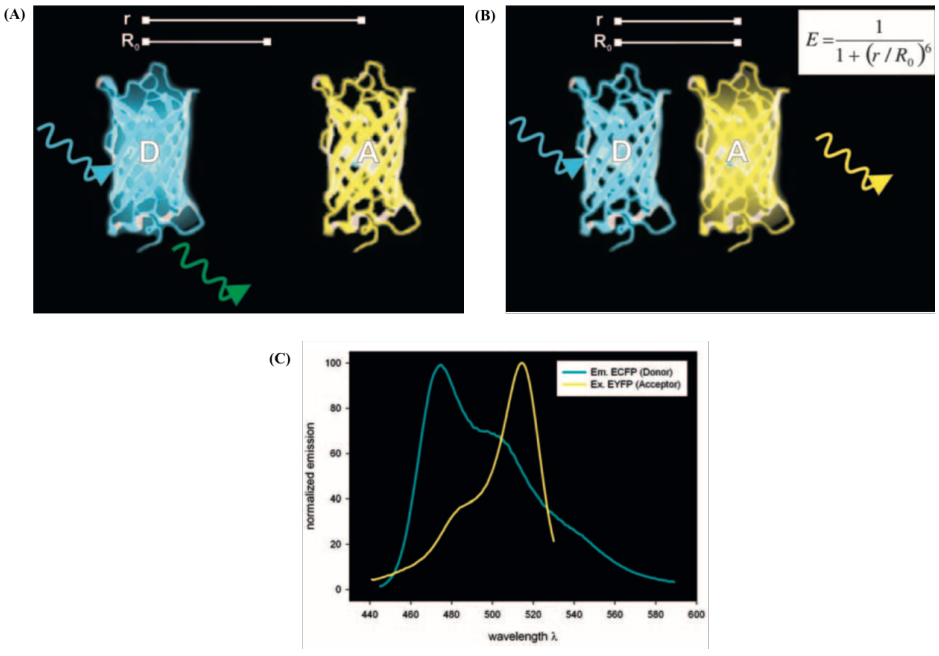


Figure 3.13 Principles of FRET.
(A) At a separating distance 'r' much larger than a threshold value R_0 (Förster radius), there is no energy transfer between the donor molecule 'D' and the acceptor molecule 'A'. **(B)** If the molecules are close to each other, energy from the exciting photon (blue arrow) is transferred to the acceptor in a non-radiative way. The excited molecule in turn emits a photon (yellow arrow). **(C)** Emission spectrum of donor (here ECFP, blue line) must overlap with excitation spectrum of acceptor (here EYFP, yellow line). (Pictures taken from Kappel, 2012)

The efficiency of the energy transfer ($\mathbf{E_{FRET}}$) is proportional to the inverse sixth power of the distance separating the donor and acceptor molecules (**r**), and is thus strongly dependent on the separation distance. It is given by, $\mathbf{E_{FRET} = 1/[1 + (r/R_0)^6]}$, where $\mathbf{R_0}$ (Förster radius), which varies for different fluorophores, is the distance at which energy transfer is 50% efficient (Lleres et al., 2007). Commonly used techniques to detect FRET in the context of microscopy are based on fluorescence intensity of either the donor (acceptor-photobleaching FRET) or the acceptor (sensitised-emission FRET) or the change in the fluorescence lifetime of the donor in the presence and absence of acceptor [FRET measured by fluorescence lifetime imaging (FLIM)] (Kappel, 2012; Vogel et al., 2006).

This FRET investigation was conducted using FLIM. A fluorescent molecule undergoes energy transitions from the ground state (S_0) to the excited state (S_1) upon absorption of a photon of a particular energy (Figure 3.14) (Kappel, 2012; Rajoria et al., 2014). The absorbed energy is stored by the fluorescent molecule for a short time before emitting as fluorescence (Kappel, 2012; Rajoria et al., 2014). The time a molecule spends in its excited state is defined as the fluorescence lifetime (Kappel, 2012), which is typically in the order of nanoseconds (10^{-9} s) for most fluorophores (Kappel, 2012; Rajoria et al., 2014). FLIM is based on the principle that FRET quenches the excited state, thereby shortening the fluorescence lifetime of the donor (Kappel, 2012; Vogel et al., 2006) (Figure 3.14). FRET can thus be determined by measuring the change in fluorescence lifetime of the donor in the presence and absence of an acceptor (Kappel, 2012; Vogel et al., 2006). FLIM generates an image based on the spatial distribution of the fluorophore (Broussard et al., 2013), therefore a major advantage of the FRET-FLIM technique

is that it both detects and spatially resolves protein-protein interactions within the context of a living cell (Kappel, 2012).

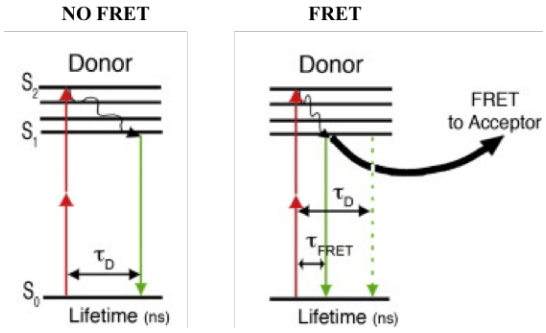


Figure 3.14 The principle of FLIM.
FLIM is based on the principle that occurrence of FRET quenches the excited state, thereby shortening the fluorescence lifetime of the donor. Left: Upon absorption of a photon of appropriate energy, the donor fluorophore changes from ground state (S_0) to the excited state (S_2). The absorbed energy is stored by the fluorescent molecule for a short time before emitting a photon (fluorescence). The time a molecule spends in its excited state is the donor fluorescence lifetime (τ_D). Right: Deactivation from the donor excited state can occur either by fluorescence (green downward-pointing arrow), or through non-radiative transfer of energy to the acceptor (FRET). FRET decreases the donor fluorescence lifetime (τ_{FRET}). (Adapted from Baird et al., 2014)

3.6.2 Construction of fluorescent fusion proteins

3.6.2.1 Construct design and cloning

This investigation was carried out with the donor-acceptor pair mClover-mRuby (Lam et al., 2012). mClover is a bright monomeric yellow-green fluorescent protein derived from the green fluorescent protein (GFP) isolated from *Aequorea victoria* (Shaner et al., 2013). The fluorescent protein mRuby is a bright monomeric red-fluorescent protein (RFP) derived from the tetrameric eqFP611 isolated from the sea anemone *Entacmaea quadricolor* (Hense et al., 2015). pp14B was labelled with the donor molecule mClover and the N-terminal domain of chCREB3 (NT-chCREB3) was labelled with the acceptor molecule mRuby. Both the nuclear and the ER-bound forms of the NT-chCREB3 (NT-chCREB3 and NT-

chCREB3-TM) were produced to probe the putative interaction. To serve as a negative control construct, the C-terminal domain of chCREB3 was produced tagged with mRuby. A short amino acid linker (Serine-Glycine-Leucine-Arginine-Serine-Arginine-Glycine-Aspartate-Proline-Proline-Valine-Alanine-Threonine) was included between the fluorescent protein and the protein of interest to provide flexibility (Figure 3.15).

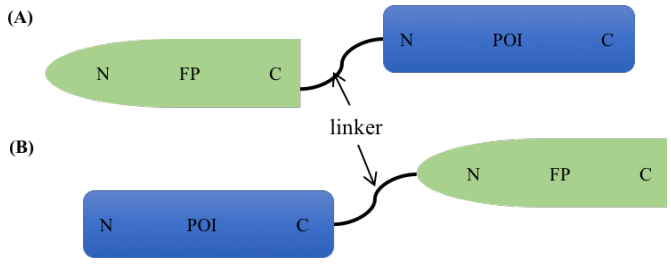


Figure 3.15 Schematic representations of the fluorescent protein fusion constructs.
The proteins of interest (POIs) were fused to either mClover or mRuby fluorescent protein (FP) at their N-terminus **(A)** or C-terminus **(B),** with a short amino acid linker sequence.

The genes encoding the fusion proteins were cloned into pcDNA3 or pcDNA3.1 for expression in DF1 cells. The GFP and RFP variants (~27kDa in size), unlike an epitope tag, are a significant addition to a protein which due to size could have steric consequences for protein folding, function, or targeting (Snapp, 2005). Hence, to maximise the likelihood of creating a functional and properly folded fluorescent fusion protein, both N-terminal and C-terminal fusions (Figure 3.15) of the proteins (except the NT-chCREB3-TM construct) were generated [mClover (N-ter) pp14B, mClover (C-ter) pp14B, mRuby (N-ter) NT-chCREB3, mRuby (C-ter) NT-chCREB3, mRuby (N-ter) CT-chCREB3, and mRuby (C-ter) CT-chCREB3]. The NT-chCREB3-TM truncation of the protein was only tagged with mRuby at its N-terminus [denoted mRuby (N-ter) NT-chCREB3-TM]. A C-terminally tagged NT-chCREB3-TM fusion protein was not made because the

fluorescent protein would be sterically constrained in such a construct with its one end attached to the ER and the other to the protein of interest.

3.6.2.2 Localisation of fluorescent fusion proteins

In order to ensure the constructed fluorescent fusion constructs localised to the expected subcellular compartments, their localisation was assessed in living DF1 cells using confocal microscopy. The pp14B construct with a N-terminal mClover tag was detected in the cytoplasm (Figure 3.16C) as expected, but the pp14B construct with a C-terminal mClover tag was detected both in the cytoplasm and nucleus (Figure 3.16D), indicating the placement of the fluorescent tag at this position affected the localisation of the protein.

The localisation of the mRuby tagged NT-chCREB3 constructs, mRuby tagged NT-chCREB3-TM construct and mRuby tagged CT-chCREB3 constructs was consistent with previous observations with the corresponding HA tagged versions by immunofluorescence (Figure 1.8). As expected, the NT-chCREB3 fusion constructs were detected predominantly in the nucleus as well as the cytoplasm (Figure 3.16E,F). The mRuby tagged NT-chCREB3-TM fusion protein colocalised with the ER-tracker dye (Figure 3.17), exhibiting ER localisation. The mRuby tagged CT-chCREB3 constructs were observed in the cytoplasm and nucleus (Figure 3.16G,H) as has been previously observed with a HA tagged version of the protein (Figure 1.8).

The localisation of unlinked mClover and mRuby was also assessed in DF1 cells. mClover and mRuby fluorescence was observed both in the cytoplasm and the

nucleus (Figure 3.16A,B) as has been previously observed (Berg and Beachy, 2008).

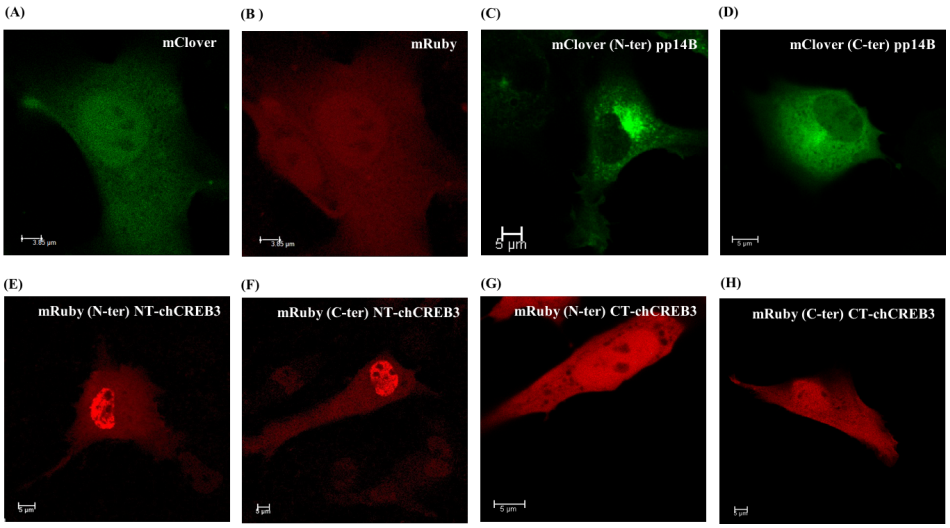


Figure 3.16 Localisation of the fluorescent fusion proteins in DF1 cells.
The localisation of the fluorescent or fluorescent fusion proteins was assessed in DF1 cells by confocal microscopy. **(A)** Unlinked mClover **(B)** Unlinked mRuby **(C)** mClover (N-ter) pp14B **(D)** mClover (C-ter) pp14B **(E)** mRuby (N-ter) NT-chCREB3 **(F)** mRuby (C-ter) NT-chCREB3 **(G)** mRuby (N-ter) CT-chCREB3 **(H)** mRuby (C-ter) CT-chCREB3 [Scale bars: 5µm (C-H) and 3.85µm (A and B)]

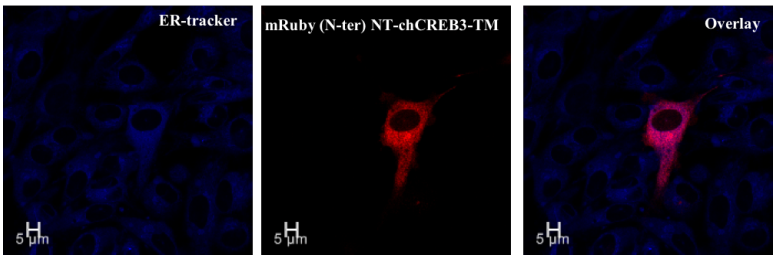


Figure 3.17 ER-localisation of mRuby tagged NT-chCREB3-TM in DF1 cells.
DF1 cells transfected with mRuby (N-ter) NT-chCREB3-TM were stained with the ER-Tracker Blue-White DPX dye, a live-cell stain that is highly selective for the ER. The merged picture shows colocalisation of the mRuby tagged NT-chCREB3-TM and the ER tracker. (Scale bars = 5µm)

3.6.2.3 Fluorescent constructs chosen for FRET-FLIM

The pp14 protein was expected to localise in the cytoplasm, therefore the pp14B construct with a N-terminal mClover was chosen for the investigation. Only one NT-chCREB3 construct—the NT-chCREB3 construct with a N-terminal mRuby—was chosen for analysis. This decision was based on the design of the constructs used for the Y2H screen. For screening, the DNA binding domain (DBD) of GAL4 (~16kDa) was fused to the N-terminus of the bait (pp14) and the activation (AD) domain of GAL4 (~12.4kDa) was fused to the N-terminus of the prey proteins (including CREB3). Since both proteins had a N-terminal tag, it was assumed that a tag placed at the N-terminus of this protein would not interfere with function. To serve as a cytoplasmic form, the mRuby (N-ter) NT-chCREB3-TM construct cloned with the TM domain was used. As a negative control construct, the mRuby (N-ter) CT-chCREB3 construct was chosen.

It is also important to note the N-terminal domain of chCREB3 was the predicted pp14-interacting domain of chCREB3 based on these three reasons: 1) HuCREB3 was identified as an interacting partner of pp14 in living yeast cells using a classical Y2H screen. Detection of protein-protein interactions in a Y2H assay relies on the expression of a reporter gene (such as lacZ or GFP) activated by the binding of a transcription factor, which has a modular structure comprising of a DNA-binding domain (DBD) and an activation domain (AD) (e.g yeast transcription factor GAL4) (Brückner et al., 2009; Miller and Stagljar, 2004). The protein of interest is fused to the DBD (known as the Bait), and the protein library is fused to the AD (referred to as the Prey). If the proteins interact, the reconstitution of a functional transcription factor will activate the expression of the reporter gene. A

disadvantage of the classical Y2H is that it requires the translocation of the interacting proteins into the nucleus, which makes it unsuitable for membrane associated proteins, integral membrane proteins and many other soluble cytosolic proteins or proteins localised in other subcellular compartments (Brückner et al., 2009; Causier and Davies, 2002). For screening, the conserved CT domain of pp14 (bait) was cloned into pGBKT7GAL4-DBD vector and expressed as a fusion C-terminal to the GAL4-DBD domain harbouring a nuclear localisation signal. The prey proteins from the HeLa cDNA library were cloned into pGADT7 GAL4-AD vector and expressed as fusions C-terminal to the GAL4 AD. The GAL4 AD fusion contains an N-terminal SV40 nuclear localisation signal that targets the protein to the yeast nucleus. Because an interaction was detected between CREB3 and pp14, both proteins would have localised in the nucleus to allow an interaction to happen. Because CREB3 is an ER-membrane bound transcription factor, it would have only been possible for the cleaved NT domain of CREB3 to localise in the nucleus. 2) The amino acid sequence identity between chCREB3 and huCREB3 is 45%, but the proteins share a high degree of similarity (73%) in the bZIP domain found within the NT domain, making NT domain the likely interacting domain. 3) The NT domain of chCREB3 is cytoplasmic, and therefore physiologically, it is the likely domain that can interact with pp14, which is a cytosolic protein.

3.6.3 FRET-FLIM analysis

3.6.3.1 Experimental and control conditions

Table 3.1 gives details about the experimental and control conditions subjected to FRET-FLIM analysis. To measure FRET, DF1 cells were transiently transfected or co-transfected with appropriate construct(s), and the mean fluoresce lifetime of

the donor molecule mClover was measured from living cells using fluorescence lifetime imaging (FLIM).

Conditions (DF1 cells transfected with indicated constructs)		
Controls		
1	Unlinked mClover only (negative control)	This donor only control was used to calculate the donor fluorescence lifetime in the absence of FRET. This was required as a system control, and for the calculation of the donor fluorescence lifetime in cells expressing the mClover-mRuby tandem fusion construct, which was used as a positive control for FRET (Explained in detail in Section 2.14).
2	mClover-mRuby tandem fusion (positve control)	This tandem fusion construct was composed of the mClover and mRuby molecules, covalently attached via a short amino acid linker sequence. This control was predicted to give maximum FRET, and hence was used to define the top end of the FRET range.
3	mClover (N-ter) pp14B only (negative control)	pp14B fused to the donor molecule was expressed without an acceptor. This donor only control was used to calculate the donor lifetime in the absence of FRET, and was used to define minimum FRET for this experiment (the bottom end of the FRET range).
4	mClover (N-ter) pp14B and mRuby (N-ter) CT-chCREB3 (negative control)	The CT domain of chCREB3 was tagged with the acceptor molecule mRuby to serve as a negative control construct. This construct was co-expressed with pp14B labelled with the donor molecule. FRET was not expected to occur.
Experimental		
5	mClover (N-ter) pp14 and mRuby (N-ter) NT-chCREB3	To verify the interaction detected between pp14B and the NT domain of chCREB3 by co-immunoprecipitation, both the nuclear and the ER-bound forms of the protein were produced appended to the acceptor molecule mRuby. These constructs were co-expressed with pp14B labelled with the donor molecule. FRET was expected if the proteins were interacting.
6	mClover (N-ter) pp14 and mRuby (N-ter) NT-chCREB3-TM	

Table 3.1 Experimental or control conditions subjected to FRET-FLIM analysis.

3.6.3.2 Optimisation of conditions

A number of experimental factors were optimised for FRET measurements. The expression of the fluorescent proteins and the fluorescent fusion constructs was initially very high in DF1 cells, and since overexpression can introduce artefacts and give rise to false positive FRET signals, the expression levels were reduced to levels recommended by previous users (the same excitation parameters were used). Multiple rounds of transfections, varying the amounts of transfected DNA and transfection reagent, were carried out to determine the amounts required to achieve the required level of expression for each construct (Table 2.9).

Determining an appropriate time point at which to obtain FRET-FLIM measurements was important. For this experiment, this was decided based on two observations. Firstly, it was found the mRuby fluorescent protein required at least 36 hours post transfection to mature, which was longer than what was required for mClover maturation. Another observation was that around 48 hours post transfection, when cells became too confluent, activation of the ER-bound NT-chCREB3-TM was observed, resulting in the relocalisation of the protein to the nucleus (probably due to cellular stress from nutrient deprivation and changes in the pH). These observations indicated the best time for acquisition was between 36 and 48 hours post transfection.

While fluorescence lifetime can be sensitive to environmental factors such as pH, temperature, autofluorescence, ion concentration, and the presence of fluorescence quenchers, it is a parameter that is independent of the fluorescence intensity or fluorophore concentration of the donor or acceptor molecule, and which is

unaffected by photobleaching (Berezin and Achilefu, 2010; Kappel, 2012). However, during optimisation, a correlation was noted between the fluorescence intensity of the donor molecule mClover and the lifetime of the donor fluorescence. When the donor lifetimes from cells expressing the donor alone (mClover alone or mClover pp14B alone) were analysed, it was found the donor fluorescence lifetime decreased with an increase in the donor fluorescence intensity (Figure 3.18). The correlation was found to be unrelated to the protein concentration or expression levels. In the case of this experiment, this was significant as the expression pp14B is not uniform throughout the cytoplasm, but a higher concentration is found around the ER that lies immediately adjacent to the cell nucleus (Figure 3.16C). What this meant was that even in the absence of acceptor, false-positive FRET signals could be detected in regions containing a comparatively higher concentration of mClover-pp14B. To overcome this issue, and to be able to compare between different experimental or control conditions, only regions of interests with equal donor intensity values (10–20 a.u) were chosen for analysis.

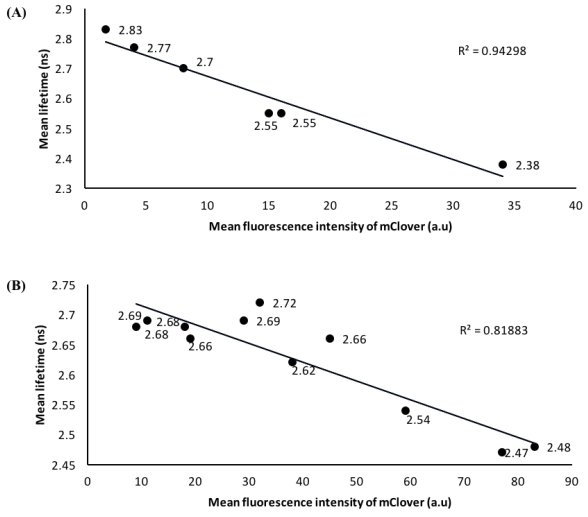


Figure 3.18 Donor fluorescence lifetime affected by the fluorescence intensity.
The mean donor fluorescence intensity values and donor lifetimes were extracted from DF1 cells expressing **(A)** unlinked mClover alone and **(B)** mClover pp14B alone. An increase in the donor fluorescence intensity correlated with a decrease in the donor fluorescence lifetime.

3.6.4 FRET-FLIM analysis of the interaction

To investigate the putative interaction between pp14B and the NT domain of chCREB3, DF1 cells were transiently transfected or co-transfected with appropriate construct(s) for each experimental or control condition (Table 3.1), and FRET measurements were obtained by FLIM from living cells 36 hours post transfection. Intermolecular interactions were quantified by the mean donor fluorescence lifetime for each FLIM image of a live cell. Fifteen cells were analysed per condition. The mean donor fluorescence lifetimes obtained from cells expressing unlinked-mClover alone and mClover-pp14B alone (the donor molecule only) were 2.75ns and 2.74ns (Figure 3.19), respectively, and the mean donor fluorescence lifetimes obtained from cells co-expressing mClover-pp14 and mRuby NT-chCREB3 or mRuby NT-chCREB3-TM were 2.73ns and 2.76ns, respectively. Therefore, no FRET was observed between pp14B and the NT domain of chCREB3.

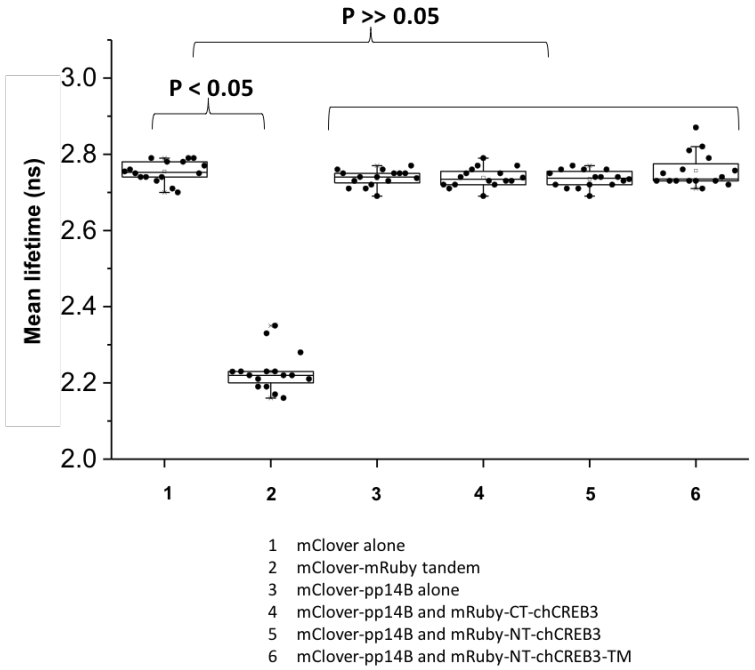


Figure 3.19 Interaction between pp14B and the NT-chCREB3 measured by FRET-FLIM.
The graph presents the distribution of mean donor fluorescence lifetimes derived from cells expressing the indicated fluorescent fusion protein(s). Fifteen cells were analysed per condition. The upper and lower quartiles of each sample are represented by the upper and lower sides of the boxes; the medians are represented by the black lines, and the whiskers indicate the variability outside the upper and lower quartiles. The 'p' (probability) values between two selected samples were determined by one-way analysis of variance (ANOVA) using SigmaPlot. (P>0.05 = no statistically significant differences between groups)

3.7 Discussion

The immediate-early 14kDa phosphoprotein, pp14, of MDV1 is associated with the neurovirulence phenotype of the virus (Tahiri-Alaoui et al., 2012). Yeast-two-hybrid screening identified the ER-bound transcription regulator, human CREB3, as an interacting partner of pp14 and bioinformatics analyses revealed chicken CREB3 and human CREB3 share a conserved domain organisation and that the sequences share 45% amino acid sequence identity, with a high degree of similarity (73%) in the conserved bZIP domain of the transcription factor. Further, fluorescence colocalisation between pp14 and chicken CREB3 in RB-1B/HA-chCREB3 infected cells indicated a potential interaction between these proteins.

This putative interaction between pp14 and chicken CREB3 was further investigated using biochemical, cellular biological and functional assays, and the results of these experiments are discussed here.

Initially in order to investigate whether pp14 interacts with chCREB3 in the context of MDV infection, immunofluorescence was carried out with RB-1B/HA-chCREB3 infected CEF cells, and fluorescence colocalisation was observed between pp14 and chCREB3, indicating a potential interaction between these proteins. To investigate if the interaction could be detected biochemically, co-immunoprecipitation was carried out with cell lysate prepared from CEF cells infected with RB-1B/HA-chCREB3 using anti-pp14, an affinity purified polyclonal antibody raised against a peptide from pp14. Co-immunoprecipitation of chCREB3 with the virally encoded pp14 was observed, indicating an interaction between pp14 and chCREB3. To further characterise the interaction, and identify the interacting domain of chCREB3, co-immunoprecipitation assays were carried out with cell lysates prepared from DF1 cells transiently co-transfected with recombinant pp14B and the NT or CT domain of chCREB3 using anti-pp14. pp14B, a small protein of 93 amino acids, was expressed without a tag because a tag attached to a small protein can sometimes interfere with folding and function. To facilitate detection, the CT and NT domains of chCREB3 were cloned with a relatively short HA epitope tag (YPYDVPDYA) at the N-terminus. The results of the co-immunoprecipitation experiments with these constructs indicated an interaction between pp14 and the NT domain of chCREB3. To confirm the detected interaction, the reverse co-immunoprecipitation was performed using an anti-HA antibody targeting the HA tagged NT-chCREB3. The reverse co-

immunoprecipitation, however, failed to detect an interaction. To further investigate these contradictory results, alternative techniques were employed.

A negative reverse co-immunoprecipitation could result from steric hindrance that may prevent an antibody from binding to the target region within the protein complex. In such cases, it would be necessary to probe using antibodies against different regions of the target protein. As this project also aimed to characterise the interaction at the structural level, a way to move forward and to verify the putative interaction was to determine if both proteins could be co-purified for SEC analysis. In parallel, however, because of the many limitations associated with cell free biochemical techniques (discussed in Section in 3.6) the putative interaction was also investigated using a cellular approach. The FRET-FLIM technique was chosen for the cellular investigation as it is a technique that provides high spatial resolution for detection of protein-protein interactions in their natural environment (Bacskai et al., 2003; Elangovan et al., 2002; Krishnan et al., 2003)

To determine if a complex of pp14B bound to the NT-chCREB3 could be isolated, pull-down experiments were performed with cell lysates prepared from insect cells co-infected with baculoviruses expressing FLAG or GST tagged pp14B and the NT-chCREB3 fused to different tags. From these lysates, the FLAG or GST pp14B tag was captured using an anti-FLAG antibody or glutathione-sepharose beads, respectively. However, as expected co-elution of the NT-chCREB3 with pp14B was not observed. When co-immunoprecipitation was attempted using anti-pp14 with cell lysate prepared from insect cells infected with baculoviruses expressing FLAG-pp14B and HA-NT-chCREB3, co-elution of the NT-chCREB3 and pp14B

was observed. These results led to the finding that there was cross-reactivity between the anti-pp14 antibody and the NT domain of chCREB3 in solution, and the earlier co-immunoprecipitation results obtained with anti-pp14 were all false positives as a result of cross-reactivity. The cross-reactivity between anti-pp14 and the NT-chCREB3 was demonstrated using immunoprecipitation assays with anti-pp14 in the presence and absence of the pp14 peptide using cell lysate prepared from DF1 cells transfected with the NT-chCREB3 alone. In the presence of the pp14 peptide, a band corresponding to the NT-chCREB3 was not observed by western blotting, indicating the pp14 peptide competitively inhibits binding of the NT-chCREB3 to the antibody. The cellular investigation using FRET-FLIM conducted in parallel also failed to detect an interaction between pp14B and the NT-chCREB3 in living DF1 cells.

In addition, the interaction was also probed using a functional assay. Initially fluorescence colocalisation and the co-immunoprecipitation data indicated an interaction between pp14B and chCREB3 in RB-1B/HA-chCREB3 infected CEF cells. In an effort to determine the functional role of the interaction of pp14 with chCREB3, a pp14 deletion mutant of pRB-1B/HA-chCREB3 (pRB-1BΔpp14/HA-chCREB3) was generated. By infecting CEF cells with either RB-1B/HA-chCREB3 or RB-1BΔpp14/HA-chCREB3, the localisation of chCREB3 was assessed in the presence and absence of pp14. CREB3, which is anchored to the ER membrane in an inactive form, is activated through regulated intramembrane proteolysis (RIP) (Chan et al., 2011). RIP liberates the N-terminal fragment of CREB3 to translocate to the nucleus (Chan et al., 2011). Since activation of CREB3 results in a change in its localisation pattern, i.e. both ER-

bound/cytoplasmic and nuclear forms can be observed upon activation (Chan et al., 2011), a hypothesis was proposed that if pp14 interacts with chCREB3 in MDV infected cells, it could affect its localisation. The role of the interaction could be either to prevent or induce its activation. However, no indication of an interaction was found as the localisation of chCREB3 was found to be unaffected by pp14 in MDV infected cells. In the majority of cells infected with either RB-1B/HA-chCREB3 (80%) or RB-1BΔpp14/HA-chCREB3 (86%), chCREB3 activation was not observed (Figure 3.2A). Activation and nuclear re-localisation of chCREB3 was only observed in approximately 20% and 14% of the RB-1B/HA-chCREB3 and RB-1BΔpp14/HA-chCREB3 infected cells, respectively (Figure 3.2A).

