Understanding the Roles of Cellular RNA-Binding Proteins in HIV-1 Infection

Kate Dicker

Hilary 2023

A thesis presented for the degree of Doctor of Philosophy for the Wellcome Trust Infection, Immunology and Translational Medicine Doctoral Training Programme.

wellcome

Acknowledgements

I would like to thank my supervisors Alfredo Castello and Quentin Sattentau for their unending guidance and support. Thank you for welcoming me into your labs, for giving me the opportunity to go to conferences, write papers, and teach students and for all your help and advice in preparing this thesis. I have not had the most straightforward DPhil and I am extremely grateful for your patience and constant optimism during a time that I am sure was just as difficult for you.

I would also like to thank Ilan Davis and Adam Fletcher for their advice and support in many, many hours of lab meetings. Your expertise and technical advice has been greatly appreciated. Thank you to all the members of the Castello lab for giving me such an enjoyable start to my DPhil and for always helping with my never-ending questions. Special thanks to Manuel Garcia-Moreno for his mentorship at the start of my project, to Marko Noerenberg for always telling me that everything will be okay and to Vincenzo Ruscica for keeping me sane in the lab during lockdowns! Also to Louisa Iselin for her patient explanations of proteomics analysis and Jeff Lee for his expertise in smFISH and microscopy. Thank you also to all the members of the Sattentau lab for welcoming me into the lab and making a very difficult time in my DPhil as easy as possible. Lachlan and Linnea, you have made being in the lab so much fun and I will miss our coffee chats.

Christy, words cannot describe how immensely grateful I am for you. You are my best friend and safe haven from DPhil stress and failed experiments. Thank you for putting up with the long hours and rants about tissue culture, and thank you for always putting a smile back on my face. You have taught me to stand up for myself and always encourage me to be the best that I can be. I honestly do not know how I would have done this without you, so thank you.

Owen, Laurence, Izzy, Annie and Thomas, I cannot thank you enough for how fun you have made my time in Oxford. The past four years have been filled with laughter, pub trips and chaos and I could not have asked for better friends.

Thank you also to my mum and dad, Lorraine and David, for making me who I am today. You have always encouraged me to do whatever I dream to do and taught me

that I can do absolutely anything if I put my mind to it. You have indulged every single one of my curiosities and put up with endless rambling about science. Not everyone gets parents like you and I am immensely grateful. And Mum, I do think you deserve an honorary degree for all the proofreading! Thank you as well to my sister, Anna, for always being there for me, even when I annoy the hell out of you.

Last but by no means least, thank you to my grandad, David, the original Dr Dicker. You have always inspired me to be a scientist and I wanted to come to Oxford because of you. I'll be putting my thesis on the shelf next to yours.

Abstract

The Human Immunodeficiency virus type 1 (HIV-1) is the causative agent of Acquired Immune Deficiency Syndrome (AIDS) which is currently estimated to affect 37.9 million people globally, representing a serious threat to global health. Developing a fundamental understanding of the HIV-1 life cycle is paramount in the search for a cure. RNA is a central molecule in HIV-1 biology due to its dual function as messenger and genome. However, the small number of proteins encoded by HIV-1 is insufficient to enable virus infection. Hence, HIV-1 hijacks cellular RNA-binding proteins (RBPs) to fulfil its life cycle. Using in virion RNA-interactome capture we identified 104 cellular RBPs interacting with HIV-1 RNA within the viral capsid. My DPhil aimed to identify which of these RBPs regulate HIV-1 infection. By screening RBP-knockout cell lines I uncovered the effects of MKRN1, FAM120A, LSM14A, MSN, LCK, CIRBP and YBX3 on HIV-1 infection. Further analysis into the role of MKRN1 revealed a novel function for MKRN1 in the positive regulation of HIV-1 viral particle release from the plasma membrane. Additionally, I have developed a tool, HIV-$1_{INscarlet}$, to enable the visualisation of the early stages of HIV-1 infection from viral particle entry to integration using super-resolution microscopy. This tool will allow for future experiments into the role of MKRN1 inside the HIV-1 capsid during the early stages of infection. Together, this work highlights novel cellular regulators of HIV-1 infection and further validates the crucial role conducted by cellular RBPs in HIV-1 infection.

List of Tables

2 Cellular RBPs acting within different stages of HIV-1 infection. 33
2 Cellular RBPs acting within different stages of HIV-1 infection. 34
2 Cellular RBPs acting within different stages of HIV-1 infection. 35
4 Machines used. 41
5 Commercial buffers. 41
6 In-house buffers 42
7 Commercial kits 43
8 Antibodies and dyes 44
9 Commercial and lab-made plasmids 45
10 Oligonucleotides 46
11 Parental cell lines 47
12 Genetically modified cell lines 48
13 Software used within this thesis 49
14 Conditions of SupT1 infection in different experiments 53
15 Components of site-directed mutagenesis 55
16 Site-directed mutagenesis thermocycling conditions 55
17 Kinase, ligase and DpnI treatment conditions 56
18 RTqPCR Components 59
19 RTqPCR Cycle Conditions 59
20 Excitation and detection wavelengths of fluorophores used in flow cytometry 60
21 Summary of effect of ivRBP-KO on HIV-1 infection 136
22 GO terms of proteins in MKRN1 interactome in HIV-1 infection. 186
23 Proteins identified in MKRN1 interactome during HIV-1 infection compared to our ivRBPome and the affinity purification of MKRN1 conducted by Hildebrandt et al, 2019. 189
23 Proteins identified in MKRN1 interactome during HIV-1 infection compared to our ivRBPome and the affinity purification of MKRN1 conducted by Hildebrandt et al, 2019. 190

23 Proteins identified in MKRN1 interactome during HIV-1 infection compared to our ivRBPome and the affinity purification of MKRN1 conducted by Hildebrandt et al, 2019. 191

List of Figures

1.1 Structure of the HIV-1 genome and mature viral particle 26
1.2 RBPs that regulate HIV-1 infection are packaged into assembling virions 37
3.1 As a, b diverge more, the RMSE increases. 57
3.2 RMSE can be dominated by larger values, whereas the normalisation in RMSRE removes this bias. 58
4.1 Schematic of the ivRIC workflow 69
4.2 ivRBPs identified by ivRIC 72
4.3 MKRN1 predicted structure and RBDs. 74
4.4 FAM120A predicted structure and RBDs. 78
4.5 LSM14A predicted structure and RBDs. 80
4.6 MSN structure and RBDs 82
4.7 LCK structure and RBDs 85
4.8 CIRBP structure and RBDs 88
4.9 YBX3 predicted structure and RBDs 90
4.10 CRISPR/Cas9 Nanoblades targeting MKRN1 93
4.11 CRISPR/Cas9 knockout of MKRN1 in SupT1 cells 94
4.12 Growth rates of SupT1-KO cells differ slightly 94
4.13 HIV-1 Nef and Gag are not detected in uninfected cells 97
4.14 MKRN1-KO increases expression of HIV-1 proteins 98
4.15 No HIV-1 RNA is detected in uninfected cells 100
4.16 MKRN1-KO increases expression of HIV-1 RNA splice variants 101
4.17 FAM120A-KO decreases expression of HIV-1 proteins 103
4.18 FAM120A-KO decreases expression of HIV-1 completely spliced RNA . 104
4.19 LSM14A-KO does not affect expression of HIV-1 proteins 106
4.20 LSM14A-KO delays expression of completely spliced vRNA 107
4.21 MSN-KO increases expression of HIV-1 Gag but not Nef 109
4.22 MSN-KO initially decreases but then increases expression of HIV-1 RNA splice variants 110
4.23 LCK-KO increases expression of HIV-1 Nef 112

4.24 LCK-KO initially upregulates vRNA expression but upregulation was not sustained . . . 113
4.25 CIRBP-KO increases expression of HIV-1 Nef but decreases expression of Gag . . . 115
4.26 CIRBP-KO downregulates completely spliced and single-spliced vRNA expression at 24 hpi and upregulates expression at 48 hpi . . . 116
4.27 YBX3-KO increases expression of HIV-1 Nef and Gag . . . 118
4.28 YBX3-KO delays expression of vRNAs . . . 119
4.29 An assay designed to determine the effect of ivRBP-KO on viral particle production . . . 120
4.30 No p24 can be detected in the supernatant of uninfected cells . . . 122
4.31 MKRN1-KO cells release fewer HIV-1 viral particles . . . 124
4.32 FAM120A-KO does not affect the number of viral particles released from HIV-$1_{mCherry\text{-}Nef}$ infected cells . . . 125
4.33 LSM14A-KO decreases the release of viral particles from infected cells 127
4.34 MSN-KO decreases the release of viral particles from infected cells . . . 128
4.35 LCK-KO decreases the release of viral particles from infected cells . . . 130
4.36 CIRBP-KO does not effect the release of viral particles from infected cells but does increase percentage of infected cells . . . 131
4.37 YBX3-KO significantly decreases the release of viral particles from infected cells and rate of infection . . . 133
5.1 Mechanism of HIV-1 budding . . . 150
5.2 Interactors and targets of MKRN1 . . . 155
5.3 CRISPR/Cas9 KO of MKRN1 in HEK293 Flp/In cells . . . 157
5.4 Viral plasmid transfection is equally efficient in WT and MKRN1-KO HEK293 cell lines . . . 158
5.5 P24 was reduced in supernatant from HEK293 Flp/In and HEK293T MKRN1-KO cells . . . 160
5.6 Fewer viral particles were released from MKRN1-KO HEK293 Flp/In and HEK293T clones . . . 162

5.7 MKRN1-KO viral particles have the same infectivity as WT viral particles 164
5.8 Immunofluorescence and smFISH of HIV-1 RNA and Gag in HIV-1 infected MKRN1-KO cells 166
5.9 Gag is enriched at the plasma membrane in MKRN1-KOs 170
5.10 MKRN1-KO cells have increased localisation of HIV-1 Gag to the plasma membrane 171
5.11 No difference in vRNA localisation in MKRN1-KO cells 172
5.12 Generation of HEK293 Flp/In GFP-MKRN1 cell line 174
5.13 Infection of HEK293 Flp/In GFP-MKRN1 for immunoprecipitation 176
5.14 GFP-MKRN1 is poorly isolated using GFP co-immunoprecipitation . . . 176
5.15 More efficient immunoprecipitation of GFP-MKRN1 is achieved when cytoplasmic fraction is isolated 177
5.16 GFP-MKRN1 is efficiently captured using an optimised GFP co-immunoprecipitation 179
5.17 Infection of HEK293 Flp/In GFP-MKRN1 for immunoprecipitation and proteomic analysis 180
5.18 177 proteins are identified in the MKRN1 interactome, 87 of which are more abundant in HIV-1 infection. 183
5.19 GFP-MKRN1 interactome overlaps with previous MKRN1 interactomes and ivRBPome datasets 185
5.20 Generation of HEK293 Flp/In GFP-MKRN1 catalytic mutant cell line . . 188
6.1 Novel pNL4-3 HIV-1 and single round viruses were engineered to express mScarlet-tagged IN 205
6.2 INscarlet is expressed in HIV-1-INscarlet viral particles and single-round viruses. 208
6.3 Individual HIV-$1_{INscarlet}$ viral particles are fluorescent 209
6.4 Addition of mScarlet does not affect the function of IN. 212
6.5 INscarlet labelled capsids can be seen in the cytoplasm of infected cells 214
6.6 INscarlet labelled capsids can be seen in the cytoplasm of infected cells at 6 hpi 215

7.1 Schematic of GFP-MKRN1 HIV-$1_{INscarlet}$ dual-fluorescent viral particles . 225
8.1 ivRBPs identified by ivRIC enriched in crosslinked and HIV-1 infected eluates 227
8.2 ivRBPs identified by ivRIC enriched in ivRIC eluates over inputs and the whole cell lysate 228
10.1 Flow cytometry raw data for Figure 6.4B and C INscarlet P24 and mScarlet expression 235
10.2 Flow cytometry raw data for Figure 6.4D, second round of infection with HIV-$1_{INscarlet}$ 236

Abbreviations

Å	Angstrom
AEBSF	4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride
AIDS	Acquired immunodeficiency syndrome
ALIX	ALG2-interacting protein X
ART	Anti-retroviral therapy
bp	base pair
BSL	Bio-safety level
CA	HIV-1 capsid
CCR5	C-C chemokine receptor type 5
cDNA	copy/complementary DNA
CFAR	Centre for AIDS reagents
CLIP	Crosslinking and immunoprecipitation
cm	centimetres
Co-IP	Co-immunoprecipitation
CXCR4	C-X-C chemokine receptor type 4
Da	Dalton
DAPI	4’,6-Diamidino-2-Phenylindole, Dihydrochloride
DENV	Dengue virus
DMEM	Dulbecco’s modified eagle medium
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleoside triphosphate
dsRNA	Double-stranded RNA
DTT	Dithiothreitol
ESCRT	endosomal sorting complex required for transport
FBS	Foetal bovine serum
GFP	Green fluorescent protein
GO	Gene ontology
gRNA	Genomic RNA

HCV	Hepatitis C Virus
HDTs	Host-directed therapies
HIV-1	Human immunodeficiency virus type 1
hnRNP	heterogeneous nuclear ribonucleoprotein
IN	HIV-1 integrase
IN	HIV-1 integrase
IP	Immunoprecipitation
ivRBPome	In virion RNA-binding proteome
ivRIC	In-virion RNA interactome capture
kbp	kilo base pairs
KD	Knockdown
kDa	kilo Dalton
KO	Knockout
lncRNA	Long non-coding RNA
LTR	Long-terminal repeat
M	Molar
MA	HIV-1 Matrix
min	Minutes
μ **m**	micrometres
ml	Millilitre
mM	Millimolar
mRNA	Messenger RNA
NC	HIV-1 nucleocapsid
NEB	New England Biolabs
nm	nanometres
NPC	Nuclear pore complex
ORF	Open reading frame
p6	HIV-1 protein 6
PBS	primer binding site
PCR	Polymerase chain reaction

PIC	Pre-integration complex
PR	HIV-1 protease
PTEFb	positive transcription elongation factor b
RBD	RNA-binding domain
RBP	RNA-binding protein
RFP	Red fluorescent protein
RIC	RNA interactome capture
RIG-I	Retinoic acid-inducible gene I
RMSE	Root-mean-square error
RMSRE	Root-mean-square relative error
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
RPMI	Roswell Park memorial institute
RRM	RNA recognition motif
RT	HIV-1 reverse transcriptase
RTC	Reverse transcription complex
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
smFISH	single molecule fluorescence in situ hybridisation
ssRNA	Single-stranded RNA
TAR	transactivation response element
TM	Transmembrane
TRIM25	Tripartite motif containing protein
tRNA	transfer RNA
TSG101	tumour susceptibility gene 101
TUBEs	Tandem ubiquitin binding entities
UTR	untranslated region
UV	Ultraviolet
vRNA	Viral RNA
VSVG	Vesicular stomatitis virus glycoprotein

WCL	Whole cell lysate
WNV	West Nile virus
WT	Wild-type
ZIKV	Zika virus
ZnF	Zinc finger

Contents

List of Tables **5**

List of Figures **7**

1 Introduction **20**

1.1 RNA-binding proteins 22

1.2 The HIV-1 life cycle 24

1.2.1 HIV-1 has an RNA genome 24

1.2.2 Virion structure of HIV-1 26

1.2.3 HIV-1 life cycle 27

1.3 RBPs within the HIV-1 life cycle 31

1.4 How do RBPs necessary for HIV-1 infection access the viral RNA? . . . 36

1.5 Aims 38

2 Materials **39**

2.1 Chemicals and Reagents 39

2.2 Machines 41

2.3 Buffers and solutions 41

2.4 Commercial kits 43

2.5 Antibodies and dyes 44

2.6 Plasmids 45

2.7 Oligonucleotides 46

2.8 Cell lines 47

2.9 Software 49

3 Methods **50**

3.1 Cell Biology 50

3.1.1 Maintenance of Cells 50

3.1.2 Cloning of MKRN1-KO SupT1 and HEK293 Flp/In cell lines . . . 50

3.1.3 Stable Transfection of HEK293 Flp/In T-REx 51

3.1.4 Generation of pseudotyped lentiviruses 51
3.1.5 Generation of HIV-1 and fluorescently tagged HIV-1 viral particles 51
3.1.6 Titration of viruses 52
3.1.7 Infection of SupT1 cells 52
3.1.8 Infection of HeLa-CD4-GFP for Immunofluorescence 53
3.1.9 Infection of HEK293 Flp/In with KO viral particles 53
3.2 Molecular Biology 54
3.2.1 Cloning GFP-MKRN1 54
3.2.2 Generation of MKRN1 mutants through site-directed mutagenesis 54
3.2.3 Cloning INscarlet tagged HIV-1 and HIV-1 single round viruses . 55
3.2.4 Plasmid preparation 56
3.2.5 Fluorescence measurement on Plate Reader 57
3.2.6 Western Blotting 58
3.2.7 RT-qPCR 59
3.2.8 Flow Cytometry 60
3.2.9 P24 ELISA 60
3.2.10 MKRN1-KO virus release assay 61
3.2.11 Tzm-bl Luciferase Assay 62
3.2.12 Synthesis of fluorescently labelled probes for smFISH 62
3.2.13 smFISH 63
3.2.14 smFISH analysis 63
3.2.15 Imaging of cell-free viral particles 64
3.2.16 Immunofluorescence 64
3.3 Proteomics 65
3.3.1 Co-immunoprecipitation 65
3.3.2 Silver Staining 66
3.3.3 Sample Preparation for Mass Spectrometry 66
3.3.4 Mass spectrometry 67
3.3.5 GO terms/STRING analysis 67
3.3.6 RBD analysis 67

4 Identifying host RBPs that play key roles in the early stages of HIV-1 infection 68

4.1 Introduction . . . 68

4.1.1 The HIV-1 in-virion RNA interactome . . . 68

4.1.2 MKRN1 . . . 73

4.1.3 FAM120A . . . 77

4.1.4 LSM14A . . . 79

4.1.5 MSN . . . 83

4.1.6 LCK . . . 84

4.1.7 CIRBP . . . 87

4.1.8 YBX3 . . . 91

4.2 Results . . . 92

4.2.1 Generation of MKRN1-KO SupT1 cell line . . . 92

4.2.2 MKRN1-KO cells have increased expression of viral proteins and RNA during HIV-$1_{\text{-R-E-mCherry-Nef}}$ infection . . . 95

4.2.3 FAM120A-KO cells have decreased expression of viral proteins and completely spliced RNA during HIV-$1_{\text{-R-E-mCherry-Nef}}$ infection . 102

4.2.4 LSM14A-KO has no overall effect on HIV-$1_{\text{-R-E-mCherry-Nef}}$ infection but may delay initial gene expression . . . 105

4.2.5 MSN-KO increases the expression of Gag and unspliced vRNA . 108

4.2.6 LCK-KO increases the initial expression of completely spliced vRNA and increases Nef protein expression . . . 111

4.2.7 CIRBP-KO increases expression of Nef and vRNAs at 48 hpi but has no effect on Gag expression . . . 114

4.2.8 YBX3-KO slightly increases Nef and Gag protein expression but delays vRNA expression . . . 117

4.2.9 MKRN1-KO cells are less permissive to infection and produce fewer viral particles . . . 119

4.2.10 FAM120A-KO does not effect the number of viral particles released from HIV-$1_{\text{mCherry-Nef}}$ infected cells . . . 123

4.2.11 LSM14A-KO reduces the number of viral particles released from HIV-$1_{mCherry\text{-}Nef}$ infected cells . . . 126
4.2.12 MSN-KO reduces the release of viral particles from infected cells 126
4.2.13 LCK-KO decreases the release of viral particles from infected cells 129
4.2.14 CIRBP-KO increases spread of HIv-1 infection . . . 129
4.2.15 YBX3-KO . . . 132
4.3 Discussion . . . 134
4.3.1 MKRN1 may be required in viral particle budding and release . . 137
4.3.2 FAM120A . . . 138
4.3.3 LSM14A . . . 139
4.3.4 MSN . . . 140
4.3.5 LCK . . . 141
4.3.6 CIRBP . . . 142
4.3.7 YBX3 . . . 143
4.3.8 Conclusions . . . 144
4.4 Future work . . . 145

5 The one RING ligase to rule them all: characterisation of the role of MKRN1 in HIV-1 infection 148

5.1 Introduction . . . 148
5.1.1 HIV-1 budding and the ESCRT pathway . . . 148
5.1.2 The role of ubiquitin and ubiquitin ligases in HIV-1 budding . . . 151
5.1.3 MKRN1 has no known connection to the ESCRT pathway . . . 152
5.2 Results . . . 156
5.2.1 Generation of MKRN1-KO HEK293 Flp/In T-rex cells . . . 156
5.2.2 MKRN1-KO in HEK293 cells inhibits HIV-1 particle release . . . 157
5.2.3 MKRN1-KO viral particles have the same infectivity as WT particles. . . . 162
5.2.4 Quantification of HIV-1 RNAs in MKRN1-KO cells . . . 165
5.2.5 Generation of HEK293 Flp/In GFP-MKRN1 cell line . . . 172
5.2.6 Optimisation of GFP-MKRN1 immunoprecipitation . . . 174

5.2.7 GFP-MKRN1 interactome in HIV-1 infection comprises 87 host cell proteins . . . 179
5.2.8 Cloning of MKRN1 catalytic mutants for further study . . . 187
5.3 Discussion . . . 192

6 Establishing a system to visualise HIV-1 infection from viral particle entry to integration 200
6.1 Introduction . . . 200
6.2 Results . . . 204
6.3 Discussion . . . 215

7 General discussion and future directions 219

8 Appendix 1 226

9 Appendix 2 229

10 Appendix 3 235

11 References 237

1 Introduction

The human immunodeficiency virus (HIV) is a retrovirus with a genome consisting of two positive-sense single stranded RNAs. Since the 1980s HIV has infected 79.3 million people globally, 37.7 million of whom are still living with HIV today [1]. The vast majority of cases are caused specifically by HIV type 1 (HIV-1). HIV-1 infects CD4+ lymphocytes causing a depletion of these cells and a weakened immune response, termed acquired immunodeficiency syndrome (AIDS). Whilst the development of anti-retroviral therapies has increased the life expectancy of AIDS sufferers, 670,000 people still died from HIV/AIDS in 2020 [1]. These numbers highlight the crucial need for a cure for HIV-1 infection and therefore developing a fundamental understanding of HIV-1 biology is critical.

Anti-retroviral therapy (ART) consists of a cocktail of compounds that target different stages of the viral life cycle. The different drugs can be classed broadly into six target areas: 1) nucleoside-analog reverse transcriptase inhibitors; 2) non-nucleoside reverse transcriptase inhibitors; 3) integrase inhibitors; 4) protease inhibitors; 5) fusion inhibitors and 6) co-receptor antagonists [2]. With the exception of co-receptor antagonists, these drugs interact with and inhibit viral proteins. In 2019, the WHO reported the prevalence of HIV-1 resistance to first-line ART is rapidly increasing in low and middle-income countries and poses a threat to global health [3]. The limited availability of ART in these countries combined with reduced ability to monitor viral load means resistant infections are often not detected and this has led to the emergence of ART-resistant viral serotypes [4].

One of the ways we can reduce resistance to ART is with host-directed therapies (HDTs). HDTs are a novel approach in the field of antivirals that interfere with host cell factors required by pathogens for infection or to enhance a protective immune response [5]. Many viruses, including HIV, utilise host factors to facilitate their life cycles and the correct function of these proteins is often essential for infection (reviewed extensively here [6–9]). By targeting these host factors, viral replication can be inhibited. Because HDT does not target viral proteins, it presents a higher genetic barrier to therapeutic resistance [5]. This is because of the slower mutation

rate of cells and the difficulty for a virus of replacing a functional host-virus which may involve multiple mutations at interaction interfaces. In addition to evading resistance, HDT also has the benefit of being applicable to a large range of viruses due to many viruses hijacking the same host proteins [9–11].

HDT does come with its own challenges, however. Whilst HDTs are less prone to resistance, it is still possible for viruses to escape the treatment by evolving to use a different cellular cofactor with a similar function, to differently interact with the host factor or to out-compete the drug in target binding [12]. For example, the HIV-1 HDT maraviroc blocks HIV-1 entry by interacting with the cellular co-receptor CCR5 and preventing its interaction with the viral envelope protein gp120. One concern was that the treatment with maraviroc could promote de novo switching from using CCR5 to CXCR4. It was shown through serial passages of HIV-1 isolates that co-receptor switching did not occur. However, some isolates acquired resistance through gp120 mutation that enabled interaction with maraviroc-bound CCR5 [13]. In addition to drug-resistance, HDT can cause serious clinical side effects when cellular factors play roles that are crucial for life processes. Although, most drugs for the treatment of non-infectious diseases target cellular proteins and are used regularly. Despite these challenges, several different HDT approaches including recombinant cytokines and cytokine antagonists have been used successfully in the treatment of COVID-19 [14], demonstrating that with the right target, HDTs are a promising anti-viral therapeutic strategy.

Perhaps the most evolutionarily conserved aspect of cellular physiology is RNA regulation. RNA metabolism is critically controlled by hundreds of RNA-binding proteins (RBPs) that interact with RNA and control its fate. RNA is also a central molecule in virus biology. In RNA viruses and retroviruses it functions as both messenger and genome and in DNA viruses it is essential for protein expression. Despite this, viruses only encode a handful of proteins able to interact with RNA and therefore require many cellular RBPs to facilitate infection [9, 15, 16]. Thus RBPs represent a valuable target area for the development of new HDTs.

1.1 RNA-binding proteins

RBPs interact dynamically with RNA forming ribonucleoprotein (RNP) complexes that control the RNA from synthesis to decay [17].They play a central role in cell biology and dysfunction in RBPs underlie a number of genetic diseases [18], cancers [19], cardiovascular disease [20], autoimmune conditions [21] and immune deficiencies [22]. Of both coding and non-coding RNA transcripts, RBPs are know as the major regulators [23] and they can be classified as either single-stranded RNA (ssRNA) binding or double-stranded (dsRNA) binding. Some RBPs involved in processes such as transcription and translation are able to bind both DNA and RNA (DRBPs) [24].

Most RBPs are modular and built from multiple repeats of canonical RNA-binding domains (RBDs) with linker regions in between. Multiple repeats of RBDs allow an RBP to have higher specificity and affinity than an individual domain. Multiple domains also endow the ability to recognise a much longer stretch of RNA than possible with a single domain, or even sequences separated by a stretch of nucleotides, or that belong to different RNAs. The linker regions can also determine the specificity of the RBDs. Long linker regions are usually disordered and allow for the recognition of a diverse set of targets whereas short linkers tend to be more ordered, often forming a short α-helix [23]. Their modular structure also enables the RBP to topologically arrange a flexible RNA for a particular function.

The conventional view of RBPs is that they act to determine the fate and function of their target RNA. By covering and uncovering the RNA at different stages, RBPs enable RNA to progress through its lifecycle and have been described as ‘mRNA’s clothes’ [25]. The concept that RBPs regulate RNA function and stability is widely applicable. However, emerging evidence is indicating that RNAs may also play diverse roles in regulating RBPs. For example: the discovery of long non-coding (lnc)RNAs revealed a number of protein-RNA interactions that break with convention. lncRNAs have been described to recruit transcription factors and chromatin-modifying complexes to chromatin to enable scaffolding or inhibition of protein complexes, suggesting that lncRNA is required for the organisation of these complexes [24, 26].

Additionally, a type of small non-coding RNA called vault (vt)RNA has been shown to mechanistically regulate selective autophagy in human and murine cells by binding to the autophagy receptor p62 and interfering with oligomerisation, a critical step for p62 function [27]. Finally, RNAs have been shown to enhance enzymatic activity of certain proteins. The E3 ubiquitin ligase tripartite motif containing protein 25 (TRIM25) functions as an effector of retinoic acid-inducible gene I (RIG-I), an innate immune sensor that recognises viral RNA. TRIM25 itself recognises RNA through a SPRY domain and it was shown that binding RNA enhances its E3 ligase activity and is a requirement for potentiation of RIG-I antiviral signalling by TRIM25 [28]. These examples are by no means exhaustive, but provide an insight into the complexity of RBPs and the crosstalk between proteins and RNA.

Another advancement in our understanding of RBPs is our knowledge of different RBDs. Historically the identification of RBPs was dependent on the presence of these signatures, the most common of which include RNA-recognition motif (RRM), K-homology domain, double-stranded (ds)RBD and RNA-binding zinc-finger (ZnF) domains [23]. However, advances in our understanding of large RNP structures such as the ribosome and spliceosome revealed the existence of complex protein-RNA interactions lacking canonical RBDs [29, 30]. More recently, methods based on organic-aqueous phase separation such as protein-crosslinked RNA extraction (XRNAX) [31], orthoganol organic phase separation (OOPS) [32] and phenol toluol extraction (PTex) [33] have enabled the detection of novel RBPs bound to non-poly(A) RNAs. Advances in proteomic techniques, like RNA interactome capture (RIC), have dramatically increased the repertoire of known human cellular RBPs to over 1500 proteins, leading to the discovery of many unconventional or non-canonical RBDs [34–36]. Additionally, RBD identification approaches like proteomic identification of RNA-binding regions (RBR-ID) have revealed dozens of novel RBDs including enzymatic cores, protein-protein interactions surfaces and DNA-binding domains [37,38]. Intrinsically disordered regions have also been shown as a prominent mode of RNA binding [39]. Disordered arginine and serine repeats, arginine and glycine repeats (called RGG-boxes or GAR repeats), and clusters of basic residues

forming positively charged surfaces are commonly found within disordered regions in RBPs and facilitate RNA-binding [35,39]. The improved understanding of RNA-protein interaction has enabled the development of several RBD prediction tools including RBDetect, which uses amino acid sequence, and BindUP, which looks at electrostatic patches on the protein surface [40].

It was also conventionally thought that RBPs with high affinity and/or specificity for their target RNAs were more likely to have specific biological functions [41]. However, this too is now being shown to be a more complex relationship due to RNA-footprinting techniques and the plasticity of RNA-binding proteomes. Methods based on high-throughput sequencing and crosslinking and immunoprecipitation assays (CLIP) like HITS-CLIP [42], iCLIP [43] and eCLIP [44] can identify, with single-nucleotide resolution, where on an RNA molecule a particular RBP binds. Whilst some RBPs bind at very specific sequences on an RNA, others appear to bind at multiple sites and to multiple RNAs suggesting the interaction is less specific.

Whilst a subset of 'housekeeping' RBPs may be constitutively and ubiquitously active, many RBPs have highly dynamic expression patterns. Moreover, RNA-binding ability can be regulated by post-translational modification, cofactor binding and protein-protein interaction. Some RBPs can remain inactive for long periods, only being activated upon binding of their target RNA [36]. For example, RIG-I and other innate immune sensors of viral (v)RNA are only activated by the presence of unusual RNA products of viral replication [45].

1.2 The HIV-1 life cycle

1.2.1 HIV-1 has an RNA genome

The genome of HIV-1 consists of two identical positive-sense ssRNA molecules, carried inside the virion as a non-covalent dimer. Each RNA molecule is approximately 9.2kb long with a 5′ cap and 3′ poly(A) tail [46]. Upon entry into a host cell, the genomic RNA is reverse transcribed into linear double-stranded DNA and irreversibly integrated into the host cell genome. Once integrated, the DNA is transcribed by

host RNA polymerase II via a promoter located in the long terminal repeat (LTR) generating a polycistronic pre-mRNA. The full length genomic (g)RNA encodes 9 open reading frames (ORFs) that generate at least 15 mature proteins through polyprotein processing and alternative splicing (Figure 1.1A).

Alternative splicing of the HIV-1 pre-mRNA produces over 40 different mRNAs which can be grouped into unspliced ($\sim$9kb), single-spliced ($\sim$4kb) and multiple-spliced ($\sim$2kb) (Figure 1.1B). Splicing of the vRNA is a highly orchestrated procedure and requires recognition of a 5′ splice site (5′ss) and a 3′ splice site (3′ss) by the cell spliceosome. HIV-1 uses different combinations of the 5′ and 3′ss to generate the alternatively-spliced species [47].

In the early stages of HIV-1 gene expression, the multiple-spliced mRNAs are produced by the removal of both 5′ and 3′ introns and translated in the cytoplasm to Tat, Rev and Nef proteins. As Rev accumulates, it facilitates nuclear export of single-spliced and unspliced mRNAs. The single-spliced mRNAs have the major 5′ intron removed and encodes the viral transmembrane envelope protein (TM) and surface proteins (SU), as well as accessory proteins Vif, Vpr and Vpu. The unspliced mRNA encodes the structural and enzymatic protein precursors Gag and Gag-Pol and serves as genomic RNA. The Gag and Gag-Pol polyproteins are proteolytically processed to produce the structural proteins Matrix (MA), capsid (CA), nucleocapsid (NC) and adaptor protein p6, and the viral enzymes protease (PR), reverse transcriptase (RT) and integrase (IN), respectively.

HIV-1 alternative splicing is a heavily regulated process. The presence of sub-optimal 5′ss, polypyrimidine tracts at 3′ss, and branch sequences next to the sites all decrease recognition by spliceosome machinery [47]. More frequently though, recognised splice sites outcompete suboptimal ones and generate a specific stoichiometry of viral RNAs [48, 49]. Additionally, cis-acting exonic splicing enhancer sequences and exonic/intronic splicing silencer sequences regulate splicing at neighbouring sites by binding cellular factors such as members of the heterogeneous nuclear ribonucleoprotein (hnRNP) family [50, 51]. Changes in expression of hnRNPs alters HIV-1 structural protein synthesis [52] and exemplifies just one of the many ways

cellular RBPs regulate HIV-1 replication.

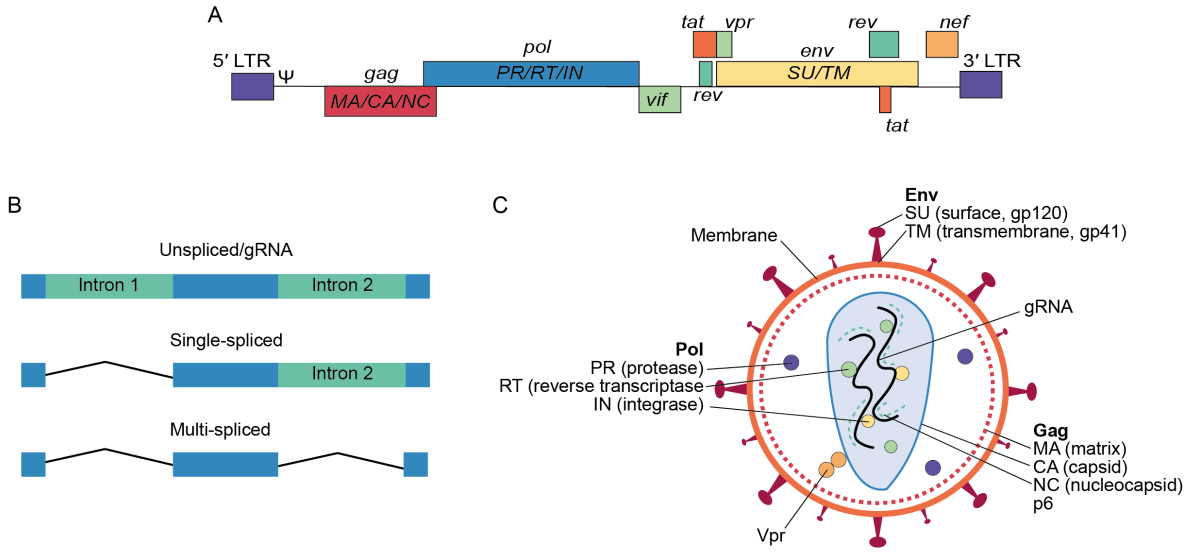

Figure 1.1: **Structure of the HIV-1 genome and mature viral particle. A.** Schematic of the HIV-1 genome structure. **B.** The three main splice variants of HIV-1 with removed introns indicated by the zig-zag line and the remaining exons shown as boxes. **C.** Structure of the mature HIV-1 particle.

1.2.2 Virion structure of HIV-1

HIV-1 virions are spherical membrane-bound particles roughly 100nm in diameter. The outer lipid membrane is derived from the plasma membrane of the host cell from which the particle was released. Embedded in this lipid membrane are the viral envelope proteins composed of the transmembrane subunit gp120 and the glycoprotein subunit gp41. These proteins are produced as a precursor, gp160, which trimerises then undergoes cleavage via a host furin-like protease into two non-covalently linked fragments: gp120 and gp41. Trimers of gp120 and gp41 stick out from the lipid membrane like spikes (Figure 1.1C) and interact with the CD4 receptor and C-C chemokine receptor type 5 (CCR5)/C-X-C chemokine receptor type 4 (CXCR4) co-receptors on host cells to facilitate fusion and viral entry [53]. Different strains of HIV-1 are classed as either R5-tropic and enter using the CCR5 co-receptor or as X4-tropic and use the CXCR4 co-receptor. In this thesis, all viruses used are derived from the NL4-3 HIV-1 strain and are X4-tropic. The NL4-3 strain was chosen specifically because it encodes all known HIV-1 proteins.

In mature viral particles, MA proteins tether to the cytoplasmic domains of the envelope proteins underneath the lipid membrane, shielding the viral capsid. The capsid has a conical structure comprised of 1500-2000 copies of the CA protein arranged in a hexameric lattice with 12 pentamers inserted within the structure to produce enough curvature to close the fullerene cone [54]. The CA hexamers are dynamic pores 8Å in diameter that open and close enabling the delivery of dNTPs into the capsid for reverse transcription [55]. Protected inside the conical capsid are the two identical copies of positive sense ssRNA genome which are bound by the proteins RT and IN (Figure 1.1C). Both of these proteins are required for reverse transcription so their encapsidation is crucial for the early steps of HIV-1 replication.

In addition to mature viral particles, HIV-1 also exists as a non-infectious immature virion. These immature particles contain a spherical core comprised of the Gag polyprotein. A conformational change is induced through cleavage of the Gag polyprotein by PR to produce the full complement of Gag and Gag-pol derived proteins. These proteins can then assemble the conical capsid which marks the maturation of the viral particle.

As well as viral proteins, many host proteins are present in HIV-1 particles. Proteomics studies of HIV-1 particles have revealed a vast number of cellular proteins are incorporated into virions [56–61]. These proteins will be discussed in more detail in Section 1.3.

1.2.3 HIV-1 life cycle

The majority of the literature defines the start of the HIV-1 life cycle as attachment of the trimeric envelope proteins gp120 and gp41 to CD4 receptors and either a CCR5 or CXCR4 co-receptor on the surface of CD4+ T lymphocytes, macrophages or dendritic cells. Upon binding to cell surface receptors, the N terminal fusion peptide of gp41 undergoes a conformational change and inserts itself into the plasma membrane of the target cell. Subsequent rearrangements of gp41 bring the virus outer membrane and target cell membrane together creating a fusion pore [53]. After membrane fusion, the capsid core of the virus is released into the cytoplasm of the target cell where it

is transported through the cytoplasm to the nucleus. This transport has been shown to be reliant on microtubules [62] and requires the association of the CA proteins with dynein and kinesin-1 motors through associated adaptor proteins [63–65].