Another important observation from the infection study was that colocalisation of pp14 and chCREB3 was not observed in the majority of RB-1B/HA-chCREB3 infected cells (Figure 3.2B). Only a subpopulation (32%) exhibited colocalisation (Figure 3.2B). As the anti-pp14 antibody was found to be cross-reactive with the NT-chCREB3 in immunoprecipitation, it is thus also likely that it cross-reacted with chCREB3 in colocalisation assays. While the anti-pp14 antibody was found to be cross-reactive with the NT-chCREB3 in solution, it was not found to be cross-reactive against denatured NT-chCREB3 on a western blot, suggesting the antibody could be interacting with a conformational epitope. It is therefore likely for the anti-pp14 antibody to cross-react with chCREB3 within cells when it is adopting a particular conformation. Hence, the apparent, partial colocalisation of pp14 and chCREB3 observed in these cells is likely to be false. Moreover, in the absence of any other evidence from the biochemical and functional assays, the possibility of an interaction in these cells is questionable.

It is also important to note the limitations associated with the cellular infection model used for the investigation. It is difficult to draw clear conclusions from this infection study as it is not known whether chCREB3 is expressed in MDV infected cells *in vivo*, and if so whether MDV infection would suppress or promote activation of chCREB3. Furthermore, it has to be recognised this experiment was designed in an effort to investigate the interaction in the context of MDV infection. An efficient *in vitro* infection system for the primary target cells of MDV was unavailable at the start of this project. The infection study was therefore carried out with CEF cells that are permissive for MDV infection, which represents an artificial infection system. Also, due to the lack of endogenous chCREB3 in CEF cells, chCREB3 introduced into the pRB-1B BAC was overexpressed from a strong promoter. Because chCREB3 was overexpressed the model may not have represented physiological conditions.

Taken together, the results of the biochemical, cell biological and functional studies provide no conclusive evidence to support an interaction between pp14 and chCREB3.

4 Structural and biophysical characterisation of pp14

4.1 Introduction

The 14kDa phosphoprotein, pp14, is associated with the neurovirulence phenotype of MDV1 (Tahiri-Alaoui et al., 2012). pp14 is a poorly characterised serotype 1 MDV specific antigen, which lacks homology to known proteins. To gain further structural and functional insights into the protein, the recombinant protein was overexpressed and purified for crystallisation trials. The purified protein was also subjected to biophysical and mass spectrometry (MS) analyses. The results of these experiments are presented in this chapter. As discussed in the Introduction (Section 1.4.1), in MDV-infected cells the mRNA variant encoding for pp14B is more abundantly expressed than the alternative variant suggesting that it may be of greater functional importance, hence the sequence of pp14B was chosen for analysis.

4.2 Protein sequence analysis

Bioinformatics analyses were carried out with the amino acid sequence of pp14B to derive information that could assist with the purification process. pp14B contains 93 amino acids. The theoretical isoelectric point (pI) and molecular weight (MW) were computed using ExPASY Compute pI/MW (http://au.expasy.org/cgi-bin/pi_tool): 4.72/10.2kDa. The 10.2kDa protein, however, migrates as a 14kDa protein on SDS-polyacrylamide gels.

The amino acid sequence of pp14B was analysed using the Regional Order Neural Network (RONN) (Yang et al., 2005), a program that estimates the disorder

probability of a given sequence. Of the 93 amino acids, RONN predicted 43% of the residues disordered, and 57% ordered (Figure 4.1). Overall, pp14B was predicted to be largely disordered.

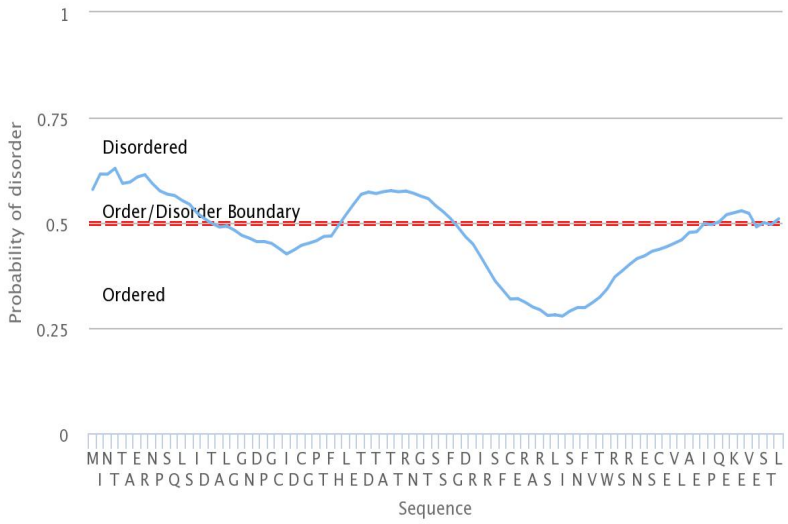


Figure 4.1 Disorder prediction for pp14B.
A RONN plot of disorder probability per residue for pp14B was generated from the website (http://www.strubi.ox.ac.uk/RONN.). The residue is predicted to be disordered if the averaged probability of disorder is >0.5. Residues 1-16, 35-49, and 85-93 were predicted disordered and residues 17-34 and 50-84 were predicted ordered.

In addition, the three-dimensional (3D) structure of pp14B (Figure 4.2A) was predicted using the Phyre2 Server (Kelley et al., 2015). Most of the protein structure (75 out of 93 residues) was modelled by *ab initio* due to lack of homology to known proteins. The protein was predicted to comprise the following secondary structure: 32% disordered regions, 53% alpha helices and 11% beta strands. The region between residues 26 and 43, where two cysteine residues (C26 and C29) and a histidine (H34) are situated, was modelled against a known template, the N-terminal zinc finger domain of mouse cell growth regulating nucleolar protein LyaR (Fold library ID: d1wjva2). A beta sheet was predicted within this region (Figure 4.2B). The template belongs to the fold, beta-beta-alpha zinc fingers; superfamily, beta-beta-alpha zinc fingers; family, Cys(2)HisCys (C2HC)-finger.

C2HC-type zinc fingers contain a short beta hairpin and an alpha helix (beta/beta/alpha structure), where a single zinc atom is held in place by C2HC residues in a tetrahedral array. Since pp14 possesses 4 cysteine residues and a histidine, this prediction raised the hypothesis that pp14 might be zinc binding.

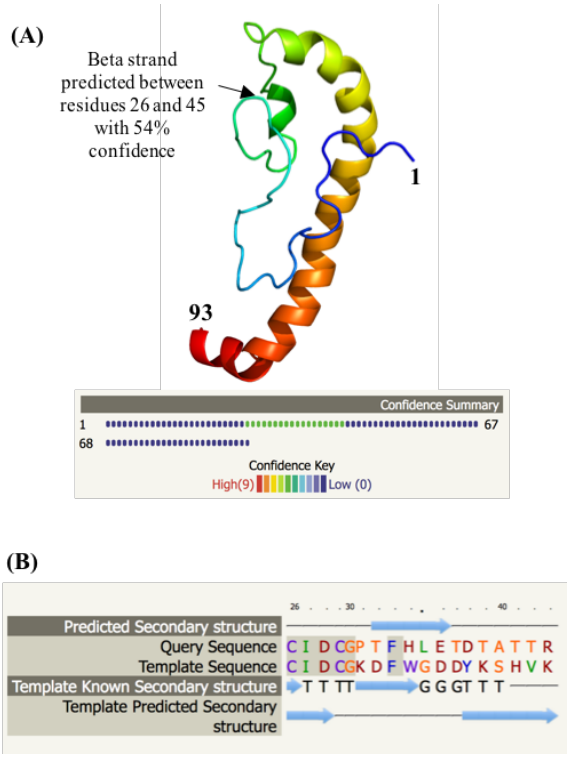


Figure 4.2 3D structure prediction for pp14B.
(A) The model generated for pp14B by Phyre2. None of the residues were modelled at >90% confidence (confidence summary shown below the model). 75 residues were modelled by *ab initio* and residues 26 to 43 were modelled against a known template, the N-terminal zinc finger domain of mouse cell growth regulating nucleolar protein LyaR (Fold library ID: d1wjva2), with 54% confidence. **(B)** Alignment of the pp14B sequence (residues 26 to 43; query sequence) with the known template sequence. (G = 3-turn helix, T = hydrogen bonded turn). For help with interpretation: http://www.sbg.bio.ic.ac.uk/phyre2/

4.3 Expression and purification of pp14

4.3.1 Expression screening of pp14B

The cloning and expression screening of pp14B constructs in *E. coli* was carried out by Louise bird of the OPPF just before the start of my DPhil. The gene encoding full-length pp14B was cloned into a suite of pOPIN vectors harbouring

different fusion tags (soluble/affinity) and the expression of the constructs was tested at small-scale (3ml culture volume) in *E. coli*. Two different *E. coli* strains, and two induction systems were tested (Figure 4.3). Soluble expression was observed for all constructs under the conditions tested in both *E. coli* strains. Based on these results, the pp14 construct with a C-terminal His tag was chosen for scaling-up and purification. His tag is ideal for recombinant protein expression due to its small size, minimal interference in protein folding and weak immunogenicity, and it allows easy purification by immobilised metal ion (nickel or cobalt) affinity chromatography.

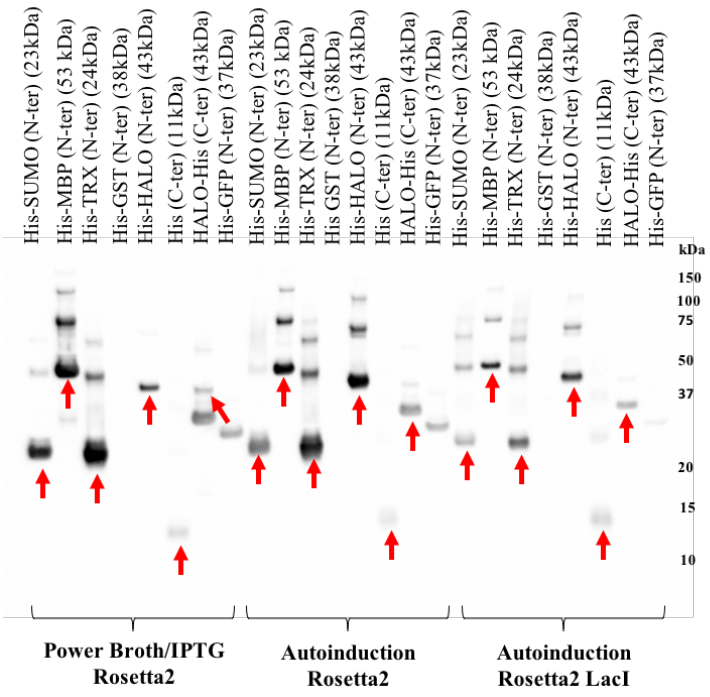


Figure 4.3 High-throughput expression screening of pp14 fusion constructs in *E. coli*.
The expression of various fusion constructs of pp14B was tested in two different *E. coli* cells (Rosetta2 and Rosetta2 LacI cells) and using two different induction conditions (induction with IPTG at 20°C overnight in power broth and with auto-induction media at 25°C for 20 hours). The anti-pp14 western blot represents the amount of soluble protein pulled down from 1ml of culture with Ni-NTA magnetic beads. The red arrows indicate where a band was expected for each fusion construct. [SUMO (Small Ubiquitin-like Modifier); MBP (Maltose-binding protein); TRX (Thioredoxin); GST (Glutathione S-transferase); HALO (modified haloalkane dehalogenase)] (Obtained from Louise Bird)

4.3.2 Scaling-up and purification of pp14B

The pp14B construct with a C-terminal His tag was expressed and extracted from 1litre (L) of Rosetta2 pLysS cells. The His-tagged protein was purified by affinity chromatography, and subjected to size exclusion chromatography (SEC) on a Superdex 75 10/300 GL column equilibrated with 50mM Tris-HCl pH 8, 500mM NaCl. The void volume of Superdex 10/300 GL column is 8.5ml and its separation range is 3 to 70kDa. The SEC elution profile exhibited a major peak immediately after the void volume (Figure 4.4) indicating aggregation of pp14B.

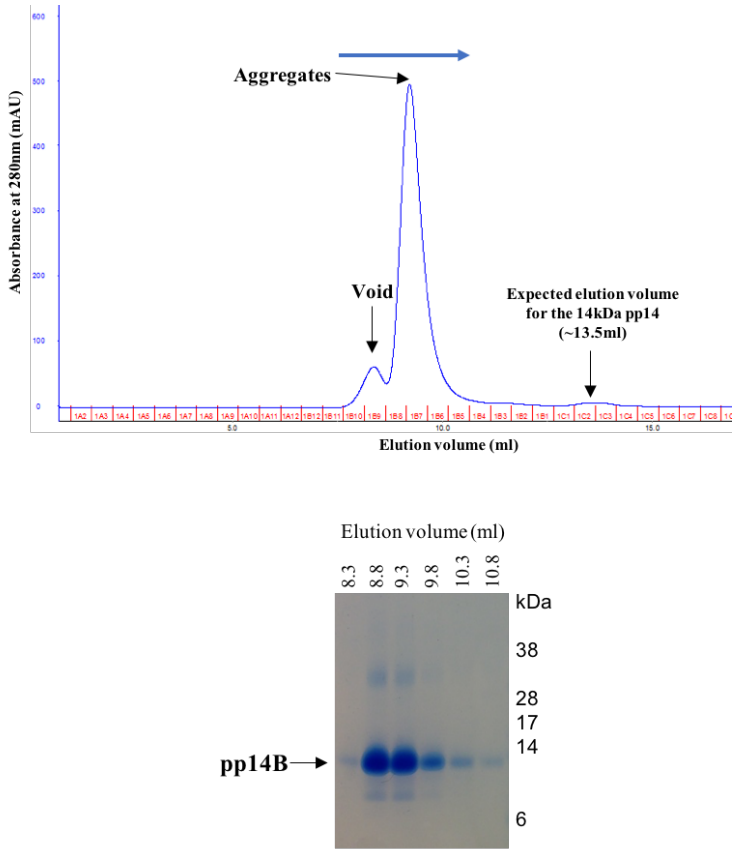


Figure 4.4 SEC analysis of pp14B extracted from *E. coli*.
Ni-NTA purified His-pp14B was subjected to SEC on a Superdex 75 10/300 GL column in 50mM Tris-HCl pH 8, 500mM NaCl and the fractions containing protein (blue arrow) were analysed by SDS-PAGE. The Coomassie blue stained gel is shown below.

4.3.3 Purification of pp14B from insect cells

To examine whether the aggregation was caused by a lack of post-translational modifications in *E. coli*, the protein was produced in insect cells. The His-tagged protein was purified by affinity chromatography, and subjected to SEC on a Superdex 75 10/300 GL column equilibrated with 50mM Tris-HCl pH 8, 500mM NaCl. Again, aggregation was observed (Figure 4.5).

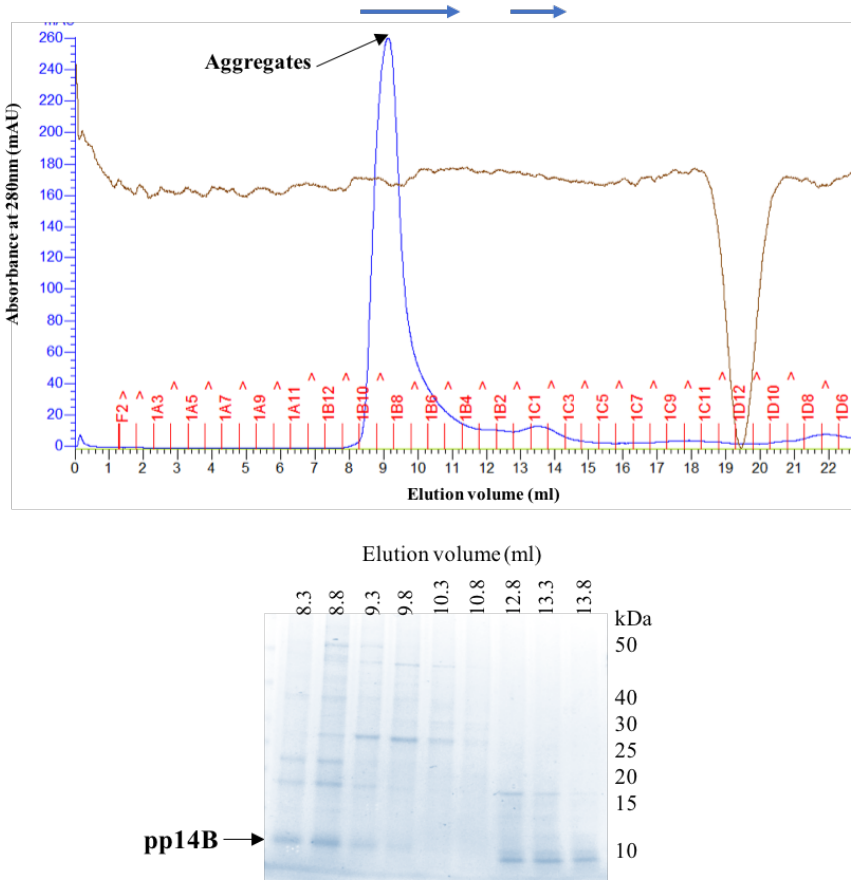


Figure 4.5 SEC analysis of pp14B extracted from insect cells.
Ni-NTA purified His-pp14B was subjected to SEC on a Superdex 75 10/300 GL column in 50mM Tris-HCl pH 8, 500mM NaCl and the fractions containing protein (blue arrows) were analysed by SDS-PAGE. The Coomassie blue stained gel is shown below.

4.3.4 Extraction of monomeric pp14B

4.3.4.1 Additive/pH screening

The expression of pp14B in *E. coli* did not lead to formation of insoluble inclusion bodies, so it appeared the protein could be forming soluble aggregates due to misfolding. To investigate whether this aggregation could be reversed by finding an appropriate formulation, the effect of different additives (Table 4.1) and a range pH (pH 5–8; Section 2.15.2.2) was tested. These additives or buffers addressed different potential causes of aggregation, e.g inappropriate cysteine-cysteine cross-links, hydrophobic interactions, electrostatic interactions, etc (Hamada et al., 2009). The effect of different conditions on the protein aggregation was assessed by SEC since pp14 formed soluble aggregates. For each buffer condition, protein was extracted from half a litre of *E. coli* culture, affinity purified and subjected to SEC analysis. A detailed description of the screen is given in Section 2.15.2.2 of Chapter 2.

Additive	**Concentration (s)**
Urea	0.5M, 1M, 2M
Arginine L-HCl	0.05M, 0.5M, 1M
BME/DTT/TCEP	1mM, 5mM, 10mM
NDSB	1M
Zinc	1mM, 10mM
Glycerol	10%
Arginine L-HCl + L-glutamic acid (L-Glu)	50mM/50mM

Table 4.1 List of additives screened

It proved difficult to solve the aggregation issue, as not one of the conditions was able to completely reverse the aggregation, including addition of reducing agents (selected results are shown in Figures 4.6 and 4.7). However, supplementing the purification buffers with arginine (L-arginine HCl) yielded monomeric pp14B at

very low abundance (Figure 4.8). Arginine is used as an aggregation suppressor during refolding experiments although its mode of action is not clearly understood (Hamada et al., 2009; Lebendiker and Danieli, 2014). A study reported that arginine forms molecular clusters in aqueous solutions, which display hydrophobic surfaces. The hydrophobic surfaces of the molecular clusters, by interacting with the hydrophobic surfaces on proteins, mask hydrophobic patches on proteins, thereby inhibiting protein aggregation (Das et al., 2007). To test the quality of the monomer obtained using this formulation, the protein was assessed by ThermoFluor. The melt curve revealed an unfolded monomer (Figure 4.11), so efforts were put into finding an improved formulation.

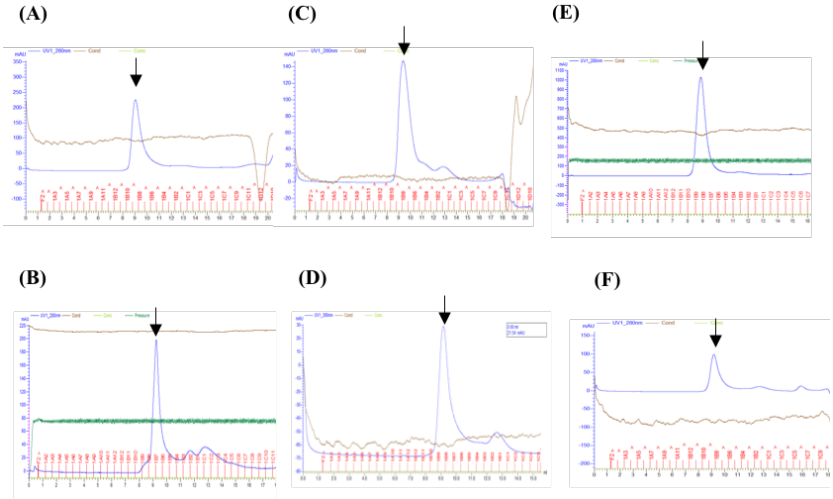


Figure 4.6 SEC analysis of pp14 treated with reducing agents.
Aggregates of pp14B in 50mM Tris-HCl pH 8, 500mM NaCl, 50mM L-arginine HCl were treated with different concentrations (1mM, 5mM and 10mM) of β-Mercaptoethanol (BME) or Dithiothreitol (DTT) or TCEP [Tris (2-carboxyethyl) phosphine] overnight, and analysed by SEC on a Superdex 75 10/300 GL column equilibrated with 50mM Tris-HCl pH 8, 500mM NaCl, 50mM L-arginine HCl supplemented with appropriate reducing agent. Chromatograms **(A)** and **(B)** were obtained from samples treated with 1mM and 5mM BME, respectively. Chromatograms **(C)** and **(D)** were obtained from samples treated with 1mM DTT and 5mM DTT, respectively. Chromatograms **(E)** and **(F)** were obtained from samples treated with 1mM and 5mM TCEP, respectively. The protein samples treated with 10mM BME/DTT/TECEP precipitated out before they could be assessed by SEC. The black arrows indicate elution of the protein in the void volume. (Only selected results are presented here. Many different experiments were carried out to assess the effect of reducing agents, e.g. lysis and purification buffers were supplemented with different reducing agents during purification, and pp14 aggregates were treated with reducing agents in different sample buffers, however nothing solved the aggregation issue.)

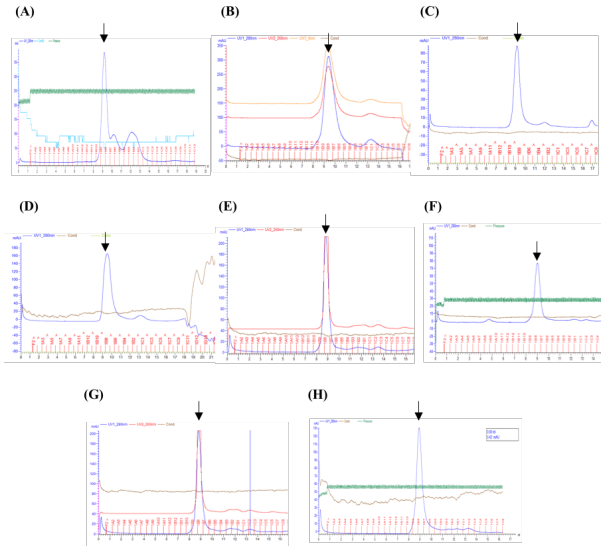


Figure 4.7 Effect of additives/pH on the aggregation of pp14.
The effect of different additives and a range of pH on the aggregation pp14 was analysed by SEC on a Superdex 75 10/300 column. The method is described in Section 2.15.2.2. Additives were tested at different concentrations. However, only selected elution profiles are shown here. Chromatograms of His-pp14B screened in the presence of different additives in 50mM Tris-HCl, 500mM NaCl: **(A)** 10% glycerol **(B)** 1M NDSB **(C)** 0.5M urea and **(D)** 0.5M L-arginine HCl. Chromatograms of His-pp14B screened in different pH/types of buffers: **(E)** 20mM PBS pH7, 500mM NaCl **(F)** 20mM MOPS pH7, 500mM NaCl **(G)** 20mM PBS pH6, 500mM NaCl **(H)** 20mM MES pH6, 500mM NaCl. When a buffer at pH 5 was tested, the protein precipitated out (pH too close its pI, 4.72).

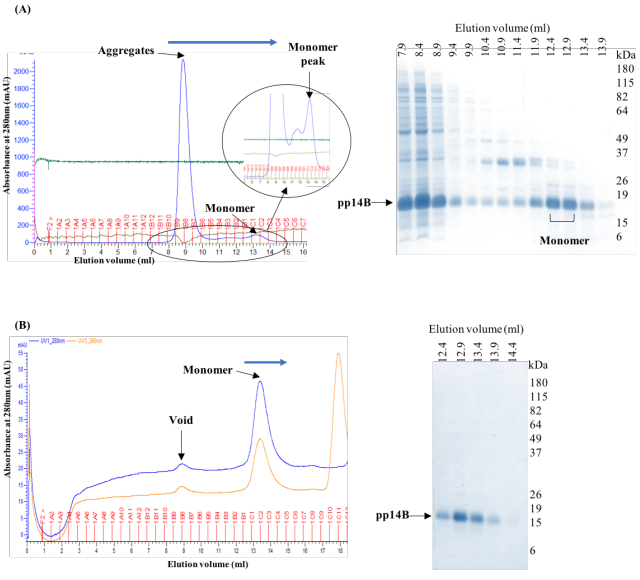


Figure 4.8 Extraction of monomeric pp14B using arginine.
(A) Affinity purified His-pp14B extracted from 3L of *E. coli* culture was subjected to SEC on a Superdex 75 10/300 GL column in 50mM Tris-HCl, pH 8, 500mM NaCl, 50mM L-arginine HCl. The fractions containing protein (blue arrow) were analysed by SDS-PAGE. The Coomassie blue stained gel is shown on the right. **(B)** The fractions containing the monomer were pooled and reanalysed by SEC, but aggregation was not observed.

4.3.4.2 Effect of zinc on solubility

Zinc was detected in very poorly ordered crystals of pp14 by X-ray fluorescence (Sections 4.5 and 4.6). Based on this finding, the effect of zinc on the solubility of pp14B was also explored. These included supplementing the expression media with zinc chloride ($ZnCl_2$), and supplementing the lysis and purification buffers with $ZnCl_2$ during extraction. A refolding experiment with urea-denatured pp14 in the presence and absence of $ZnCl_2$ was also attempted (Section 2.15.2.3). These experiments, however, failed to recover monomeric pp14 (results not shown).

4.3.4.3 Mutation of cysteine residues in pp14B

Mutating free cysteines of proteins has been a strategy to increase the solubility/stability of proteins especially those with multiple cysteines that are difficult to be solubilise using reducing agents. To test this possibility, all cysteine residues of pp14B (C26A, C29A, C57A and C77A; Table 2.1) were mutated to alanine, which has a hydrophobic side chain as that of a cysteine. As the wild-type pp14B construct with a C-terminal His tag expressed well in *E. coli,* the cysteine-null pp14B (Cys-null pp14B) was also cloned with a C-terminal His tag, by cloning into pOPINE, for expression and purification.

For SEC analysis, the protein was extracted from 1L of *E. coli* cells. The same buffer formulation as the one used for co-immunoprecipitation assays (Sections 2.12, 3.2.1 and 3.3) was used for extraction of this protein: 50mM Tris-HCl pH 7.5, 150mM NaCl, 0.5% NP40, 5% glycerol. This buffer contained a mild non-ionic detergent, NP40, which was different from the one that had been used in

previous purifications (Tween 20). To test the effect of this detergent, the cells were resuspended and lysed in this buffer. After incubating the cell lysate with Ni-NTA beads, the captured protein was eluted with 50mM Tris-HCl pH 7.5, 150mM NaCl, 300mM imidazole. The sample was then subjected to SEC on a Superdex 75 10/300 GL column in Tris-HCl pH 7.5, 150mM NaCl. The SEC profile exhibited two peaks: a major peak immediately after the void volume, and a small peak around 13.5ml where the monomeric form of the 14kDa protein was expected to elute (Figure 4.9). The SEC profile showed only a small proportion of the purified protein eluted as monomers, while the majority still formed soluble aggregates, indicating disulphide bond formation was not playing a role in the formation of aggregates, consistent with lack of effect of reducing agents (Figure 4.6). However, these results suggested it may be possible to recover some monomeric protein by supplementing the lysis buffer with NP40.

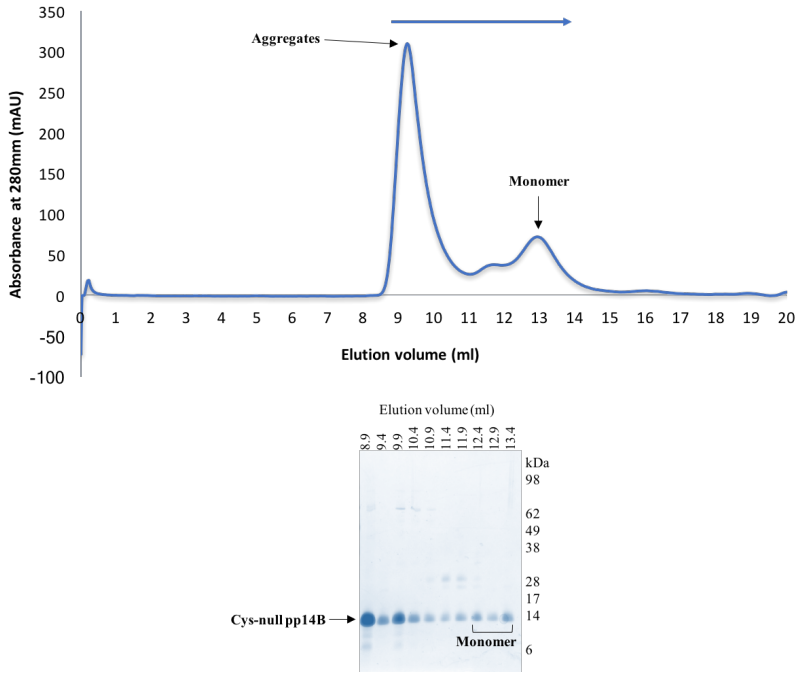


Figure 4.9 Extraction of monomeric Cys-null pp14B.
Affinity purified His-Cys-null pp14B extracted from 1L of *E. coli* culture was subjected to SEC on a Superdex 75 10/300 GL column in 50mM Tris-HCl pH 7.5, 150mM NaCl. The fractions containing protein (blue arrow) were analysed by SDS-PAGE. The Coomassie blue stained gel is shown below.

4.3.4.4 Extraction of monomeric pp14B using NP40

The purification and isolation of monomeric Cys-null pp14B using the protocol described in Section 4.3.4.3 indicated it may be possible to obtain monomeric pp14B by supplementing the lysis buffer with NP40. The same purification conditions were therefore tested with the wild type pp14B, which yielded some monomeric pp14B. Subsequently, to further improve solubility, the concentrations of NP40 and glycerol in the lysis buffer were increased to 2% and 10%, respectively. This slightly improved the yield of the monomeric protein that could be extracted (Figure 4.10). It was found using this protocol approximately 0.3milligrams (mg) of monomeric pp14B could be obtained from a 10L culture of *E. coli*.

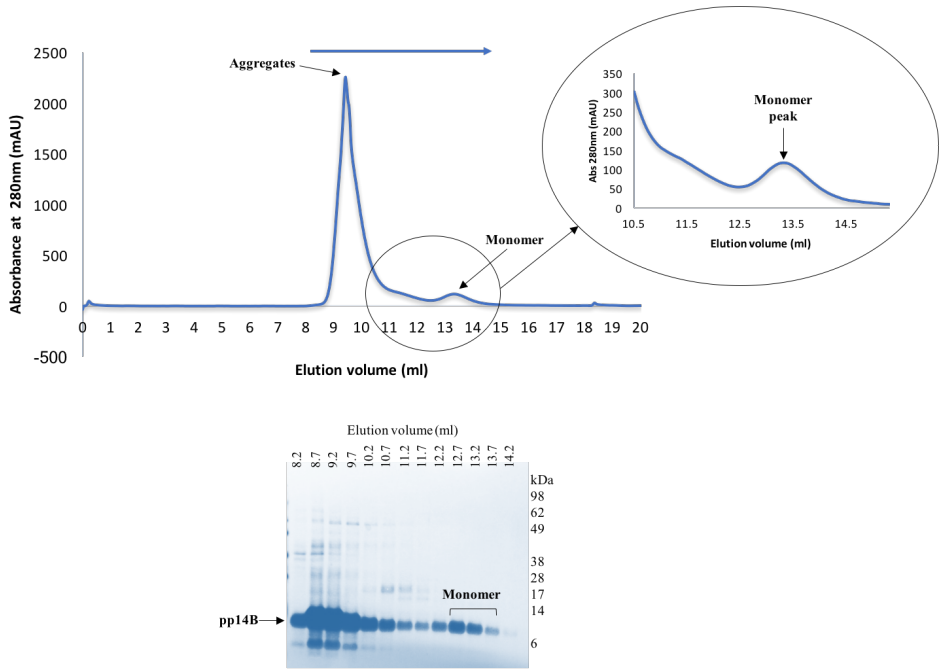


Figure 4.10 Extraction of monomeric wild type pp14B using NP40.
Affinity purified His-pp14B extracted from 5L of *E. coli* culture was subjected to SEC on a Superdex 75 10/300 GL column in 50mM Tris-HCl pH 7.5, 150mM NaCl. The fractions containing protein (blue arrow) were analysed by SDS-PAGE. The Coomassie blue stained gel is shown below.

4.4 ThermoFluor analysis

ThermoFluor is a method useful for optimising conditions for crystallisation trials. It is a temperature-based assay that can be used to assess the stability and quality of proteins, and screen for conditions that stabilise proteins. ThermoFluor uses a hydrophobic fluorophore to distinguish between the folded and unfolded state of the protein. If a properly folded protein is subjected to ThermoFluor, no or low fluorescence signal will be detected at low temperature as no hydrophobic surface patches are presented on the surface of the protein. However, as the temperature increases the protein starts to unfold (the so called "melt") leading to the exposure of hydrophobic regions, which are now accessible to the fluorophore. The fluorophore upon binding to the exposed hydrophobic regions give rise to a fluorescence signal. The melt curve is used to determine the melting temperature (Tm) of the protein. To monitor the folded state of the extracted pp14B monomers in different formulations, ThermoFluor was used. In these experiments, the protein sample was mixed with SYPRO orange, and the temperature was ramped from 25°C to 99°C taking a fluorescence reading every 1°C.

The melt curve exhibited by the monomeric wild type pp14B extracted using arginine analysed in 50mM Tris-HCl pH 8, 500mM NaCl, 50mM L-arginine HCl revealed the unfolded state of the monomers (Figure 4.11). The fluorescent dye SYPRO orange exhibited high background fluorescence below the transition temperature (63°C), and the shape of the curve after the transition was non-interpretable. This curve was indicative of an already partially unfolded protein, having surface exposed hydrophobic patches, which could have been due to sub-optimal buffer composition or due to a lack of interaction partner.

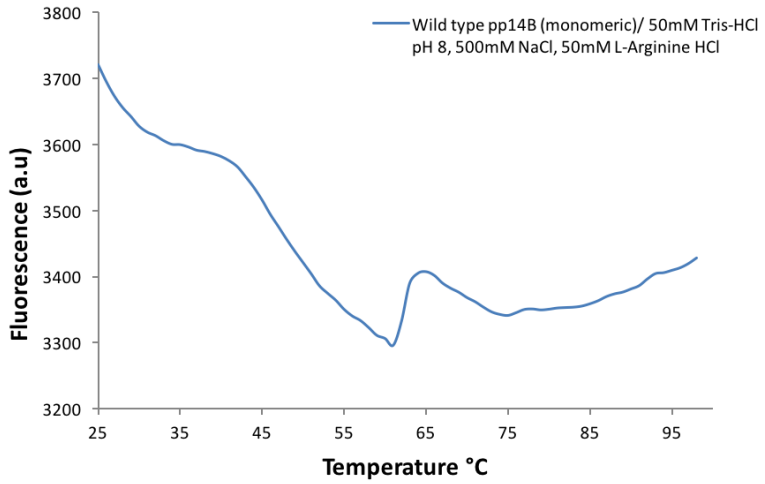


Figure 4.11 ThermoFluor profile obtained for the monomeric pp14B extracted using arginine.
4µg of wild type pp14B (monomeric) in 50mM Tris-HCl pH 8, 500mM NaCl, 50mM L-arginine HCl was mixed with SYPRO orange and the melt curve was generated by ramping the temperature from 25°C to 99°C taking a fluorescence reading every 1°C.

The melt curve exhibited by the monomeric Cys-null pp14B in 50mM Tris-HCl pH 8, 150mM NaCl appeared more promising (Figure 4.12) although the curve did not display an ideal sigmoidal shape indicative of good quality (properly folded protein). The curve had low fluorescence intensity before the unfolding point and a prominent but a broad transition was observed, which occurred from 77°C till the end point of the experiment at 99°C. However, subsequently when the stability of the monomers was tested in a range of pH (pH 4–9) with increased salt concentration of 500mM NaCl, and in the sample buffer again (50mM Tris-HCl pH 8, 150mM NaCl), the melt curves looked different and notably displayed multiple transitions (Figure 4.13).

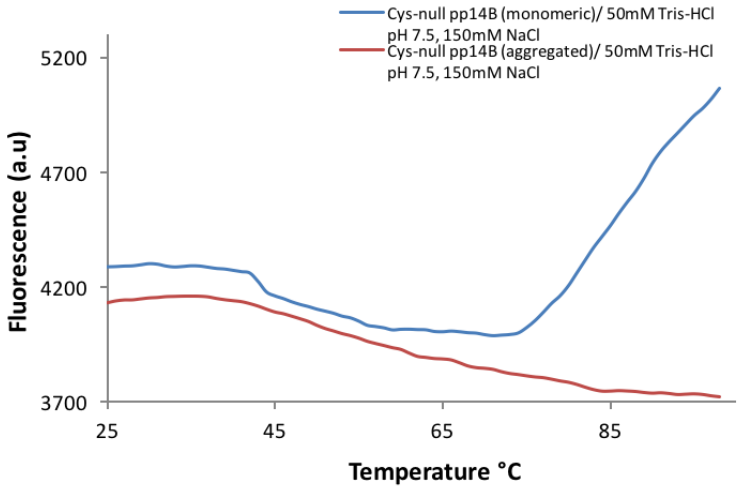


Figure 4.12 ThermoFluor profiles obtained for the Cys-null pp14B.
5µg of Cys-null pp14B (monomeric or aggregated protein taken from the void volume) in 50mM Tris-HCl pH 7.5, 150mM NaCl was mixed with SYPRO Orange. The melt curves were generated by ramping the temperature from 25°C to 99°C taking a fluorescence reading every 1°C.

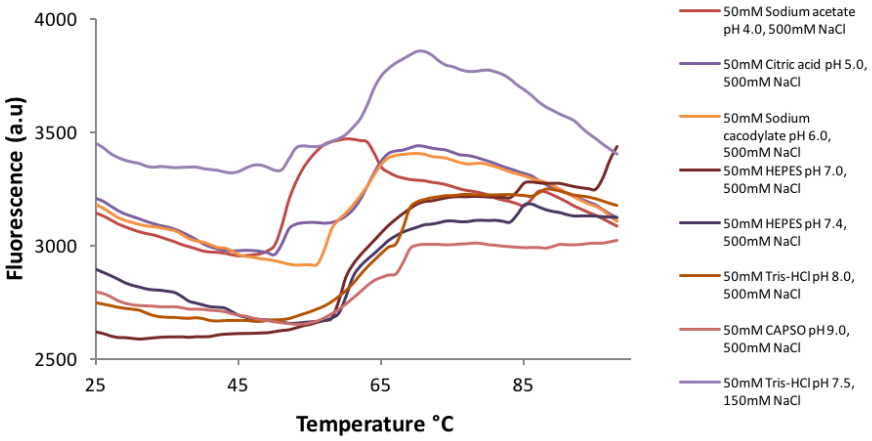


Figure 4.13 pH screen with the monomeric Cys-null pp14B.
3µg of Cys-null pp14B (monomeric) was mixed with indicated buffers and SYPRO Orange. The melt curves were generated by ramping the temperature from 25°C to 99°C taking a fluorescence reading every 1°C.

The monomeric wild type pp14B extracted using NP40 was also found to be unfolded in 50mM Tris-HCl pH 7.5, 150mM NaCl (Figure 4.14); however, in the presence of BME, this protein appeared to adopt a more folded state displaying a sigmoidal-like melting curve. These results were, however, irreproducible with monomers extracted from a different preparation (using the same buffer

conditions). The melt curves obtained looked very different and notably displayed multiple transitions (Figure 4.15) as observed for the monomeric Cys-null pp14B in the pH screen (Figure 4.13).

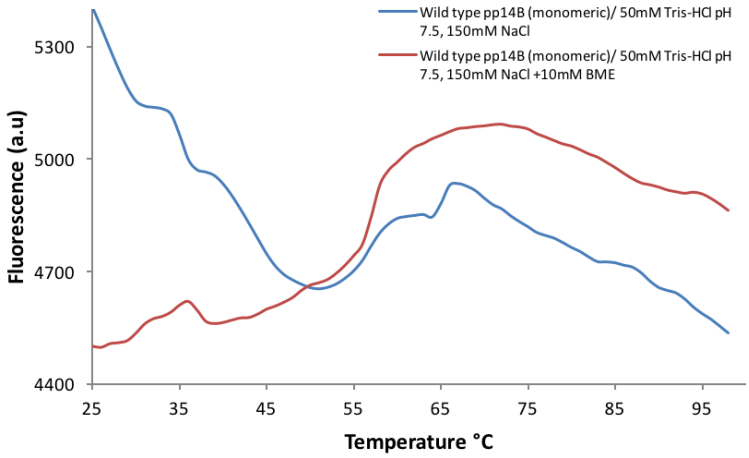


Figure 4.14 ThermoFluor profiles obtained for the monomeric wild type pp14B extracted with NP40 (preparation 1).
5µg of wild type pp14B (monomeric) in 50mM Tris-HCl pH 7.5, 150mM NaCl, +/- 10mM BME was mixed with SYPRO orange. The melt curves were generated by ramping the temperature from 25°C to 99°C taking a fluorescence reading every 1°C.

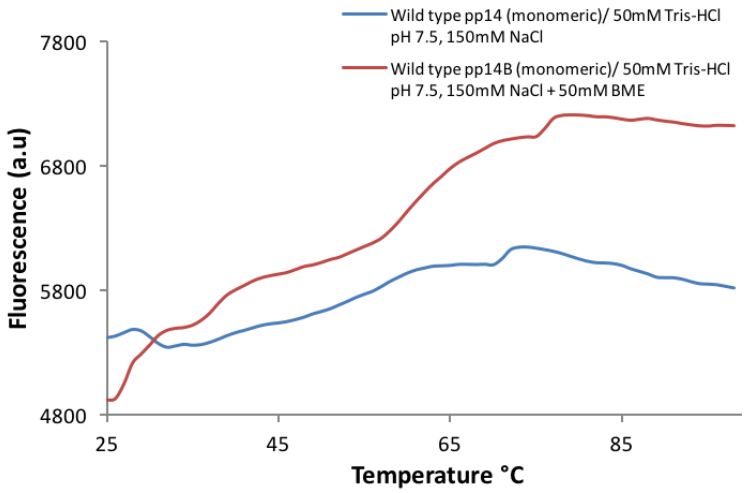


Figure 4.15 ThermoFluor profiles obtained for the monomeric wild type pp14B extracted with NP40 (preparation 2).
10µg of wild type pp14B (monomeric) in 50mM Tris-HCl pH 7.5, 150mM NaCl, +/- 50mM BME was mixed with SYPRO orange. The melt curves were generated by ramping the temperature from 25°C to 99°C taking a fluorescence reading every 1°C.

The changing thermal unfolding behaviour of the monomers (i.e. not exhibiting the same melt curve consistently) raised the hypothesis that pp14 may adopt heterogeneous molten globule-like structures consisting of an ensemble of loosely folded structures.

4.5 Crystallisation of pp14B

Initially, sitting drop vapour-diffusion crystallisation trials were carried out with the purified soluble aggregates of pp14B (in 50mM Tris-HCl, pH 7.5 500mM NaCl). Very thin needle-like or amorphous crystals were obtained (Figure 4.16). However, crystallisation trials set up with the purified monomeric Cys-null pp14B (in 50mM Tris-HCl pH 7.5, 150mM NaCl) and monomeric wild type pp14B (in 50mM Tris-HCl pH 7.5 150mM NaCl +/-10mM BME) yielded no crystals.

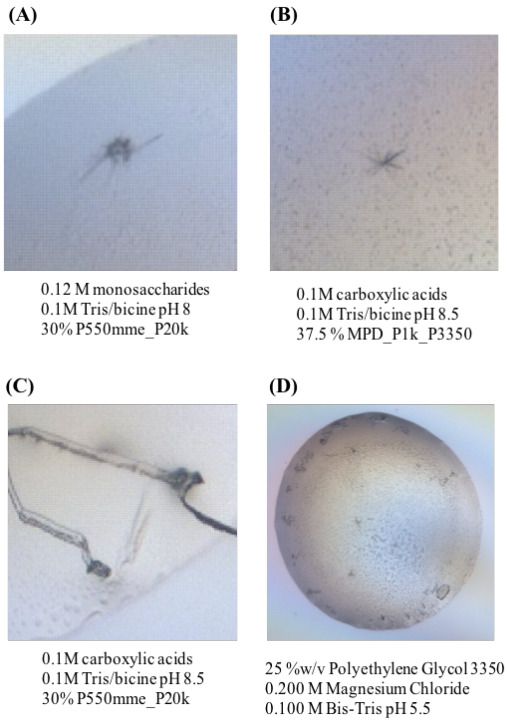


Figure 4.16 Crystallisation experiments with the aggregated form of pp14B. (**A-D**) Crystals obtained from crystallisation trials with the aggregated form of pp14B.

4.6 X-ray and X-ray fluorescence analyses of pp14B crystals

The crystals obtained from crystallisation trials with the purified soluble aggregates of pp14B (Figure 4.16) were tested on beamline I02 at Diamond Light Source (Didcot, UK). Two crystals diffracted and were identified as paracrystalline (Figure 4.17A and B, left), hence with the aim of obtaining diffraction quality crystals efforts were focused towards isolating stable monomeric pp14B.

To test if these crystals contained zinc, elemental analysis was carried out using synchrotron radiation X-ray fluorescence (XRF) at the I02 microfocus beamline. XRF is based on the principle that atoms in a sample excited by incident X-ray photons emit X-ray photons of specific energies as the system relaxes. The specific energies of the emitted photons are characteristic of the elements in the sample (Sherman, 1955), which are observed as peaks in an XRF spectrum. The XRF spectra from the pp14B crystals revealed the presence of zinc (Figure 4.17A and B, right).

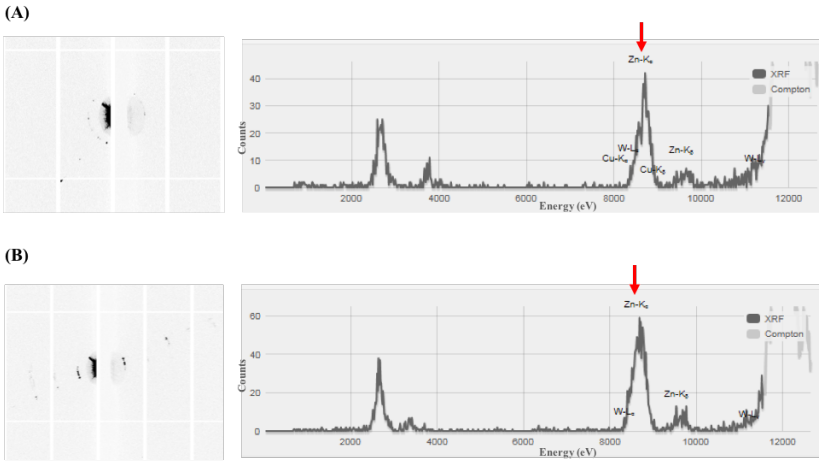


Figure 4.17 X–ray and X-ray fluorescence analysis of pp14B crystals.
Representative X-ray diffraction patterns obtained from crystals harvested from 0.12M monosaccharides, 0.1M Tris/bicine pH 8, 30% P550mme_P20k (**A**, left) and 0.12M monosaccharides, 0.1M Tris/bicine pH 8, 30% P550mme_P20k (**B**, left). Shown on the right are the XRF spectra obtained from these crystals. The red arrows indicate the zinc peaks that were detected.

4.7 Secondary structure analysis using Circular Dichroism (CD)

The ThermoFluor melt curves of monomeric pp14B indicated the protein exists as a heterogeneous population (partially unfolded and partially folded structures) (Figures 4.11, 4.14 and 4.15). These results suggested the protein may be lacking significant secondary structure. To investigate this further, far-ultraviolet (UV) circular dichroism (CD) spectroscopy was used. CD spectroscopy measures the difference in absorption of left- and right-handed circularly polarised light as it travels through chiral molecules such as proteins (Miles and Wallace, 2006). CD spectra in the far-UV wavelength range (~190 to 250nm) are reflective of the polypeptide backbone conformations of proteins (Miles and Wallace, 2006). The spectra provide information about their secondary structural composition and can be used to quantify the secondary structural elements (such as alpha helix, beta sheet, turns, and disordered regions) (Micsonai et al., 2015).

Far-UV CD analyses were performed using electronic circular dichroism (ECD) at beamline B23 of Diamond Light Source with different pp14B samples. Both wildtype and Cys-null pp14B were analysed. To investigate the differences in secondary structure between the monomeric and aggregated forms of the wild type and Cys-null pp14B, CD measurements were performed with both forms of the proteins. BME was found to have an effect on the ThermoFluor melt curve of the monomeric wild type pp14B (Figures 4.14 and 4.15). Thus, to determine if a conformational change could be detected, a CD measurement was obtained from monomeric wild type pp14B after addition of BME. In addition, to gain information about the folding of the wild type pp14B, a CD-based thermal denaturation experiment was also undertaken. CD spectra were obtained in the far

UV range, 260nm–185nm. Data were processed and analysed using proprietary software, CDApps (Hussain et al., 2015), developed at beamline B23. The secondary structure estimation was conducted using the CONTIN/LL algorithm (Sreerama and Woody, 2000) in CDApps.

The results revealed that the monomeric form of the wild type pp14B contained approximately 25% alpha helices and 17% beta-sheets but was predominantly disordered consisting of 35% disordered regions and 23% turns (Figure 4.18). Although upon addition of BME, a change in the melt curve of the wild type pp14B was observed by ThermoFluor (Figures 4.14 and 4.15), its secondary structure was found to be unaffected by BME (Figure 4.18). The Cys-null pp14B monomer also had a similar secondary structural composition to that of the wild type pp14B monomer (Figure 4.18). Further, the structural compositions were found to be similar for both the aggregated and monomeric forms of the proteins (Figure 4.18).

Based on the finding of the XRF analysis of pp14B crystals which revealed the presence of zinc in pp14B (Figure 4.17), the effect of zinc was also assessed during this investigation. Zinc chloride ($ZnCl_2$) was added to monomeric wild type pp14B in 50mM Tris-HCl, 150mM NaCl (+/- BME) to a final concentration of 1mM (molar ratio of protein: Zn^{2+}, 1:1.4). Upon addition of $ZnCl_2$, precipitation of the protein was observed with and without BME. To determine if this was specific for Zn^{2+}, the sample was tested with magnesium chloride ($MgCl_2$; molar ratio of protein:Mg^{2+}, 1:1.4), but no precipitation was observed. Further, to determine if the precipitation could be related to the cysteine residues, $ZnCl_2$ was added to monomeric Cys-null pp14B in 50mM Tris-HCl, 150mM NaCl. The Cys-null

pp14B also precipitated out of solution indicating the interaction was not specific for the cysteines and other residues may be involved. Crosslinking of protein molecules by the metal leads to precipitation (Yang et al., 2000). Common ligands for the zinc ion include histidine, aspartate, glutamate, and cysteine (Zhang et al., 2012). Further mutational analysis will therefore be required to determine which residues are involved in the zinc-protein complex formation leading to precipitation. CD analysis was deterred because of zinc-induced precipitation (Yang et al., 2000). Further experiments (such as titrating protein and zinc concentrations or varying the solution pH) will need to be performed to determine if soluble pp14-zinc complexes could be obtained.

To infer information about the folded state of the pp14B, a CD-based thermal denaturation experiment was conducted with monomeric wild type pp14B. In this experiment, CD spectra were obtained in the far UV range (260nm–185nm) as a function of temperature (every 10°C from 10°C to 90°C and a final scan after cooling down to 10°C) and the secondary structural changes were inferred from the spectra (Figure 4.19). The apparent midpoint of the unfolding transition (T_m), calculated using CDApps, was 49.8°C. However, around this temperature, only a subtle change in the shape of the spectrum was observed (Figure 4.19A); a shift from an alpha helical to beta structure was observed, but there was no significant increase of the unordered content (Figure 4.19B). At the beginning of the experiment (at 10°C), the protein possessed 19% alpha helices, 26% beta sheets, 22% turns, and 33% disordered regions, and during the thermal denaturation no drastic change in the secondary structure of the protein was observed in the range 10°C to 70°C (Figure 4.19B). However, at 80°C, the unordered/turns content

increased significantly to 49%/31%. Interestingly, analysis of the CD spectrum recorded at 10°C after the thermal denaturation revealed that the monomer regained its native secondary structural features upon cooling (Figure 4.19).

Further experiments will need to be performed in order to be able to draw conclusions from the results of the CD-thermal denaturation experiment. Far-UV CD data revealed the protein is natively largely disordered, and during the thermal denaturation, only a subtle change in the shape of the spectrum was observed at the apparent midpoint of the unfolding transition, indicating the protein does not possess regions that are significantly structured, consistent with the ThermoFluor data. Although these results indicate the protein exists as a very flexible, partially unfolded protein, to draw definite conclusions, information about the tertiary structure of the protein is required. This can be obtained using near-UV CD, which is sensitive to the tertiary structure of proteins (Mina and Kumar, 2014). If a protein is folded into a well-defined structure, significant near-UV signals can be observed, whereas if a protein possesses secondary structure but lacks a defined 3D structure (e.g. an incorrectly folded or "molten-globule" structure), the signals in the near-UV region will be nearly zero (Mina and Kumar, 2014).

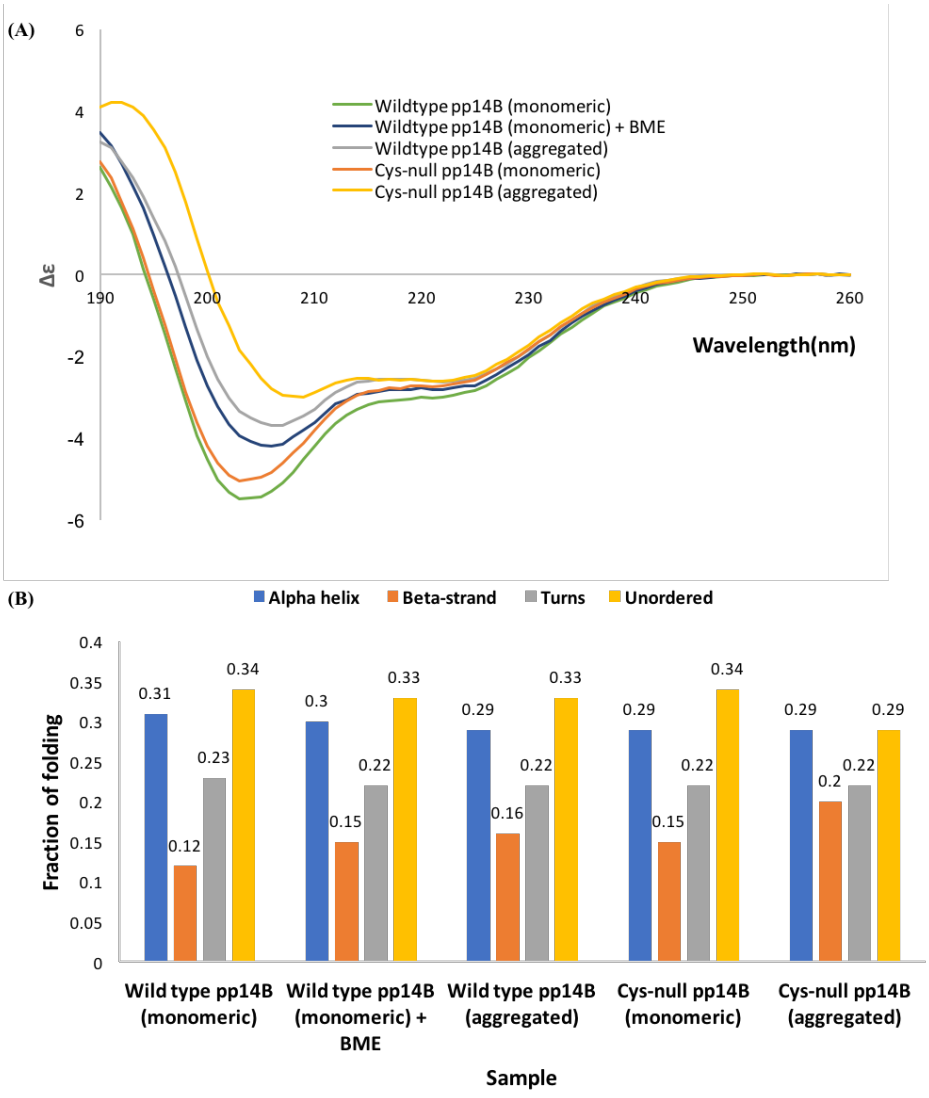


Figure 4.18 Secondary structure analysis using far-UV CD spectroscopy. **(A)** Far-UV spectra obtained from the indicated samples in 50mM Tris-HCl, 150mM NaCl (+/-10mM BME). **(B)** The estimated secondary structure content of the samples derived from the far-UV spectra.

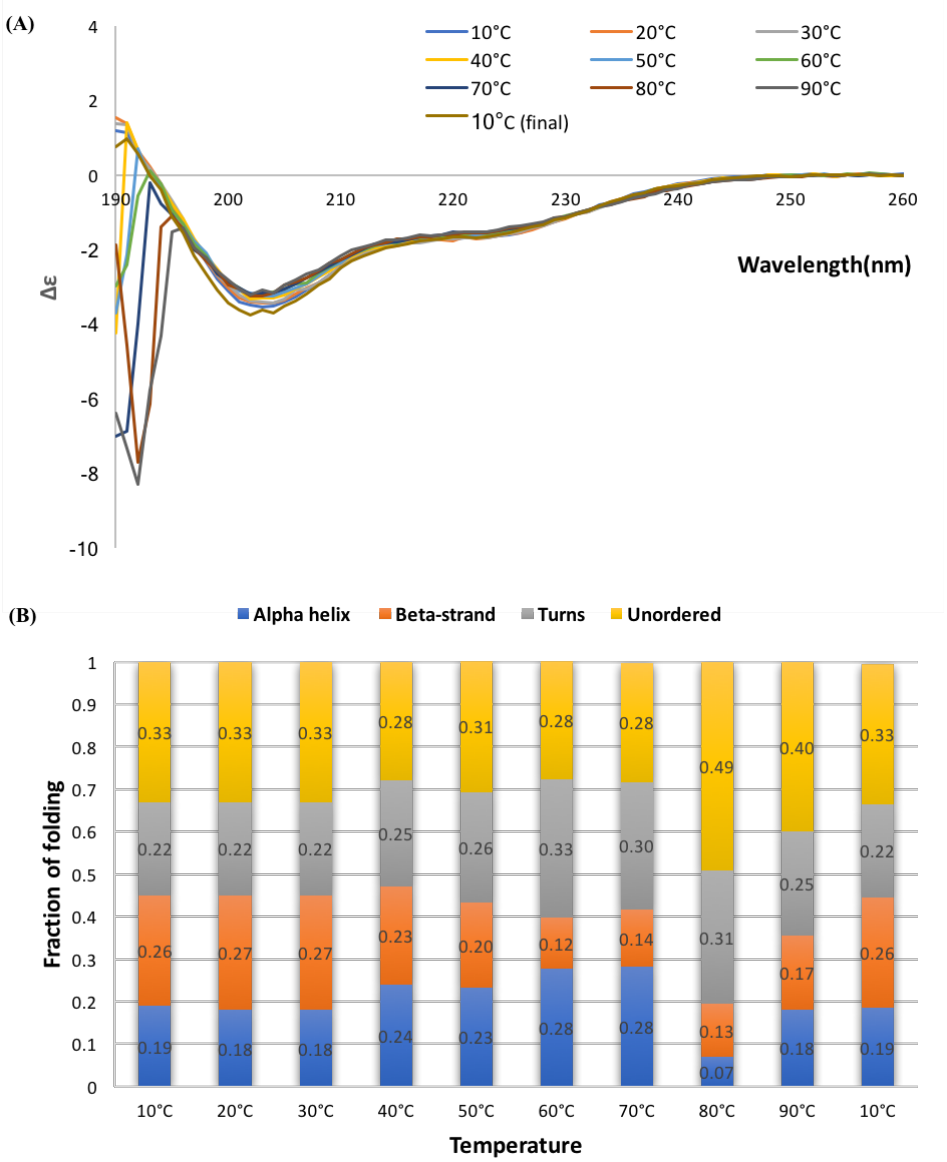


Figure 4.19 Far-UV CD thermal denaturation experiment with monomeric wild type pp14B.
(A) Far-UV CD spectra were recorded every 10°C from 10°C to 90°C, and a final scan was taken after cooling down to 10°C. **(B)** Secondary structural changes inferred from the far-UV spectra.

4.8 Discussion

MDV pp14 is associated with the neurovirulence phenotype of the virus. It is a poorly characterised MDV1 specific antigen, which lacks homology to any known protein. To gain further structural and functional insights into the protein, the recombinant protein was overexpressed and purified for crystallisation trials. The purified protein was also subjected to biophysical and mass spectrometry (MS) analyses. This discussion summarises the findings of these investigations. A detailed discussion is given in Section 6.1.2 of Chapter 6.