As the capsid travels towards the nucleus, the vRNA inside the core is reverse transcribed by RT into proviral dsDNA. In order to initiate reverse transcription, a cellular $tRNA^{Lys3}$ is bound to the vRNA near the 5 end to prime the first steps of reverse transcription. This vRNA RNP complex is known as the reverse transcription complex (RTC) and features other viral proteins such as NC, which acts as a chaperone. Cellular proteins such as the histone deacetylases SAP18 and HDAC1 have also been suggested as part of the RTC [66, 67]. HIV-1 RT also has RNase H activity and degrades the 5 end of the vRNA leaving the newly synthesised proviral dsDNA. There are approximately 50 copies of RT within the virion and it is unclear whether the same RT that synthesises the DNA also degrades the RNA [68]. Through the enzymatic action of viral IN, the RTC transitions into the pre-integration complex (PIC), a large RNP complex containing viral and cellular proteins [69]. This event is also marked by changes in CA core structure [70].

Whilst reverse transcription is thought to occur inside the capsid in the cytoplasm, the point at which the capsid core disassembles is under debate. Historically, capsid uncoating was believed to happen in the cytoplasm as the size of the conical capsid ($\sim$120nm by 60nm by 40nm) exceeds the maximum opening size of nuclear pore complexes ($\sim$39nm) [70]. The immediate uncoating model suggested a complete loss of capsid structure in the cytoplasm [71] and was supported by the inability to co-purify CA with any other viral proteins in the RTC due to technology at the time. However, it is understood that the role of the capsid during infection is to protect the vRNA and vDNA from detection by cytoplasmic innate immune sensors and other restriction factors. It is unclear how this is achieved if the capsid disassembles within the first few hours of infection.

In the last decade more evidence has been found to suggest either partial uncoating of the capsid in the cytoplasm or disassembly of the capsid at nuclear pores [55, 72–74]. More recently, papers suggest that the capsid uncoats once inside

the nucleus [75], near the sites of DNA integration [76] and that reverse transcription may be completed in the nucleus [76–80]. Furthermore, CA has been shown to interact with members of the nuclear pore complex (NPC) on both the cytoplasmic (NUP358) and nuclear sides (NUP153) of the pore suggesting CA plays a role in nuclear import [81]. Recent studies have also shown morphological changes in CA which allows the CA-associated PIC to fit through NPCs [70].

Once inside the nucleus, the proviral dsDNA is integrated into the host genome by the viral protein IN and a range of cellular proteins. Large-scale sequencing has revealed millions of retroviral DNA integration sites across the human genome [82]. However, integration is not entirely random, there is a low level of sequence specificity [83] and HIV-1 has been shown to favour integration in bodies of actively expressed genes [84]. After a period of indeterminate latency, host RNA polymerase II transcribes the proviral DNA to produce the differently spliced vRNAs discussed in Section 1.2.1. Export of completely (multi-) spliced viral mRNAs is facilitated via the canonical NXF1 nuclear export pathway [85], enabling the synthesis of the viral protein Rev. Rev then facilitates the export of unspliced and single-spliced viral mRNA from the nucleus through recognition of the cis-acting Rev response element which is found in the introns of unspliced viral mRNAs. Interaction of Rev with the nuclear export receptor Crm1 enables transport of the RNAs into the cytoplasm before splicing is completed [86, 87].

Whilst the multi-spliced and single-spliced mRNAs are translated in the cytoplasm to produce their encoded proteins, the full-length unspliced RNA acts as both mRNA for translation and genomic (g)RNA to be packaged into new virions. It is unclear how it is decided which function the RNA will conduct. However, it has been shown that the individual RNA molecules carry out only one function and that only the nontranslating RNAs serving as gRNA are selectively packaged into assembling virions [88].

The NC domain of Gag is the primary driver of RNA packaging into the virion through recognition of dimerised HIV-1 gRNA [89]. NC contains two ZnF-like domains and directs packaging of the gRNA by binding to a packaging signal, the ψ element, located near the 5 end of the 5 untranslated region (UTR). The HIV-1 RNP initiates

in the perinuclear region [90] and is then transported to the plasma membrane for assembly.

HIV-1 viral particles assemble in the inner leaflet of the plasma membrane. However, the exact mechanism of how Gag proteins are targeted to the plasma membrane is still under investigation. Nuclear magnetic resonance (NMR) data revealed that interaction of MA with a derivative of phosphatidylinositol-4,5-bisphosphate exposes a myristate group on MA, facilitating its insertion into the inner leaflet of the plasma membrane, thereby stabilising Gag in the membrane [91]. Upon reaching the plasma membrane, Gag induces the recruitment of cholesterol and sphingolipid enriched microdomains, called lipid rafts, which are thought to serve as platforms for virus assembly [92].

Whilst the Gag-RNA RNP nucleates assembly of viral particles, the interaction between CA domains of the Gag polyprotein is responsible for driving Gag multimerisation. HIV-1 CA is structured in an N-terminal domain (NTD) and a C-terminal domain (CTD). The CA CTD and the N terminus of p6 contribute to the six-fold interaction at the interface of the hexameric CA pores. This interaction forms an immature Gag lattice of hexameric rings [93].

The method of incorporation of envelope glycoproteins into the membrane of the assembling particles is not completely understood. Mutation-based experiments suggest a central role for MA binding the cytoplasmic tail of gp41 [94]. However, several non-mutually exclusive mechanisms have been additionally suggested that may participate in Env incorporation: passive incorporation of surface-expressed Env; Gag-Env co-targeting to lipid rafts; and indirect binding of MA and gp41 through CA or NC [92].

After assembly of the immature Gag lattice, membrane fission must occur to release the nascent particle from the cell surface. The p6 domain of Gag harbours two highly conserved motifs: PTAP and YPXL, which are responsible for recruiting tumour susceptibility gene 101 (TSG101) [95] and ALG2-interacting protein X (ALIX) [96], respectively. TSG101 and ALIX are both part of the endosomal sorting complex required for transport (ESCRT) which is responsible for several membrane budding

and scission processes within the cell. ESCRT-mediated membrane fission releases the newly assembled virions from the plasma membrane [97]. Inside the immature viral particles HIV-1 PR cleaves the Gag polyprotein to form the individual proteins (MA, CA, NC and p6) which spontaneously assemble into the mature capsid structure forming a mature, infectious viral particle [92].

Despite having an RNA-centric life cycle, only 5 of the HIV-1 proteins are capable of binding RNA (NC, Rev, MA, RT and IN), which is insufficient to facilitate its complex life cycle. To circumvent this limitation, HIV-1, and viruses in general, usurps cellular RBPs for its own use. These RBPs interact with vRNA forming dynamic RNPs, which are critical for mediating gene expression [98]. It is well known that HIV-1 hijacks cellular RBPs for use in reverse transcription and RNA processing [99–108]. However, it is becoming apparent that cellular RBPs are involved in virtually all stages of viral infection.

1.3 RBPs within the HIV-1 life cycle

RBPs play important roles in many viral infections, the effect is not limited to HIV-1. Recent work by Garcia-Moreno et al [15] illustrated a remodelling of the cell RBPome during Sindbis virus infection with over 200 host RBPs exhibiting enhanced or inhibited RNA-binding activity. Similarly, in severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, a third of the cell RBPome was found to remodel upon infection. Breakdown of the SARS-CoV-2 RNP revealed 139 cellular RBPs interacting with the vRNA [109]. Several groups have used genomic screening approaches to establish the repertoire of host proteins involved in HIV-1 replication via RNA-knockdown [110–113]. Additionally, many studies looking at individual RBPs have determined roles in HIV-1 replication. However, less extensive research has been conducted identifying the whole cellular RBPome for HIV-1.

In 2011, two groups used HIV-1 RNA modified with MS2 bacteriophage-binding sites to identify the HIV-1 RNA-interactome [114, 115]. By adding flag- or maltose binding protein (MBP)-tagged MS2 coat protein to the lysates of cells that had been transfected with the modified HIV-1 RNA, the HIV RNP could be immunoprecipitated.

These two strategies identified many previously known mRNA binders and identified novel functions in HIV-1 replication for the splicing activator Matrin-3 (MATR3) and the pre-mRNA splicing regulator hnRNPK. Whilst these results are valuable for the discovery of novel RBPs interacting with vRNA, they are limited by their use of modified RNA transfection instead of HIV-1 infection. To address this issue, Knoener et al developed a hybridization and mass spectrometry approach to identify the host interactome of unspliced HIV-1 RNA using native HIV-1 infection [116].

Hybridization purification of RNA-protein complexes followed by mass spectrometry (HyPR-MS) can be broken into four steps. Firstly, HIV-1 infected Jurkat T cells are formaldehyde-crosslinked to preserve RNA-protein interactions. Secondly, the infected cell lysate is incubated with biotinylated capture oligonucleotides which recognise unspliced vRNA. Thirdly, the oligonucleotide-bound vRNA is isolated using streptavidin beads and subject to stringent washing. Finally, the RNP complexes are released from the beads and analysed by mass spectrometry. Using this approach, Knoener et al identified 189 RBPs interacting with vRNA. Of these 189, 24 had been previously shown to bind HIV-1 RNA and 90 had known interactions with viral RBPs, which demonstrates the credibility of the technique. Functional effects of several proteins they identified (MOV10, hnRNPC, HSPD1 and RBBP4) were confirmed through infection of siRNA knockdown cells [116].

After the initial success of HyPR-MS, Knoener et al went on to develop HyPR-MS_{sv} to determine the host RBPs differentially associated with unspliced, single-spliced and multi-spliced HIV-1 RNA [117]. By using capture oligonucleotides complementary to three distinct regions of HIV-1 RNA they identified over 900 proteins in the interactomes of all three splice variants and 200 proteins differentially associated with the different splice variant pools. They again performed siRNA knockdown to determine functional effects but this time they screened 121 host proteins, 80 of which were shown to be effectors of HIV-1 RNA regulation. Again, many of the proteins identified were also identified in the previous HIV-1 RNA-interactome studies which validated the strategy. Of the 900 proteins associated with all splice variants, many were well-established RNA regulatory factors and interactors shown to play important roles in HIV-1 RNA

regulation (e.g. the dead box helicases DHX9 and DDX3). Regarding the differentially associated proteins, it was hypothesised that RBPs associated with unspliced and single-spliced RNAs were likely to be involved in Rev-driven nuclear export pathways. Additionally, proteins preferentially associated with unspliced RNA but not single- or multi-spliced may be involved in genome packaging or virus assembly. Some of the RBPs identified in these studies and how they act in the viral life cycle are introduced in Table 2. This is by no means an exhaustive discussion of RBPs involved in infection but serves to demonstrate that almost all stages of the HIV-1 life cycle are regulated by cellular RBPs.

Table 2: Cellular RBPs acting within different stages of HIV-1 infection.

Stage of Life Cycle	**Example cellular RBPs involved**
Reverse transcription	UPF1 [99], Y-box binding protein 1 (YBX1) [100], DHX9 [101], EEF1A [102, 103], PPIA [104] and aminoacyl-tRNA synthetases [105] promote reverse transcription. Whereas Mov10 RISC complex RNA helicase (MOV10) [106], pyruvate kinase M1/2 (PKM) [107] and enolase 1 (ENO1) [108] are known to inhibit reverse transcription. Additionally, APOBEC proteins exert antiviral activity on the products of reverse transcription by inducing cytidine deamination of the reverse-transcribed DNA strand. Changes from C to U cause complementary G to A conversion during second strand synthesis and this hypermutation inhibits HIV-1 replication [118]. In addition, APOBEC3 proteins also restrict HIV-1 by a deamination-independent mechanism that consists of altering reverse transcription template switching frequency [119]. To exert their antiviral effects in HIV-1, these proteins are recruited into viral particles by binding to the vRNA and nucleocapsid as new virions assemble [120].
Integration	Polypyrimidine tract binding protein (PTBP) and its associated splicing factor PSF have been shown to interact with HIV-1 IN to promote integration of viral cDNA [121]

Table 2: Cellular RBPs acting within different stages of HIV-1 infection.

Stage of Life Cycle	**Example cellular RBPs involved**
Transcription	Transcription of HIV-1 provirus is achieved by a complex of viral and cellular factors including RNA polymerase II and positive transcription elongation factor complex (P-TEFb). HIV-1 Tat liberates P-TEFb from sequestration by inactive snRNP complexes [122] to phosphorylate and form a complex with stalled RNA polymerase II, enabling the release of the transcriptional complex [123].
Nuclear Export	Multi-spliced vRNA is exported from the nucleus via the canonical NXF1 nuclear export pathway [85]. Whilst the Crm1 nuclear export pathway facilitates export of HIV-1 unspliced and single-spliced RNAs in a Rev-dependent manner [86]. Additionally, the RNA helicase DDX3 is required for nucleo-cytoplasmic shuttling of HIV-1 vRNAs and is required for Rev function [124].
Translation	Translation of HIV-1 RNAs can occur via a cap-dependent mechanism which is initiated by the eukaryotic translation initiation factors eIF4E, eIF4G and eIF4A binding to 5' capped RNA [125]. Alternatively, translation can occur via an IRES-driven mechanism in which an IRES present in Gag can bind the 40S ribosomal subunit, eIF3 and eIF4A [126]. In addition to positive cofactors, cellular RBPs also negatively regulate viral translation. Protein kinase RNA-activated (PKR) is an interferon-stimulated gene activated by dsRNA. It blocks the initiation of translation resulting in the inhibition of viral protein synthesis [127]. However, the PKR-mediated antiviral response is suppressed by other cellular RBPs including TAR RNA binding protein (TRBP) and adenosine deaminase enzyme that acts on dsRNA 1 (ADAR1). Both of these proteins bind to dsRNA and sequester it from PKR [128].

Table 2: Cellular RBPs acting within different stages of HIV-1 infection.

Stage of Life Cycle	**Example cellular RBPs involved**
Chaperoning of vRNP	Unspliced HIV-1 RNA forms an RNP with staufen double-stranded RNA-binding protein 1 (STAU1) [129–131]. Compositional characterisation of these RNP complexes revealed the presence of many other cellular RBPs including the chaperones HSPD1 (HSP60), HSPA1 (HSP70) and HSPA8 (HSC70) [131]. It has been proposed that the stress-inducible chaperone HSP70 and its constitutive form, HSC70, interact with nascent HIV-1 Gag-vRNA complexes and hold them in an assembly-competent conformation during their transport towards the plasma membrane [132].
Viral particle assembly and release	ESCRT driven budding from the plasma membrane is initiated by the recruitment of ALIX. RIC studies have revealed that ALIX is endowed with RNA-binding activity [36]. The interaction of HIV-1 Gag with ALIX relies on the presence of vRNA. The N-terminal Bro1 domain and the central V-shaped domains of ALIX interact with HIV-1 nucleocapsid and p6 domains of Gag respectively [133, 134]. Interestingly, the interaction between ALIX Bro1 domain and HIV-1 nucleocapsid is disrupted by RNase treatment, suggesting that the vRNA molecule forms a bridge between the two proteins. Both ALIX Bro1 and nucleocapsid are highly positively charged and are believed to establish electrostatic interactions with the phosphate backbone of the vRNA [134].

1.4 How do RBPs necessary for HIV-1 infection access the viral RNA?

Many of the cellular RBPs discussed above interact with the vRNA during the early stages of HIV-1 infection. However, with recent research showing capsid disassembly occurring at NPCs, one question is left unanswered: how do the RBPs necessary for HIV-1 infection access the viral RNA?

In order for the NPC uncoating model to work, the viral and cellular factors necessary for vRNA metabolism must be encapsidated inside the viral particles during assembly. Whilst the capsid is made up of dynamic pores, when in the fully open conformation these pores are only 8Å in diameter [55]. This is sufficient for dNTPs to pass through, but not for the cellular RBPs necessary for reverse transcription and other early infection steps. Therefore, the capsid represents a barrier that should impede cellular RBPs from interacting with the vRNA in the cytoplasm of the newly infected cell. The Castello lab hypothesised that cellular factors required by HIV-1 for the initial steps of infection are packaged into assembling HIV-1 viral particles before they bud from the plasma membrane (Figure 1.2). Also included in this group of proteins would be cellular innate immune restriction factors present in the cytoplasm that would recognise the vRNA. The incorporation of APOBEC3 proteins into HIV-1 capsids through recognition of the viral gRNA provides an excellent example of this [118, 120]. No previous analysis of RBPs within HIV-1 viral particles has been conducted. Therefore, I performed a deep review of published HIV-1 particle proteomes to determine whether RBPs were present within these proteomes. This revealed that 187 cellular RBPs were found collectively within the HIV-1 particle proteomes [9]. These analyses identified all proteins within HIV-1 virions so it cannot be concluded that the RBPs would necessarily be interacting with the vRNA. However, it provided strong support for our hypothesis that cellular factors are deliberately packaged inside assembling virions.

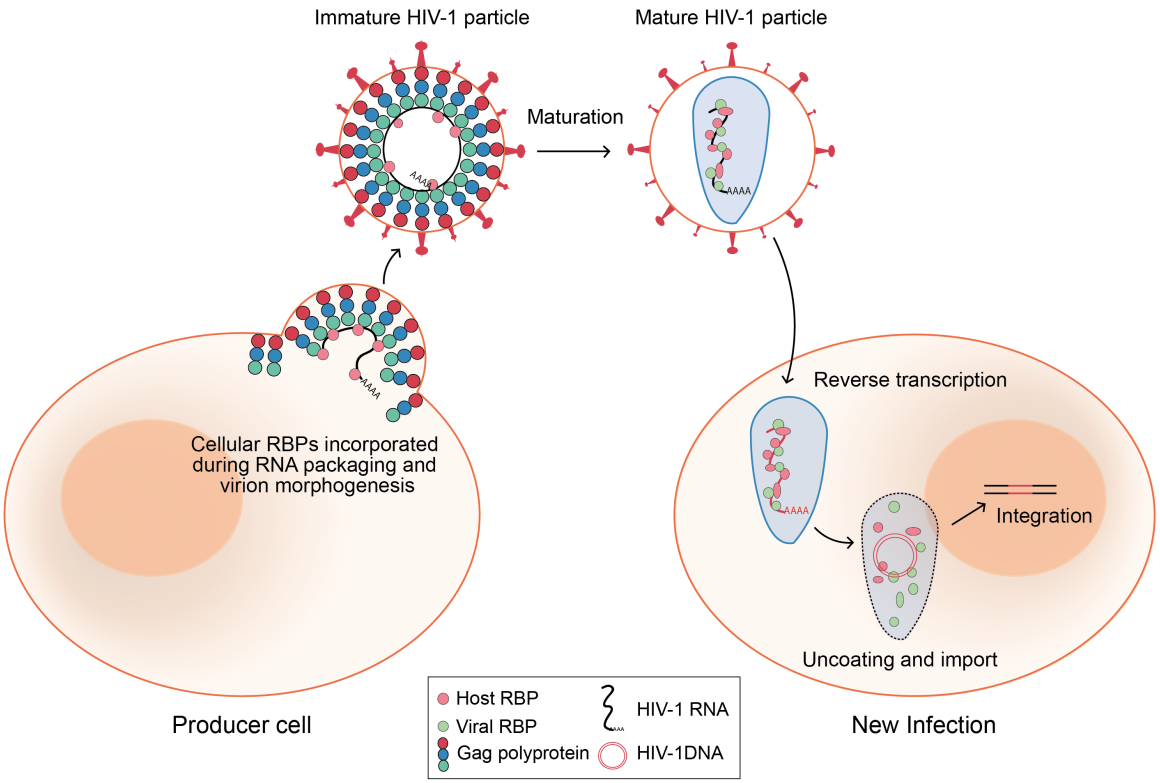

Figure 1.2: **RBPs that regulate HIV-1 infection are packaged into assembling virions.** We hypothesise that the cellular RBPs required by HIV-1 for the initial stages of infection are packaged into assembling viral particles before they bud from the plasma membrane. These proteins are then present in the capsid for new infection. Also included in this group of proteins would be cellular innate immune restriction factors.