The purification and biophysical investigations (using ThermoFluor and CD spectroscopy) of the monomeric wildtype pp14B reveal pp14 exhibits hallmarks of a molten globule. Molten globules are partially folded proteins that exhibit some characteristics of both folded (or native) and unfolded proteins (Redfield, 1999). The classic molten globule is compact possessing native secondary structure, such as α-helices and β-sheets, while lacking fixed tertiary interactions as of an unfolded protein (Redfield, 1999). Consistent with the molten globule hypothesis the following properties were identified: 1) As measured by size exclusion chromatography the monomeric form of pp14B elutes where a 14kDa species would be expected to elute indicating it may adopt a compact shape although further characterisation using Small Angle X-ray Scattering (SAXS) would be required to determine the exact compactness of the protein. 2) The protein exists as a flexible, unfolded protein lacking a stable tertiary structure, which is evident from the ThermoFluor melt curves exhibited by the monomeric wild type pp14B. The ThermoFluor melt curves of two different preparations of pp14B displayed different shapes, which were representative of a partially unfolded and a partially

folded form of the protein. These curves indicate the existence of transient folding intermediates. To gain further information about the tertiary structure of the protein, near-UV CD can be used. Near-UV CD is sensitive to the tertiary structure of proteins (Mina and Kumar, 2014). If a protein is folded into a well-defined structure, significant near-UV signals can be observed, whereas if a protein possesses secondary structure but lacks a defined 3D structure (e.g. an incorrectly folded or "molten-globule" structure), the signals in the near-UV region will be nearly zero (Mina and Kumar, 2014). 3) pp14 possesses substantial secondary structure, as Far-UV CD spectroscopy revealed the monomeric form of the wild type pp14B comprised of 25% alpha helices, and 17% beta-sheets. However, the protein is largely disordered containing 35% unordered regions and 23% turns, which would contribute to the overall structural flexibility. 4) Molten globules have been reported to exhibit significant exposed hydrophobic surface area compared to the native state which leads to a propensity for aggregation (Fink, 2001). Consistent with this, pp14 shows a strong tendency to aggregate in solution. Two mechanisms could contribute to the aggregation of pp14. As ThermoFluor reveals, pp14 does not possess a fixed folded structure, and interconverts between (partially) folded and unfolded states. One driving factor could therefore be hydrophobic surfaces exposed by the unfolded intermediates; hydrophobic residues make intermolecular interactions to form protein aggregates. The disordered content of the protein will be another contributing factor, as disordered regions typically expose hydrophobic residues. The extraction of monomeric form of the protein requires NP40, a mild non-ionic detergent. Detergents are known to prevent aggregation by binding to hydrophobic patches of proteins that prevents protein-protein interactions (Lebendiker and Danieli, 2014). This also suggests the

aggregation could be occurring through intermolecular interactions between surface exposed hydrophobic residues. In this study, it was also shown that the aggregation issue is not related to inappropriate disulphide bond formation between the cysteine residues—a Cys-null mutant of pp14B was as aggregation-prone as the wild type protein.

Only poorly ordered crystals were obtained from crystallisation trials conducted with pp14B. The molten globule nature of the protein means pp14 is an unsuitable candidate for X-ray crystallography, and hence the structural characterisation of pp14 will require alternative techniques such as heteronuclear multidimensional nuclear magnetic resonance (NMR), which is an extremely powerful technique for protein 3D structure determination in solution (Uversky, 2002). Alternatively, co-crystallisation of pp14 with an interaction partner could be attempted as the binding of specific partners to some intrinsically disordered proteins or regions involves a disorder-to-order transition, which can result in more structured conformations (Jakob et al., 2014).

Another interesting finding deduced from two separate experiments is that pp14 is a putative zinc-binding protein. Based upon the three-dimensional (3D) model predicted for pp14B by the Phyre2 server, pp14 was predicted to contain a zinc-binding motif. In support of this hypothesis, zinc ions were detected in poorly ordered crystals of pp14B by X-ray fluorescence. Another sign of zinc binding activity of pp14 was observed when addition of zinc chloride induced precipitation of the wild type pp14B. The induction of precipitation was not found to be specific for the cysteine residues as zinc chloride also induced precipitation of the Cys-null

pp14B protein. Common ligands for the zinc ion include histidine, aspartate, glutamate, and cysteine (Zhang et al., 2012). Further mutational analysis will therefore be required to determine which residues are involved in the zinc-protein complex formation leading to precipitation. Moreover, although precipitation indicates possible zinc binding by pp14, further experiments (such as titrating protein and zinc concentrations or varying the solution pH) will need to be performed to determine if soluble pp14-zinc complexes could be obtained, for structural and functional analyses.

Protein phosphorylation is an important molecular mechanism that regulates protein function. pp14 has been known to be a heavily phosphorylated protein (Hong et al., 1995), and has been predicted to possess potential serine and threonine phosphorylation sites. Fourteen potential serine and threonine phosphorylation sites were identified within pp14B using the NetPhos 3.1 Server (Blom et al., 1999): serines 11, 47, 48, 55, 62, 65, 72, 76, 91 and threonines 4, 41, 42, 46, 69. However, so far the specific sites of phosphorylation have not been identified and the experimental evidence that has been reported so far comes from western blotting analysis of phosphatase treated or untreated cell lysates prepared from MDV-infected cells (Hong et al., 1995). The authors reported the treated sample showed a slight shift in size (1–2kDa smaller) relative to the untreated sample (Hong et al., 1995). During this study, in an effort to characterise the sites of phosphorylation, recombinant pp14B was extracted from DF1 cells and subjected to phosphorylation analysis by mass spectrometry (MS) (Section 2.16). This analysis identified two phosphorylation sites in the shared C-terminal region of pp14—Ser 72 and Ser 76 of pp14B—providing conclusive evidence in support

of previous reports. The role and possible structural and functional implications of phosphorylation on pp14 are discussed in detail in Section 6.1.2 of Chapter 6.

5 Structural characterisation of chicken CREB3

5.1 Introduction

Our current knowledge on the structure and function of CREB3 comes from various studies conducted with the mammalian CREB3. CREB3 is a type II membrane-associated protein, which has its N-terminal domain in the cytoplasm, its transmembrane segment spanning the ER membrane and the C-terminal domain in the ER lumen (Chan et al., 2011). It is a bZIP transcription factor that plays a role in the ER-stress induced unfolded protein response (UPR) (Chan et al., 2011). In addition to playing a role in ER-stress induced mechanisms, CREB3 has been reported to play a role in cell proliferation and migration (Jang et al., 2007; Kim et al., 2010), tumour suppression (Jin et al., 2000), inflammatory gene expression (Eleveld-Trancikova et al., 2010; Sung et al., 2008), and has been shown to be a target of viruses such as HSV, HCV and HIV (Blot et al., 2006; Jin et al., 2000; Lu and Misra, 2000).

To date, no crystal structure is available for CREB3. In an effort to gain structural and functional insights into the transcription factor, the N-terminal, C-terminal and bZIP domains of the protein were generated, overexpressed and purified for crystallisations experiments. The results of these experiments are presented in this chapter.

5.2 Bioinformatic analyses of chCREB3

To assist with expression and purification, bioinformatic analyses were conducted with the amino acid sequence of chCREB3. The majority of the sequence was

predicted to be disordered (Figure 5.1A). An alpha helix was predicted between residues 204 and 274 (Figure 5.1B), the region predicted to be the bZIP domain of the transcription factor.

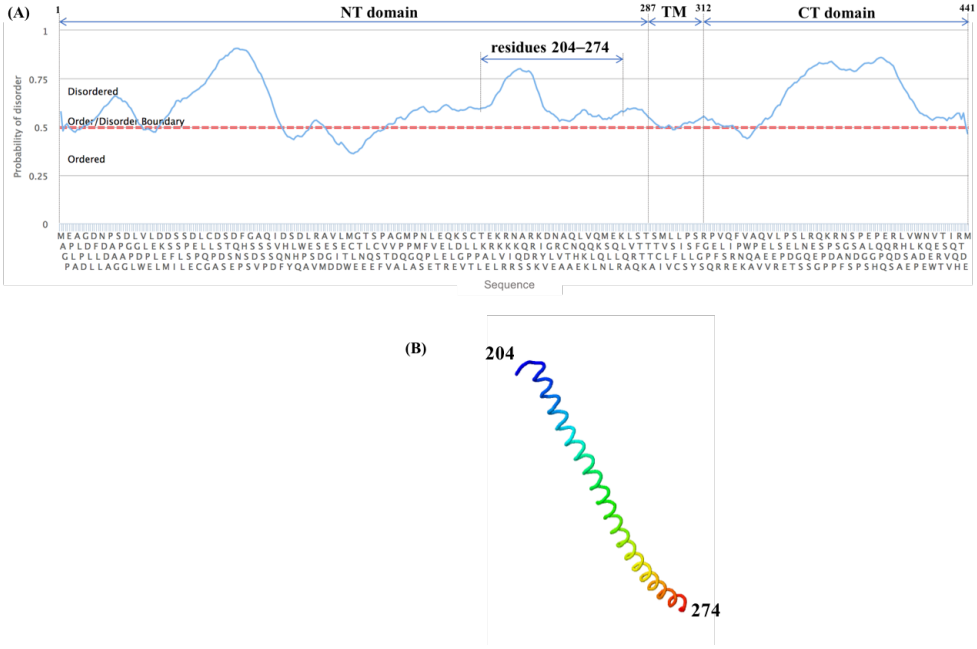


Figure 5.1 Disorder and structure prediction for chCREB3.
(A) RONN disorder prediction for chCREB3. **(B)** Phyre2 3D structure prediction for chCREB3. While 76% of the sequence was predicted disordered, residues 204 to 274 were modelled using the highest scoring template, the crystal structure of the heterodimeric bZIP complex MafB:cFos bound to DNA (Template ID c2wt7B), with 99.4% confidence. This region corresponds to the bZIP domain of the transcription factor.

5.3 Characterisation of the N-terminal domain of chCREB3

5.3.1 Purification of the N-terminal domain of chCREB3

The N-terminal domain of CREB3 is an important functional domain that is essential for its role as a transcriptional regulator. In response to cellular stress, CREB3 is cleaved by proteases which liberates the NT domain from the membrane allowing it to translocate into the nucleus to act as a transcription factor (Saito, 2014). The NT domain of CREB3 harbours important functional units (Figure 1.11) including the S2P protease cleavage site required for its activation; domains

that serve a role in sensing stress conditions through interaction with host or pathogen factors; the bZIP and transcriptional activation (TA) domains required for its nuclear function; and putative nuclear receptor boxes (Blot et al., 2006; Chan et al., 2011).

For crystallisation, the NT domain of chCREB3 (residues 1–288) was cloned into pOPIN expression vectors containing various tags. Expression screens were conducted with these fusion constructs in *E. coli* and insect cells. None of the constructs expressed in *E. coli*, but as described in Section 3.3 of Chapter 3, small-scale expression screening identified three fusion constructs that expressed in insect cells [HA (C-ter) NT-chCREB3, FLAG (C-ter) NT-chCREB3 and His (C-ter) NT-chCREB3]. These constructs, however, failed to express when the expression was scaled-up, and hence the protein could not be purified from insect cells.

To overcome the issue with expression and as a strategy to improve crystallisation success, the structured regions of the domain were considered as alternative targets for crystallisation. The conserved bZIP domain was the only predicted structured region within the NT domain (Figure 5.1). Knowing that bZIP proteins bind target DNA as dimers, and that transcription factors of the CREB3 subfamily are predicted to favour homodimerisation within the family (Vinson et al., 2002), obtaining a homodimer of the chCREB3 bZIP seemed a possibility. The bZIP domain of chCREB3 was therefore cloned for expression and purification.

5.3.2 Purification of the bZIP domain of chCREB3

5.3.2.1 Expression and purification

The bZIP is the simplest known protein-DNA recognition motif (Miller, 2009). The bZIP domain comprises of a basic DNA binding region and a leucine zipper region that facilitates dimerisation (Vinson et al., 2002). The bZIP domain of chCREB3 (residues 211–274) was identified using the Uniprot database (http://www.uniprot.org). For expression and purification, the gene encoding the chCREB3 bZIP (Table 2.1) was cloned into pOPINF expression vector, which harbours an N-terminal His6-3C cleavable tag (MAHHHHHHSS GLEVLFQGP). The His-tagged protein, with a predicted molecular weight of ~11.1kDa, was expressed in Rosett2a pLysS cells. It was purified by affinity chromatography and subjected to SEC on a Superdex 75 10/300 GL equilibrated with 50mM Tris-HCl pH 8, 150mM NaCl. The SEC elution profile displayed three unresolved peaks at elution volumes ~12.2ml, ~13ml, and ~13.8ml and the SDS-PAGE analysis of the protein-containing fractions showed the presence of the bZIP protein in all the fractions (Figure 5.2). The molecular weights of species expected to elute at these volumes range between 43kDa and 10kDa. These peaks therefore indicated the presence of monomeric, dimeric and oligomeric forms of the protein. For crystallisation trials, all fractions were pooled and concentrated using a 3kDa molecular weight cut off (MWCO) concentrator. Crystallisation trials were set up in the presence and absence of DNA targets with the purified protein.

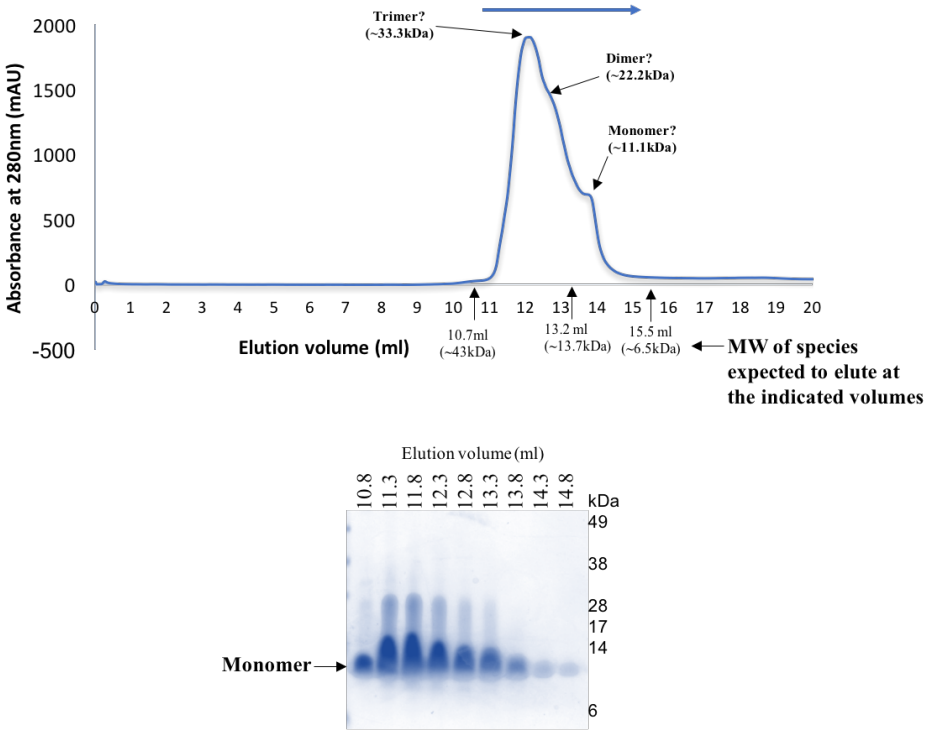


Figure 5.2 Purification of the chCREB3 bZIP.
SEC chromatogram of the chCREB3 bZIP analysed on a Superdex 75 10/300 GL column in 50mM Tris-HCl pH8, 150mM NaCl. The fractions containing protein (blue arrow) were analysed by SDS-PAGE. The Coomassie blue stained gel is shown below.

5.3.2.2 Crystallisation of the chCREB3 bZIP

Sitting drop vapour-diffusion crystallisation trials were set up with the purified His-tagged chCREB3 bZIP (at ~44 mg/ml in 50mM Tris-HCl pH 8, 150mM NaCl). Crystal clusters consisting of thin, rod-shaped crystals (Figure 5.3) were obtained overnight at 4°C with the reservoir condition 50% v/v 2-Methyl-2,4-pentanediol (MPD), 100mM Tri-sodium citrate pH 5.6, 10mM Magnesium chloride (MORPHEUS screen). These crystals were tested on beamline I04 (Diamond Light Source), but these diffracted to resolutions lower than 6Å.

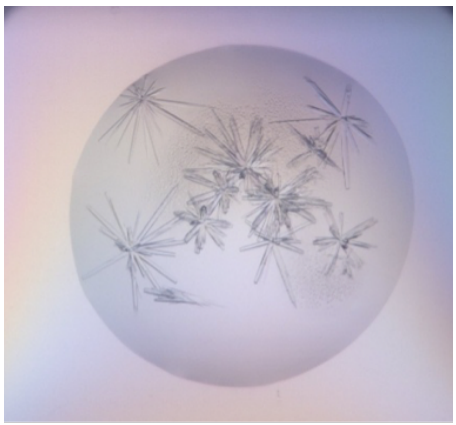


Figure 5.3 Crystals of the chCREB3 bZIP.
Initial crystals obtained in 50% v/v 2-Methyl-2,4-pentanediol, 100mM tri-Sodium citrate pH 5.6, 10mM Magnesium chloride (MORPHEUS screen).

To obtain larger crystals, the following optimisation screens were set up: a pH screen around the condition, and a three-row optimisation screen, where the rows had different protein: screen drop ratios (1:1, 2:1 and 2:1) with a 100% to 67% dilution gradient of the screen across all twelve columns. Larger crystals (Figure 5.4) were obtained from a drop containing 1:1 protein: screen, 43 % v/v 2-Methyl-2,4-pentanediol, 100mM Tri-sodium citrate; pH 5.6, 10mM Magnesium chloride. As the reservoir condition contained MPD, the crystals were cryoprotected by immersing in the reservoir solution without addition of any extra cryoprotectant. Subsequently, the crystals were cooled to liquid nitrogen temperature. Diffraction data were collected at 100 K at Diamond Light Source, beamline I04. 180° of data, up to 3.95Å, collected on a single crystal of the chCREB3 bZIP were auto-processed using Xia2 (Winter et al., 2013a), integrated using DIALS (Waterman et al., 2013) and scaled and merged using Aimless (Evans, 2006).

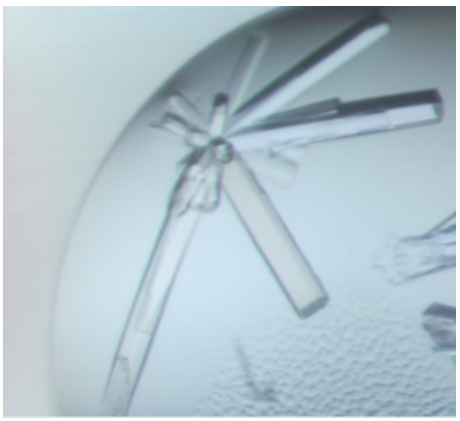


Figure 5.4 Crystals obtained after optimisation.
Optimised crystals obtained from a drop containing 1:1 protein:screen, 43 % v/v 2-Methyl-2,4-pentanediol, 100mM Tri-sodium citrate pH 5.6, 10mM Magnesium chloride.

5.3.2.3 Structure solution and refinement of the chCREB3 bZIP

The structure of the bZIP domain of chCREB3 was solved by molecular replacement (MR) with the program PHASER (McCoy et al., 2007) using a dimer extracted from the CREB bZIP-CRE complex structure (PDB code: 1DH3; Figure 1.12A) after deletion of the bound duplex DNA. The sequence of CREB bZIP has 39% identity to the sequence of chCREB3 bZIP. The structure was refined by iterative rounds of manual model building in COOT (Emsley and Cowtan, 2004) and using BUSTER (Smart et al., 2012). The geometry of the final model was validated with MOLPROBITY (Chen et al., 2010). Data collection and refinement statistics are summarised in Table 5.1.

Model refinement was limited by the resolution of the diffraction data (~3.95 Å). The asymmetric unit of the crystal contained seven copies of the bZIP monomer (3 dimers and a monomer on the crystallographic 2-fold axis). Hence, electron density averaging was applied to improve the quality of the electron density map. The solvent content of the crystals was calculated to be 85%. Hence, as a way to

improve on the diffraction limit of the crystals, crystals were dehydrated. However, it proved difficult to flash-freeze the dehydrated crystals because they dissolved readily in cryoprotectant solutions. Moreover, crystals cryoprotected with different cryoprotectants or precipitants were tested by X-ray, but the diffraction limit could not be improved. The chCREB3 bZIP was crystallised without removing the N-terminal His-3C cleavable tag (MAHHHHHHSS GLEVLFQGP). It is possible removing this flexible tag may have resulted in higher-quality crystals.

Data collection	chCREB3 bZIP
Space group	C 2 2 2
Cell dimensions	
a, *b*, *c* (Å)	137.897 167.057 115.46
α, β, γ (°)	90, 90, 90
Wavelength (Å)	0.9795
Resolution range (Å)	51.63 - 3.95 (4.091 - 3.95)
CC (1/2)	0.932 (0.763)
Rmerge	0.279 (1.633)
$I / \sigma I$	3.3 (1.0)
Completeness (%)	100.0 (100.0)
Redundancy	6.5 (6.3)
Total reflections	77770 (5573)
Unique reflections	12046 (887)
Refinement	
R_{work} / R_{free}	0.2741 /0.3063
No of. atoms	
Protein	3563
B-factors	
Protein	184.88
R.m.s. deviations	
Bond lengths (Å)	0.020
Bond angles (°)	2.18
Ramachandran plot quality	
Favoured (%)	99.28
Allowed (%)	0.72
Outliers (%)	0.00

Table 5.1 Data collection and refinement statistics for the chCREB3 bZIP. Values in parentheses correspond to the highest resolution shell. (RMSD, root mean square deviation from ideal geometry; CC (1/2), cross-correlation between random half-datasets) R_{free} was calculated for a 5% subset of reflections.

5.3.2.4 Structure of the chCREB3 bZIP

This bZIP domain of chCREB3 forms a stable homodimer even in the absence of DNA (Figure 5.5). The solved 3D structure of the homodimeric chCREB3 bZIP is similar to the crystal structure of the heterodimer formed by the bZIP domains of activating transcription factor-4 (ATF4) and CCAAT box/enhancer-binding protein β (C/EBPβ) in the absence DNA (Podust et al., 2001) and other previously published heterodimeric or homodimeric bZIP/DNA complexes (Glover and Harrison, 1995; König and Richmond, 1993; Schumacher et al., 2000).

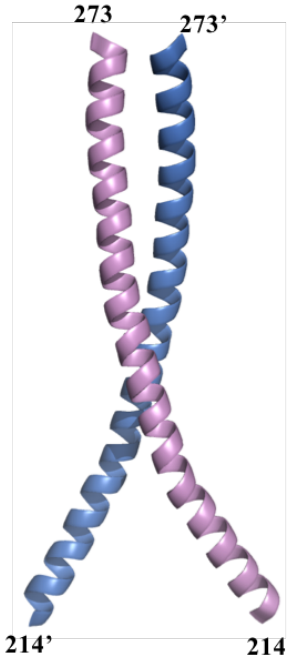


Figure 5.5 The bZIP domain of chCREB3 forms a stable homodimer.

The homodimeric chCREB3 bZIP forms a continuous α-helix in which the C-terminal leucine zipper region forms a parallel coiled-coil dimerisation interface, and the N-terminal basic region also forms a continuous α-helix even in the absence of DNA (Figure 5.6).

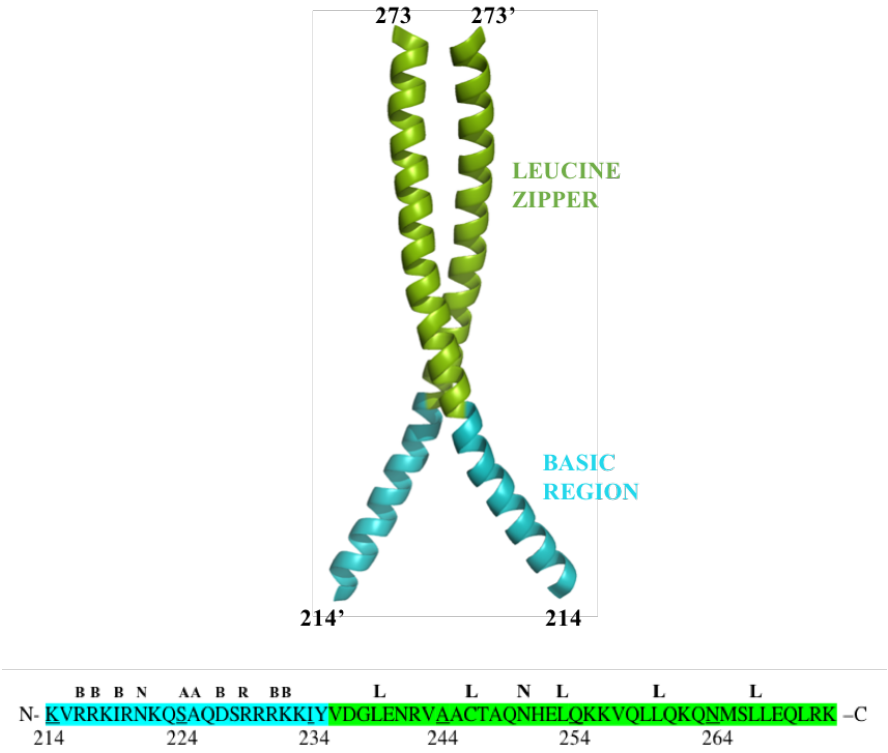


Figure 5.6 The homodimeric chCREB3 bZIP has a structured basic region even in the absence of DNA.
The cartoon representation of the chCREB3 bZIP homodimer shows the structured basic region (blue) and the leucine zipper region (green). The amino acid sequence of the chCREB3 bZIP is shown below, where the basic region residues are highlighted in blue, and the leucine zipper residues are highlighted in green. The consensus sequence is shown above the sequence.

A striking feature of the structure of the homodimeric chCREB3 bZIP is the presence of a disulphide bond in the hydrophobic core of the dimer interface (Figure 5.7). bZIP family members form dimers via their characteristic leucine zipper helices which consist of four to five heptad (seven-residue) repeats of amino acids. The seven positions of the heptads are labelled (**a-b-c-d-e-f-g**) in one helix, and (**a'-b'-c'-d'-e'-f'-g'**) in the other. Hydrophobic residues occur preferentially in the **a** and **d** positions of the leucine zipper, with a variety of amino acids in the **a** positions and leucine in the **d** positions majority (~84%) of the time (Vinson et al., 2002). Hydrophobic interactions by leucine and other hydrophobic amino acids in **a** and **d** positions in the helix form the hydrophobic core of the bZIP dimer that

facilitates dimerisation (Vinson et al., 2002). Interestingly, a cysteine residue (C246) occupies the **d** position of the first heptad of the chCREB3 bZIP (Figure 5.7A). Strong electron density can be observed connecting the cysteines from the parallel helices, indicating this cysteine is involved in forming an intermolecular disulphide bond (Figure 5.7B,C). To assess the presence of cysteine-cysteine cross-linking between the cysteine residues of the chCREB3 bZIP, purified chCREB3 bZIP protein was analysed by SDS PAGE without addition of any reducing agent in the loading buffer. The Coomassie blue stained gel revealed the presence of a predominantly dimeric species (Figure 5.7D), supporting the presence of a disulphide bond in the crystal structure of the homodimeric chCREB3 bZIP.

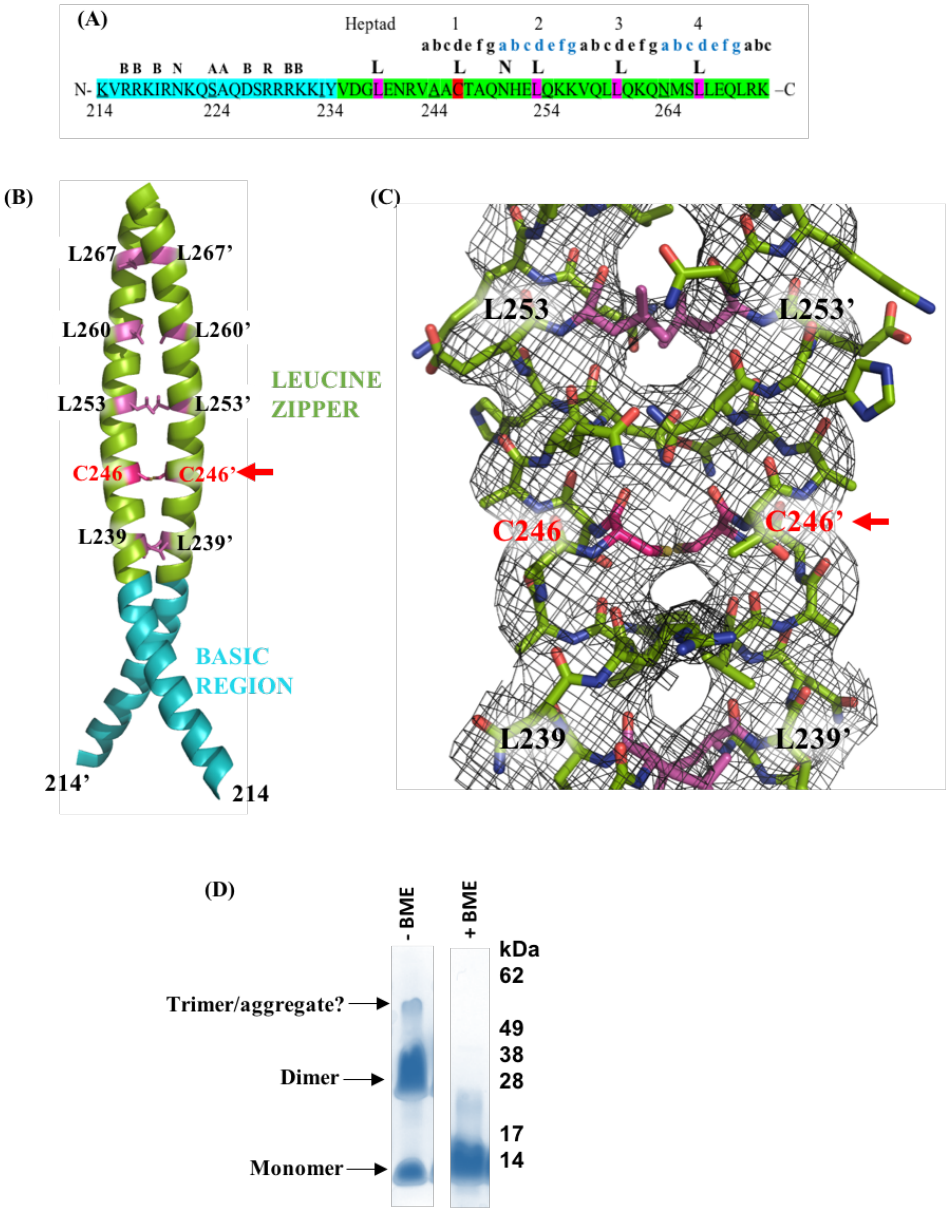


Figure 5.7 chCREB3 has a disulphide bond in the hydrophobic core of the bZIP.
(A) The amino acid sequence of the chCREB3 bZIP composed of heptad repeats with the seven positions of the heptads labelled **a,b,c,d,e,f,g**. The consensus sequence is shown above the sequence. The leucine residues in position **d** of the heptads are highlighted in pink. The cysteine residue (C246) at position **d** of the first heptad is highlighted in red. **(B)** The main chains of the two peptide chains of the chCREB3 bZIP are represented as ribbons. The side chains of the leucine residues and the cysteine participating in the dimer association are represented as stick models. The position of the intermolecular disulphide bond is indicated by a red arrow. **(C)** 2Fo-Fc electron density map at 3.95Å resolution, contoured at 1σ, showing electron density connecting the cysteine residues from the parallel helices. The location of the disulphide bond is indicated by a red arrow. **(D)** SDS-PAGE analysis of purified chCREB3 bZIP protein in the presence and absence of BME.

Another important observation from the crystal structure is the presence of electron density connecting equivalent asparagine residues in position **a** of the second and fourth heptads (N250 and N64) of the parallel helices (Figure 5.8), indicative of strong interactions between these residues. Typically hydrophobic residues occur in the **a** and **d** positions of the leucine zipper (Vinson et al., 2002), which facilitate dimerisation through hydrophobic interactions. However, it has been shown that polar residues that sometimes replace hydrophobic residues at these positions drive dimerisation specificity (Hakoshima, 2014; Vinson et al., 2002). For example, in most homodimerising bZIP transcription factors position **a** of the second heptad is occupied by an asparagine residue (Vinson et al., 2002). Asparagine in this position has been shown to hydrogen bond interhelically with asparagine in the equivalent position of the second monomer (Harbury et al. 1993). A second asparagine in position **a** of the fourth heptad (Figure 5.8A), found only in members of the CREB3 subfamily, has been predicted to encourage homodimerisation within the family, excluding heterodimerisation with homodimerising families without an asparagine in this position (Vinson et al., 2002). The strong electron density observed in this X-ray structure suggests these residues are not only involved in promoting homodimerisation, but may also be important for inducing stability, which could be functionally important for these bZIP factors.