To test this hypothesis we pioneered a technique called in-virion RNA interactome capture (ivRIC). Based on the previously discussed RIC method, ivRIC determines the complement of host RBPs interacting with vRNA inside viral particles - the 'in-virion (iv)RBPome' [135]. This approach takes HIV-1 particles from the supernatant of infected cells and purifies them via a sucrose density gradient. UV irradiation at 254nm is used to induce crosslinks between protein-RNA interactions at < 2Å distance. After lysis under denaturing conditions, oligo(dT) capture is used to isolate the vRNA before treating with RNases to leave just the bound RBPs. These proteins were then identified using quantitative mass spectrometry.

Using ivRIC we identified 104 cellular RBPs bound to the vRNA which we called 'ivRBPs'. Whilst some of the ivRBPs identified had known roles in HIV-1 infection,

many were novel HIV-1 interactions and some had no known interaction with viruses at all. My thesis aims to validate these interactions and elucidate the role of some of these proteins in HIV-1 infection.

1.5 Aims

As highlighted by Dicker et al 2021 [9], we know many RBPs are involved in viral infection and incorporated into viral particles. However, we know very little about the roles these RBPs play inside the viral particle or why they are packaged. My DPhil aims to broaden our knowledge of the RBPs present in HIV-1 viral particles and for a chosen few proteins, understand their role in HIV-1 infection. A greater understanding of the host proteins that are necessary for HIV-1 infection is the first step in the development of new host-targeted therapeutics against HIV-1. The specific goals of my research are as follows:

1. Identify host ivRBPs that play key roles in the early stages of HIV-1 infection by assessing the affect of ivRBP knockout (KO) on viral protein, RNA and particle production.

2. Characterise the role of one of these ivRBPs in HIV-1 infection and establish a mechanism of action.

3. Develop a system to visualise the initial stages of HIV-1 infection from viral particle entry to integration using super-resolution microscopy.

2 Materials

2.1 Chemicals and Reagents

Reagent	Supplier	Identifier
1 kb Plus DNA Ladder	NEB	N3200S
1-step Ultra TMB-ELISA	Pierce	34028
AEBSF hydrochloride	BioChemica	A14210100
AgeI-HF	NEB	R3552S
Benzonase Nuclease	Millipore	E1014
BglII	NEB	R0144S
Binding control agarose beads	Chromotek	bab-20
Blasticidin	InvivoGen	ant-bl-05
Chloroquine	ACROS Organics	50-63-5
Ciprofloxacin	Thermo Fisher	J61317.06
DAPI	Thermo Fisher	D1306
DH5α *E. coli*	NEB	C2987H
DMEM	Gibco	11965092
Doxycyline	Sigma Aldrich	D9891
DTT	Sigma Aldrich	D1532
EcoRI-HF	NEB	R3101S
Empigen BB	fluka biochemika	45165
FBS	Gibco	16000044
GFP-Trap agarose beads	Chromotek	gta-20
Hygromycin B	Thermo Fisher	J60681.MD
KpnI-HF	NEB	R3142S
NotI-HF	NEB	R3189S
NuPAGE™4-12% Bis-Tris 1mm gel	Invitrogen	NP0322BOX
One Shot™Stbl3 Chemically Competent *E. coli*	Invitrogen	C737303
Opti-MEM reduced serum Medium	Gibco	31985062
Paraformaldehyde, 16% w/v aq. soln., methanol free	Thermo Fisher	043368.9M

Reagent	Supplier	Identifier
PBS	Gibco	10010023
Penicillin/streptomycin	Gibco	15140122
Precision Plus Protein Dual Colour Standards	Bio-Rad	1610374
ProLong Diamond Antifade Mountant	Invitrogen	P36961
Ribonucleoside Vanadyl Complex	NEB	S1402S
RNAsecure RNase Inactivation Reagent	Invitrogen	AM7006
RPMI 1640	Gibco	11530586
SlowFade Diamond antifade mountant	Thermo Fisher	S36967
Solution 13 AO-DAPI	Chemometec	910-3013
T4 DNA ligase	NEB	M0202S
Triton X-100	Promega	H5141
Trypan Blue Stain (0.4%)	Invitrogen	T10282
TrypLE Express Enzyme	Gibco	12604013
Tween-20	Sigma Aldrich	P1379-100ML
VECTASHIELD Hardset antifade mounting medium	Vector laboratories	H-1400-10
X-tremeGENE 9 DNA transfection reagent	Roche	XTG9-RO
X-tremeGENE 360 DNA transfection reagent	Roche	XTG360-RO
Zeocin	Invitrogen	R25001

2.2 Machines

Table 4: Machines used.

Machine name	**Supplier**
QuantStudio5	Applied Biosystems
Nanodrop 2000 Spectrophotometer	Thermo Fisher
Trans-Blot Turbo Transfer System	Bio-Rad
LI-COR Odyssey Fc Imaging System	LI-COR
Countess II FL Automated Cell Counter	Thermo Fisher
SpectraMaxM5 plate reader	Molecular Devices
CLARIOstar plate reader	BMG Labtech
Olympus SpinSR10 spinning disc confocal microscope	Olympus
Zeiss 880 Airyscan confocal microscope	Zeiss
Olympus SoRa super-resolution spinning disc con...	Olympus
Evos M7000 micrscope	Evos
Guava EasyCyte	Cytek Biosciences

2.3 Buffers and solutions

Table 5: Commercial buffers.

Buffer	**Supplier**	**Identifier**
Glo Lysis Buffer 1X	Promega	E2661
NuPAGE LDS Sample Buffer 4X	Invitrogen	NP0007
Stellaris® RNA FISH Hybridization Buffer	LGC Biosearch	SMF-HB1-10
Stellaris® RNA FISH Wash Buffer A	LGC Biosearch	SMF-WA1-60
Stellaris® RNA FISH Wash Buffer B	LGC Biosearch	SMF-WB1-20
NuPAGE MOPS SDS running buffer	Invitrogen	NP0001
Restore western blot stripping buffer	Thermo Fisher	21059
rCutSmart	NEB	B6004S

Table 6: In-house buffers

In-house buffers	**Composition**
1% NP40 lysis buffer	10 mM Tris HCl, 150 mM NaCl, 0.5 mM EDTA, 1% NP40 pH 7.5
0.5% NP40 wash Buffer	10 mM Tris HCl, 150 mM NaCl, 0.5 mM EDTA, 0.5% NP40 pH 7.5
IP Wash Buffer	10 mM Tris HCl, 150 mM NaCl, 0.5 mM EDTA
1M Tris	In MilliQ water, pH 10.4
200mM glycine	In MilliQ water, pH 2.5
PBST	1x PBS, Tween at stated %
PBSTx	1x PBS, Triton-X100 at stated %
TMT/SS	1x TBS, 2% BSA, 20% FCS
TBS/E/S	10% FBS, 1x TBS, 0.5% Empigen
2x HBS	50 mM HEPES, 1.5 mM Na2HPO4, 280 mM NaCl, 10 mM KCl, 6 mM sucrose pH 7.09
RIPA lysis buffer	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% Triton-X100, 0.5 mM EDTA, 25 U benzonase, 1x AEBSF
10% Sucrose	10% sucrose, 50 mM Tris HCl pH 7.4, 100 mM NaCl, 0.5 mM EDTA

2.4 Commercial kits

Table 7: Commercial kits

Kit	**Supplier**	**Identifier**
Luna Universal One-step RT-qPCR kit	NEB	E3005E
Monarch Genomic DNA Purification kit	NEB	T3010S
RNeasy mini kit	Qiagen	74004
QIAquick Gel Extraction Kit	Qiagen	28706X4
QIAquick PCR Purification Kit	Qiagen	28104
PureLink™HiPure Plasmid Filter Maxiprep Kit	Invitrogen	K210016
QIAprep Spin Miniprep Kit	Qiagen	27104
Bright-Glo Luciferase Assay System	Promega	E2620
HIV-1 p24 ELISA pair set	SinoBiological	SEK11695
HIV-1 Gag p24 DuoSet ELISA	R&D Systems	DY7360-05
SilverQuest staining kit	Invitrogen	LC6070
Q5® Site-directed mutagenesis kit	NEB	E0554S

2.5 Antibodies and dyes

Table 8: Antibodies and dyes

Antibody/dye	**Supplier**	**Identifier**
CellMask Green	Thermo Fisher	C37608
DAPI	Invitrogen	D1306
GAPDH	Invitrogen	PA1-988
GFP booster	chromotek	gb2AF488
Goat anti-human IgG Alexa Fluor 488	Life technologies	A11013
HIV Env 2G12	CFAR	3064
HIV GP120	CFAR	3056
HIV IN (p32)	Santa Cruz Biotech	SC-69721
HIV p24 37G12	CFAR	ARP3279
HIV p24/55	CFAR	ARP0314
HIV RT (p66)	Santa Cruz Biotech	SC-57829
IRDye 680RD Donkey anti-Rabbit IgG	LI-COR Biosciences	926-68073
IRDye 680RD Goat anti-Mouse IgG	LI-COR Biosciences	926-68070
IRDye 800CW Donkey anti-Rabbit IgG	LI-COR Biosciences	926-32213
IRDye 800CW Goat anti-Human IgG	LI-COR Biosciences	926-32232
IRDye 800CW Goat anti-Rat IgG	LI-COR Biosciences	926-32219
KC57-FITC	Beckman Coulter	6604665
LIVE/DEAD fixable Near-IR dead cell stain kit	Invitrogen	L34975A
MKRN1	Bethyl Laboratories	A300-990A
Phalloidin-Alexa Fluor 488	Thermo Fisher	A12379
RFP	chromotek	5F8-100
RFP booster	chromotek	rba594
VSVG	Santa Cruz Biotech	SC-365019
beta-actin	Sigma	A1978
HIV Nef 2H12	Invitrogen	MA1-71505
Goat anti-human IgG Alexa Fluor 647	Invitrogen	A48279

2.6 Plasmids

Table 9: Commercial and lab-made plasmids

Plasmid	**Backbone**	**Source**
pcDNA3-FLAG-MKRN1	pcDNA3-FLAG	Addgene
pcDNA5-FRT-TO	pcDNA5-FRT-TO	Thermo Fisher
pcDNA5-FRT-TO-eGFP-linker	pcDNA5-FRT-TO	Made in Castello Lab
pcDNA5-FRT-TO-MKRN1-C281A-GFP	pcDNA5-FRT-TO	Made in this project
pcDNA5-FRT-TO-MKRN1-GFP	pcDNA5-FRT-TO	Made in this project
pcDNA5-FRT-TO-MKRN1-H307E-GFP	pcDNA5-FRT-TO	Made in this project
pcDNA5-FRT-TO-MKRN1-M285E-GFP	pcDNA5-FRT-TO	Made in this project
pHEF-VSV-G	pHEF-VSV-G	Addgene
pNL4.3	pNL4.3	CFAR
pNL4.3.Luc.R-E	pNL4.3.Luc.R-E	CFAR
pNL4.3-Gag-mCherry	pNL4.3	Made in Castello lab
pNL4.3-IN-scarlet	pNL4.3	Made in this project
pNL4.3-Nef-mCherry	pNL4.3	Made in Castello lab
pNL4.3-R-E-Gag-mCherry	pNL4.3.Luc.R-E	Made in Castello lab
pNL4.3-R-E-IN-scarlet	pNL4.3.Luc.R-E	Made in this project
pNL4.3-R-E-T2A-Nef-mCherry	pNL4.3.Luc.R-E	Made in Castello lab
pOG44 Flp-recombinase expression vector	pOG44	Thermo Fisher
pTWIST-AMP-INscarlet	pTWIST	Twist Bioscience

2.7 Oligonucleotides

Table 10: Oligonucleotides

Name	**Sequence**
HIVprotease_F	GTAGGACCTACACCTGTCAAC
HIVtat_R	GCTGACTTCCTGGATGCTTC
SDM_H307E_Fw	CAACTGCAACGAGACCTACTGTCTCAAG
SDM_H307E_Rv	GAGAGGATCCCGAAGCGG
SDM_C281A_Fw	GGACATGGTGGCCGGGATCTGCATGGAGGTGG
SDM_C281A_Rv	TTGCTGCGCTGCACGGCA
SDM_M285E_Fw	TGGGATCTGCGAGGAGGTGGTCTAT
SDM_M285E_Rv	CACACCATGTCCTTGCTG
UnsplicedHIV-f	GACGCTCTCGCACCCATCTC
UnsplicedHIV-r	CTGAAGCGCGCACGGCAA
Single-splicedHIV-f	GGCGGCGACTGGAAGAAGC
Single-splicedHIV-r	CTATGATTACTATGGACCACAC
Multi-splicedHIV-f	GACTCATCAAGCTTCTCTATCAAA
Multi-splicedHIV-r	AGTCTCTCAAGCGGTGGT
Actin 3’	TGACCGGAGTCCATCACCAT
Actin 5’	CGGGAGAAGATGACCCAGAT
TS576	CACCGCTGCTTCTTTGTCTCCACGG
TS577	AAACCCGTGGAGACAAAGAAGCAGC
TS578	CACCGAAACCCACAATGCTGTGAAA
TS579	AAACTTTCACAGCATTGTGGGTTTC
eGFP-MKRN1-Fw	CGATTGAAGCTTACCATGGTGAGCAAGGGCGAGGAG
eGFP-MKRN1-Rv	CTACTCGACCTTCTAAAAATACTGAACCTAGATATCATCCGCC

2.8 Cell lines

Table 11: Parental cell lines

Cell line	**Cell type**	**Source**
HEK293T	Human embryonic kidney cells	Castello lab
HEK293 Flp/In T-Rex	Human embryonic kidney cells	Castello lab
SupT1	Human T lymphocytes	Castello lab

Table 12: Genetically modified cell lines

Cell line	**Cell type**	**Source**
SupT1 MKRN1-KO	Human T lymphocytes	Cloned in this project
SupT1 FAM120A-KO	Human T lymphocytes	Castello lab
SupT1 LSM14A-KO	Human T lymphocytes	Castello lab
SupT1 MSN-KO	Human T lymphocytes	Castello lab
SupT1 LCK-KO	Human T lymphocytes	Castello lab
SupT1 SUB1-KO	Human T lymphocytes	Castello lab
SupT1 CIRBP-KO	Human T lymphocytes	Castello lab
SupT1 YBX3-KO	Human T lymphocytes	Castello lab
SupT1 ZAP-KO	Human T lymphocytes	Castello lab
HEK293 Flp/In T-Rex MKRN1-GFP	Human embryonic kidney cells	Cloned in this project
HEK293 Flp/In T-rex MKRN1-KO	Human embryonic kidney cells	Cloned in this project
HEK293 Flp/In T-Rex MKRN1-M285E-GFP	Human embryonic kidney cells	Cloned in this project
HEK293 Flp/In T-Rex MKRN1-C281A-GFP	Human embryonic kidney cells	Cloned in this project
HEK293 Flp/In T-Rex MKRN1-H307E-GFP	Human embryonic kidney cells	Cloned in this project
HeLa-CD4-GFP	Human cervical carcinoma cells (HeLa)	Castello lab
TZM-bl	Human cervical carcinoma cells (HeLa)	Sattentau lab
HEK293T MKRN1-KO	Human embryonic kidney cells	Kind gift from König lab

2.9 Software

Table 13: Software used within this thesis

Software	Version	Source
AlphaFold	2	[136, 137]
BindUP	N/A	[40]
Cytoscape	3.9.1	[138]
FlowJo	10.8.0	FlowJo
GO enrichment analysis	2023-03-06	[139]
Graphpad Prism	9	Graphpad
ImageJ	1.53t	National Institutes of Health
Image Studio Lite	5.2	LI-COR imaging systems
Microsoft Excel	2302	Microsoft
NEBaseChanger	1.3.3	https://nebasechanger.neb.com/
Olympus cellSens Dimension	3.2	Olympus Life Science
Omero.figure	3.2.0	[140]
Pymol	2	[141]
RBDetect	N/A	https://nishuai.shinyapps.io/RBDetect/
Stellaris Probe Designer	4.2	https://www.biosearchtech.com/stellaris-designer
STRING	11.5	[142]

3 Methods

3.1 Cell Biology

3.1.1 Maintenance of Cells

All cells were maintained under sterile conditions at 37°C with 5% CO_2. Adherent cell lines were cultured in Dulbecco's modified eagle medium (DMEM) with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin. Suspension cell lines were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium with 10% FBS and 0.1% ciprofloxacin. In addition to standard culture conditions, all HEK293 Flp/In T-REx cell lines were cultured in 5 µg mL^{-1} blasticidin and 150 µg mL^{-1} zeocin. Cells were regularly passaged to avoid over-confluency. For adherent cells this was achieved by removing media, washing with PBS, incubating briefly with TrypLE Express Enzyme 1X at 37°C and resuspended in fresh media. For suspension cells, cells were mixed, counted using AO-DAPI stain and a cell counter and then 2.5x10^5 cells were resuspended in fresh media. For all experiments requiring a certain seeding density of cells, cells were counted using AO-DAPI stain and a cell counter.

3.1.2 Cloning of MKRN1-KO SupT1 and HEK293 Flp/In cell lines

CRISPR/Cas9 nanoblades targeting two sequences within the MKRN1 gene were generated by Thibault Sohier in Emiliano Ricci's lab, University of Lyon. Two guide (g)RNAs were synthesised using the oligonucleotides TS576 and TS577, and TS578 and TS579 (Table 10). The gRNAs targeted a sequence overlapping exon 4 of the MKRN1 gene, which resulted in a double-cleavage and excision of the sequence between the target sites. CD4+ T lymphocytes (SupT1) and HEK293 Flp/In cells were seeded at 1x10^5 cells/well in 200 µL appropriate media with 10% FBS and transduced with 15ul nanoblades for 4 hours at 37°C. After 3 days cells were counted and diluted to enable the isolation of a single cell per well in a 96 well plate. After expansion, successful KO clones were confirmed via PCR for gene disruption and western blotting to detect MKRN1 protein expression. Clones were selected based on the absence of

MKRN1 protein expression.

3.1.3 Stable Transfection of HEK293 Flp/In T-REx

HEK293 Flp/In T-Rex cells were seeded at $2x10^5$ cells per well in 6 well plates. At 24 hours, 2 µg of the relevant purified pcDNA5-FRT-TO plasmid was mixed with 0.5 µg pOG44 and 6 µL Xtreme Gene 9 transfection reagent in 100 µL OptiMem media and added to cells after 15 minutes. After 24 hours media was supplemented with 150 µg mL^{-1} hygromycin B and 5 µg mL^{-1} blasticidin. Induction of the relevant RBP-GFP fusion protein was achieved by addition of 1 µg mL^{-1} doxycycline for 24 hours.

3.1.4 Generation of pseudotyped lentiviruses

Pseudotyping was conducted following the protocol by Kutner et al [143]. HEK293T cells were seeded to 30% confluency and transfected with 2M $CaCl_2$, 2x HBS buffer pH 7.09 and 84 µg of pHEF-VSVG plasmid and 252 µg of the relevant pNL4.3.Luc.-R-E based plasmid. Transfected cells were incubated at 37°C overnight and the media was replaced with complete DMEM the following morning. 48 hours post-transfection the supernatant was centrifuged at 500xg for 10 minutes and passed through a 0.45 µm filter to remove cell debris. Pseudotyped virus was precipitated by the addition of 1 volume of Lenti-X concentrator to 3 volumes of supernatant and incubated at 4°C overnight. The viral pellet was collected via centrifugation at 4500xg for 1 hour at 4°C and resuspended in PBS. Aliquots were stored at -80°C and freeze-thawed once only.