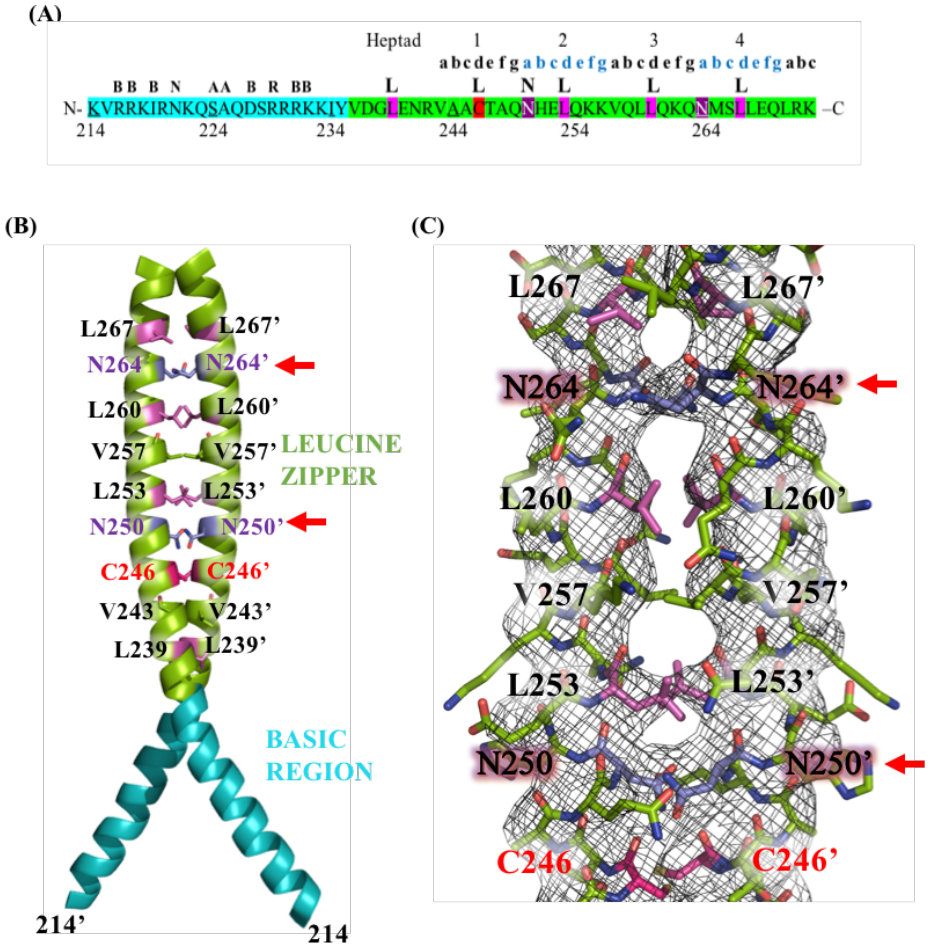


Figure 5.8 Asparagine residues at position 'a' of heptads promote homodimerisation and stability.
(A) The asparagine residues in position **a** of the second and fourth heptads are highlighted in purple in the amino acid sequence of the chCREB3 bZIP. **(B)** The main chains of the two peptide chains of the chCREB3 bZIP are represented as ribbons. The side chains of residues in the **a** and **d** positions that are participating in the dimer association are represented as stick models. The locations of the asparagine residues are indicated by the red arrows. **(C)** 2Fo-Fc electron density map at 3.95Å resolution, contoured at 1σ, showing electron density connecting the equivalent asparagine residues from the parallel helices. The red arrows indicate the location of the residues.

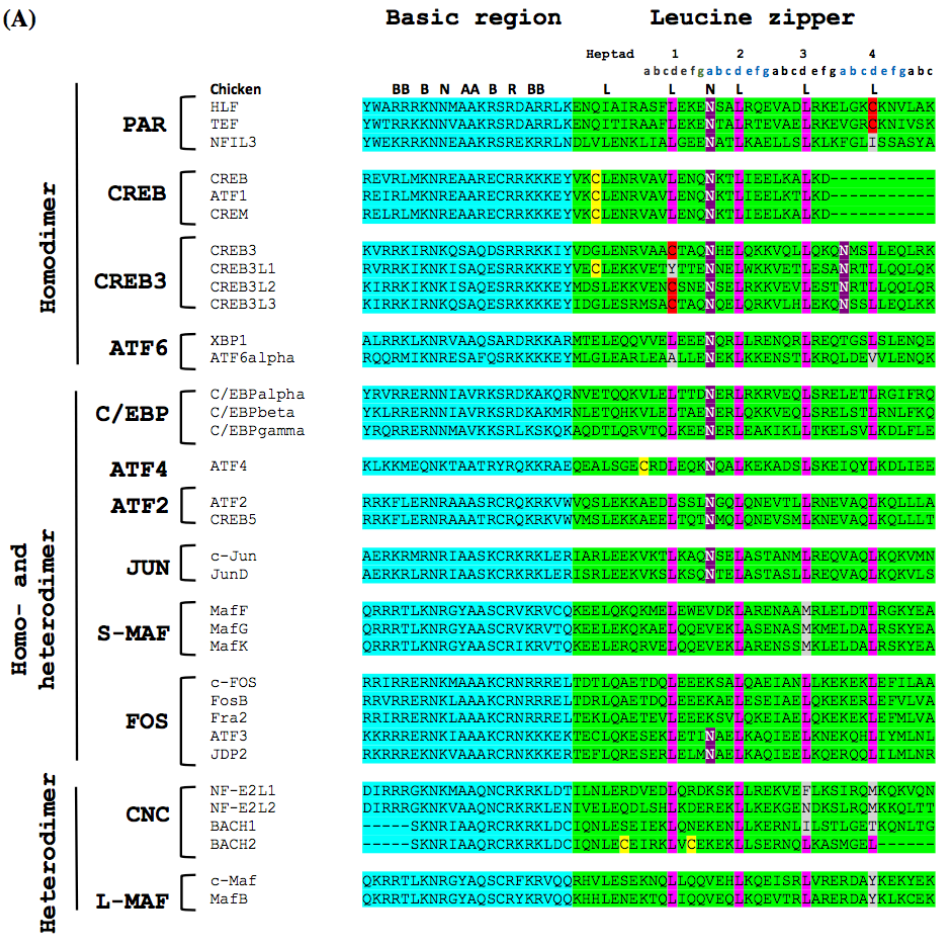
5.3.2.5 Sequence alignment of bZIP sequences

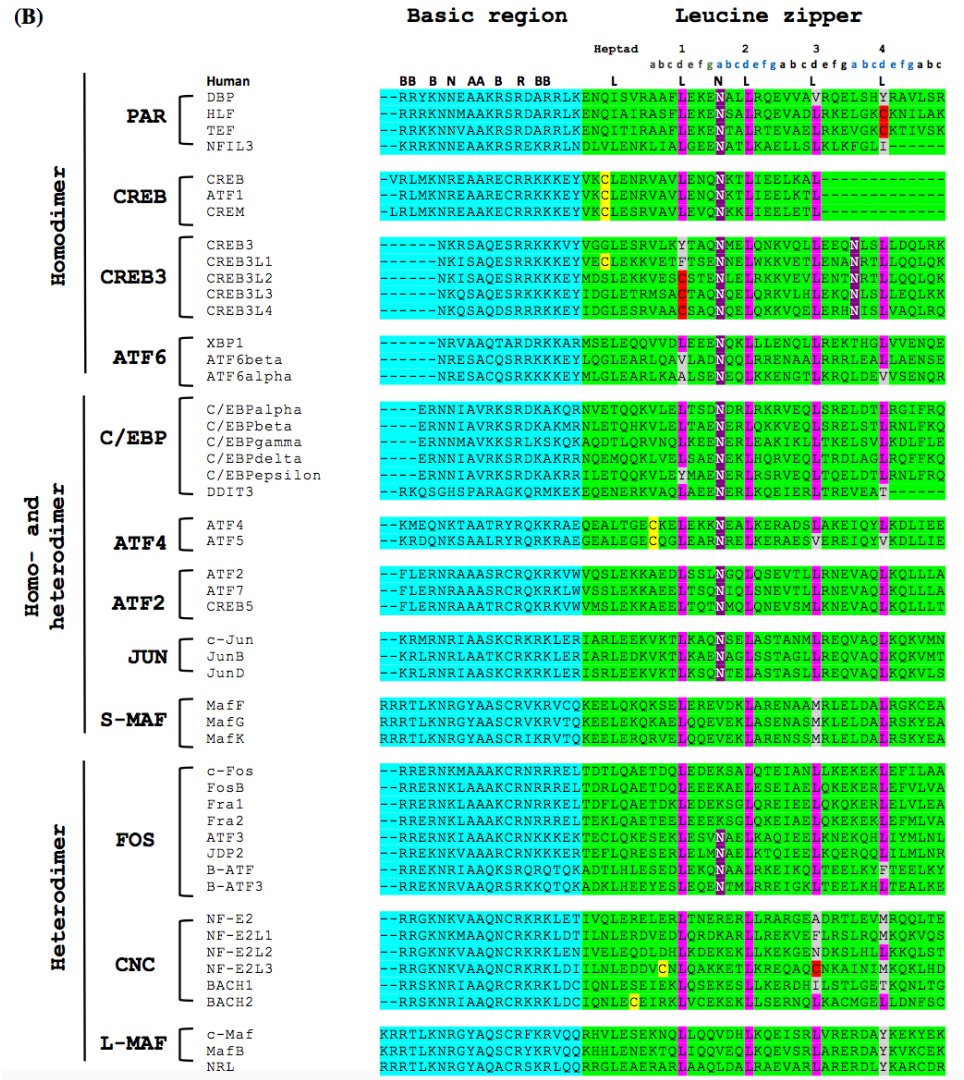
An interesting finding from the crystal structure of the chCREB3 bZIP was the presence of a covalent disulphide bond in the hydrophobic core of the dimer. This is the first time a disulphide has been observed in a crystal structure of a bZIP domain, although previously Podust et al., 2001 and Ciaccio and Laurence, 2009 have reported the presence of an intermolecular disulphide bond in the bZIP domains of ATF4 and ATF5 using NMR and CD studies. ATF4 and ATF5 possess a cysteine residue at position **a** of the first heptad in the leucine zipper region. Podust et al., 2001 performed a CD study with the bZIP domain of ATF4 which had this cysteine residue (Cys 310) of ATF4 substituted with serine to prevent oxidation and formation of high molecular weight aggregates. From this study, they reported the bZIP domain of ATF4 does not form a stable homodimer in the unbound state and that it forms a relatively weak homodimer even in the presence of the DNA target (Podust et al., 2001). However, based on the structure of the bZIP domain of ATF4 in the crystal structure of the C/EBPβ:ATF4 bZIP heterodimer, they predicted the bZIP domain of ATF4 may be able to form a stable homodimer under oxidising conditions as in the crystal structure this cysteine residue is positioned such that equivalent cysteine residues from interacting monomers can make a potential disulphide bond (Podust et al., 2001). The NMR studies on the bZIP domain of ATF5 in the presence and absence of this intermolecular disulphide bond, however, found disulphide bond formation at this position does not confer formation of a stable coiled-coil (Ciaccio and Laurence, 2009). Both studies, however, reported that disulphide bond formation at the cysteine residue in the **a** position of the first heptad increased the α-helical content

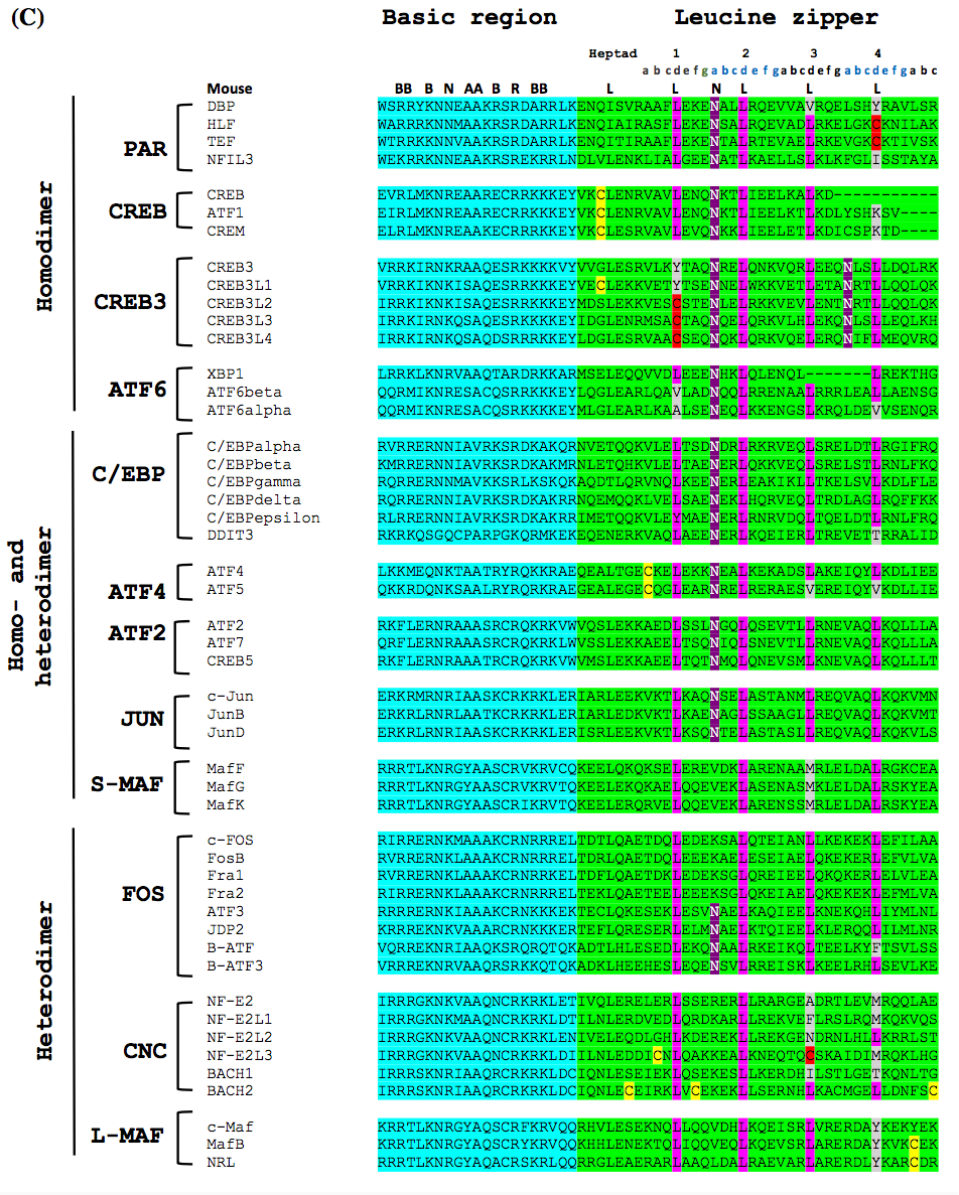
of the bZIP proteins, hence an intermolecular disulphide bond could act to stabilise a coiled-coil (Ciaccio and Laurence, 2009; Podust et al., 2001).

To investigate whether a cysteine residue at position **d** of the first heptad could be unique for chCREB3 or members of the CREB3 subfamily, a multiple sequence alignment analysis was conducted with the bZIP domain sequences of chicken, human and mouse bZIP transcription factors. As described in C Vinson et al, 2002, proteins were placed into 12 groups based on predicted dimerisation properties (predicted to form a homodimer, homodimer and heterodimer, and heterodimer). The alignment results revealed that while the majority of the bZIP domains possess a leucine residue at position **d** of the first heptad, only CREB3, CREB3L2, and CREB3L3 of the chicken CREB3 subfamily (Figure 5.9A), and CREB3L2, CREB3L3 and CREB3L4 of the human and mouse CREB3 subfamily (Figure 5.9B,C) possess a cysteine residue at this position. These results indicate this cysteine residue at this position is unique for members of the CREB3 subfamily, suggesting a function role for this residue, *in vivo.*

Figure 5.9 Sequence alignment of bZIP sequences.
The basic and leucine zipper regions of **(A)** chicken, **(B)** human and **(C)** mouse bZIP transcription factors were aligned using Clustal. The consensus sequence is shown above the sequence alignment. As described in C Vinson et al, 2002, proteins were placed into 12 groups based on predicted dimerisation properties. The basic and leucine zipper regions are highlighted in cyan and green, respectively. Within the leucine zipper region, leucine residues at position **d** of the heptads are highlighted in magenta; cysteine residues at position **d** of the heptads are highlighted in red; any residue other than leucine or cysteine at position **d** of the heptads is highlighted in grey; cysteine residues at any position other than **d** are highlighted in yellow; and as asparagine residues that occupy position **a** of the 2nd and 4th heptads are highlighted in purple.




(B)
Basic region
Leucine zipper
Heptad
Human
BB B N AA B R BB
Homodimer
Homo- and heterodimer
Heterodimer
PAR
DBP
HLF
TEF
NFIL3
CREB
CREB
ATF1
CREM
CREB3
CREB3
CREB3L1
CREB3L2
CREB3L3
CREB3L4
ATF6
XBP1
ATF6beta
ATF6alpha
C/EBP
C/EBPalpha
C/EBPbeta
C/EBPgamma
C/EBPdelta
C/EBPepsilon
DDIT3
ATF4
ATF4
ATF5
ATF2
ATF2
ATF7
CREB5
JUN
c-Jun
JunB
JunD
S-MAF
MafF
MafG
MafK
FOS
c-Fos
FosB
Fra1
Fra2
ATF3
JDP2
B-ATF
B-ATF3
CNC
NF-E2
NF-E2L1
NF-E2L2
NF-E2L3
BACH1
BACH2
L-MAF
c-Maf
MafB
NRL



5.3.2.6 Crystallisation of the chCREB3 bZIP/DNA complex

Crystallisation trials were carried out with the bZIP domain of chCREB3 in the presence of the following palindromic 20-mer or 12-mer sequences: 5′-CTTGGC**TGACGTCA**GCCAAG-3′ (CRE site in bold) and 12-mer 5′-G**ATGACGTCAT**C-3′ (ATF/CREB site in bold). These sequences were adapted from previously published sequences (Fujii et al., 1998; Schumacher et al., 2000). The molar ratio of bZIP dimer (0.5mM) to CRE duplex (0.5mM) was 1. Crystals

appeared in two different conditions that were different from the one the chCREB3 bZIP protein crystallised in (Figure 5.10). The conditions were: 0.005M Magnesium chloride hexahydrate, 0.05M MOPS pH 7, 25% v/v 1,4-Dioxane, 0.001M Spermin (Hampton Natrix2 screen), and 0.01M Magnesium chloride hexahydrate, 0.002M Barium chloride dehydrate, 0.05M MOPS pH 7.0, 30% v/v 1,4-Dioxane (Hampton Natrix2 screen). These crystals were very unstable and dissolved back in solution during the freezing step. Hence, they could not be tested to determine whether they contained a complex of bZIP/DNA.

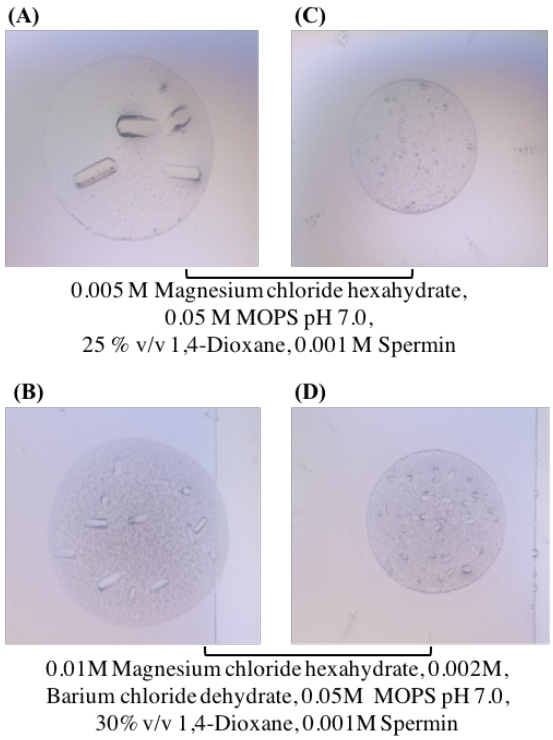


Figure 5.10 Crystals obtained from crystallisation trials with the chCREB3 bZIP in the presence of DNA.
(A) and **(B)** Crystals that appeared from crystallisation trials with the 12-mer DNA oligomer, 5′-G**ATGACGTCAT**C-3′ (ATF/CREB site in bold) and the chCREB3 bZIP. **(C)** and **(D)** Crystals that appeared from crystallisation trials with the 20-mer DNA oligomer, 5′-CTTGGC**TGACGTCA**GCCAAG-3′ (CRE site in bold) and the chCREB3 bZIP.

5.4 Purification of the C-terminal domain of chCREB3

5.4.1 Expression screening in insect cells

The C-terminal domain of CREB3 resides in the ER lumen. It consists of a proline-rich region. The ER luminal domain of CREB3 does not have an ER-stress sensing domain. It has been found that the ER stress-sensing function in the ER luminal domain of chCREB3 is not required for its activation, indicating it may be activated by other cell-specific mechanisms or signal in addition to ER stress (Saito, 2014).

The CT domain of chCREB3 (residues 313–441) was expressed and purified for crystallisation trials. Before the start of my DPhil, the gene encoding the CT-chCREB3 was cloned into a suite of pOPIN vectors harbouring different fusion tags and the expression of these fusion constructs was tested in *E. coli* (by Louise Bird of the OPPF). However, suitable candidates were not identified for scaling-up. Therefore, recombinant baculoviruses expressing the fusion constructs were generated and the expression was tested at small-scale in Sf9 cells. The cells were infected with either 5µl or 50µl virus and harvested at 48 hours and 72 hours post infection (hpi). Good soluble expression of the following constructs was observed: His (C-ter) CT-chCREB3, His-TRX (N-ter) CT-chCREB3, His-GST (N-ter) CT-chCREB3 and HALO-His (C-ter) CT-chCREB3 (the tags are defined in the legend of Figure 5.11). Following this, the CT-chCREB3 construct with a C-terminal his tag was chosen for scaling-up and purification.

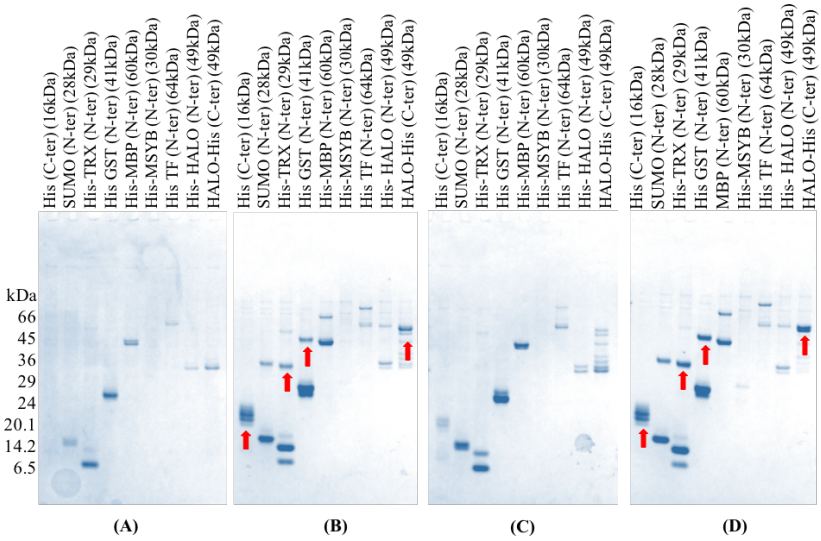


Figure 5.11 High-throughput expression screening of the CT-chCREB3 constructs in insect cells using the baculovirus system.
The expression of the CT-chCREB3 fusion constructs was tested in 5ml Sf9 insect cells. Sf9 cells were infected with either 5µl or 50µl virus and harvested at 48 hours and 72 hours post infection (hpi). The polyacrylamide gels, **(A)** 5µl/48hpi **(B)** 5µl/72hpi **(C)** 50µl/48hpi **(D)** 50µl/72hpi, represent the amount of soluble protein pulled down from 1ml of culture with Ni-NTA magnetic beads. The red arrows indicate where soluble expression was observed. [Tags: SUMO (Small Ubiquitin-like Modifier); TRX) (Thioredoxin); GST (Glutathione S-transferase); MBP (Maltose-binding protein); MSYB (MysB); TF (Transferrin); HALO (modified haloalkane dehalogenase)]

5.4.2 Scale-up and purification

The CT-chCREB3 construct with a C-terminal His tag (~15.5kDa) was expressed and purified from Sf9 cells. The protein was purified by affinity chromatography and subjected to SEC on a Superdex 75 10/300 GL column equilibrated with 50mM Tris-HCl pH 8, 500mM NaCl. The SEC chromatogram revealed the elution of the protein at the expected volume (~12ml) (Figure 5.12). Crystallisation trials were set up with the purified protein (4mg/ml in 50mM Tris-HCl pH 8, 500mM NaCl). However, no crystals were obtained. Due to the disordered nature of this domain, co-crystallisation with an interaction partner could be attempted.

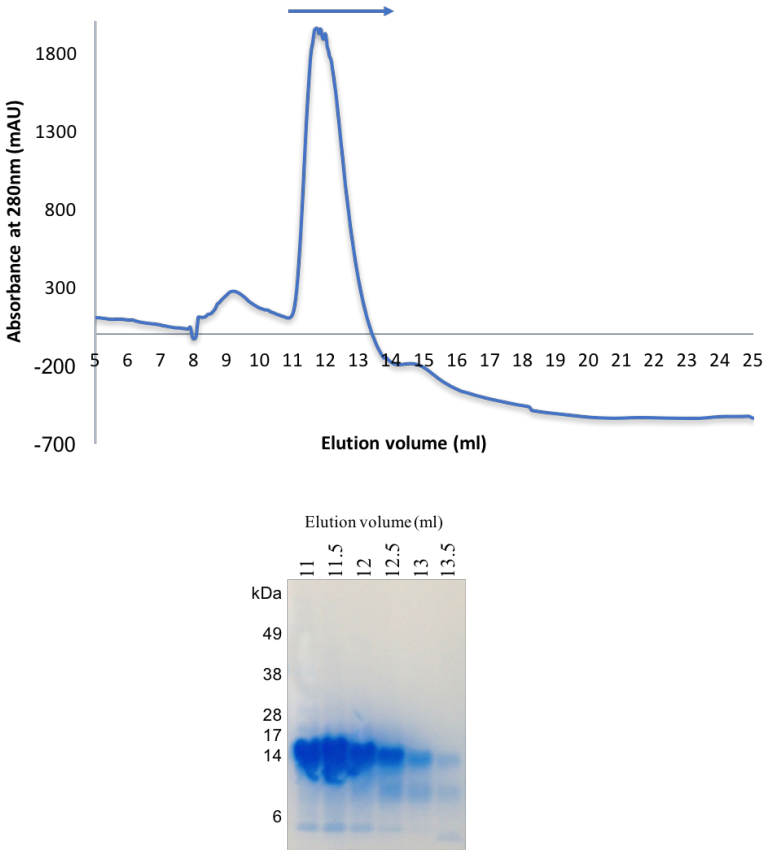


Figure 5.12 Purification of the CT-domain of chCREB3 from insect cells. Affinity purified His CT-chCREB3 was subjected to SEC on a Superdex 75 10/300 GL column in 50mM Tris-HCl pH8, 500mM NaCl. The fractions containing protein (blue bar) were analysed by SDS-PAGE. The Coomassie blue stained gel is shown below.

5.5 Discussion

CREB3 has been characterised as a type II membrane-associated protein with its N-terminal domain in the cytoplasm, its transmembrane segment spanning the ER membrane and the C-terminal domain in the ER lumen (Chan et al., 2011). Predicted to be highly disordered, to date, no crystal structure is available for CREB3. In an effort to gain structural and functional insights into the transcription factor, the N-terminal, C-terminal and the bZIP domains of the protein were generated, overexpressed and purified for crystallisations experiments.

Crystallisation trials could not be conducted with the N-terminal domain of chCREB3, as the expression of this construct could not be scaled-up. Crystallisation trials with the C-terminal domain of chCREB3 yielded no crystals. Both the N-terminal and C-terminal domains were predicted to be highly disordered and crystallisation of disordered proteins, particularly in their unbound (disordered) states, often is not possible (Gill et al., 2016). Hence, to improve crystallisation success, the only predicted structured region of the transcription factor, the bZIP domain, found within the N-terminal domain of chCREB3, was generated and purified from bacteria and subjected to crystallisation trials in the presence and absence of DNA targets. The crystals obtained from crystallisation trials with DNA targets were unstable, which dissolved in cryoprotectant solutions before they could be frozen for X-ray analysis. However, diffraction quality crystals were obtained from crystallisation trials conducted with the chCREB3 bZIP in the absence of DNA, and the structure was solved by molecular replacement. The structural and functional insights gained from the structure of the homodimeric chCREB3 bZIP is discussed in detail in Section 6.1.3 of Chapter 6.

6 Conclusions and future work

6.1 Conclusions and future work

Marek's Disease (MD) is a major infectious disease of poultry (Boodhoo et al., 2016), caused by serotype 1 strains of MDV (Osterrieder et al., 2006). MD is controlled by extensive vaccination (Osterrieder et al., 2006), which has been suggested to drive the evolution of MDV1 towards greater virulence (Gong et al., 2013; Read et al., 2015). Over time, various pathotypes of MDV1 have emerged, which are defined by their ability to break through vaccine induced protection against MD (Petherbridge et al., 2003). Significantly, highly neurovirulent strains of MDV1 are frequently associated with vaccine failures that result in huge economic losses (Gimeno, 2008), and are a continuing to threat to the poultry industry. Despite the continuing threat posed by neurovirulent strains, the mechanisms underlying MDV-mediated neuropathology are still not well understood (Tahiri-Alaoui et al., 2012). It is, however, important to identify and characterise both viral and host factors that contribute to neuropathology as advancing our understanding of MDV neuropathogenesis will facilitate the development of improved intervention strategies against MD.

The 14kDa phosphoprotein, pp14, is the only identified, MDV1 specific neurovirulence factor (Tahiri-Alaoui et al., 2012). Tahiri-Alaoui et al (2012) showed deletion of pp14 from the virulent strain of MDV1, RB-1B, impairs the ability of the virus to cause gross lesions to the peripheral nerves of chickens (Tahiri-Alaoui et al., 2012). This study indicated pp14 is necessary for successful lymphocytic infiltration and manifestation of lesions in the nerves of infected

chickens (Tahiri-Alaoui et al., 2012). These peripheral lesions, that lead to demyelination of peripheral nerves, are associated with the clinical paralysis observed in the classical form of the disease (Payne and Biggs, 1967). The precise cellular and molecular mechanisms that mediate cellular infiltration into nerves, leading to the establishment of lesions, are still not understood. Hence, unravelling the mechanisms by which pp14 mediates neurovirulence in chickens will shed light on our understanding of MDV neuropathogenesis.

Yeast-two-hybrid screening identified the ER-bound transcription regulator, human CREB3, as an interacting partner of pp14 and bioinformatics analyses revealed that chicken CREB3 and human CREB3 share a conserved domain organisation and the sequences share 45% amino acid sequence identity, with a high degree of similarity (73%) in the conserved bZIP domain of the transcription factor. Further, fluorescence colocalisation between pp14 and chicken CREB3 in RB-1B/HA-chCREB3 infected cells indicated a potential interaction between these proteins. In search of interacting partners for pp14, Tahiri-Alaoui et al. identified human CREB3 by yeast two-hybrid screen (unpublished data). Bioinformatic analyses revealed that chicken CREB3 and human CREB3 share a conserved domain organisation, and the sequences share 45% amino acid sequence identity, with a high degree of similarity (73%) in the conserved bZIP domain found within the N-terminal domain of the transcription factor. Subsequently, pp14 was shown to co-localise with recombinant chicken CREB3 in MDV infected CEF cells (unpublished data) indicating a potential interaction between pp14 and chicken CREB3. These results led to the hypothesis that pp14 might be interacting with chicken CREB3, and the interaction could be linked to the neurovirulence of

MDV. The primary goal of this DPhil research project was to further investigate this putative interaction between pp14 and chicken CREB3 using *in vitro* studies, with a view to determining if the interaction is linked to the neurovirulence of MDV. Other objectives of the project were to gain structural and functional insights into the MDV neurovirulence factor pp14 and the host transcription factor chCREB3, through structural and biophysical studies. The main findings of this DPhil project, and directions for future research are discussed below.

6.1.1 Interaction between pp14 and chCREB3

To probe the putative interaction between pp14 and chCREB3, this project employed a combination of biochemical, cellular and functional assays. Experiments were conducted with recombinant proteins expressed in different cell lines and using a cellular infection model designed to examine the interaction in the context of MDV infection. The main conclusions drawn from these various studies of the interaction are discussed below.

A combination of cellular and biochemical techniques was used to investigate the putative interaction using recombinant proteins, which were expressed in the biologically relevant avian cell line, DF1 or insect cells. However, no evidence was obtained to support an interaction between pp14 and chCREB3 in these cells. The biochemical experiments, co-immunoprecipitation or co-purification assays, conducted using cell lysates prepared from DF1 or insect cells expressing pp14B and the CT or NT domains of chCREB3 detected no interaction between pp14B and chCREB3 (Sections 3.3–3.4). The FRET-FLIM technique, which was used to

investigate the putative interaction in the cellular context, also detected no interaction between pp14 and chCREB3 in DF1 cells (Figure 3.19).

The putative interaction between pp14 and chCREB3 was also investigated in the context of MDV infection. An efficient *in vitro* infection system for the primary target cells of MDV (B or T cells) was unavailable at the beginning of this project (Schermuly et al., 2015), hence primary chicken embryo fibroblast (CEF) cells that are permissive for MDV replication were used. Since endogenous chCREB3 was not detected in CEF cells, as a strategy to introduce chCREB3 expression in infected cells, the bacterial artificial chromosome of RB-1B, a virulent strain of MDV1, was engineered to express HA-tagged chCREB3 (pRB-1B/HA-chCREB3). The interaction was probed using a combination of biochemical, and cellular assays with RB-1B/HA-chCREB3 infected CEF cells. No interaction was detected between the virally encoded pp14, and the over-expressed recombinant chCREB3 in RB-1B/HA-chCREB3 CEF cells, by co-immunoprecipitation. Significantly, colocalisation of pp14 and chCREB3 was not observed in the majority of RB-1B/HA-chCREB3 infected CEF cells. Only a subpopulation (32%) exhibited colocalisation (Figure 3.2B). The apparent, partial colocalisation of pp14 and chCREB3 observed in these cells is likely to be false, as the anti-pp14 antibody was found to be cross-reactive with the NT-chCREB3 in immunoprecipitation and it is thus also likely that it cross-reacted with chCREB3 in colocalisation assays. While the anti-pp14 antibody was found to be cross-reactive with the NT-chCREB3 in solution, it was not found to be cross-reactive against denatured NT-chCREB3 on a western blot, suggesting the antibody could be interacting with a conformational epitope (Section 3.5). It is therefore likely for the anti-pp14

antibody to cross-react with chCREB3 within cells when it is adopting a particular conformation. Moreover, in the absence of any other evidence from the biochemical and functional assays (discussed below), the possibility of an interaction in these cells is questionable.

Furthermore, no evidence was found from the functional assay designed to investigate the functional role of the interaction of pp14 with chCREB3 in MDV infected cells. This assay examined the localisation of chCREB3 in the presence or absence of pp14 in CEF cells infected with either RB-1B/HA-chCREB3 or RB-1BΔpp14/HA-chCREB3, a pp14 deletion mutant, using immunofluorescence. CREB3, which is anchored to the ER membrane in an inactive form (Chan et al., 2011), is activated through regulated intramembrane proteolysis (RIP). RIP liberates the N-terminal fragment of CREB3 from the membrane allowing it to translocate to the nucleus (Chan et al., 2011) to perform its transcriptional activation function. Since activation of CREB3 results in a change in its localisation pattern, i.e. both ER-bound/cytoplasmic and nuclear forms can be observed upon activation (Chan et al., 2011), a hypothesis was proposed that if pp14 interacts with chCREB3 in MDV infected cells, it could affect its localisation. The role of the interaction could be either to prevent or induce its activation. However, no indication of an interaction was found as the localisation of chCREB3 was found to be unaffected by pp14 in MDV infected cells (Figure 3.2).

Overall, no conclusive evidence was obtained from the interaction studies to support an interaction between pp14 and chCREB3. Consequently, these results raise questions about the initial interaction identified between pp14 and huCREB3

using Y2H screening. The interaction detected between pp14 and huCREB3 had not been validated using either a biochemical or a high-resolution cellular technique such as FRET, therefore further experiments will need to be performed to determine if the identified interaction could be reproduced in another system. It is also possible that the initial Y2H result could be a false positive since Y2H screens are prone to false positives (Brückner et al., 2009). False positives in Y2H are physical interactions detected in the screening in yeast, which are not reproducible in an independent system (Brückner et al., 2009). Possible reasons for false positive interactions in yeast include: a high expression level of bait and prey and their localisation in a compartment, which does not correspond to their natural cellular environment; interaction of prey with the reporter proteins, e.g. GAL4 in the classic Y2H; proteins which allow yeast to overcome nutritional selection when overexpressed; and unspecific interactions that can occur with proteins that are known to be "sticky" or that are not correctly folded (Brückner et al., 2009).

6.1.2 Insights into the structure of MDV pp14

MDV pp14 is associated with the neurovirulence phenotype of the virus (Tahiri-Alaoui et al., 2012). Lacking homology to any know protein, pp14 represents a MDV serotype 1 specific antigen (Tahiri-Alaoui et al., 2012). For structural analysis, the protein was overexpressed and purified to homogeneity from bacteria and subjected to crystallisation trials. The purified protein was also characterised using Thermofluor, CD spectroscopy and MS. Important structural and functional features identified from these investigations are discussed below.

One important finding of this study is that pp14 exhibits hallmarks of a molten globule in solution. Molten globules are partially folded proteins that exhibit some characteristics of both folded (or native) and unfolded proteins (Redfield, 1999). The classic molten globule is compact possessing native secondary structure, such as α-helices and β-sheets, while lacking fixed tertiary interactions as of an unfolded protein (Redfield, 1999). Supporting the molten globule hypothesis, many properties were identified. The changing unfolding behaviour of bacterially expressed and purified pp14B assessed by ThermoFluor reveals the protein exists as a flexible, unfolded protein lacking a stable tertiary structure. The melt curves of two different preparations of pp14B (extracted using the same protocol) displayed different shapes, one representative of a partially unfolded and the other, a partially folded form of the protein, indicating the existence of different folding intermediates. Lacking a sigmoidal shape, these ThermoFluor melt curves of pp14B also prove that pp14 does not possess a compact, globular structure. The instability of pp14 can be explained by the largely disordered nature of the protein. Secondary structure analysis of purified monomeric pp14B using far-UV circular dichroism reveals the protein, while possessing substantial secondary structure consisting of approximately 25% alpha helices, and 17% beta-sheets, is largely disordered containing 35% unordered regions and 23% turns. The protein also shows a strong tendency to aggregate in solution. Two factors could contribute to the aggregation of pp14. As ThermoFluor reveals, pp14, lacking a fixed 3D structure, interconverts between (partially) folded and unfolded states. One driving factor could therefore be hydrophobic surfaces exposed by the unfolded intermediates, as intermolecular interactions between surface exposed hydrophobic residues result in protein aggregation. The disordered content of the

protein would be another contributing factor. Disordered regions are solvent exposed (Zahran et al., 2015) and binding motifs in disordered regions are often enriched in hydrophobic amino acids (Wallin, 2017).

The structural features identified provide important functional insights into the protein. The presence of disordered regions indicates pp14 may have a broad range of binding partners, and therefore could be multifunctional. Several reports suggest intrinsically disordered proteins are enriched in signalling and regulatory functions as flexibility of disordered regions permit interaction with a range of binding partners (Babu et al., 2011; Babu, 2016), which means the same protein can carry out several functions in multiple pathways (Jakob et al., 2014). Experimental and computational approaches have also shown that, disordered regions are enriched for sites that can be phosphorylated as conformational flexibility of disordered protein regions provides advantages over structured regions (Van Der Lee et al., 2014). These studies thus suggest structural disorder could be functionally very important for pp14, a natively phosphorylated protein.

Diffraction quality crystals of pp14 could not be obtained from crystallisation trials. pp14, being largely disordered and structurally flexible, is an unsuitable candidate for crystallisation. A strategy to stabilise the structure of pp14 for crystallisation experiments would be to screen for additives using ThermoFluor (Reinhard et al., 2013) or far-UV CD spectroscopy to identify ligands that may induce disorder-to-order transition in the protein. Alternatively, disordered proteins may be co-crystallised with an interacting partner as the binding of specific partners to some intrinsically disordered proteins or regions involves a

disorder-to-order transition, which can result in more structured conformations (Jakob et al., 2014). Apart from crystallisation studies, several other techniques are available that can be applied to learn more about the partially unfolded molten globule pp14. Information about its compactness and tertiary structure can be derived using small-angle X-ray scattering (SAXS) and CD spectroscopy in the near-UV spectral region (190–250nm) (Uversky, 2002), respectively. Heteronuclear multidimensional NMR, which is a powerful technique for protein 3D structure determination in solution (Uversky, 2002), can be used to obtain a solution structure of pp14.

Another finding, emerging from multiple lines of evidence, is that pp14 is a putative zinc-binding protein. The hypothesis that pp14 might contain a zinc-binding domain was drawn from the 3D structure predicted by the Phyre Server using the amino acid sequence of pp14B. Using the zinc finger domain of a C2HC-type zinc finger protein (cell growth regulating nucleolar protein LyaR) as the template, a beta sheet was modelled between residues 26 to 43 of pp14B, a region possessing a histidine residue (H34), and two cysteine residues (C26 and C29). This prediction indicated it might possess a C2HC-type zinc finger motif. In support of the hypothesis, zinc ions were detected in poorly ordered crystals of pp14B by X-ray fluorescence. Another sign of possible zinc binding by pp14 was observed when addition of zinc chloride to monomeric pp14B (molar ratio of protein: Zn^{2+}, 1:1.4) resulted in its precipitation. Zinc-induced precipitation of pp14B was found to specific, as magnesium chloride (molar ratio of protein:Mg^{2+}, 1:1.4) did not induce precipitation. However, zinc-induced precipitation was not found to be specific for the cysteine residues, as zinc ions were also able to induce

precipitation of the Cys-null mutant of pp14B. This indicated other residues might be involved. Further mutational analyses can be performed to identify the specific residues involved in forming a complex with Zn^{2+} promoting precipitation. Although precipitation indicates possible zinc binding by pp14, it would be interesting to determine whether it would be possible to obtain soluble complexes of zinc-pp14 by titrating protein and zinc concentrations, or varying the pH of the solution, for further analysis (Nafissi Varcheh et al., 2011). Binding of zinc ions to proteins has been shown to cause diverse structural effects from inducing disorder to order transitions (Amartely et al., 2016; Rahman et al., 2011), and structural re-arrangements (Amartely et al., 2016), to promoting dimerisation (Li et al., 2007), oligomerisation (Wineman-Fisher and Miller, 2016; Zahid et al., 2015), and precipitation and aggregation (Huang et al., 1997), consequently affecting their function. These studies thus suggest the zinc binding activity of pp14 could be functionally relevant.

Moreover, this study provides additional evidence to support previous reports that pp14 is post-translationally modified by phosphorylation. In previous reports, pp14 has been described as heavily phosphorylated (Hong et al., 1995); however the experimental evidence came from western blotting analysis of phosphatase treated or untreated cell lysates prepared from MDV-infected cells (Hong et al., 1995). The authors reported the treated sample showed a slight shift in size (1–2kDa smaller) relative to the untreated sample (Hong et al., 1995). This study, however, provides conclusive supporting evidence for the presence of two phosphorylation sites in the shared C-terminal region of pp14. Using MS, the phosphorylation analysis of recombinant pp14B extracted from DF1 cells identified Ser 72 and Ser

76 as phosphorylated. These results open the possibility of investigating the importance of phosphorylation on the structure and function of pp14 by mutagenesis. Phosphorylation is an important molecular mechanism that modifies protein structure and function (Jakob et al., 2014). Structurally, it can modify the surface properties of a protein to affect its self-association or propensity to interact with other molecules (Jakob et al., 2014). Both disorder to order and order to disorder transitions have been observed to follow the phosphorylation event (Jakob et al., 2014). In addition, phosphorylation can have functional effects, resulting in protein activation or inhibition (Jakob et al., 2014). Several studies have reported phosphorylation-induced functional unfolding of proteins (Koudelka et al., 2009; Tait et al., 2010). An example of this includes p27, an intrinsically disordered protein that folds upon binding to cylin-dependent kinase (Cdk)/cyclin complexes, inhibiting the kinase activity of Cdk (Grimmler et al., 2007; Jakob et al., 2014). It was shown that phosphorylation of Y88 in the kinase inhibitory domain (KID) of p27 exerts local unfolding on the bound conformation of KID relieving it from its inhibitory activity (Grimmler et al., 2007; Jakob et al., 2014). Based on these studies, it would be interesting to investigate if dephosphorylation of critical residues of pp14 could induce any disorder to order transitions in its structure. Moreover, mutational analyses can be performed to determine the functional importance of the identified phosphorylated sites. For example, the critical serine residues of pp14 can be mutated in the BAC clone of RB-1B and a bird challenge experiment can be carried out with this pp14 mutant virus to determine if, with these phosphorylation sites of pp14 mutated, the mutant virus would still be able to mediate neurovirulence.

Overall, the findings that in addition to adopting a flexible structure, its structure is susceptible to modifications by mechanisms such as phosphorylation and metal binding, indicate pp14 is capable of binding multiple binding partners, and may perform diverse functions under different cellular conditions or at different stages of infection. This is highly likely due to the expression of pp14 in all the different stages of infection. Although pp14 has been identified as an immediate early protein of MDV1 (Hong and Coussens, 1994), it has been detected in lytically infected, and latently infected and transformed cells of MDV1 (Hong et al., 1995).

6.1.3 Insights into the structure of chCREB3

Our current knowledge on the structure and function of CREB3 comes from various studies conducted with the mammalian CREB3. CREB3, a bZIP transcription factor, is an ER-bound protein that plays a role in the ER-stress induced unfolded protein response (UPR) (Chan et al., 2011). In addition to playing a role in ER-stress induced mechanisms, CREB3 has been reported to play a role in cell proliferation and migration (Jang et al., 2007; Kim et al., 2010), tumour suppression (Jin et al., 2000), inflammatory gene expression (Eleveld-Trancikova et al., 2010; Sung et al., 2008), and has been shown to be a target of viruses such as HSV, HCV and HIV (Blot et al., 2006; Jin et al., 2000; Lu and Misra, 2000).

CREB3 has been characterised as a type II membrane-associated protein with its N-terminal domain in the cytoplasm, its transmembrane segment spanning the ER membrane and the C-terminal domain in the ER lumen (Chan et al., 2011). CREB3, like many transcription factors, is highly disordered, and to date, no

crystal structure is available for CREB3. In an effort to gain structural and functional insights into the transcription factor, attempts were made to purify various domains of chCREB3 (the C-terminal, N-terminal and bZIP domains) for crystallisations experiments.

Crystallisation trials could not be conducted with the N-terminal domain of chCREB3, as the expression of this construct could not be scaled-up. Crystallisation trials with the C-terminal domain of chCREB3 yielded no crystals. The DNA binding domain (bZIP), composed of a leucine zipper dimerisation interface and a basic DNA binding structure, was the only predicted structured region of the protein. Predicted to favour homodimerisation within its family (Vinson et al., 2002), the bZIP domain of chCREB3 was purified from bacteria and subjected to crystallisation trials in the presence and absence of DNA targets. Unfortunately, the crystals obtained from crystallisations trials conducted with DNA targets were unstable, which dissolved in cryoprotectant solutions before they could be frozen for X-ray analysis. However, diffraction quality crystals were obtained from crystallisation trials conducted with the chCREB3 bZIP in the absence of DNA. The structure, determined by molecular replacement, provides important structural and functional insights into the homodimeric bZIP as discussed below.

This study, presenting the crystal structure, provides evidence that the bZIP domain of chCREB3 is capable of forming a stable homodimer even in the absence of DNA (Figure 6.1A). The solved 3D structure of the homodimeric chCREB3 bZIP is similar to the crystal structure of the heterodimer formed by the bZIP

domains of ATF4 and C/EBPβ in the absence DNA (Podust et al., 2001) and other previously published heterodimeric or homodimeric bZIP/DNA complexes (Figure 6.1) (Glover and Harrison, 1995; König and Richmond, 1993; Schumacher et al., 2000). This is the first crystal structure of a homodimeric bZIP containing structured basic regions in the absence of a DNA target.

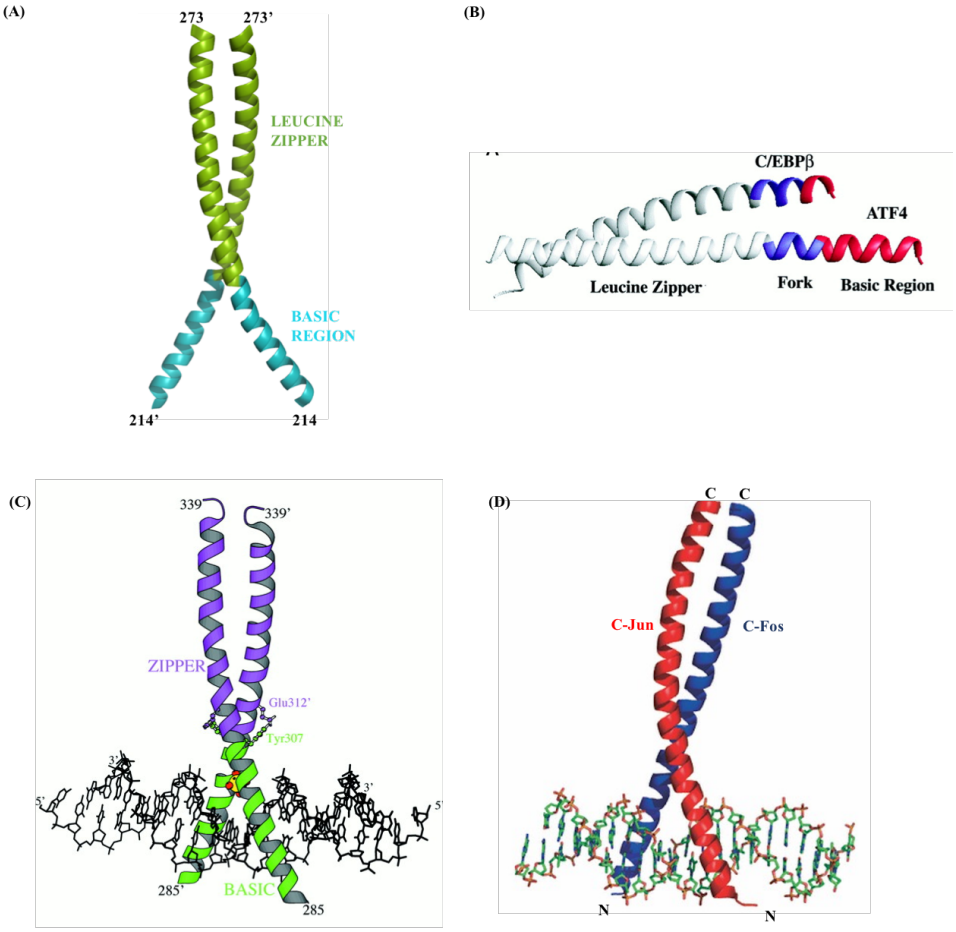


Figure 6.1 Crystal structures of bZIP domains.
(A) Crystal structure of the homodimeric chCREB3 bZIP. **(B)** Crystal structure of the heterodimer formed by the bZIP domains of ATF4 and C/EBPβ without a DNA target (PDB code 1CI6; Taken from Podust et al., 2001) **(C)** Crystal structure of the homodimeric CREB bZIP bound to a DNA oligomer containing the CRE site (PBD code: 1DH3; Taken from Schumacher et al., 2000) **(D)** Crystal structure of the heterodimeric bZIP c-Fos:c-Jun bound to a DNA oligomer containing the AP1 site (PDB code: 1FOS; From König and Richmond, 1993 and Worrall and Mason, 2011)

The homodimeric chCREB3 bZIP forms a continuous α-helix in which the C-terminal leucine zipper region forms a parallel coiled-coil dimerisation interface, and the N-terminal basic region also forms a continuous α-helix even in the absence of DNA (Figure 6.1A). The presence of a structured DNA binding region in the absence of DNA has been observed with the bZIP domain of ATF4 in the crystal structure of the heterodimeric bZIP C/EBPβ:ATF4 unbound to DNA, determined by Podust LM et al (2000) (Figure 6.1B). In this structure, the basic region of C/EBPbeta is disordered, whereas the basic region of ATF4 adopts a stable helical conformation (Podust et al., 2001). The authors of this study proposed the presence of a structured alpha helical DNA binding region could augment the DNA-binding specificity of a bZIP transcription factor (Podust et al., 2001). Generally, the basic region of most bZIP proteins is largely unfolded in the unbound state (Hollenbeck et al., 2002). Upon binding to DNA, the unfolded basic region becomes α-helical bringing into position the DNA-binding residues essential for making sequence-specific interactions with the target DNA (Miotto and Struhl, 2006). Since structural disorder imparts plasticity, which allows a protein to adopt more than one conformation (Babu et al., 2011), disorder in the basic region of bZIP domains indicates it can adapt to interact with a broad range of sequences (Podust et al., 2001). In contrast, a well-defined conformation that is already complimentary for the target DNA sequence will restrict interactions with sequences other than the target site, imposing binding specificity (Podust et al., 2001). In support of this hypothesis, the crystal structure the heterodimeric bZIP C/EBPβ:ATF4 unbound to DNA reveals that most of the conserved DNA-binding residues orient in position for specific DNA (CRE) binding in the basic DNA binding region possessing an ordered α-helical conformation (Podust et al., 2001).

It was also shown that this heterodimeric bZIP binds to the ATF binding site, CRE element, but not the canonical C/EBPβ DNA site, CCAAT box DNA, with high affinity (Podust et al., 2001).These results suggest the structured ATF4 DNA binding region might be determining the binding specificity of the heterodimeric bZIP. In addition to imposing specificity, possessing ordered DNA binding region could also enhance the DNA-binding affinity of bZIP proteins as previous studies have revealed an increase in helical content of bZIP proteins correlates with increased DNA-binding affinity (Szilák et al., 1997a, 1997b; Woolley et al., 2006). It has also been proposed that having a high intrinsic helicity in the bZIP basic region will increase the overall stability of the bound protein-DNA complex, which in turn will lead to low dissociation rates, and the converse will be true for sequences that have low intrinsic helicity (Das et al., 2012). Taken together, this structural feature of the chCREB3 bZIP suggests chCREB3 is likely to interact with its specific target site, CRE element, with high affinity (Figure 6.2).

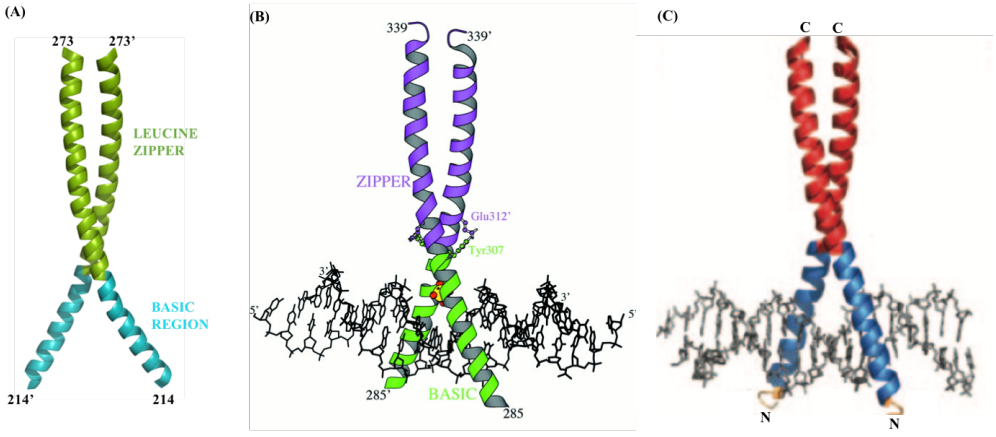


Figure 6.2 Comparison of the basic regions of bZIP domains.
The chCREB3 bZIP **(A)** has a structured basic region in the absence of DNA, which appears positioned for DNA binding, as observed in the crystal structures of the homodimeric bZIPs of CREB **(B)** and GCN4 **(C)** bound to DNA oligomers containing the CRE or ATF/CREB site (TGACGTCA). (PBD codes: 1DH3 and 1DGC). (König and Richmond, 1993; McDonald et al., 2006; Schumacher et al., 2000)

Another striking feature of the structure of the homodimeric chCREB3 bZIP is the presence of a disulphide bond in the hydrophobic core of the dimer interface (Figure 6.7). bZIP family members form dimers via their characteristic leucine zipper helices which consist of seven-residue (heptad) repeats of amino acids. The seven positions of the heptads are labelled (**a**-**b**-**c**-**d**-**e**-**f**-**g**) in one helix, and (**a'**-**b'**-**c'**-**d'**-**e'**-**f** '-**g'**) in the other. Hydrophobic residues occur preferentially in the **a** and **d** positions, with a variety of amino acids in the **a** positions and leucine in the **d** positions majority (84%) of the time (Vinson et al., 2002). Hydrophobic interactions by leucine and other hydrophobic amino acids in **a** and **d** positions in the helix form the hydrophobic core of the bZIP dimer, facilitating dimerisation (Vinson et al., 2002). Interestingly, a cysteine residue occupies the **d** position of the first heptad (Cys 246) of the chCREB3 bZIP. The structure reveals this cysteine is involved in disulphide bond formation as strong electron density can be observed connecting the cysteines from the parallel helices. As previous mutational studies have demonstrated amino acids in positions **a** and **d** regulate leucine zipper oligomerisation, dimerisation stability and specificity (Vinson et al., 2002), this cysteine residue replacing the consensus leucine at this position, by forming a covalent bond, may play a role in promoting stability. The bZIP domain of ATF5 which possess a cysteine residue at position **a** of the first heptad has been reported to form an unstable homodimer in the absence and presence of a disulphide bond (Ciaccio and Laurence, 2009). This suggests an intermolecular disulphide bond formed between cysteine residues at position **d** of the first heptad increases stability, thereby promoting homodimerisation, whereas an intermolecular disulphide bond formed between cysteine residues at position **a** of the first heptad does not promote stability/homodimerisation.

The disulphide bond in the chCREB3 bZIP could also play a role in extending the stabilising α-helical conformation from the leucine zipper region to the basic region. In a previous study, Ciaccio NA et al (2009) investigated the effects of disulphide bond formation on the aggregation of the bZIP domain of ATF5, which is only partially structured and displays marginal stability in solution under physiological conditions (Ciaccio and Laurence, 2009). This domain consists of a single α-helix and possesses a single cysteine residue. The authors reported the percent helicity of the ATF5 bZIP increased from 19.0% to 23.4% in the presence of a intermolecular disulphide bond, and the covalent cross-linking helped to retain helicity, which inhibited the formation of large aggregates (Ciaccio and Laurence, 2009). The authors proposed disordered regions of ATF5 bZIP would play a role in initiating the process of aggregation (Ciaccio and Laurence, 2009). To determine if the disulphide bond is essential for promoting stability and preventing aggregation, the chCREB3 bZIP can be purified in the presence of a reducing agent or mutational analysis, by substituting the cysteine with a leucine, can be performed.

Another important observation from the crystal structure is the presence of electron density connecting equivalent asparagine residues in position **a** of the second and forth heptads (N250 and N64) of the parallel helices (Figure 6.8), indicative of strong interactions between these residues. Typically hydrophobic residues occur in the **a** and **d** positions of the leucine zipper (Vinson et al., 2002) which facilitate dimerisation through hydrophobic interactions. However, it has been shown that polar residues, which sometimes replace hydrophobic residues at these positions drive dimerisation specificity (Hakoshima, 2014; Vinson et al., 2002). For

example, in most homodimerising bZIP transcription factors position **a** of the second heptad is occupied by an asparagine residue (Vinson et al., 2002). Asparagine in this position has been shown to hydrogen bond interhelically with asparagine in the equivalent position of the second monomer (Harbury et al. 1993). A second asparagine in position **a** of the forth heptad (Figure 5.8A), found only in members of the CREB3 subfamily, has been predicted to encourage homodimerisation within the family, excluding heterodimerisation with homodimerising families without an asparagine in this position (Vinson et al., 2002). The strong electron density observed in this X-ray structure suggests these residues are not only involved in promoting homodimerisation, but may also be important for inducing stability, which could be functionally important for these bZIP factors.

Another important finding deduced from the multiple sequence alignment analysis conducted with the bZIP domain sequences of chicken, human and mouse bZIP transcription factors is that while the majority of bZIP domains contain a leucine at position **d** of the first heptad, only members of the CREB3 subfamily possess a cysteine at this position (Figure 5.9). Taken together with the crystal structure of the chCREB3 bZIP providing evidence that this cysteine has the capacity to form a disulphide bond with the equivalent cysteine in the parallel helix, these results suggest this cysteine at this position may serve a functional role regulating activity of these transcription factors, *in vivo.* A number of transcription factors have been reported to contain redox-sensitive cysteine residues at their DNA-binding sites, which affect their DNA binding ability (Trachootham et al., 2008). In most cases, oxidation of critical cysteine residues abolishes DNA-binding activity of these

transcription factors (Trachootham et al., 2008). For example, under non-reducing conditions formation of an intermolecular disulphide bridge between the cysteine residues of each basic motif within the Fos/Jun dimer prevents DNA binding (Bannister et al., 1991). Other examples of factors whose DNA-binding activity is negatively regulated by oxidation of critical cysteine residues include the nuclear factor kappa B (NF-κB), Hypoxia-inducible factor 1-alpha (HIF-1α), P53 and the nuclear receptor proteins, oestrogen and glucocorticoid receptors (Amoutzias et al., 2006; Arrigo, 1999; Morel and Barouki, 1999). Specific proteins are involved in facilitating the DNA-binding activity of several of these factors by reducing the cysteine residue under oxidative conditions (Trachootham et al., 2008). For example, redox effector factor 1 (Ref-1) could stimulate the DNA binding activity of Fos-Jun dimers by reducing the cysteine residue (Xanthoudakis et al., 1992). In contrast, an example of a transcription factor whose DNA-binding activity is up-regulated by oxidative stress is HoxB5, a member of the Hox family of proteins (Arrigo, 1999; Galang and Hauser, 1993). Oxidative conditions trigger the dimerisation of HoxB5, a structural reorganisation required for the cooperative binding of this factor to tandem DNA target sites(Arrigo, 1999; Galang and Hauser, 1993). A critical cysteine residue (Cys 232) located in the highly conserved homeobox encoded DNA-binding domain is essential for this redox regulation (Arrigo, 1999; Galang and Hauser, 1993). In the case of the chCREB3 bZIP, the crystal structure reveals this cysteine, located downstream of the DNA binding region, takes part in dimerisation. Hence, as has been observed with HoxB5 oxidation of this cysteine is likely to upregulate the DNA binding activity of chCREB3 by promoting dimerisation and structural changes required for DNA-binding (Figure 6.2). However, further experiments will need to be performed to

determine the DNA-binding affinity of the homodimeric chCREB3 bZIP under reducing and non-reducing conditions and to define the exact role of this cysteine located in the first heptad of the chCREB3 bZIP.