3.1.5 Generation of HIV-1 and fluorescently tagged HIV-1 viral particles

HEK293T cells were seeded to 70% confluency in DMEM only and transfected with 25 µg of the relevant pNL4.3 plasmid in 0.2 µM PEI and DMEM that had incubated at room temperature for 20 minutes. Cells were incubated at 37°C for 4 hours before the transfection mix was replaced with fresh complete DMEM and then incubated for a further 48 hours at 37°C. Viral supernatant was centrifuged at 500xg for 10 minutes and passed through a 0.45 µm filter to remove cell debris. The viruses were

then precipitated by the addition of 1 volume of Lenti-X concentrator to 3 volumes of supernatant and incubated at 4°C overnight. A viral pellet was collected via centrifugation at 4500xg for 1 hour at 4°C and resuspended in PBS. Aliquots were stored at -80°C and freeze-thawed once only.

3.1.6 Titration of viruses

Virus or pseudotyped lentivirus was titrated in quadruplet in 1:3 dilutions in complete DMEM, starting at a dilution of 1:12, to a final volume of 100 µL per well of a 96 well sterile flat-bottomed plate. 10,000 Tzm-bl cells were added to each well in 100 µL complete DMEM. Cells were incubated at 37°C for 30-36 hours then washed with PBS and lysed in 1x Glo lysis buffer for 2 minutes at room temperature. Bright Glo luciferase substrate was added to the lysed cells 1:1 and incubated at room temperature for 2 minutes. Luminescence was read across all wavelengths with 1000ms integration time on a SpectraMaxM5 plate reader. Analysis of luminescence readings was conducted in Microsoft Excel and infectious dose 50% (ID_{50} calculated using the National Institutes of Health ID-50 programme [144]. For bio-safety level 3 viruses, the plate was wrapped in parafilm, frozen at -80°C for 24 hours then removed from the CL3 facility and completely thawed before the luciferase assay was conducted.

3.1.7 Infection of SupT1 cells

For infection with infectious fluorescently tagged or WT HIV-1, SupT1 cells were always spinoculated. Virus was added to cells at the appropriate MOI in a small volume of media (50 µL/well in a 96 well plate or 200 µL/well in a 24 well plate) (Table 14) and centrifuged at 1200xg for 2 hours at 37°C then incubated at 37°C for a further 2 hours. Media was then replaced with fresh complete media (200 µL/well for a 96 well plate or 500 µL/well for a 24 well plate) and the cells were returned to a 37°C incubator with 5% CO_2.

For infection of SupT1 cells with pseudotyped single round viruses, pseudotyped viruses were added to cells in their culture volume of media (2 mL/well in a 6 well plate, 500 µL/well in a 24 well plate or 200 µL/well for a 96 well) and pipetted up and down to

mix. Infected cells were incubated at 37°C with 5% CO_2.

Table 14: Conditions of SupT1 infection in different experiments

Experiment	**Seeding Density**	**Media Volume**	**Plate Type**	**MOI**
Western blotting and RT-qPCR	$2x10^6$	2 mL	6 well	1
Fluorescence over 72 hpi	$5x10^4$	200 µL	96 well	1
Flow cytometry in SupT1-KOs	$1x10^4$	200 µL	96 well	1
Flow cytometry of HIV-INscarlet	$2x10^5$	500 µL	24 well	variable
P24 ELISA	$1x10^4$	200 µL	96 well	0.1

3.1.8 Infection of HeLa-CD4-GFP for Immunofluorescence

HeLa-CD4-GFP cells were seeded at a density of $1x10^5$ cells/well on sterile glass coverslips in a 24 well plate in 500 µL DMEM with 10% FBS, no phenol red, and induced with 1 mg mL^{-1} doxycycline. After 24 hours, different fluorescent viruses were added at an MOI of 10 in 500 µL DMEM with 10% FBS, no phenol red. The cells were incubated at 37°C for 6 hours before washing with PBS and fixing in 4% PFA.

Fixed cells were permeablised in 0.3% PBSTx and saturated with 5% BSA in PBS. The cells were incubated in anti-p24 (37G12) for 1 hour before washing and incubation with anti-human IgG AlexaFluor-647, GFP-booster and RFP-booster for 1 hour in the dark. The cells were washed again in PBS and a nuclei stain was carried out using DAPI. After a final wash, the coverslips were mounted on glass slides using VECTASHIELD hardset mounting medium. Slides were stored at -20°C. Imaging was conducted on a Zeiss 880 confocal inverted microscope equipped with Airyscan super-resolution detector. A 40x objective was used with laser excitation at 405 nm, 488 nm, 561 nm and 633 nm.

3.1.9 Infection of HEK293 Flp/In with KO viral particles

WT HEK293 Flp/In cells were seeded at a density of $5x10^4$ cells/well in a 96 well plate in 200 µL DMEM with 10% FBS, no phenol red. After 24 hours, KO viruses were added at an MOI of 10 in 200 µL DMEM with 10% FBS, no phenol red. The cells were incubated at 37°C for 48 hours before washing with PBS, trypsinising and fixing in

4% PFA. Cells were analysed for mCherry fluorescence using a Guava EasyCyte flow cytometer counting 15000 events.

3.2 Molecular Biology

3.2.1 Cloning GFP-MKRN1

The DNA sequence for MKRN1 was amplified from a pcDNA3-FLAG-MKRN1 vector using eGFP-MKRN1 WT Fw and eGFP-MKRN1 WT Rv primers in PCR. The PCR product was separated on a 0.8% agarose gel and the MKRN1 band was cut out and purified using Qiagen QIAquick gel extraction kit. The purified MKRN1 fragment and pcDNA5-FRT-TO-eGFP vector were digested using NotI-HF and KpnI-HF restriction enzymes as per manufacturers instructions. The digested products were separated on a 0.8% agarose gel then purified using the QIAquick gel extraction kit. Purified vector was then dephosphorylated using antarctic phosphatase and 36.2 ng MKRN1 fragment was ligated into 50 ng vector backbone using T4 DNA ligase as per manufacturers instructions. The assembled plasmid was transformed into DH5α *E. coli* and resulting colonies were screened for successful transformation via PCR. After confirmation with sequencing, the pcDNA5-FRT-TO-eGFP-MKRN1 plasmid was prepared using the Qiagen Plasmid Midiprep kit. A total of 2 µg of pcDNA5-FRT-TO-eGFP-MKRN1 plasmid and pOG44 plasmid at a 1:9 ratio was co-transfected into HEK293 Flp/In T-rex cells using X-tremeGENE 9 DNA transfection reagent. Transfected cells were grown in selective DMEM with 50 mg mL^{-1} hygromycin B and 10 mg mL^{-1} blasticidin to select for successfully transfected colonies. Expression of GFP-MKRN1 was then confirmed via western blot and microscopy.

3.2.2 Generation of MKRN1 mutants through site-directed mutagenesis

To generate catalytic mutants of MKRN1 site-directed mutagenesis was used to introduce point mutations in the MKRN1 gene. Primers were designed using NEBaseChanger to target three different codons in MKRN1 H307E, M285E and C281A. A Q5 site-directed mutagenesis kit (NEB) was used to introduce the mutations

into the MKRN1 gene within the pcDNA5-FRT-TO-eGFP-MKRN1 plasmid using the primers SDM_H307E_Fw, SDM_H307E_Rv, SDM_M285E_Fw, SDM_M285E_Rv, SDM_C281A_Fw and SDM_C281A_Fw. This was conducted according to the reaction conditions in Table 15 and 16. After mutagenesis a kinase, ligase and DpnI treatment was applied as in Table 17 and incubated for 5 minutes at room temperature to ligate the mutated plasmid. The assembled plasmids were transformed into DH5α *E. coli* and resulting colonies were sequenced to confirm correct mutation of MKRN1. The mutated MKRN1 constructs were then transfected into HEK293 Flp/In T-rex MKRN1-KO cells as described in the previous method.

Table 15: Components of site-directed mutagenesis

	25 µL reaction	**Final concentration**
Q5 Hot Start High Fidelity 2x master mix	12.5 µL	1x
10 µM Fw primer	1.25 µL	0.5 µM
10 µM Rv primer	1.25 µL	0.5 µM
pcDNA5-FRT-TO-eGFP-MKRN1	1 µL	25 ng
Nuclease free water	9 µL	

Table 16: Site-directed mutagenesis thermocycling conditions

Step	**Temperature (°C)**	**Time**
Initial denaturation	98	30 s
25 cycles	98	10 s
	60	30 s
	72	30 s
Final extension	72	3 min 45 s

3.2.3 Cloning INscarlet tagged HIV-1 and HIV-1 single round viruses

The DNA sequence for the INscarlet construct was cut from a pTWIST-AMP-INT-scarlet vector using AgeI-HF, EcoRI-HF and BglI restriction digest as manufacturers instructions. The HIV-1 and HIV-1 single round backbone HIV-$1_{\text{-R-E}}$ were digested

Table 17: Kinase, ligase and DpnI treatment conditions

	Volume	**Final concentration**
PCR product	1 µL	
2x KLD reaction buffer	5 µL	1x
10x KLD enzyme mix	1 µL	1x
Nuclease free water	3 µL	

from pNL4.3 and pNL4.3-R-E vectors with EcoRI and AgeI. The digested products were separated on a 0.8% agarose gel and then cut and purified using a Qiagen QIAquick gel extraction kit. The purified INscarlet construct was ligated into each backbone using T4 DNA ligase per the manufacturers instructions. Assembled plasmids were transformed into Stbl3 *E. coli* and the resulting colonies were screened for successful transformation via PCR. After confirmation with sequencing, pNL4.3-INscarlet and pNL4.3-R-E-INscarlet plasmid stocks were prepared using the Qiagen Plasmid Maxiprep kit.

3.2.4 Plasmid preparation

Constructs validated by sequencing were transformed into stable competent *E. coli* according to the manufacturer's protocol. The transformed cells were spread onto selective LB agar containing 100 µg mL^{-1} of ampicillin or kanamycin (depending on plasmid resistance gene) and incubated overnight at 37°C. The following day, a single colony was transferred into 250mL LB containing 50 µg mL^{-1} of ampicillin or kanamycin and incubated overnight at 37°C with shaking at 225rpm. 500 µL of culture was mixed 1:1 with 50% (v/v) glycerol to create a bacterial glycerol stock that was stored at -80°C for future 250 mL culture preparations. Plasmid DNA was maxiprepped following the manufacturer's protocol and quantified using a Nanodrop 2000 spectrophotometer. All plasmid stocks were regularly sequenced to maintain sequence integrity.

3.2.5 Fluorescence measurement on Plate Reader

Cells were seeded in triplicate in a black, flat, clear-bottomed 96 well plate in 200 µL RPMI with no phenol red and 10% FBS. The outer wells of the plate were never used for experimental readings and filled with 200 µL PBS to reduce evaporation. Fluorescently labelled pseudotyped lentiviral replicons were added to cells at an MOI of 1, mixed and then immediately placed in a CLARIOstar plate reader. Cells were incubated at 37°C and 5% CO_2 in the plate reader for 72 hours. Fluorescence was measured every 15 minutes at 570nm with a gain of 2200 and a focal height of 5mm.

Given the length of the time series and the possibility of only very small differences between two curves, root mean square relative error (RMSRE) was chosen to determine the statistical difference between fluorescence time series. RMSRE is not dominated by larger error values, which can skew root mean square error (RMSE) and area under curve scores. Root mean square error (RMSE) is calculated as the standard deviation of the residual errors of two series:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i}^{N}(\mathbf{a}_i - \mathbf{b}_i)^2}$$

for two real valued series $\mathbf{a}$ and $\mathbf{b}$ both of length N.

Two identical series $\mathbf{a}$ and $\mathbf{b}$ would have a RMSE of 0. However, given two non-identical series, as they diverge, the residual error increases and RMSE increases.

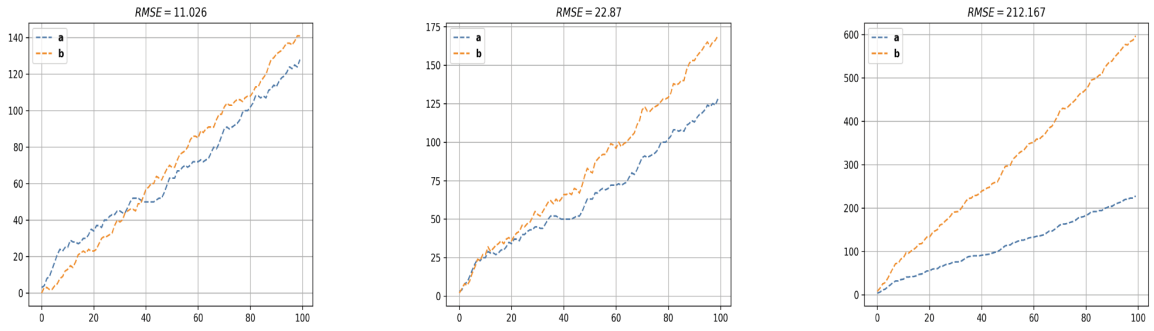

Figure 3.1: As $\mathbf{a}$, $\mathbf{b}$ diverge more, the RMSE increases.

However, if the deviation between each point stays constant, but the series values increase over time, then later (larger) residual errors will dominate earlier values. This

can be problematic for interpreting RMSE values, as larger deviations will skew the result. To avoid this the relative component is added to the error descriptor. By normalising each residual, the error is relative, and larger values will not dominate the error. This can be expressed as the Root Mean Square Relative Error, defined as:

$$RMSRE = \sqrt{\frac{1}{N}\sum_{i}^{N}\left(\frac{\mathbf{a}_i - \mathbf{b}_i}{\mathbf{a}_i}\right)^2}$$

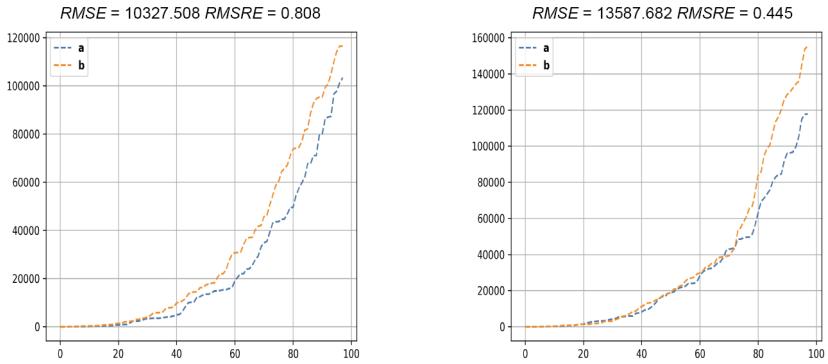

Figure 3.2: RMSE can be dominated by larger values, whereas the normalisation in RMSRE removes this bias.

3.2.6 Western Blotting

Cells to be lysed for western blotting were centrifuged at 400xg for 5 minutes to pellet then washed with PBS and incubated on ice in RIPA lysis buffer for 45 minutes. Protein concentration was measured by reading absorbance at 280nm on a nanodrop. Samples were prepared for loading by adding NuPAGE™4X LDS sample buffer 1:4 and 10mM DTT 1:10 to 30 µg protein and heated at 95°C for 5 minutes. Proteins were separated on NuPAGE™4-12% Bis-Tris 1mm gels at 180v in 1x MOPS buffer. A transblot turbo transfer system was used to transfer gels to nitrocellulose membranes for 12 minutes at 25v. Membranes were blocked in 5% BSA with 0.1% Tween-20 in PBS for 1 hour rocking at room temperature then incubated in primary antibody in PBS 5% BSA 0.1% Tween-20 overnight at 4°C. Membranes were washed in PBS 0.1% Tween-20 then incubated in secondary antibody in PBS 5% BSA 0.1% Tween-20 for 1 hour at room temperature or 4°C overnight, both on a rocker in the dark. Membranes

were then exposed at the relevant wavelengths using a LI-COR Odyssey Fc imaging system.

3.2.7 RT-qPCR

Cells to be lysed for RT-qPCR were centrifuged at 400xg for 5 minutes to pellet then washed with PBS and resuspended in buffer RLT from Qiagen RNeasy mini kit. The kit protocol was then followed to purify RNA and RNA concentration was measured on a nanodrop. For RT-qPCR 20 µL reactions were set up in triplicate with the conditions below using a Luna One-step RT-qPCR kit and cycled on a QuantStudio5 thermocycler.

Table 18: RTqPCR Components

Component	**Final Concentration**
Luna universal one-step reaction mix (2x)	1x
Luna warmstart RT enzyme mix (20x)	1x
Forward primer 10 µM	0.4 µM
Reverse primer 10 µM	0.4 µM
Sample RNA	0.2 µg
Nuclease free water	

Table 19: RTqPCR Cycle Conditions

Cycle step	**Temperature**	**Time**	**Cycles**
Reverse transcription	55	10 min	1
Initial denaturation	95	1 min	1
Denaturation	95	10 sec	40
Extension	60	30 sec	40
Melt curve	60-95	various	1

3.2.8 Flow Cytometry

Cells were washed in PBS then stained with near IR fluorescent reactive dye for 20 minutes at room temperature before being fixed in 4% paraformaldehye in PBS for 30 minutes. Where intracellular antibodies were used, cells were permeabilised for 10 minutes with 0.3% PSBTx then washed in PBS. For the detection of HIV-1 p24 cells were incubated with KC57-FITC antibody at 1:200 dilution in PBS for 15 minutes in the dark. For the application of a live/dead stain cells were incubated with LIVE/DEAD™ fixable Near-IR dead cell stain kit diluted at 1:1000 in PBS for 30 minutes in the dark. Cells were washed in PBS then resuspended in fresh PBS. Flow cytometry was carried out on a Cytoflex counting 30,000 events and measuring fluorescence as described in Table 20

Table 20: Excitation and detection wavelengths of fluorophores used in flow cytometry

Dye/antibody/fluorophore	**Excitation (nm)**	**Detection (nm)**
mCherry	580	610
mScarlet	569	594
KC57-FITC	491	516
LIVE/DEAD™ Near-IR	633/635	780

3.2.9 P24 ELISA

Viral supernatant was inactivated through incubating in PBS 5% empigen BB at 56°C for 1 hour. ELISA plates were coated overnight at room temperature with 100 µg mL^{-1} anti-HIV-1-P24 capture antibody from the Sinobiological p24 ELISA kit (SEK11695). The following day, plates were washed with TBS 0.05% Tween-20 and blocked with ELISA blocking buffer for 30 minutes at room temperature. Serial dilutions of P24 standards and samples were prepared in TBS/E/S buffer and added to the ELISA plate then incubated for 2 hours at 37°C with gentle agitation. The plates were washed 6x and incubated with biotinylated anti-HIV-1-P24 antibody in TMT/SS buffer for 2 hours at room temperature. Plates were washed 6x again and incubated with streptavidin-horseradish peroxidase conjugate in TMT/SS buffer for 1 hour at room temperature

in the dark. Plates were washed 6 times then incubated in TMB until the highest concentration P24 standard developed a dark blue colour. The reaction was stopped with 0.5 M H_2SO_4 and absorbance was measured at 450 nm and 570 nm using a SpectraMaxM5 plate reader. Absorbance readings were analysed using Microsoft Excel and GraphPad Prism.

For determination of P24 concentration of virus purified from transfected HEK293T and Flp/In MKRN1-KO cell supernatant, the R&D systems HIV Gag P24 DuoSet ELISA kit (DY7360-05) was used as per the manufacturer's instructions. Prior to the ELISA, sucrose purified virus samples were inactivated with 1% Triton-X100 for 1 hour on ice then diluted 1:1000 in reagent diluent buffer from the kit.

3.2.10 MKRN1-KO virus release assay

HEK293T WT and MKRN1-KO cells or HEK293 Flp/In WT and MKRN1-KO cells were seeded at a density of $1x10^7$ cells per 150 mm^2 dish in 20 mL DMEM without phenol red with 10% FBS, 1 mM sodium pyruvate and 1X glutamax. The following day cells were transfected using X-tremeGENE 360 transfection reagent. A reaction mix of 2 mL Opti-mem, 20 µg plasmid DNA and 60 µL X-tremeGENE 360 was prepared at room temperature and incubated for 15 minutes before being added dropwise to the cells. For GFP controls, 20 µg of pcDNA5-eGFP-linker plasmid was used. For HIV-$1_{\text{-R-E-Nef-mCherry}}$ the pNL4.3-R-E-T2A-Nef-mCherry plasmid was used at a 4:1 ratio with pHEF-VSVG. After 48 h the cells were imaged for brightfield, GFP and mCherry fluorescence using an EVOS M7000 microscope.

Supernatant was centrifuged at 400xg for 5 minutes to remove cells then filtered through a 0.45 µm syringe filter. Small aliquots were set aside for later analysis whilst the remaining filtered supernatant was gently layered on top of a 10% sucrose cushion. This was centrifuged at 10,000xg for 4 h at 4°C with minimal acceleration and deceleration. Purified viral pellets were then resuspended in 500 mL PBS and aliquoted before snap freezing on dry ice and storing at -80°C. P24 concentration of virus stock was determined using a P24 ELISA.

Transfected cells were trypsinised and the cell pellet was washed twice with PBS.

Cell pellets were stored at -80°C or immediately lysed in 50 µL 1x RIPA buffer with 1:1000 AEBSF and 1:1000 benzonase for 45 minutes on ice. For western blot analysis, lysates were prepared as previously described. Supernatant to be analysed by western blot was prepared by adding NuPAGE™4X LDS sample buffer 1:4 and 10mM DTT 1:10 and heating at 95°C for 5 minutes.

3.2.11 Tzm-bl Luciferase Assay

Viral supernatant was titrated in duplicate in 1:3 dilutions in complete DMEM, starting with undiluted, to a final volume of 100 µL/well of a 96 well sterile flat-bottomed plate. 10,000 Tzm-bl cells were added to each well in 100 µL complete DMEM. Cells were incubated at 37°C for 30-36 hours then washed with PBS and lysed in 1x Glo lysis buffer for 2 minutes at room temperature. The plates were then wrapped in parafilm and frozen at -80°C then rethawed to ensure complete cell lysis. Bright Glo luciferase substrate was added to the lysed cells 1:1 and incubated at room temperature for 2 minutes. Luminescence was read across all wavelengths with 1000ms integration time on a SpectraMaxM5 plate reader. Analysis of luminescence readings was conducted in Microsoft Excel and infectious dose 50% (ID_{50} calculated using the National Institutes of Health ID-50 programme [144].

3.2.12 Synthesis of fluorescently labelled probes for smFISH

32 probes for smFISH were designed using Stellaris Probe Designer with the following parameters: organism, human; masking level, 5; oligo length, 20 nt; minimum spacing length, 3 nt. The 3' region of HIV-1 pNL4-3 reference sequence was used as a target sequence. Oligonucleotides were singly labelled with Alexa Fluor 647 at 3' ends according to a published protocol [145] and were concentration normalised to 25µM. All probe sets used in this study had degree of labelling $>$0.94.

3.2.13 smFISH

Cells were seeded on sterile glass coverslips at a density of 2x10^{4}cells/well in 24 well plates in 500 µL of DMEM with 10% FBS and no phenol red. Cells were incubated overnight at 37°C and 5% CO_2. The following day, the media was removed from the cells and replaced with 500 µL fresh media with HIV-1-Gag-mCherry$_{rep}$ to give an MOI of 1. At 24, 48 and 72 hpi, cells adhered to coverslips were washed with PBS and fixed in 4% PFA for 30 minutes at room temperature. Cells were washed again before permeabilisation with 0.1% Triton-X100 in PBS (PBSTx) for 10 minutes at room temperature. Cells were washed in Stellaris RNA FISH buffer A then 125 nM Stellaris FISH probes were added in Stellaris RNA FISH hybridization buffer. The plate was placed inside a humidified chamber and incubated overnight at 37°C. After hybridization, cells were washed with Stellaris RNA FISH buffer A for 30 minutes at 37°C then incubated in 2 µg mL^{-1} DAPI and a cytoplasm stain (Phalloidin-Alex Fluor 488 was used for the first repeat, CellMask Green for the second and third) in Stellaris RNA FISH buffer A for 30 minutes at 37°C in the dark. The cells were washed for a final time in Stellaris RNA FISH buffer B for 5 minutes at room temperature. Coverslips were then mounted onto glass slides using SlowFade Diamond antifade mountant and sealed with clear nail varnish. Slides were left to dry then stored at -20°C before imaging.

Imaging for smFISH was carried out on an Olympus SpinSR10 spinning disk confocal system equipped with Prime BSI and Prime 95B sCMOS cameras. Objectives used were ×100 oil and manual image acquisition was performed with Olympus cellSens Dimension software. Images were uploaded and stored in the University of Oxford OMERO server [140], and OMERO.figure (3.2.0) was used to generate presented image visualisations.

3.2.14 smFISH analysis

Cell segmentation was performed manually in ImageJ using 2D maximum intensity projected images of phalloidin or CellMask Green stains. Total number of cells per image was counted using a custom ImageJ macro script written by Dr Jeff Lee

(Davis lab, Department of Biochemistry, University of Oxford). Single-molecule-level quantification of smFISH images was performed by Dr Jeff Lee using a custom Python pipeline using Bigfish [146], skimage, and numpy libraries. Tif files were converted to a numpy array, and individual cells were segmented as described above. Images where cells were labelled with the CellMask stain were pre-processed with a median filter, radius = 50. Background signal in the smFISH channel was subtracted using ImageJ. Threshold setting for smFISH spot detection was set specifically for each set of images collected in each session. ImageJ was also used to manually draw linear cross-sections on HIV-1 infected cells and capture the signal intensity of the line profile for each channel.

3.2.15 Imaging of cell-free viral particles

Viral particles were produced in HEK293T cells and collected from cell supernatant at 48 hpi. Poly-L-lysine coated coverslips were found to generate high background during imaging therefore uncoated sterile coverslips were used. 20 µL of virus concentrate was incubated in 500 µL PBS on glass coverslips at room temperature for 30 minutes. The viral suspension was removed and the coverslips were gently washed with PBS before fixing for 20 minutes in 4% paraformaldehyde in PBS. Viral particles were then washed in 50 mM glycine in PBS to quench unreacted aldehyde groups. 5% BSA was used to block the coverslips and then viral particles were incubated with 2G12 anti-HIV-Env at 1:1000 in 5% BSA for 1 hour. Viral particles were then washed before incubating in anti-human-IgG 488 at 1:10000 for 1 hour. Coverslips were mounted on glass slides using VECTASHIELD hardset antifade mounting medium. Images were captured on an Olympus SoRa spinning disc microscope at 60x magnification.

3.2.16 Immunofluorescence

HeLa-CD4-GFP cells were seeded on sterilised high precision coverslips, induced with 1 µg µL^{-1} doxycycline and incubated for 24 hours at 37°C. They were then infected at the relevant MOI and induced again with 1 µg µL^{-1} doxycycline. At the relevant time point, cells were washed then fixed using 4% paraformaldehye, 30 µM in PBS for 30

minutes. Cells were washed then permeabilised using 0.3% PBSTx for 10 minutes. PBS with 5% BSA was used to saturate cells for 30 minutes at room temperature then cells were incubated with primary antibodies in 5% BSA in 0.003% PBSTx in the dark for 1 hour. Cells were washed 3x in 0.3% PBSTx then incubated with secondary antibodies in 5% BSA in 0.003% PBSTx in the dark for 1 hour. Cells were washed again then incubated with DAPI in 5% BSA in 0.003% PBSTx for 10 minutes. A final wash was conducted and coverslips were then mounted on glass slides using VECTASHIELD hardset antifade mounting medium. Cells were imaged on a Zeiss 880 Airyscan super-resolution confocal microscope at 100x magnification.

3.3 Proteomics

3.3.1 Co-immunoprecipitation

HEK293 Flp/In WT, GFP and GFP-MKRN1 cells were seeded at a density of $4x10^6$ cells/dish in 10 cm dishes with 10 mL of DMEM with 10% FBS. HEK293 Flp/In GFP-MKRN1 cells were induced with 10 µL of 1 mg mL^{-1} doxycycline for 24 hours. Cells were incubated overnight at 37°C and 5% CO_2 before infection with HIV-1-mCherry-Nef_{rep} at an MOI of 1. GFP-MKRN1 cells were induced for a second time and all cells were incubated for another 24 hours at 37°C and 5% CO_2. HEK293 Flp/In GFP cells were induced for only 2 hours due to the strength of GFP expression.

For collection of cells, the plates were washed twice with 5 mL ice cold PBS then lysed in 1 mL 1% NP40 lysis buffer with complete protease inhibitor for 30 minutes on ice. One set of plates were also lysed in the presence of 25U benzonase to degrade all DNA and RNA from the samples. The cell lysate was scraped from plates and centrifuged at 800 xg for 5 minutes to pellet the nuclear fraction. The lysate was then centrifuged again at 10000 xg for 5 minutes to pellet the mitochondrial fraction. The cytoplasmic fraction was collected and stored at -80°C until immunoprecipitation was performed.

For the immunoprecipitation (IP), all incubation steps were performed at 4°C rotating end-over-end and all centrifugation was carried out for 5 minutes at 2500 xg

at 4°C. GFP-Trap and control binding agarose beads were equilibrated in 1% NP40 lysis buffer with complete protease inhibitor as per the Chromotek GFP-Trap protocol. An aliquot of each lysate was stored at -20°C for later analysis before pre-clearing the remaining lysate using control binding beads for 10 minutes. An aliquot of pre-cleared lysate was stored at -20°C for later analysis and the remaining lysate was incubated with 20 µL GFP-Trap agarose beads for 1 hour at 4°C. The flowthrough was collected and stored at -20°C for later analysis. The beads were washed twice with 1% NP40 lysis buffer, four times with 0.5% NP40 lysis buffer and twice more with IP wash buffer. In the final washing step, the beads were transferred to a new tube. Proteins were eluted from the beads using 25 µL 200 mM glycine at pH 2.5 with forecful pipetting for 60 seconds before centrifuging to sediment the beads. The eluted fraction was then immediately neutralised with 2.5 µL 1 M Tris pH 10.4. The elution step was repeated for a second time and IP eluates were stored at -20°C.

3.3.2 Silver Staining

Samples were denatured with NuPAGE 4x LDS sample buffer and 10 mM DTT at 95°C for 5 minutes. Denatured samples were separated on 4-12% Bis-Tris gels then silver staining was performed following the Invitrogen SilverQuest staining kit protocol.

3.3.3 Sample Preparation for Mass Spectrometry

IP eluates were prepared for mass spectrometry analysis by trypsin digestion followed by desalting. Samples were denatured using 4 M urea in 100 mM ammonium bicarbonate buffer pH 7.8 for 10 minutes shaking at 650 rpm at room temperature. Cysteine reduction was conducted with 10 mM Tris(2-carboxyethyl)phosphine for 30 minutes at room temperature. Following this, cysteines were alkylated with 50 mM 2-chloroacetamide prepared in 8 M urea and 100 mM ammonium bicarbonate buffer for 30 minutes at room temperature. Samples were then predigested using LysC at 1 µg per 100 µg of protein for 2 h at 37°C shaking at 800 rpm. Urea concentration was then diluted to 2 M in 100 mM ammonium bicarbonate buffer and calcium chloride was added to a final concentration of 2 mM. Trypsin was added at a concentration of 1 µg

per 40 µg of protein and incubated at 37°C shaking at 800 rpm for 16 h. Trypsination was stopped using 5% formic acid and the samples were centrifuged for 30 minutes at 14000 rpm at 4°C to remove aggregates. The following centrifugation steps were all conducted at 4000 rpm at room temperature. C18 resin was activated using 100% acetonitrile, centrifuged to remove liquid, washed in 0.1% trifluoroacetic acid and centrifuged to remove liquid again. The trypsin-digested samples were loaded onto the C18 resin by centrifugation and washed twice in 0.1% trifluoroacetic acid. Peptides were eluted with 50% acetonitrile in 0.1% trifluoroacetic acid by centrifugation at 4000 rpm then 8000 rpm for 2 minutes each at room temperature. Peptides were dried in protein lobind eppendorf tubes and stored at -20°C before injection.

3.3.4 Mass spectrometry

Label-free liquid chromatography-mass spectrometry was conducted by the Advanced Proteomics Facility at the University of Oxford. MaxQuant protein identification and quantification software was used to identify proteins present in the samples. Only proteins identified by MaxQuant with a FDR$<$1% were included in the analysis. NSAF100 scores were calculated for each protein and analysed in Excel.

3.3.5 GO terms/STRING analysis

To generate interaction networks protein IDs were imported into string-db [142] and analyses were based on experiments and databases with a confidence score of at least 0.4. These networks were downloaded and arranged in Cytoscape. String-db was also used to perform GO term analyses and KEGG pathway analyses.

3.3.6 RBD analysis

Protein structures were downloaded from the Protein Databank or AlphaFold if no structure had been published. Structures were annoated in Pymol with known RBDs described in the Protein Databank and with predicted RNA-binding sites based on BindUP prediction and RBDetect.

4 Identifying host RBPs that play key roles in the early stages of HIV-1 infection

4.1 Introduction

Several RBPs have been identified in the proteome of the HIV-1 viral particle. Whilst we know RBPs are important in the intracellular steps of the viral life cycle, we know little about their potential functionality (if any) of RBPs inside the viral particles. However, a few RBPs such as APOBEC, MOV10 and UPF1 have been reported to regulate the initial steps of infection, particularly reverse transcription [99, 106, 147]. Because HIV virions incorporate a portion of the cytoplasm during budding from the plasma membrane, it is possible that many cellular proteins are passively incorporated into virion and lack functional roles. To increase the likelihood of identifying proteins with specific functions in regulating the HIV-1 life cycle, the Castello lab developed a new approach, referred to as in-virion RNA Interactome Capture (ivRIC) that isolates and identifies the RBPs that are actively interacting with the gRNA inside viral particles [135].

4.1.1 The HIV-1 in-virion RNA interactome

To define the HIV-1 gRNA interactome in viral particles, Dr Manuel Garcia-Moreno (Castello Lab) built on the previously established RNA interactome capture (RIC) method [15, 148] and adapted it to viral particles purified from the supernatant of infected CD4+ T lymphocytes (SupT1). An identical protocol is applied in parallel to uninfected cell supernatants to control for cellular vesicles with virus-like sizes (eg exosomes). These controls are subject to the same sucrose density gradient purification as the viral particles. Next UV light at 254 nm is applied to the purified particles to induce RNA-to-protein crosslinking at 'zero distances' ($<$2Å) [149]. A non-irradiated control is also included to ensure that only UV-dependent interactions are considered in the analysis. UV irradiation was used as opposed to formaldehyde crosslinking as it is irreversible and has a much higher specificity. UV crosslinking

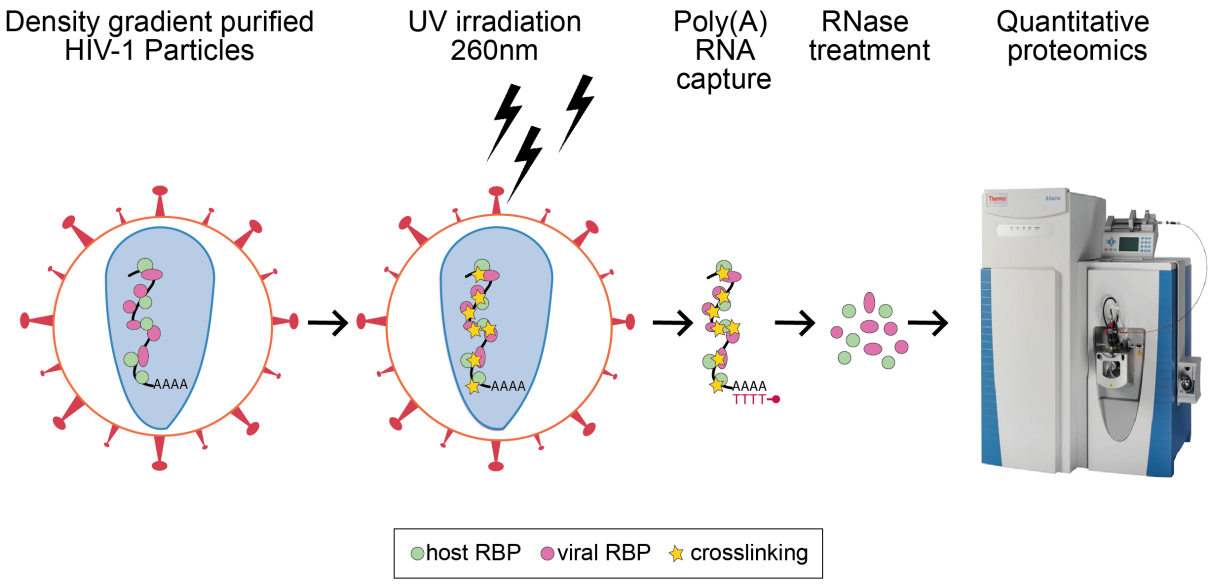

Figure 4.1: **Schematic of the ivRIC workflow**

induces nucleotide-protein crosslinking and does not promote detectable protein-protein crosslinks. It thus leads to very specific and irreversible immobilisation of proteins in direct contact with the RNA. The downside of UV crosslinking is its low efficiency when compared to formaldehyde. The use of UV crosslinking is thus more prone to false negatives while minimising the occurrence of false positives. I am of the opinion that having lower number of candidates of high quality is better than larger datasets populated by a high number of irrelevant proteins. The crosslinked particles are lysed under denaturing conditions and the proteins crosslinked to the polyadenylated vRNA are then isolated using oligo(dT) magnetic beads. Stringent washing is conducted to ensure all non-covalent binders are efficiently removed [35, 148]. The eluates are then treated with RNases to degrade the vRNA and the proteins are analysed by quantitative proteomics (Figure 4.1).

Garcia-Moreno et al identified a total of 104 RBPs bound to the vRNA in virions (referred to as in virion (iv)RBPs) in an infection and crosslink-dependent manner (Figure 4.2A and B). Few exosomal proteins were identified in both infected and mock-infected controls with very low abundance. However, only proteins significantly enriched in the UV crosslinked and HIV-infected samples over the non-irradiated and mock infected samples were classified as ivRBPs. Of the 104 ivRBPs, 87% were annotated with the Gene Ontology term ”RNA-binding” (GO:0003723). This suggests

that bonafide RBPs were successfully isolated, amongst which novel RNA-interactors were also uncovered. Analysis of the ivRIC inputs (i.e. the proteome of the viral particles prior to RNA isolation) and eluates by quantitative proteomics showed that a few ivRBPs were enriched in the inputs, suggesting that they establish short-lived or low-affinity interactions with the vRNA (Figure 4.2C). By contrast, most ivRBPs were enriched in the eluates over the inputs reflecting long-lived, geometrically optimal interactions with the vRNA. The pool of identified ivRBPs were also enriched in ivRIC eluates over the whole proteome of CD4+ T lymphocytes (Figure 4.2D). This was established through comparison of protein intensity (determined by MaxQuant) within ivRIC eluates compared to protein intensity of whole cell lysates collected from the infected lymphocytes. This indicated that ivRBPs were not incorporated into viral particles because they were highly abundant in the cell cytoplasm, but likely because of specific incorporation mechanisms remaining to be defined [135]. Interestingly, a large proportion of the ivRBPs are nuclear while viral particle formation occurs at the plasma membrane. This suggests that the concentration of these proteins at the assembly sites will be low. A paper including all of these discoveries, on which I am an author, is currently deposited in BioRxiv and can be read in the following link: Garcia-Moreno et al, 2023.