Taken together, these findings suggest members of the CREB3 subfamily, by possessing a putative redox-sensitive cysteine residue within the bZIP domain, might be involved in cell responses to oxidative stress. Insights from the crystal structure of chCREB3 bZIP—covalent intermolecular disulphide bond inducing homodimerisation stability and possession of a structured DNA binding region that could potentially enhance both DNA binding specificity and affinity—suggest activity of these factors is likely to be enhanced during oxidative stress responses. Oxidative stress occurs when the level of reactive oxidative species (ROS) exceeds the antioxidant capacity of the cell and can be observed during viral infections and in different pathologic states, such as cancer, atherosclerosis, rheumatoid arthritis, acquired immunodeficiency syndrome (AIDS), Parkinson's, and Alzheimer's diseases or when cells are exposed to different oxidative agents (Arrigo, 1999). In general, ROS at low levels act as signalling molecules that promote cell proliferation and cell survival. In contrast, a severe increase in ROS can induce cell death (Trachootham et al., 2008).

6.2 Summary and final remarks

This DPhil project provides new structural and functional insights into the 14kDa phosphoprotein of MDV1 associated with the neurovirulence phenotype of the virus. Biophysical characterisation of recombinant pp14B identifies pp14 as a molten globule. The results reveal the protein, while possessing substantial

secondary structure, is largely disordered lacking a stable tertiary structure. The results of this project also indicate pp14 is a putative zinc-binding protein. In addition, supporting previous reports, this project provides conclusive evidence for the presence of two phosphorylation sites in the shared C-terminal region of pp14—serines 72 and 76 of pp14B—opening the possibility of investigating the importance of phosphorylation on the structure and function of pp14. Structural flexibility, through a lack of a definite ordered tertiary structure, and functional features that can induce structural modifications, indicate pp14 might interact with a number of binding partners and therefore could play multiple roles during MDV infection—a strong possibility due to the expression of the protein in all the different stages of virus infection.

Furthermore, findings from this project provide structural and functional insights into the bZIP transcription factor chCREB3. The X-ray structure of the chCREB3 bZIP provides evidence that chCREB3 is capable of forming a stable homodimer. The crystal structure reveals the bZIP domain of chCREB3 possesses a structured DNA binding region even in the absence of DNA, a feature that could potentially enhance both the DNA binding specificity and affinity of chCREB3. Significantly, the homodimeric bZIP possesses an intermolecular disulphide bond connecting equivalent cysteine residues at position **d** of the first heptad of the parallel helices. This covalent disulphide bond in the hydrophobic core of the bZIP may play a role in promoting stability. The sequence alignment of bZIP sequences from chicken, human and mouse reveals only members of the CREB3 subfamily possess this cysteine residue, indicating it could act as a redox-sensor. These results indicate members of the CREB3 subfamily, by possessing a putative redox-sensitive

cysteine with the capacity to form an intermolecular disulphide bond, may be activated in response to oxidative stress. Insights from the crystal structure of chCREB3 bZIP—covalent intermolecular disulphide bond inducing homodimerisation stability and possession of a structured DNA binding region that could potentially enhance both DNA binding specificity and affinity—suggest activity of these transcription factors is likely to be potentiated during oxidative stress responses.

A combination of techniques (biochemical, cellular and functional assays), using recombinant proteins from different cell lines or MDV infected cells, was used to investigate the putative interaction between pp14 and chicken CREB3. This investigation, unfortunately, found no conclusive evidence in support of the predicted interaction. The current knowledge on viral and host cell factors, and the underlying mechanisms that contribute to MDV neurovirulence is limited. While pp14 is the only viral factor that has been associated with MDV neurovirulence to date, its mechanisms of action are not known. In this regard, identifying host cell factors that interact with pp14 and understanding the role and effect of interactions between pp14 and specific host cell factors on MD would be a step forward towards gaining a better understanding of the mechanisms involved in neuropathology. Several novel approaches to screen for binding partners are available, which can be used to probe for new interacting partners for pp14. These include the dual expression recombinase based (DERB) single vector system (Lu et al., 2008), a cellular tool for high-throughput screening of protein-protein interactions, and cell free *in vitro* approaches such as the tandem affinity purification (TAP) method, chemical crosslinking in combination with MS or co-

affinity purification coupled with MS (Berggård et al., 2007). It is advantageous that a specific antibody against pp14 is available as it can be used to capture and identify biologically relevant interaction partners from MDV infected or transformed cells. If novel binding partners can be identified, co-crystallisation with a binding partner may make it possible to obtain diffraction quality crystals of pp14 for high-resolution structure determination. The knowledge gained from further characterisation of the viral factor pp14, and its interactions with host factors may one day be exploited to develop interventions against MDV, which continues to evolve into increasingly neurovirulent forms.

References

Abdul-Careem, M.F., Haq, K., Shanmuganathan, S., Read, L.R., Schat, K.A., Heidari, M., Sharif, S., 2009. Induction of innate host responses in the lungs of chickens following infection with a very virulent strain of Marek's disease virus. Virology 393, 250–257.

Abdul-Careem, M.F., Hunter, B.D., Sarson, A.J., Mayameei, A., Zhou, H., Sharif, S., 2006. Marek's Disease Virus–Induced Transient Paralysis Is Associated with Cytokine Gene Expression in the Nervous System. Viral Immunol. 19, 167–176. https://doi.org/10.1089/vim.2006.19.167

Ahmed, M.S., Hiroki, O., Sasaki, J., Ochiai, K., Goryo, M., 2016. Persistence of chicken anemia virus antigen and inclusions in spontaneous cases of Marek's disease visceral lymphomas in broiler chickens at slaughterhouses. J. Vet. Med. Sci. 78, 825–829.

Amartely, H., David, A., Shamir, M., Lebendiker, M., Izraeli, S., Friedler, A., 2016. Differential effects of zinc binding on structured and disordered regions in the multidomain STIL protein. Chem. Sci. 7, 4140–4147.

Amor, S., Strassheim, S., Dambrine, G., Remy, S., Rasschaert, D., Laurent, S., 2011. ICP27 protein of Marek's disease virus interacts with SR proteins and inhibits the splicing of cellular telomerase chTERT and viral vIL8 transcripts. J. Gen. Virol. 92, 1273–1278.

Amoutzias, G.D., Bornberg-Bauer, E., Oliver, S.G., Robertson, D.L., 2006. Reduction/oxidation-phosphorylation control of DNA binding in the bZIP dimerization network. BMC Genomics 7, 107.

Arrigo, A.-P., 1999. Gene expression and the thiol redox state. Free Radic. Biol. Med. 27, 936–944. https://doi.org/10.1016/S0891-5849(99)00175-6

Arvin, A., Campadelli-Fiume, G., Mocarski, E., Moore, P.S., Roizman, B., Whitley, R., Yamanishi, K., 2007. Human herpesviruses: biology, therapy, and immunoprophylaxis. Cambridge University Press.

Atkins, K.E., Read, A.F., Savill, N.J., Renz, K.G., Islam, A., Walkden-Brown, S.W., Woolhouse, M.E., 2013. Vaccination and reduced cohort duration can drive virulence evolution: Marek's disease virus and industrialized agriculture. Evolution 67, 851–860.

Babu, M.M., 2016. The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. Biochem. Soc. Trans. 44, 1185. https://doi.org/10.1042/BST20160172

Babu, M.M., van der Lee, R., de Groot, N.S., Gsponer, J., 2011. Intrinsically disordered proteins: regulation and disease. Curr. Opin. Struct. Biol. 21, 432–440. https://doi.org/10.1016/j.sbi.2011.03.011

Bacskai, B.J., Skoch, J., Hickey, G.A., Allen, R., Hyman, B.T., 2003. Fluorescence resonance energy transfer determinations using multiphoton fluorescence lifetime imaging microscopy to characterize amyloid-beta plaques. J. Biomed. Opt. 8, 368–375.

Baigent, S.J., Smith, L.P., Nair, V.K., Currie, R.J., 2006. Vaccinal control of Marek's disease: current challenges, and future strategies to maximize protection. Vet. Immunol. Immunopathol. 112, 78–86.

Baird, L., Swift, S., Llères, D., Dinkova-Kostova, A.T., 2014. Monitoring Keap1–Nrf2 interactions in single live cells. Plants Pharm. Shelf Recent Trends Leads Find. Bioprod. 32, 1133–1144. https://doi.org/10.1016/j.biotechadv.2014.03.004

Bannister, A., Cook, A., Kouzarides, T., 1991. In vitro DNA binding activity of Fos/Jun and BZLF1 but not C/EBP is affected by redox changes. Oncogene 6, 1243–1250.

Barbosa, S., Fasanella, G., Carriera, S., Llarena, M., Fox, R., Barreca, C., Andrew, D., O'Hare, P., 2013. An orchestrated programme regulating secretory pathway genes and cargos by the transmembrane transcription factor CREB-H. Traffic Cph. Den. 14, 382–398. https://doi.org/10.1111/tra.12038

Ben-Porat, T., Rixon, F.J., 1979. Replication of herpesvirus DNA IV: Analysis of concaterners. Virology 94, 61–70.

Berezin, M.Y., Achilefu, S., 2010. Fluorescence Lifetime Measurements and Biological Imaging. Chem. Rev. 110, 2641–2684. https://doi.org/10.1021/cr900343z

Berg, R.H., Beachy, R.N., 2008. Fluorescent Protein Applications in Plants, in: Methods in Cell Biology. Academic Press, pp. 153–177. https://doi.org/10.1016/S0091-679X(08)85008-X

Berrow, N.S., Alderton, D., Owens, R.J., 2009. The Precise Engineering of Expression Vectors Using High-Throughput In-Fusion™ PCR Cloning. High Throughput Protein Expr. Purif. Methods Protoc. 75–90.

Bird, L.E., Rada, H., Flanagan, J., Diprose, J.M., Gilbert, R.J., Owens, R.J., 2014. Application of In-Fusion™ cloning for the parallel construction of E. coli expression vectors. DNA Cloning Assem. Methods 209–234.

Blom, N., Gammeltoft, S., Brunak, S., 1999. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites1 1Edited by F. E. Cohen. J. Mol. Biol. 294, 1351–1362. https://doi.org/10.1006/jmbi.1999.3310

Blot, G., Lopez-Vergès, S., Treand, C., Kubat, N.J., Delcroix-Genête, D., Emiliani, S., Benarous, R., Berlioz-Torrent, C., 2006. Luman, a New Partner of HIV-1 TMgp41, Interferes with Tat-mediated Transcription of the HIV-1 LTR. J. Mol. Biol. 364, 1034–1047. https://doi.org/10.1016/j.jmb.2006.09.080

Boodhoo, N., Gurung, A., Sharif, S., Behboudi, S., 2016. Marek's disease in chickens: a review with focus on immunology. Vet. Res. 47, 119. https://doi.org/10.1186/s13567-016-0404-3

Broussard, J.A., Rappaz, B., Webb, D.J., Brown, C.M., 2013. Fluorescence resonance energy transfer microscopy as demonstrated by measuring the activation of the serine/threonine kinase Akt. Nat. Protoc. 8, 265–281. https://doi.org/10.1038/nprot.2012.147

Brückner, A., Polge, C., Lentze, N., Auerbach, D., Schlattner, U., 2009. Yeast Two-Hybrid, a Powerful Tool for Systems Biology. Int. J. Mol. Sci. 10, 2763–2788. https://doi.org/10.3390/ijms10062763

Brunovskis, P., Velicer, L.F., 1995. The Marek's disease virus (MDV) unique short region: alphaherpesvirus-homologous, fowlpox virus-homologous, and MDV-specific genes. Virology 206, 324–338.

Burbelo, P.D., Gabriel, G.C., Kibbey, M.C., Yamada, Y., Kleinman, H.K., Weeks, B.S., 1994. LZIP-1 and LZIP-2: two novel members of the bZIP family. Gene 139, 241–245.

Burnside, J., Bernberg, E., Anderson, A., Lu, C., Meyers, B.C., Green, P.J., Jain, N., Isaacs, G., Morgan, R.W., 2006. Marek's disease virus encodes MicroRNAs that map to meq and the latency-associated transcript. J. Virol. 80, 8778–8786.

Cantello, J.L., Anderson, A.S., Morgan, R.W., 1994. Identification of latency-associated transcripts that map antisense to the ICP4 homolog gene of Marek's disease virus. J. Virol. 68, 6280–6290.

Carpenter, J.E., Grose, C., 2014. Varicella-zoster virus glycoprotein expression differentially induces the unfolded protein response in infected cells. Front. Microbiol. 5, 322. https://doi.org/10.3389/fmicb.2014.00322

Causier, B., Davies, B., 2002. Analysing protein-protein interactions with the yeast two-hybrid system. Plant Mol. Biol. 50, 855–870.

Chan, C.-P., Kok, K.-H., Jin, D.-Y., 2011. CREB3 subfamily transcription factors are not created equal: Recent insights from global analyses and animal models. Cell Biosci. 1, 6. https://doi.org/10.1186/2045-3701-1-6

Chan, S.-W., 2014. The unfolded protein response in virus infections. Front. Microbiol. 5, 518. https://doi.org/10.3389/fmicb.2014.00518

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 66, 12–21. https://doi.org/10.1107/S0907444909042073

Cho, K., Endoh, D., Qian, J., Ochiai, K., Onuma, M., Itakura, C., 1998. Central nervous system lesions induced experimentally by a very virulent strain of Marek's disease virus in Marek's disease-resistant chickens. Avian Pathol. 27, 512–517.

Churchill, A., Payne, L., Chubb, R., 1969. Immunization against Marek's disease using a live attenuated virus. Nature 221, 744–747.

Ciaccio, N.A., Laurence, J.S., 2009. Effects of Disulfide Bond Formation and Protein Helicity on the Aggregation of Activating Transcription Factor 5 (ATF5). Mol. Pharm. 6, 1205–1215. https://doi.org/10.1021/mp900058t

Cohen, C., Parry, D.A., 1990. α-Helical coiled coils and bundles: how to design an α-helical protein. Proteins Struct. Funct. Bioinforma. 7, 1–15.

Das, R.K., Crick, S.L., Pappu, R.V., 2012. N-terminal segments modulate the α-helical propensities of the intrinsically disordered basic regions of bZIP proteins. J. Mol. Biol. 416, 287–299.

Das, U., Hariprasad, G., Ethayathulla, A.S., Manral, P., Das, T.K., Pasha, S., Mann, A., Ganguli, M., Verma, A.K., Bhat, R., 2007. Inhibition of protein aggregation: supramolecular assemblies of arginine hold the key. PloS One 2, e1176.

Davison, F., Nair, V., 2005. Use of Marek's disease vaccines: could they be driving the virus to increasing virulence? Expert Rev. Vaccines 4, 77–88.

Davison, F., Nair, V., 2004. Marek's disease: an evolving problem. Academic Press.

Denard, B., Seemann, J., Chen, Q., Gay, A., Huang, H., Chen, Y., Ye, J., 2011. The Membrane-Bound Transcription Factor CREB3L1 Is Activated in Response to Virus Infection to Inhibit Proliferation of Virus-Infected Cells. Cell Host Microbe 10, 65–74. https://doi.org/10.1016/j.chom.2011.06.006

DenBoer, L.M., Hardy-Smith, P.W., Hogan, M.R., Cockram, G.P., Audas, T.E., Lu, R., 2005. Luman is capable of binding and activating transcription from the unfolded protein response element. Biochem. Biophys. Res. Commun. 331, 113–119.

Deng, X., Li, X., Shen, Y., Qiu, Y., Shi, Z., Shao, D., Jin, Y., Chen, H., Ding, C., Li, L., Chen, P., Ma, Z., 2010. The Meq oncoprotein of Marek's disease virus interacts with p53 and inhibits its transcriptional and apoptotic activities. Virol. J. 7, 348–348. https://doi.org/10.1186/1743-422X-7-348

Dorange, F., Tischer, B.K., Vautherot, J.-F., Osterrieder, N., 2002. Characterization of Marek's disease virus serotype 1 (MDV-1) deletion mutants that lack UL46 to UL49 genes: MDV-1 UL49, encoding VP22, is indispensable for virus growth. J. Virol. 76, 1959–1970.

Elangovan, M., Day, R.N., Periasamy, A., 2002. Dynamic imaging using fluorescence resonance energy transfer. Biotechniques 32, 1260–1265.

Eleveld-Trancikova, D., Sanecka, A., van Hout-Kuijer, M.A., Looman, M.W., Hendriks, I.A., Jansen, B.J., Adema, G.J., 2010. DC-STAMP interacts with ER-resident transcription factor LUMAN which becomes activated during DC maturation. Mol. Immunol. 47, 1963–1973.

Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. Sect. D 60, 2126–2132.

Evans, P., 2006. Scaling and assessment of data quality. Acta Crystallogr. D Biol. Crystallogr. 62, 72–82.

Fink, A.L., 2001. Molten Globule, in: ELS. John Wiley & Sons, Ltd. https://doi.org/10.1038/npg.els.0003007

Foster, D., 1998. Development of a spontaneously immortalized chicken embryo fibroblastic cell line. Virology 248, 305–311.

Fox, R.M., Andrew, D.J., 2015. Transcriptional regulation of secretory capacity by bZip transcription factors. Front. Biol. 10, 28–51.

Fujii, Y., Ohira, T., Kyougoku, Y., Toda, T., Yanagida, M., Hakoshima, T., 1998. Crystallographic characterization of Pap1–DNA complex. Acta Crystallogr. D Biol. Crystallogr. 54, 1014–1016.

Galang, C.K., Hauser, C.A., 1993. Cooperative DNA binding of the human HoxB5 (Hox-2.1) protein is under redox regulation in vitro. Mol. Cell. Biol. 13, 4609–4617.

Gill, M.L., Byrd, R.A., Palmer III, A.G., 2016. Dynamics of GCN4 facilitate DNA interaction: a model-free analysis of an intrinsically disordered region. Phys. Chem. Chem. Phys. 18, 5839–5849.

Gimeno, I.M., 2008. Marek's disease vaccines: A solution for today but a worry for tomorrow? Evol. Consequences Vaccin. 26, Supplement 3, C31–C41. https://doi.org/10.1016/j.vaccine.2008.04.009

Gimeno, I.M., Witter, R.L., Hunt, H.D., Reddy, S.M., Neumann, U., 2001. Differential attenuation of the induction by Marek's disease virus of transient paralysis and persistent neurological disease: a model for pathogenesis studies. Avian Pathol. 30, 397–409.

Glover, J.M., Harrison, S.C., 1995. Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. Nature 373, 257.

Gong, Z., Zhang, L., Wang, J., Chen, L., Shan, H., Wang, Z., Ma, H., 2013. Isolation and analysis of a very virulent Marek's disease virus strain in China. Virol. J. 10, 155–155. https://doi.org/10.1186/1743-422X-10-155

Graham, F., Smiley, J., Russell, W., Nairn, R., 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J. Gen. Virol. 36, 59–72.

Grimmler, M., Wang, Y., Mund, T., Cilenšek, Z., Keidel, E.-M., Waddell, M.B., Jäkel, H., Kullmann, M., Kriwacki, R.W., Hengst, L., 2007. Cdk-Inhibitory Activity and Stability of p27Kip1 Are Directly Regulated by Oncogenic Tyrosine Kinases. Cell 128, 269–280. https://doi.org/10.1016/j.cell.2006.11.047

Gupta, M., Deka, D., Ramneek, 2016. Sequence analysis of Meq oncogene among Indian isolates of Marek's disease herpesvirus. Meta Gene 9, 230–236. https://doi.org/10.1016/j.mgene.2016.07.009

Hakoshima, T., 2014. Leucine Zippers. eLS.

Hamada, H., Arakawa, T., Shiraki, K., 2009. Effect of additives on protein aggregation. Curr. Pharm. Biotechnol. 10, 400–407.

Heidari, M., Wang, D., Delekta, P., Sun, S., 2016. Marek's disease virus immunosuppression alters host cellular responses and immune gene expression in the skin of infected chickens. Vet. Immunol. Immunopathol. 180, 21–28.

Hense, A., Prunsche, B., Gao, P., Ishitsuka, Y., Nienhaus, K., Ulrich Nienhaus, G., 2015. Monomeric Garnet, a far-red fluorescent protein for live-cell STED imaging. Sci. Rep. 5, 18006.

Hildebrandt, E., Dunn, J.R., Perumbakkam, S., Niikura, M., Cheng, H.H., 2014. Characterizing the molecular basis of attenuation of Marek's disease virus via in vitro serial passage identifies de novo mutations in the helicase-primase subunit gene UL5 and other candidates associated with reduced virulence. J. Virol. https://doi.org/10.1128/JVI.03869-13

Hollenbeck, J.J., McClain, D.L., Oakley, M.G., 2002. The role of helix stabilizing residues in GCN4 basic region folding and DNA binding. Protein Sci. Publ. Protein Soc. 11, 2740–2747.

Hong, Y., Coussens, P.M., 1994. Identification of an immediate-early gene in the Marek's disease virus long internal repeat region which encodes a unique 14-kilodalton polypeptide. J. Virol. 68, 3593–3603.

Hong, Y., Frame, M., Coussens, P.M., 1995. A 14-kDa immediate-early phosphoprotein is specifically expressed in cells infected with oncogenic Marek's disease virus strains and their attenuated derivatives. Virology 206, 695–700. https://doi.org/10.1016/S0042-6822(95)80091-3

Huang, X., Atwood, C.S., Moir, R.D., Hartshorn, M.A., Vonsattel, J.-P., Tanzi, R.E., Bush, A.I., 1997. Zinc-induced Alzheimer's Aβ1–40 aggregation is mediated by conformational factors. J. Biol. Chem. 272, 26464–26470.

Hussain, R., Benning, K., Myatt, D., Javorfi, T., Longo, E., Rudd, T.R., Pulford, B., Siligardi, G., 2015. CDApps: integrated software for experimental planning and data processing at beamline B23, Diamond Light Source. Corrigendum. J. Synchrotron Radiat. 22, 862–862. https://doi.org/10.1107/S1600577515007602

Ishikura-Kinoshita, S., Saeki, H., Tsuji-Naito, K., 2012. BBF2H7-Mediated Sec23A Pathway Is Required for Endoplasmic Reticulum-to-Golgi Trafficking in Dermal Fibroblasts to Promote Collagen Synthesis. J. Invest. Dermatol. 132, 2010–2018. https://doi.org/10.1038/jid.2012.103

Jakob, U., Kriwacki, R., Uversky, V.N., 2014. Conditionally and Transiently Disordered Proteins: Awakening Cryptic Disorder To Regulate Protein Function. Chem. Rev. 114, 6779–6805. https://doi.org/10.1021/cr400459c

Jang, S.-W., Kim, Y.S., Kim, Y.R., Sung, H.J., Ko, J., 2007. Regulation of Human LZIP Expression by NF-κB and Its Involvement in Monocyte Cell Migration Induced by Lkn-1. J. Biol. Chem. 282, 11092–11100. https://doi.org/10.1074/jbc.M607962200

Jenkins, F.J., Hoffman, L.J., 2000. Overview of herpesviruses, in: Infectious Causes of Cancer. Springer, pp. 33–49.

Jin, D.-Y., Wang, H.-L., Zhou, Y., Chun, A.C.S., Kibler, K.V., Hou, Y.-D., Kung, H., Jeang, K.-T., 2000. Hepatitis C virus core protein-induced loss of LZIP function correlates with cellular transformation. EMBO J. 19, 729–740. https://doi.org/10.1093/emboj/19.4.729

Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., 2013. DNA-binding specificities of human transcription factors. Cell 152, 327–339.

Joy, I., Lee, H., Luxton, G.G., Smith, G.A., 2006. Identification of an essential domain in the herpesvirus VP1/2 tegument protein: the carboxy terminus directs incorporation into capsid assemblons. J. Virol. 80, 12086–12094.

Kanemoto, S., Kobayashi, Y., Yamashita, T., Miyamoto, T., Cui, M., Asada, R., Cui, X., Hino, K., Kaneko, M., Takai, T., Matsuhisa, K., Takahashi, N., Imaizumi, K., 2015. Luman is involved in osteoclastogenesis through the regulation of DC-STAMP expression, stability and localization. J. Cell Sci. 128, 4353–4365. https://doi.org/10.1242/jcs.176057

Kang, H., Kim, Y.S., Ko, J., 2009. A novel isoform of human LZIP negatively regulates the transactivation of the glucocorticoid receptor. Mol. Endocrinol. 23, 1746–1757.

Kappel, C., 2012. FRET with FLIM: Quantitative In-vivo Biochemistry.

Kato, K., Izumiya, Y., Tohya, Y., Takahashi, E., Hirai, K., Kawaguchi, Y., 2002. Identification and characterization of Marek's disease virus serotype 1 (MDV1) ICP22 gene product: MDV1 ICP22 transactivates the MDV1 ICP27 promoter synergistically with MDV1 ICP4. Vet. Microbiol. 85, 305–313. https://doi.org/10.1016/S0378-1135(01)00522-3

Kato, S., Hirai, K., 1985. Marek's disease virus. Adv. Virus Res. 30, 225–277.

Kaufer, B.B., Jarosinski, K.W., Osterrieder, N., 2011. Herpesvirus telomeric repeats facilitate genomic integration into host telomeres and mobilization of viral DNA during reactivation. J. Exp. Med. 208, 605–615.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., Sternberg, M.J.E., 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 10, 845–858.

Kim, H.-C., Choi, K.-C., Choi, H.-K., Kang, H.-B., Kim, M.-J., Lee, Y.-H., Lee, O.-H., Lee, J., Kim, Y.J., Jun, W., 2010. HDAC3 selectively represses CREB3-mediated transcription and migration of metastatic breast cancer cells. Cell. Mol. Life Sci. 67, 3499–3510.

Kondo, S., Murakami, T., Tatsumi, K., Ogata, M., Kanemoto, S., Otori, K., Iseki, K., Wanaka, A., Imaizumi, K., 2005. OASIS, a CREB/ATF-family member, modulates UPR signalling in astrocytes. Nat. Cell Biol. 7, 186–194.

Kondo, S., Saito, A., Hino, S., Murakami, T., Ogata, M., Kanemoto, S., Nara, S., Yamashita, A., Yoshinaga, K., Hara, H., Imaizumi, K., 2007. BBF2H7, a Novel Transmembrane bZIP Transcription Factor, Is a New Type of Endoplasmic Reticulum Stress Transducer. Mol. Cell. Biol. 27, 1716–1729. https://doi.org/10.1128/MCB.01552-06

Kong, Y., Janssen, B.J., Malinauskas, T., Vangoor, V.R., Coles, C.H., Kaufmann, R., Ni, T., Gilbert, R.J., Padilla-Parra, S., Pasterkamp, R.J., Jones, E.Y., 2016. Structural Basis for Plexin Activation and Regulation. Neuron 91, 548–560. https://doi.org/10.1016/j.neuron.2016.06.018

König, P., Richmond, T.J., 1993. The X-ray Structure of the GCN4-bZIP Bound to ATF/CREB Site DNA Shows the Complex Depends on DNA Flexibility. J. Mol. Biol. 233, 139–154. https://doi.org/10.1006/jmbi.1993.1490

Koudelka, T., Hoffmann, P., Carver, J.A., 2009. Dephosphorylation of αs-and β-caseins and its effect on chaperone activity: a structural and functional investigation. J. Agric. Food Chem. 57, 5956–5964.

Krishnan, R.V., Masuda, A., Centonze, V., Herman, B., 2003. Quantitative imaging of protein–protein interactions by multiphoton fluorescence lifetime imaging microscopy using a streak camera. J. Biomed. Opt. 8, 362–367.

Lam, A., St-Pierre, F., Gong, Y., Marshall, J.D., Cranfill, P.J., Baird, M.A., McKeown, M.R., Wiedenmann, J., Davidson, M.W., Schnitzer, M.J., Tsien, R.Y., Lin, M.Z., 2012. Improving FRET dynamic range with bright green and red fluorescent proteins. Nat. Methods 9, 1005–1012. https://doi.org/10.1038/nmeth.2171

Lebendiker, M., Danieli, T., 2014. Production of prone-to-aggregate proteins. FEBS Lett. 588, 236–246.

Lee, L.F., Wu, P., Sui, D., Ren, D., Kamil, J., Kung, H.J., Witter, R.L., 2000. The complete unique long sequence and the overall genomic organization of the GA strain of Marek's disease virus. Proc Natl Acad Sci USA 97, 6091–6096.

Li, H., Zhao, Y., Guo, Y., Li, Z., Eisele, L., Mourad, W., 2007. Zinc Induces Dimerization of the Class II Major Histocompatibility Complex Molecule That Leads to Cooperative Binding to a Superantigen(). J. Biol. Chem. 282, 5991–6000. https://doi.org/10.1074/jbc.M608482200

Liang, G., Audas, T.E., Li, Y., Cockram, G.P., Dean, J.D., Martyn, A.C., Kokame, K., Lu, R., 2006. Luman/CREB3 Induces Transcription of the Endoplasmic Reticulum (ER) Stress Response Protein Herp through an ER Stress Response Element. Mol. Cell. Biol. 26, 7999–8010. https://doi.org/10.1128/MCB.01046-06

Lleres, D., Swift, S., Lamond, A.I., 2007. Detecting protein-protein interactions in vivo with FRET using multiphoton fluorescence lifetime imaging microscopy (FLIM). Curr. Protoc. Cytom. 12–10.

Lonze, B., Ginty, D., 2002. Function and regulation of CREB family transcription factors in the nervous system. Neuron 35, 605–623. https://doi.org/10.1016/S0896-6273(02)00828-0

Lu, R., Misra, V., 2000. Potential Role for Luman, the Cellular Homologue of Herpes Simplex Virus VP16 (α Gene trans-Inducing Factor), in Herpesvirus Latency. J. Virol. 74, 934–943.

Lu, R., Yang, P., O'Hare, P., Misra, V., 1997. Luman, a new member of the CREB/ATF family, binds to herpes simplex virus VP16-associated host cellular factor. Mol. Cell. Biol. 17, 5117–5126.

Lupiani, B., Lee, L., Reddy, S., 2001. Protein-coding content of the sequence of Marek's disease virus serotype 1. Curr. Top. Microbiol. Immunol. 255, 159–190.

Mahy, B.W., Van Regenmortel, M.H., 2010. Desk Encyclopedia Animal and Bacterial Virology. Academic Press.

Marek, J., 1907. Multiple nervenentzuendung (polyneuritis) bei huehnern. Dtsch. Tierärztl. Wochenschr. 15, 417–421.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674. https://doi.org/10.1107/S0021889807021206

McDonald, R.J., Kahn, J.D., Maher III, L.J., 2006. DNA bending by bHLH charge variants. Nucleic Acids Res. 34, 4846–4856.

McPherson, M., Delany, M., 2016. Virus and host genomic, molecular, and cellular interactions during Marek's disease pathogenesis and oncogenesis. Poult. Sci. 95, 412–429.