To validate these results, I selected nine ivRBPs (indicated in Figure 4.2) to assess for functionality in HIV-1 infection. These were chosen with the following reasoning:1) they must exhibit high 'crosslink-ability' to HIV-1 RNA, since UV crosslinking favours long-lived protein-RNA interactions these ivRBPs are likely to establish strong and specific interactions with vRNA. 2) These RBPs must be enriched in ivRIC eluates over the proteome of the producer cell, indicating they are actively incorporated into virions and are more likely to have a specific function in HIV-1 particles. 3) They are all expressed in primary T CD4+ lymphocytes and macrophages, hence they hold physiological relevance. A deep review of the literature was performed to further select candidates based on hypothesis-driven choices.

Initially, I included SUB1 (aka positive cofactor 4, PC4) and ZC3HAV1 (aka zinc finger anti-viral protein, ZAP) in my pool of ivRBP candidates. However,

SUB1-KO lymphocytes were only viable for short periods of time which is not conducive to the length of experiments required to study HIV-1 infection. Additionally, ZC3HAV1 KO could only be achieved for one of the two known isoforms at a time in CD4+ T lymphocytic lines. Since ZC3HAV1 isoform functions are (at least partially) overlapping, ZC3HAV1-partial KOs were excluded from further analysis. The remaining ivRBPs, MKRN1, FAM120A, LSM14A, MSN, LCK, CIRBP and YBX3 are introduced in the following subsections.

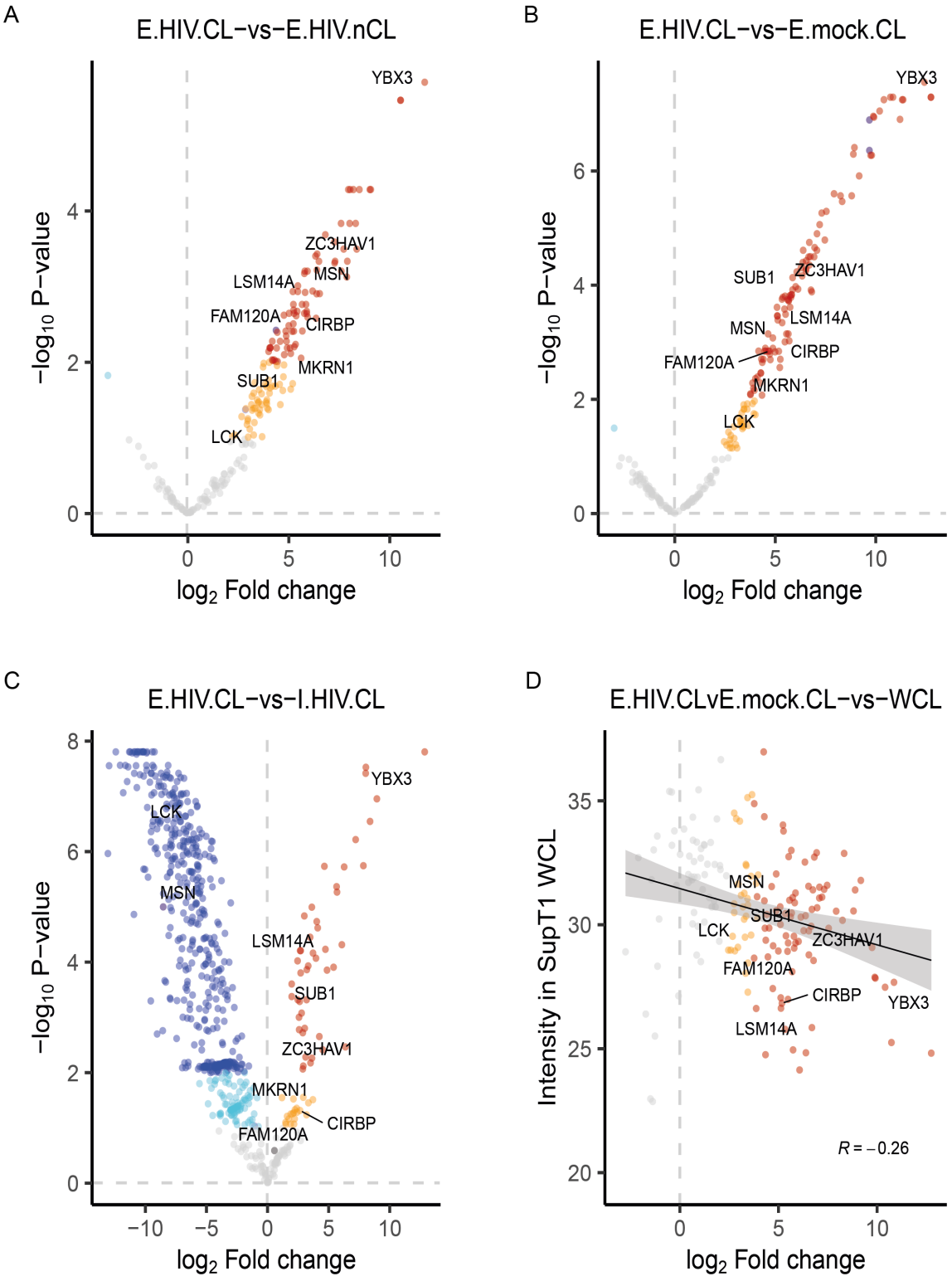

Figure 4.2: **ivRBPs identified by ivRIC [135]. A.** Volcano plot showing the log2 fold change and the significance (p value) of each RBP between crosslinked viral particles and non-crosslinked viral particles using data from three biological replicates. False discovery rate (FDR) <1% shown in red, <10% shown in yellow.**B.** Same as in **A** but for 72hpi HIV-1 infected and uninfected conditions. **C.** Same as in **A** but for crosslinked ivRIC eluates vs crosslinked ivRIC inputs. FDR <1% shown in blue and <10% shown in cyan for inputs. **D.** Scatterplot comparing the intensity of each RBP in the ivRIC eluate and whole cell lysate of lymphocytes.

4.1.2 MKRN1

Makorin ring finger protein 1 (MKRN1) is an E3 ubiquitin ligase belonging to a highly evolutionarily conserved family of four paralogs (MKRN1-4) [150]. It has RING ligase domain and four CCCH ZnF domains which are thought to provide its RNA binding ability. MKRN1 has mainly been described in metabolism [151], as a cell cycle regulator involved in the control of p53 and p21 [152] and is therefore implicated in multiple cancers [153, 154]. However, it has also been reported to regulate telomere length [155, 156] and RNA polymerase II transcription [157].

No structure has been established for MKRN1 but the Alphafold prediction shows a compact core formed by the ZnF domains and multiple disordered regions (Figure 4.3). Interestingly, BindUP prediction of the positive patches on the protein surface does not align with the canonical RBDs and RBDetect prediction based on amino acid sequence only highlights two of the ZnF domains. This uncertainty in RBDs perhaps reflects how little is known about the RNA-binding activity of MKRN1. A study in mouse embryonic stem cells revealed MKRN1 as an interactor with hundreds of cellular mRNAs and multiple RBPs including poly(A)-binding protein (PABP) C1 [158], which was further corroborated in HEK293 cells [159]. Hildebrant et al showed MKRN1 is a novel factor in the ribosome-associated quality control complex. MKRN1 is recruited to A-rich sequences in mRNA in a PABPC1-dependent manner to prevent the production of erroneous proteins arising from prematurely polyadenylated transcripts, thereby maintaining proteome integrity [160]. However, none of the literature has revealed how MKRN1 interacts with RNA.

Whilst MKRN1 has never been connected with HIV-1 infection, it has been implicated in viral and bacterial infection. In both porcine circovirus and West Nile virus (WNV) infection, MKRN1 ubiquitinates the viral capsid proteins, causing premature degradation of the capsid and inhibition of viral replication [161, 162]. Additionally, MKRN1 was found to ubiquitinate whole *Mycobacterium tuberculosis* in vitro suggesting it may play a role in ubiquitin-mediated autophagy [163]. Interestingly, the precursor protein VII of human adenoviruses has been shown to bind MKRN1 and enhance self-ubiquitination of MKRN1, leading to its proteasomal degradation. It

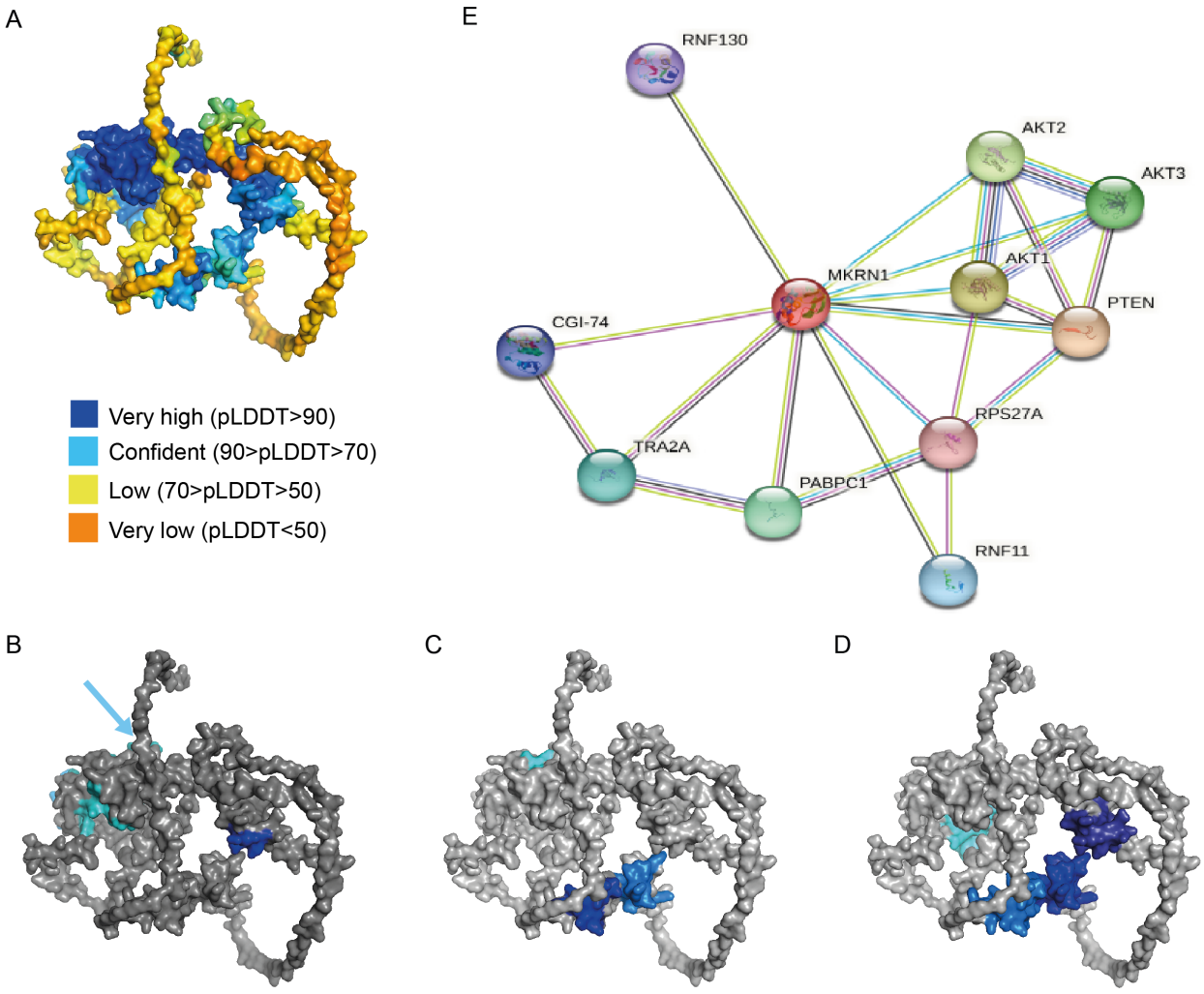

Figure 4.3: **MKRN1 predicted structure and RBDs.A.** Alphafold predicted structure of MKRN1 with confidence indicated by colour (AF-Q9UHC7). **B.** Blue regions indicate the top three positive patches on the surface of MKRN1 predicted by BindUP. **C.** RBDetect prediction of RBDs based on amino acid sequence, shown in blue. **D.** Canonical RBDs identified by uniprot, shown in blue. **E.** STRING analysis of MKRN1 interacting proteins. The colours of the connecting lines indicate the type of interaction: teal is a known interaction from a curated database; pink is an experimentally determined interaction; green is predicted based on gene neighbourhood; red is predicted on gene fusion; blue is predicted from gene co-occurence; yellow predicted from textmining; black predicted from co-expression and lilac predicted from protein homology.

was hypothesised that this is an immune evasion mechanism designed to circumvent MKRN1-driven degradation [164]. Each of these studies focused on the ubiquitin ligase function of MKRN1 and did not investigate the role of RNA-binding activity. It is plausible to speculate that the RBDs would mediate the interaction of MKRN1 with the RNAs of these viruses. However, this has not been tested and is not clear how this would occur for encapsidated vRNAs. The authors showed that the C terminus of MKRN1 was necessary for its effect in infection. Both the N and C terminus of MKRN1 contains classical RBDs that are known to mediate RNA binding nearby the poly(A) tail of mRNAs. Thus, it is not possible to rule out the need of RNA binding for MKRN1 regulatory role in infection.

It was also shown that MKRN1 expression is decreased in measles virus and vesicular stomatitis virus infected cells but not HIV-1, hepatitis C (HCV) or hepatitis B virus infection. It should be noted, however, that each of these infections were studied in different reporter cancer cell lines. MKRN1 also has been related to the inhibition of apoptosis during classical swine fever virus infection through interaction with the viral non-structural protein 4B [165], but no mechanism has been established. Additionally, MKRN1 may not have been shown to interact with HIV-1 but other E3 ubiquitin ligases are known to initiate degradation of the HIV-1 capsid. In particular, the tripartite motif protein TRIM5 assembles in a hexameric lattice on retroviral capsids and builds an N terminal ubiquitin chain which triggers both innate immune signalling and proteasomal degradation of the capsid [166].

The STRING analysis of MKRN1 (Figure 4.3) unsurprisingly shows interaction with several cell cycle regulators as well as PABPC1 and RPS27A, a Ubiquitin-40S ribosomal protein. STRING is a biological database that builds networks of both known and predicted protein-protein interactions collected from online databases and high-throughput screening. The colours of connecting lines between proteins highlight the source from which the interaction has been determined eg pink indicates proteins that have been shown to interact experimentally. The limitations of STRING analyses are the fact that some of the lines of evidence shown do not involve direct protein-protein interactions but indirect observations such as co-occurrence. Moreover, the

experimental data feeding the database is not always at the same quality standards. Finally, the observed interactions may be limited to specific cell types or physiological conditions. STRING networks are, therefore, a hypothesis building tool as one can make use of high-confidence interactions to raise questions about the protein's function. STRING networks should always be interpreted against this background. Interestingly, we identify both PABPC1 and RPS27A in our HIV-1 ivRBPome. Cassar et al and Hildebrandt et al also identified a pool of MKRN1 interacting proteins which are not included in the STRING analysis [158, 160]. Of these interactors, ELAVL1, hnRNPD, IGF2BP3, MOV10, PABPC1, PABPC4, SYNCRIP, UPF1, YBX1 and YBX3 are also found in our ivRBPome. Additionally, ELAVL1, hnRNPD, IGF2BP3, MOV10, PABPC1, PABPC4 and SYNCRIP were specifically shown to be ubiquitination targets of MKRN1, as well as another ivRBPome protein, FAM120A [160]. This, together with our results, indicates that MKRN1 could be involved in an RNP complex that is packaged into HIV-1 virions. These ivRBPS are all strongly enriched in the ivRBPome over whole cell lysate (appendix 1, Figure 8.2B) which supports the idea of a complex being incorporated into viral particles. MKRN1 does not actually feature on this plot because it is expressed at such a low level in SupT1 cells that it was not identified in the WCL. This makes it particularly intriguing that we find MKRN1 inside the viral particles.

The interaction of MKRN1 in multiple infections and with multiple proteins in the ivRBPome makes it an interesting candidate to study in HIV-1 infection. In addition to this, it is well established that HIV-1 hijacks the ubiquitin-proteasome to aid replication and that multiple HIV-1 proteins are ubiquitinated [167–170]. Ubiquitin is also found inside HIV-1 viral particles [171] and we identify it in our ivRBPome. Because of this, I hypothesise that MKRN1 could be doing several things in HIV-1 infection: 1) recognising polyadenylated vRNA and regulating viral protein translation; 2) ubiquitinating cellular RBPs within the capsid with roles in the early steps of the HIV-1 life cycle and/or 3) ubiquitinating HIV-1 capsid (or other viral protein). Regarding the third suggestion, MKRN1 associates with cellular mRNAs nearby the poly(A) tail. The known substrates of MKRN1 are RNA-binding proteins suggesting that its RNA-

binding activity recruits the E3 ubiquitin ligase to the 'right environment' [160]. The interaction of MKRN1 with HIV-1 RNA can potentially place its E3 ubiquitin ligase at close range from capsid molecules (as part of the Gag polyprotein) that nucleate on the viral RNA at the assembly sites, leading to incorporation of MKRN1 into HIV-1 particles.

4.1.3 FAM120A

The constitutive coactivator of PPARγ-like protein 1, FAM120A, also known as oxidative stress-associated src activator (Ossa), is a poorly characterised cytoplasmic protein that has been suggested as an activator of several intracellular kinases. Following oxidative stress, FAM120A has been shown to activate src family kinases and acts as a scaffolding protein for src family kinases to phosphorylate and activate phosphatidylinositol-3-kinase [172].

No structure has been published for FAM120A. However, Alphafold prediction (Figure 4.4) suggests it has a globular N terminus and central region, with a highly disordered C terminus. Whilst FAM120A does not have any canonical RBDs, the globular centre has a positively charged surface. Additionally, the highly disordered C terminus has been experimentally shown to bind RNAs such as insulin-like growth factor II (IGF-II) mRNA and promotes extracellular secretion of IGF-II [173].