McPherson, M.C., Cheng, H.H., Delany, M.E., 2016. Marek's disease herpesvirus vaccines integrate into chicken host chromosomes yet lack a virus-host phenotype associated with oncogenic transformation. Vaccine 34, 5554–5561.

Micsonai, A., Wien, F., Kernya, L., Lee, Y.-H., Goto, Y., Réfrégiers, M., Kardos, J., 2015. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. Proc. Natl. Acad. Sci. 112, E3095–E3103.

Miles, A.J., Wallace, B., 2006. Synchrotron radiation circular dichroism spectroscopy of proteins and applications in structural and functional genomics. Chem. Soc. Rev. 35, 39–51.

Miller, J., Stagljar, I., 2004. Using the yeast two-hybrid system to identify interacting proteins. Protein-Protein Interact. Methods Appl. 247–262.

Miller, M., 2009. The Importance of Being Flexible: The Case of Basic Region Leucine Zipper Transcriptional Regulators. Curr. Protein Pept. Sci. 10, 244–269.

Miller, M., Shuman, J.D., Sebastian, T., Dauter, Z., Johnson, P.F., 2003. Structural basis for DNA recognition by the basic region leucine zipper transcription factor CCAAT/enhancer-binding protein α. J. Biol. Chem. 278, 15178–15184.

Mina, U., Kumar, P., 2014. Fundamentals and Techniques of Biophysics and Molecular Biology.

Minarovits, J., Gonczol, E., Valyi-Nagy, T., 2007. Latency strategies of herpesviruses. Springer.

Miotto, B., Struhl, K., 2006. Differential Gene Regulation by Selective Association of Transcriptional Coactivators and bZIP DNA-Binding Domains. Mol. Cell. Biol. 26, 5969–5982. https://doi.org/10.1128/MCB.00696-06

Montclare, J.K., Sloan, L.S., Schepartz, A., 2001. Electrostatic control of half-site spacing preferences by the cyclic AMP response element-binding protein CREB. Nucleic Acids Res. 29, 3311–3319.

Morel, Y., Barouki, R., 1999. Repression of gene expression by oxidative stress. Biochem. J. 342, 481–496.

Moriuchi, H., Moriuchi, M., Smith, H.A., Cohen, J.I., 1994. Varicella-zoster virus open reading frame 4 protein is functionally distinct from and does not complement its herpes simplex virus type 1 homolog, ICP27. J. Virol. 68, 1987–1992.

Murakami, T., Saito, A., Hino, S., Kondo, S., Kanemoto, S., Chihara, K., Sekiya, H., Tsumagari, K., Ochiai, K., Yoshinaga, K., 2009. Signalling mediated by the endoplasmic reticulum stress transducer OASIS is involved in bone formation. Nat. Cell Biol. 11, 1205–1211.

Nafissi Varcheh, N., Luginbuehl, V., Aboofazeli, R., Merkle, H.P., 2011. Preparing poly (lactic-co-glycolic acid)(PLGA) microspheres containing lysozyme-zinc precipitate using a modified double emulsion method. Iran. J. Pharm. Res. 203–209.

Nakagawa, Y., Satoh, A., Yabe, S., Furusawa, M., Tokushige, N., Tezuka, H., Mikami, M., Iwata, W., Shingyouchi, A., Matsuzaka, T., 2014. Hepatic CREB3L3 controls whole-body energy homeostasis and improves obesity and diabetes. Endocrinology 155, 4706–4719.

Nazerian, K., Solomon, J., Witter, R., Burmester, B., 1968. Studies on the etiology of Marek's disease. II. Finding of a herpesvirus in cell culture. Exp. Biol. Med. 127, 177–182.

Neerukonda, S., 2015. Transcriptional analysis of the unfolded protein response (UPR) and lymphoma microenvironment during Marek's disease virus (MDV) infection.

O'Neil, K.T., Shuman, J.D., Ampe, C., DeGrado, W.F., 1991. DNA-induced increase in the. alpha.-helical content of C/EBP and GCN4. Biochemistry (Mosc.) 30, 9030–9034.

Oslowski, C.M., Urano, F., 2011. Measuring ER stress and the unfolded protein response using mammalian tissue culture system. Methods Enzymol. 490, 71–92. https://doi.org/10.1016/B978-0-12-385114-7.00004-0

Osterrieder, N., Kamil, J.P., Schumacher, D., Tischer, B.K., Trapp, S., 2006. Marek's disease virus: from miasma to model. Nat Rev Micro 4, 283–294. https://doi.org/10.1038/nrmicro1382

Padilla-Parra, S., Audugé, N., Coppey-Moisan, M., Tramier, M., 2008. Quantitative FRET Analysis by Fast Acquisition Time Domain FLIM at High Spatial Resolution in Living Cells. Biophys. J. 95, 2976–2988. https://doi.org/10.1529/biophysj.108.131276

Parcells, M.S., Arumugaswami, V., Prigge, J.T., Pandya, K., Dienglewicz, R.L., 2003. Marek's disease virus reactivation from latency: changes in gene expression at the origin of replication. Poult. Sci. 82, 893–898.

Pasumarthy, M.K., Murhammer, D.W., 1994. Clonal Variation in the Spodoptera frugiperda IPLB-SF21-AE Insect Cell Population. Biotechnol. Prog. 10, 314–319.

Patel, L., Abate, C., Curran, T., 1990. Altered protein conformation on DNA binding by Fos and Jun. Nature 347, 572–575.

Payne, L., Biggs, P., 1967. Studies on Marek's disease. II. Pathogenesis. J. Natl. Cancer Inst. 39, 281–302.

Payne, L., Gillespie, A., Howes, K., 1991. Induction of myeloid leukosis and other tumours with the HPRS-103 strain of ALV. Vet. Rec. 129, 447–448.

Payne, L., Venugopal, K., 2000. Neoplastic diseases: Marek's disease, avian leukosis and reticuloendotheliosis. Rev. Sci. Tech. Int. Off. Epizoot. 19, 544–564.

Payne, L.N., 2012. Marek's Disease: Scientific Basis and Methods of Control. Springer Science & Business Media.

Petherbridge, L., Brown, A.C., Baigent, S.J., Howes, K., Sacco, M.A., Osterrieder, N., Nair, V.K., 2004. Oncogenicity of Virulent Marek's Disease Virus Cloned as Bacterial Artificial Chromosomes. J. Virol. 78, 13376–13380. https://doi.org/10.1128/JVI.78.23.13376-13380.2004

Petherbridge, L., Howes, K., Baigent, S.J., Sacco, M.A., Evans, S., Osterrieder, N., Nair, V., 2003. Replication-Competent Bacterial Artificial Chromosomes of Marek's Disease Virus: Novel Tools for Generation of Molecularly Defined Herpesvirus Vaccines. J. Virol. 77, 8712–8718. https://doi.org/10.1128/JVI.77.16.8712-8718.2003

Podust, L.M., Krezel, A.M., Kim, Y., 2001. Crystal Structure of the CCAAT Box/Enhancer-binding Protein β Activating Transcription Factor-4 Basic Leucine Zipper Heterodimer in the Absence of DNA. J. Biol. Chem. 276, 505–513. https://doi.org/10.1074/jbc.M005594200

Powell, P., 1986. Marek's Disease–A world poultry problem. Worlds Poult. Sci. J. 42, 205–218.

Pratt, W.D., Cantello, J., Morgan, R.W., Schat, K.A., 1994. Enhanced expression of the Marek's disease virus-specific phosphoproteins after stable transfection of MSB-1 cells with the Marek's disease virus homologue of ICP4. Virology 201, 132–136.

Qi, H., Fillion, C., Labrie, Y., Grenier, J., Fournier, A., Berger, L., El-Alfy, M., Labrie, C., 2002. AIbZIP, a novel bZIP gene located on chromosome 1q21. 3 that is highly expressed in prostate tumors and of which the expression is up-regulated by androgens in LNCaP human prostate cancer cells. Cancer Res. 62, 721–733.

Rahman, L.N., Bamm, V.V., Voyer, J.A., Smith, G.S., Chen, L., Yaish, M.W., Moffatt, B.A., Dutcher, J.R., Harauz, G., 2011. Zinc induces disorder-to-order transitions in free and membrane-associated Thellungiella salsuginea dehydrins TsDHN-1 and TsDHN-2: a solution CD and solid-state ATR-FTIR study. Amino Acids 40, 1485–1502.

Rajoria, S., Zhao, L., Intes, X., Barroso, M., 2014. FLIM-FRET for Cancer Applications. Curr. Mol. Imaging 3, 144–161. https://doi.org/10.2174/22115552036661411172211 11

Read, A.F., Baigent, S.J., Powers, C., Kgosana, L.B., Blackwell, L., Smith, L.P., Kennedy, D.A., Walkden-Brown, S.W., Nair, V.K., 2015. Imperfect Vaccination Can Enhance the Transmission of Highly Virulent Pathogens. PLoS Biol. 13, e1002198. https://doi.org/10.1371/journal.pbio.1002198

Reddy, S.M., Lupiani, B., Gimeno, I.M., Silva, R.F., Lee, L.F., Witter, R.L., 2002. Rescue of a pathogenic Marek's disease virus with overlapping cosmid DNAs: use of a pp38 mutant to validate the technology for the study of gene function. Proc. Natl. Acad. Sci. 99, 7054–7059.

Redfield, C., 1999. Molten globules. Curr. Biol. 9, R313. https://doi.org/10.1016/S0960-9822(99)80199-7

Reiling, J.H., Olive, A.J., Sanyal, S., Carette, J.E., Brummelkamp, T.R., Ploegh, H.L., Starnbach, M.N., Sabatini, D.M., 2013. A CREB3–ARF4 signalling pathway mediates the response to Golgi stress and susceptibility to pathogens. Nat. Cell Biol. 15, 1473–1485.

Reinhard, L., Mayerhofer, H., Geerlof, A., Mueller-Dieckmann, J., Weiss, M.S., 2013. Optimization of protein buffer cocktails using Thermofluor. Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun. 69, 209–214. https://doi.org/10.1107/S1744309112051858

Ren, D., Lee, L.F., Coussens, P.M., 1994. Identification and characterization of Marek's disease virus genes homologous to ICP27 and glycoprotein K of herpes simplex virus-1. Virology 204, 242–250.

Saito, A., 2014. Physiological functions of endoplasmic reticulum stress transducer OASIS in central nervous system. Anat. Sci. Int. 89, 11–20. https://doi.org/10.1007/s12565-013-0214-x

Schermuly, J., Greco, A., Härtle, S., Osterrieder, N., Kaufer, B.B., Kaspers, B., 2015. In vitro model for lytic replication, latency, and transformation of an oncogenic alphaherpesvirus. Proc. Natl. Acad. Sci. 112, 7279–7284.

Schumacher, D., McKinney, C., Kaufer, B.B., Osterrieder, N., 2008. Enzymatically inactive U S 3 protein kinase of Marek's disease virus (MDV) is capable of depolymerizing F-actin but results in accumulation of virions in perinuclear invaginations and reduced virus growth. Virology 375, 37–47.

Schumacher, M.A., Goodman, R.H., Brennan, R.G., 2000. The structure of a CREB bZIP· somatostatin CRE complex reveals the basis for selective dimerization and divalent cation-enhanced DNA binding. J. Biol. Chem. 275, 35242–35247.

Shaner, N.C., Lambert, G.G., Chammas, A., Ni, Y., Cranfill, P.J., Baird, M.A., Sell, B.R., Allen, J.R., Day, R.N., Israelsson, M., Davidson, M.W., Wang, J., 2013. A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum. Nat. Methods 10, 407–409. https://doi.org/10.1038/nmeth.2413

Sherman, J., 1955. The theoretical derivation of fluorescent X-ray intensities from mixtures. Spectrochim. Acta 7, 283–306.

Shuman, J.D., Vinson, C.R., McKnight, S.L., 1990. Evidence of changes in protease sensitivity and subunit exchange rate on DNA binding by C/EBP. Science 249, 771–774.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539.

Smart, O.S., Womack, T.O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C., Bricogne, G., 2012. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. Acta Crystallogr. D Biol. Crystallogr. 68, 368–380. https://doi.org/10.1107/S0907444911056058

Snapp, E., 2005. Design and Use of Fluorescent Fusion Proteins in Cell Biology. Curr. Protoc. Cell Biol. Editor. Board Juan Bonifacino Al CHAPTER, Unit-21.4. https://doi.org/10.1002/0471143030.cb2104s27

Sreerama, N., Woody, R.W., 2000. Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set. Anal. Biochem. 287, 252–260. https://doi.org/10.1006/abio.2000.4880

Strassheim, S., Stik, G., Rasschaert, D., Laurent, S., 2012. mdv1-miR-M7-5p, located in the newly identified first intron of the latency-associated transcript of Marek's disease virus, targets the immediate-early genes ICP4 and ICP27. J. Gen. Virol. 93, 1731–1742.

Sung, H.J., Kim, Y.S., Kang, H., Ko, J., 2008. Human LZIP induces monocyte CC chemokine receptor 2 expression leading to enhancement of monocyte chemoattractant protein 1/CCL2-induced cell migration. Exp. Mol. Med. 40, 332. https://doi.org/10.3858/emm.2008.40.3.332

Szilák, L., Moitra, J., Krylov, D., Vinson, C., 1997a. Phosphorylation destabilizes α-helices. Nat. Struct. Biol. 4, 112–114.

Szilák, L., Moitra, J., Vinson, C., 1997b. Design of a leucine zipper coiled coil stabilized 1.4 kcal mol− 1 by phosphorylation of a serine in the e position. Protein Sci. 6, 1273–1283.

Tahiri-Alaoui, A., Matsuda, D., Xu, H., Panagiotis, P., Burman, L., Lambeth, L.S., Petherbridge, L., James, W., Mauro, V., Nair, V., 2009. The 5′ Leader of the mRNA Encoding the Marek's Disease Virus Serotype 1 pp14 Protein Contains an Intronic Internal Ribosome Entry Site with Allosteric Properties. J. Virol. 83, 12769–12778. https://doi.org/10.1128/JVI.01010-09

Tahiri-Alaoui, A., Smith, L.P., Kgosana, L., Petherbridge, L.J., Nair, V., 2012. Identification of a Neurovirulence Factor from Marek's Disease Virus. Avian Dis. 57, 387–394. https://doi.org/10.1637/10322-080912-Reg.1

Tait, S., Dutta, K., Cowburn, D., Warwicker, J., Doig, A.J., McCarthy, J.E.G., 2010. Local control of a disorder–order transition in 4E-BP1 underpins regulation of translation via eIF4E. Proc. Natl. Acad. Sci. U. S. A. 107, 17627–17632. https://doi.org/10.1073/pnas.1008242107

Thompson, K.S., Vinson, C.R., Freire, E., 1993. Thermodynamic characterization of the structural stability of the coiled-coil region of the bZIP transcription factor GCN4. Biochemistry (Mosc.) 32, 5491–5496.

Tischer, B.K., Kaufer, B.B., 2012. Viral bacterial artificial chromosomes: generation, mutagenesis, and removal of mini-F sequences. BioMed Res. Int. 2012.

Tischer BK, von Einem J, Kaufer B, Osterrieder N, 2006. Two-step red-mediated recombination for versatile high-efficiency markerless DNA manipulation in Escherichia coli. Biotechniques. 40(2):

Trachootham, D., Lu, W., Ogasawara, M.A., Valle, N.R.-D., Huang, P., 2008. Redox Regulation of Cell Survival. Antioxid. Redox Signal. 10, 1343–1374. https://doi.org/10.1089/ars.2007.1957

Tulman, E.R., Afonso, C.L., Lu, Z., Zsak, L., Rock, D.L., Kutish, G.F., 2000. The Genome of a Very Virulent Marek's Disease Virus. J. Virol. 74, 7980–7988. https://doi.org/10.1128/JVI.74.17.7980-7988.2000

Uversky, V.N., 2002. Natively unfolded proteins: A point where biology waits for physics. Protein Sci. Publ. Protein Soc. 11, 739–756.

Van Der Lee, R., Buljan, M., Lang, B., Weatheritt, R.J., Daughdrill, G.W., Dunker, A.K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D.T., 2014. Classification of intrinsically disordered regions and proteins. Chem. Rev. 114, 6589–6631.

Vaughn, J., Goodwin, R., Tompkins, G., McCawley, P., 1977. The establishment of two cell lines from the insect Spodoptera frugiperda (Lepidoptera; Noctuidae). Vitro Cell. Dev. Biol.-Plant 13, 213–217.

Vinson, C., Hai, T., Boyd, S., 1993. Dimerization specificity of the leucine zipper-containing bZIP motif on DNA binding: prediction and rational design. Genes Dev. 7, 1047–1058.

Vinson, C., Myakishev, M., Acharya, A., Mir, A.A., Moll, J.R., Bonovich, M., 2002. Classification of Human B-ZIP Proteins Based on Dimerization Properties. Mol. Cell. Biol. 22, 6321–6335. https://doi.org/10.1128/MCB.22.18.6321-6335.2002

Vogel, S.S., Thaler, C., Koushik, S.V., 2006. Fanciful fret. Sci STKE 2006, re2-re2.

Wallin, S., 2017. Intrinsically disordered proteins: structural and functional dynamics. Res. Rep. Biol. 8, 7–16.

Waterman, D.G., Winter, G., Parkhurst, J.M., Fuentes-Montero, L., Hattne, J., Brewster, A., Sauter, N.K., Evans, G., Rosenstrom, P., 2013. The DIALS framework for integration software. CCP4 Newslett Protein Crystallogr 49, 13–15.

Wexler, P., 2005. Encyclopedia of toxicology. Academic Press.

Wineman-Fisher, V., Miller, Y., 2016. Effect of Zn 2+ ions on the assembly of amylin oligomers: insight into the molecular mechanisms. Phys. Chem. Chem. Phys. 18, 21590–21599.

Winter, G., Lobley, C.M., Prince, S.M., 2013. Decision making in xia2. Acta Crystallogr. D Biol. Crystallogr. 69, 1260–1273.

Witter, R., Schat, K., 2003. DISEASES OF POULTRY, CHAPTER 15: SUBCHAPTER--MAREK'S DISEASE. Dis. Poult. 407–465.

Witter, R.L., Gimeno, I.M., Reed, W.M., Bacon, L.D., 1999. An acute form of transient paralysis induced by highly virulent strains of Marek's disease virus. Avian Dis. 43, 704–720.

Woolley, G.A., Jaikaran, A.S., Berezovski, M., Calarco, J.P., Krylov, S.N., Smart, O.S., Kumita, J.R., 2006. Reversible photocontrol of DNA binding by a designed GCN4-bZIP protein. Biochemistry (Mosc.) 45, 6075–6084.

Xanthoudakis, S., Miao, G., Wang, F., Pan, Y.C., Curran, T., 1992. Redox activation of Fos-Jun DNA binding activity is mediated by a DNA repair enzyme. EMBO J. 11, 3323–3335.

Xu, M., Fitzgerald, S.D., Zhang, H., Karcher, D.M., Heidari, M., 2012. Very virulent plus strains of MDV induce an acute form of transient paralysis in both susceptible and resistant chicken lines. Viral Immunol. 25, 306–323.

Yang, T., Cleland, J.L., Lam, X., Meyer, J.D., Jones, L.S., Randolph, T.W., Manning, M.C., Carpenter, J.F., 2000. Effect of zinc binding and precipitation on structures of recombinant human growth hormone and nerve growth factor. J. Pharm. Sci. 89, 1480–1485. https://doi.org/10.1002/1520-6017(200011)89:11<1480::AID-JPS10>3.0.CO;2-M

Yang, Z.R., Thomson, R., Mcneil, P., Esnouf, R.M., 2005. RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. Bioinformatics 21, 3369–3376.

Yu, D., Ellis, H.M., Lee, E.-C., Jenkins, N.A., Copeland, N.G., 2000. An efficient recombination system for chromosome engineering in Escherichia coli. Proc. Natl. Acad. Sci. 97, 5978–5983.

Zahid, H., Miah, L., Lau, A.M., Brochard, L., Hati, D., Bui, T.T.T., Drake, A.F., Gor, J., Perkins, S.J., McDermott, L.C., 2015. Zinc-induced oligomerization of zinc α2 glycoprotein reveals multiple fatty acid-binding sites. Biochem. J. 473, 43. https://doi.org/10.1042/BJ20150836

Zahran, S., Pan, J.S., Liu, P.B., Hwang, P.M., 2015. Combining a PagP fusion protein system with nickel ion-catalyzed cleavage to produce intrinsically disordered proteins in E. coli. Protein Expr. Purif. 116, 133–138.

Zeng, X., Herndon, A.M., Hu, J.C., 1997. Buried asparagines determine the dimerization specificities of leucine zipper mutants. Proc. Natl. Acad. Sci. 94, 3673–3678.

Zhang, J., Yang, W., Piquemal, J.-P., Ren, P., 2012. Modeling Structural Coordination and Ligand Binding in Zinc Proteins with a Polarizable Potential. J. Chem. Theory Comput. 8, 1314–1324. https://doi.org/10.1021/ct200812y

Zhang, K., Shen, X., Wu, J., Sakaki, K., Saunders, T., Rutkowski, D.T., Back, S.H., Kaufman, R.J., 2006. Endoplasmic reticulum stress activates cleavage of CREBH to induce a systemic inflammatory response. Cell 124, 587–599.

Zhang, Y., Zhou, J., Wang, L., 2013. Mini review roles of the bZIP gene family in rice. Genet. Mol. Res. GMR 13, 3025–3036.

Zhu, G.-S., Iwata, A., Gong, M., Ueda, S., Hirai, K., 1994. Marek's disease virus type 1-specific phosphorylated proteins pp38 and pp24 with common amino acid termini are encoded from the opposite junction regions between the long unique and inverted repeat sequences of viral genome. Virology 200, 816–820.

Appendix A – Vector maps

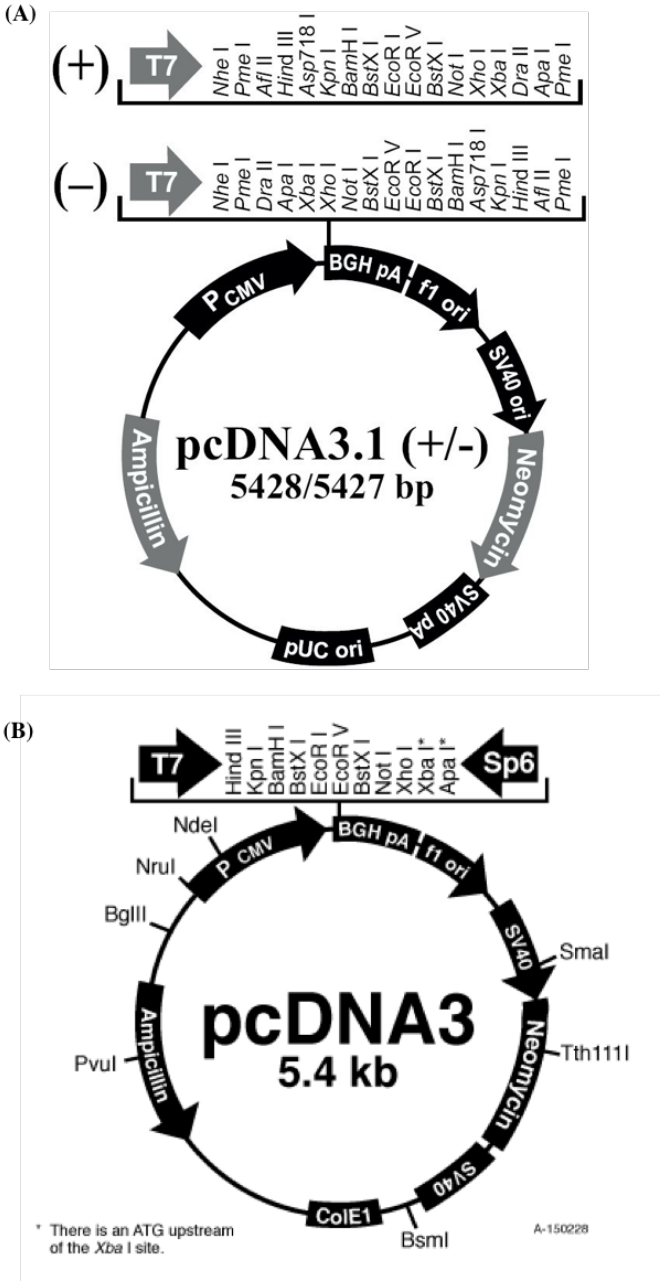
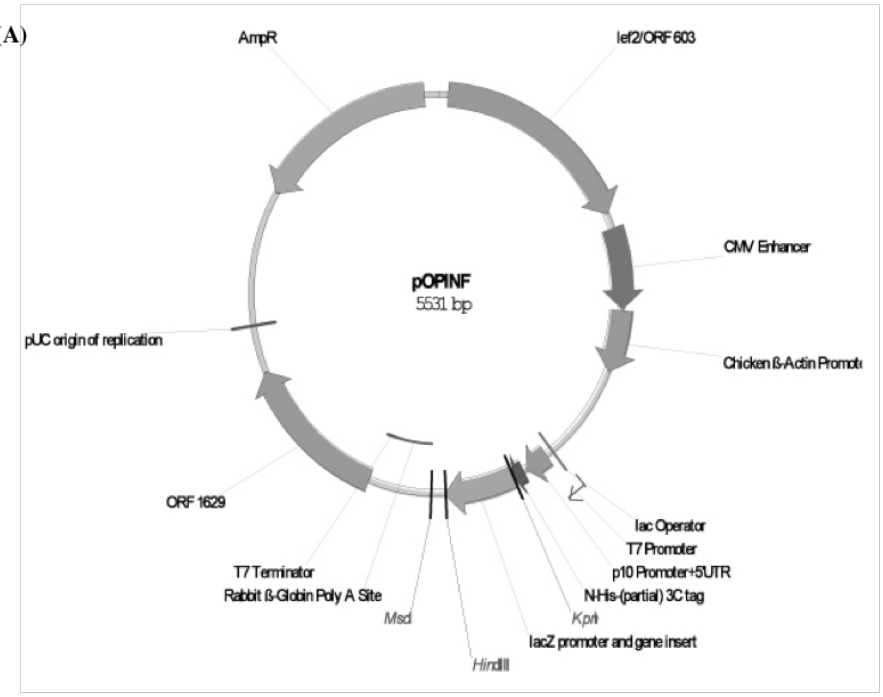
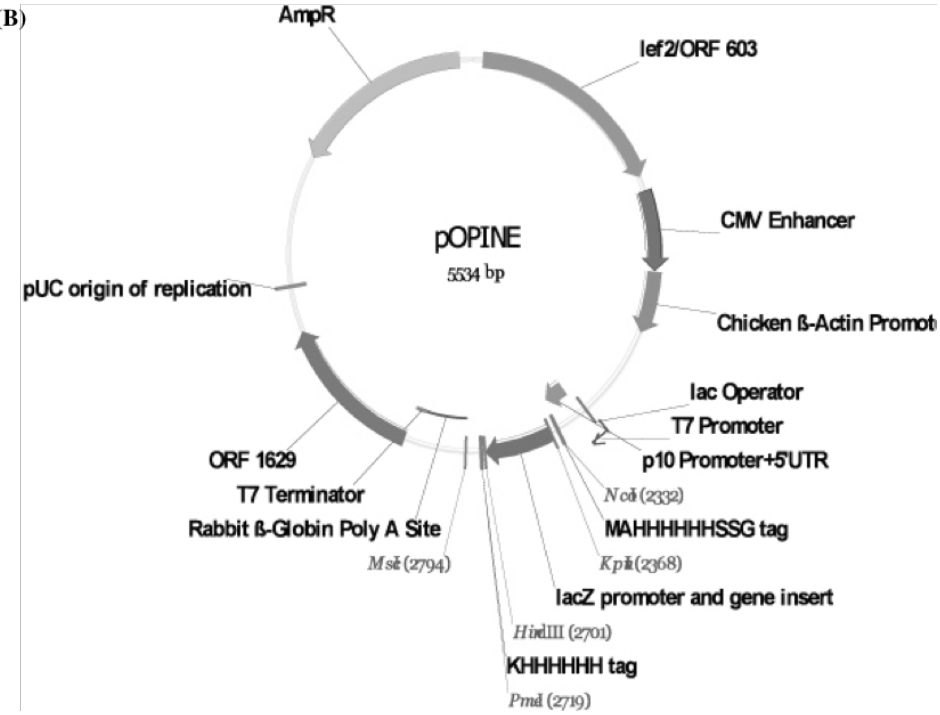


Figure A 1 Vector maps (pcDNA3 vectors).
Schematic representation of the **(A)** pcDNA3.1 and **(B)** pcDNA3 vectors showing key features of the plasmids and details of the multiple cloning site. (Obtained from: https://www.addgene.org).

(A)


AmpR
lef2/ORF 603
CMV Enhancer
pOPINF
5531 bp
pUC origin of replication
Chicken ß-Actin Promot
ORF 1629
lac Operator
T7 Promoter
p10 Promoter+5'UTR
T7 Terminator
Rabbit ß-Globin Poly A Site
N-His-(partial) 3C tag
Msd
Kpn
lacZ promoter and gene insert
HindIII

(B)


AmpR
lef2/ORF 603
CMV Enhancer
pOPINE
5534 bp
pUC origin of replication
Chicken ß-Actin Promot
lac Operator
T7 Promoter
p10 Promoter+5'UTR
ORF 1629
T7 Terminator
Rabbit ß-Globin Poly A Site
NcoI (2332)
MAHHHHHHSSG tag
MseI (2794)
KpnI (2368)
lacZ promoter and gene insert
HindIII (2701)
KHHHHHH tag
PmeI (2719)

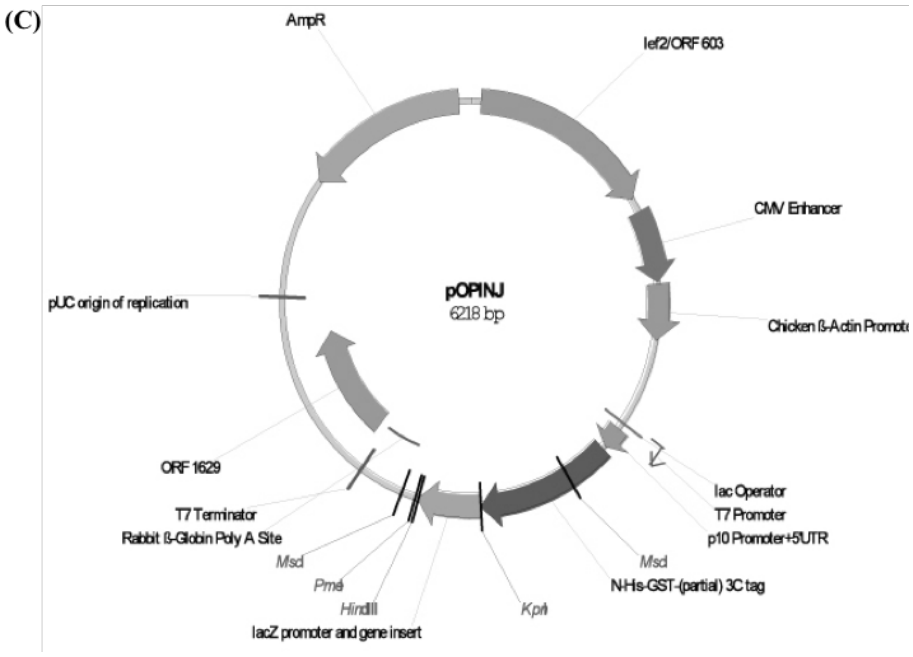


Figure A 2 Vector maps (pOPIN vectors).
Schematic representation of the pOPIN vectors, pOPINF **(A)**, pOPINE **(B)** and pOPINJ **(C)**, showing key features of the plasmids and details of the In-Fusion cloning region. (Obtained from: https://www.addgene.org)