Whilst no direct relationship has been described between FAM120A and viral infection, it is involved in several processes required for HIV-1 infection. Kelly et al revealed FAM120A interacts with the micro (m)RNA silencing protein argonaute 2 (Ago2) and binds to homopolymeric tracts in the 3 UTRs of approximately 2000 mRNAs. Interestingly, one third of mRNAs bound by Ago2 were co-bound by FAM120A suggesting they may act together in miRNA-mediated silencing pathways [174]. Activation of CD4+ T lymphocytes is essential for productive HIV-1 infection and requires substantial transcriptional changes within the cell, including miRNAs. Several miRNAs shown to be downregulated upon T cell activation are anti-HIV miRNAs that directly target HIV-1 RNA [175, 176] or inhibit cellular co-factors of infection [177–179]. Cellular miRNAs that downregulate host anti-viral proteins during HIV-1 infection have

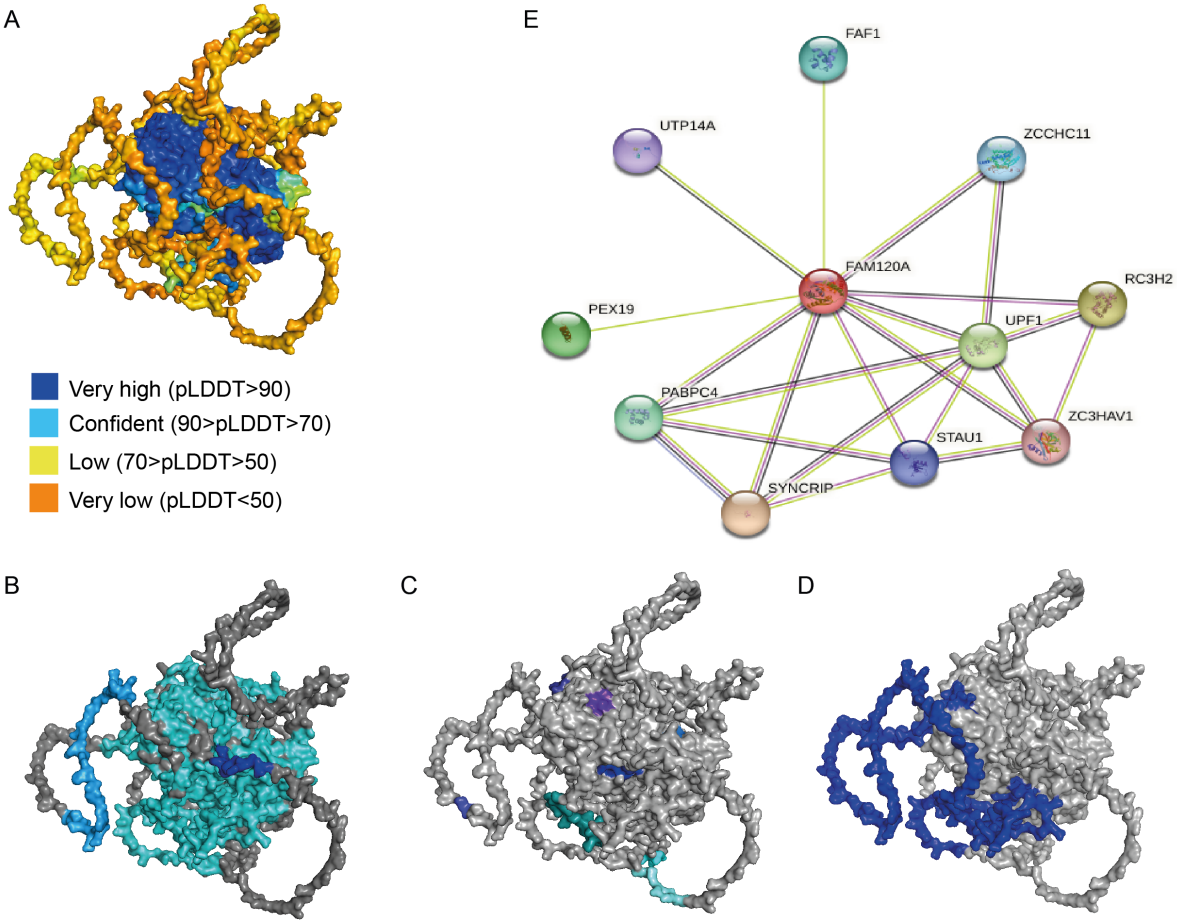

Figure 4.4: **FAM120A predicted structure and RBDs. A.** Alphafold predicted structure of FAM120A with confidence indicated by colour (AF-Q9NZB2-F1). **B.** Blue regions indicate the top three positive patches on the surface of FAM120A predicted by BindUP. **C.** RBDetect prediction of RBDs based on amino acid sequence, shown in blue. **D.** RBD as identified by Tanaka et al 2009, shown in blue. **E.** STRING analysis of FAM120A interacting proteins. The colours of the connecting lines indicate the type of interaction: teal is a known interaction from a curated database; pink is an experimentally determined interaction; green is predicted based on gene neighbourhood; red is predicted on gene fusion; blue is predicted from gene co-occurence; yellow predicted from textmining; black predicted from co-expression and lilac predicted from protein homology.

also been reported [180,181]. HIV-1 could potentially be recruiting complexes involving FAM120A to control cellular gene expression through miRNA-mediated silencing pathways.

FAM120A has also been implicated in cargo movement. Bartolome et al found FAM120A associates with the motor protein kinesin to modulate cell-surface trafficking of the interleukin (IL)-13 receptor [172]. HIV-1 has been shown to engage kinesin and other microtubule-associated motor proteins to promote trafficking of the capsid to the nucleus [62, 64, 65] and trafficking of the viral RNP to the plasma membrane [182]. Additionally, STRING analysis revealed interaction of FAM120A with STAU1, a core member of the HIV-1 RNP [129], as well as three other RBPs identified in our ivRBPome: SYNCRIP, PABPC4, UPF1 and ZC3HAV1. Thus, FAM120A may be involved in trafficking of the HIV-1 RNP.

4.1.4 LSM14A

LSM14A (also known as RAP55) is a member of the like-Sm (LSm) family of proteins, named because of their similarity to the Smith (Sm) antigen, which systemic lupus erythematosus patients commonly produce antibodies against [183]. These proteins are defined by their three-dimensional (3D) structure as they form characteristic rings of six or seven subunits. However, their function is not well characterised and each family member appears to play distinct roles in RNA metabolism [184].

The structure of each LSm subunit is shown in Figure 4.5, the Sm domain being shown specifically in 4.5B. Despite having known RNA-binding activity, LSM14A lacks canonical RBDs. It has several disordered regions which align with predicted RBDs from BindUP and RBDetect analysis (Figure 4.5C and D). Experimentally, the C terminal of LSM14A has been shown to bind RNA [185].

LSM14A was originally identified as a major component of processing (P)-bodies which are discrete cytoplasmic structures linked to storage and mRNA degradation. P-bodies typically contain the decapping enzymes DCP1a and DCP2, the exoribonuclease XRN1 and the DEAD-box family helicase Rck/p54. The size and number of these complexes fluctuates with the demand for mRNA degradation [185].

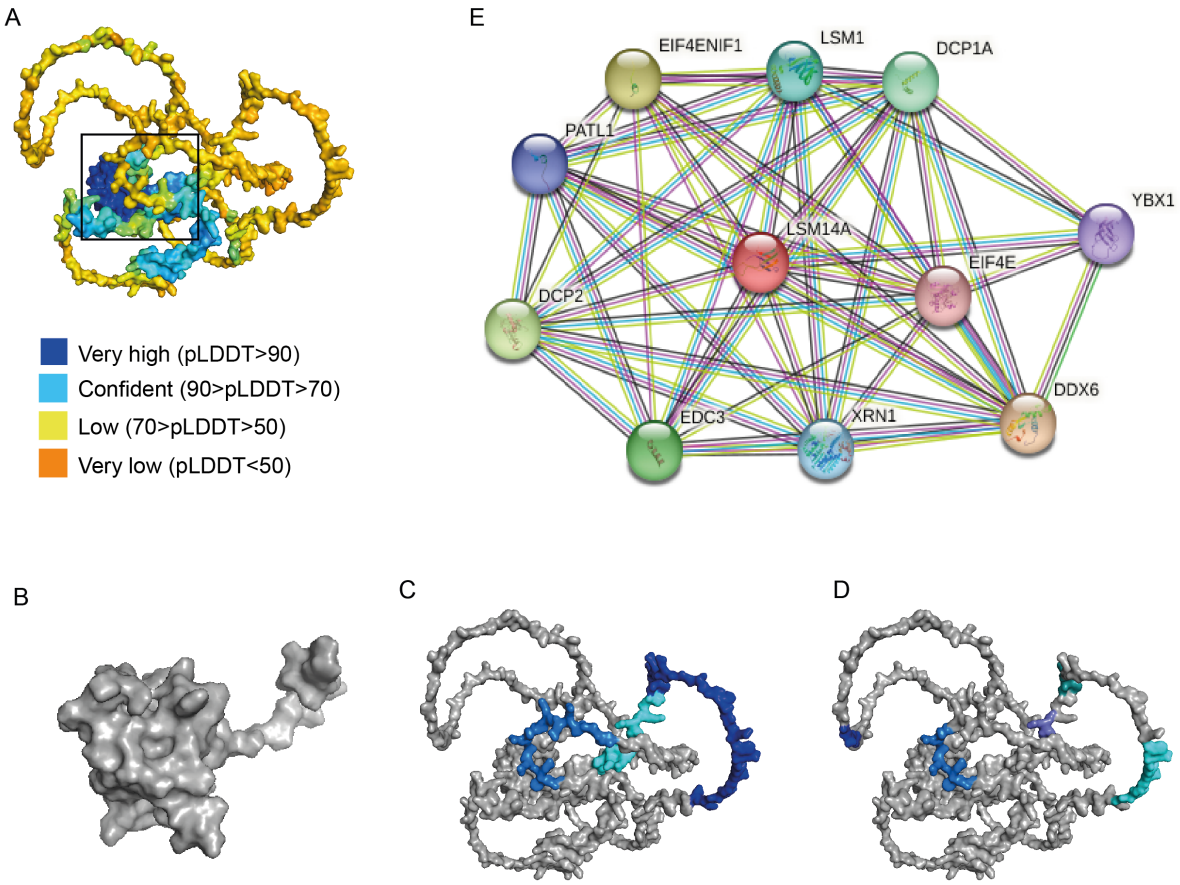

Figure 4.5: **LSM14A predicted structure and RBDs. A.** Alphafold predicted structure of LSM14A with confidence indicated by colour (AF-Q8ND56-F1). **B.** 2.62Å resolution Xray structure of amino acids 1-84 (PDB 6FYW), relation to whole protein indicated by black box in A. **C.** Blue regions indicate the top three positive patches on the surface of LSM14A predicted by BindUP. **D.** RBDetect prediction of RBDs based on amino acid sequence, shown in blue. **E.** STRING analysis of LSM14A interacting proteins. The colours of the connecting lines indicate the type of interaction: teal is a known interaction from a curated database; pink is an experimentally determined interaction; green is predicted based on gene neighbourhood; red is predicted on gene fusion; blue is predicted from gene co-occurence; yellow predicted from textmining; black predicted from co-expression and lilac predicted from protein homology.

Whilst we do not identify DCP1/2, XRN1 or Rck in our ivRBPome, we do find the P-body protein [186] and predicted interactor of LSM14A, Y box-binding protein 1 (YBX1) (Figure 4.5E). LSM14A was also reported to localise to stress granules. These are cytoplasmic, phase-separated, membraneless RNP complexes which function to promote cell survival by condensing translationally stalled mRNAs, ribosomal components and other factors associated with translation [187]. It is speculated that LSM14A functions as a shuttle between P-bodies and stress granules, perhaps targeting irreversibly damaged mRNA present in stress granules, for degradation in P-bodies [185]. In influenza A virus (IAV) infection, the viral nonstructural protein 1 has been shown to interact with LSM14A to prevent targeting of the viral nucleoprotein to P-bodies and stress granules [188].

A second role for LSM14A has been described in the context of the cellular anti-viral response. It has been shown to act as a sensor for both viral DNA and RNA, activating interferon (IFN) pathways upon detection of foreign nucleotides. It is thought that LSM14A-induced activation of IFNβ leads to increase of RIG-I expression to amplify the cellular anti-viral response [189]. Additionally, LSM14A-deficient mice exhibited downregulation of the anti-viral signalling protein STING, indicating LSM14A activation subsequently promotes STING signalling or regulates STING expression [190].

A direct link between HIV-1 infection and LSM14A has not been established. However, HIV-1 mRNA is targeted to P-bodies by cellular miRNA and viral replication is increased when this interaction is disrupted [176]. I hypothesise that LSM14A may recognise and bind HIV-1 RNA resulting in degradation of vRNA in P-bodies and/or activating the interferon response. It is unclear why being present inside the capsid would assist these functions. Therefore it is possible that LSM14A is incorporated as a by-product of its recognition of the vRNA. However, it is also possible that LSM14A is incorporated into viral particles to conduct an entirely novel function. Regardless of function inside or outside of the capsid, LSM14A presents an interesting and novel candidate to investigate in the context of HIV-1 infection.

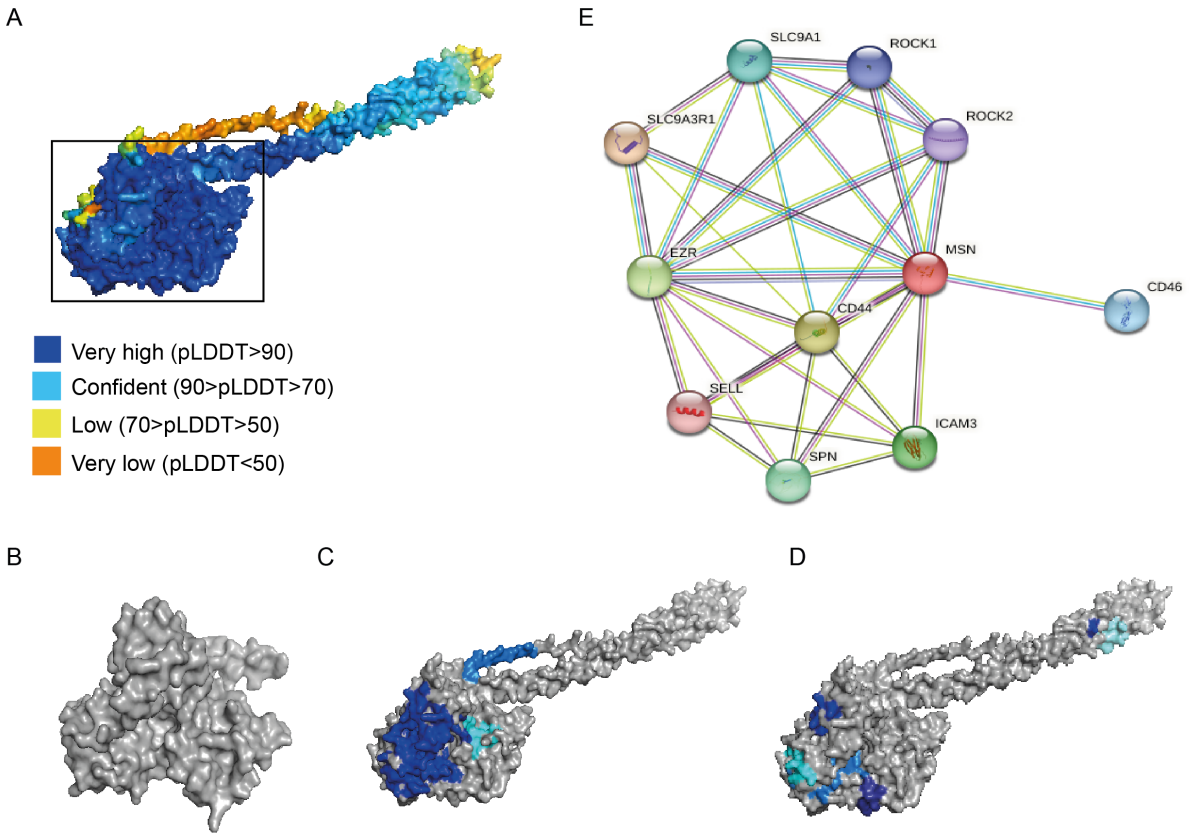

Figure 4.6: **MSN structure and RBDs. A.** Alphafold predicted structure of MSN with confidence indicated by colour (AF-P26038-F1). **B.** 1.73Å resolution Xray structure of amino acids 1-346 (PDB 6TXQ), relation to whole protein indicated by black box in A. **C.** Blue regions indicate the top three positive patches on the surface of MSN predicted by BindUP. **D.** RBDetect prediction of RBDs based on amino acid sequence, shown in blue. **E.** STRING analysis of MSN interacting proteins. The colours of the connecting lines indicate the type of interaction: teal is a known interaction from a curated database; pink is an experimentally determined interaction; green is predicted based on gene neighbourhood; red is predicted on gene fusion; blue is predicted from gene co-occurence; yellow predicted from textmining; black predicted from co-expression and lilac predicted from protein homology.

4.1.5 MSN

Membrane-organising extension spike protein, Moesin (MSN), is an ezrin-radixin-moesin (ERM) family member that connects the actin cytoskeleton to the plasma membrane. Rearrangement of the actin cytoskeleton drives many processes within the cell, including T cell activation. ERM family members are defined by N-terminal negatively charged FERM domains (band four-point-one ezrin radixin moesin homology domains), highlighted in Figure 4.6B, and a C-terminal F-actin binding tail. In the inactive conformation, the tail interacts with the FERM domain, tucking the tail into the protein. Phosphorylation of C-terminal threonine 558 weakens the interaction between the FERM domain and the tail, allowing the tail to extend, consequently exposing the membrane-protein and actin binding sites. The membrane protein binding site within the N terminus enables MSN to interact with the plasma membrane and anchor actin to the membrane [191]. This activated conformation is shown in Figure 4.6A.

The ERM proteins all contain novel globular RBDs [37]. In *Drosophila*, MSN was uncovered as a crucial component of nuclear mRNA export complex but was only shown to interact with the RNP through mRNA export factor PCID2, not directly with mRNA [192, 193]. There are several disordered regions predicted by amino acid sequence that correspond to the BindUP and RBDetect predicted RBDs highlighted in Figure 4.6C and D. The lack of strongly defined RNA-binding activity may be reflected by the stronger enrichment of MSN in ivRIC inputs over eluates, suggesting it does not have a strong interaction with HIV-1 RNA.

The role of MSN as a cytoskeletal regulator is better understood than its role in RNA metabolism. It is well established that ERM proteins are necessary for clustering of CD4 and assembly of T cell receptor complex molecules on microvilli, which is required for efficient T cell adhesion and homing to lymphoid organs [194–196]. MSN particularly has been shown to be essential for F-actin reassembly and clustering of CD4 with the CXCR4 co-receptor during membrane fusion and at cell-cell junctions during viral spreading [195]. MSN has also been implicated in the re-distribution of the intracellular adhesion molecule (ICAM)-3, CD44, CD46 and SPN (CD43) which

induces polarisation of T lymphocytes [197], a switch that is required for productive HIV-1 infection [198]. Figure 4.6E highlights the relationship of MSN with these proteins. Interestingly, we detect SPN in our ivRBPome and CD44 was identified in HIV-1 particles by Chertova et al [56]. In contrast to the pro-viral role MSN plays in T cell activation, MSN has also been found to block very early stages of HIV-1 infection [199, 200]. However, these studies were conducted in rat and human embryonic fibroblasts which are physiologically irrelevant to HIV-1 infection.

It would be plausible to think MSN is incorporated into viral particles purely because of its localisation at the plasma membrane in sites of high kinetic activity. However, Figure 4.2D suggests MSN is enriched in crosslinked HIV-1 particles over the total cell proteome suggesting a direct interaction with vRNA, at least in the immature viral particle. This, along with its contrasting pro- and anti-viral behaviours warrant further investigation into the role of MSN in HIV-1 viral particles.

4.1.6 LCK

Lymphocyte-specific protein tyrosine kinase (LCK) (also known as p56) is a non-receptor Src family tyrosine kinase involved in T cell receptor signalling and dampening the anti-viral innate immune response [201–203]. It can be found free in the cytoplasm or localised to the plasma membrane, through palmitoylation and myristoylation, and interacts with the plasma membrane co-receptors CD4 and CD8. However, co-receptor-bound LCK has been found to be less active than cytoplasmic LCK [204].

Like all Src-family kinases, LCK consists of regulatory SH2 and SH3 domains, a tyrosine kinase domain and a short C-terminal tail. The N terminus mediates the association with the cytoplasmic tails of CD4 and CD8 through a CxxC motif [205]. LCK has no known RBDs and has no described RNA-binding role within the literature. However, BindUP software predicts three regions surrounding the core of the protein that have a positive surface charge and are likely to bind RNA (Figure 4.7B-D). Entirely different regions of the protein are predicted to be RNA-binding by RBDetect, suggesting that LCK may harbour completely novel RBDs.

LCK is a well-understood negative regulator of innate immune signalling. Together

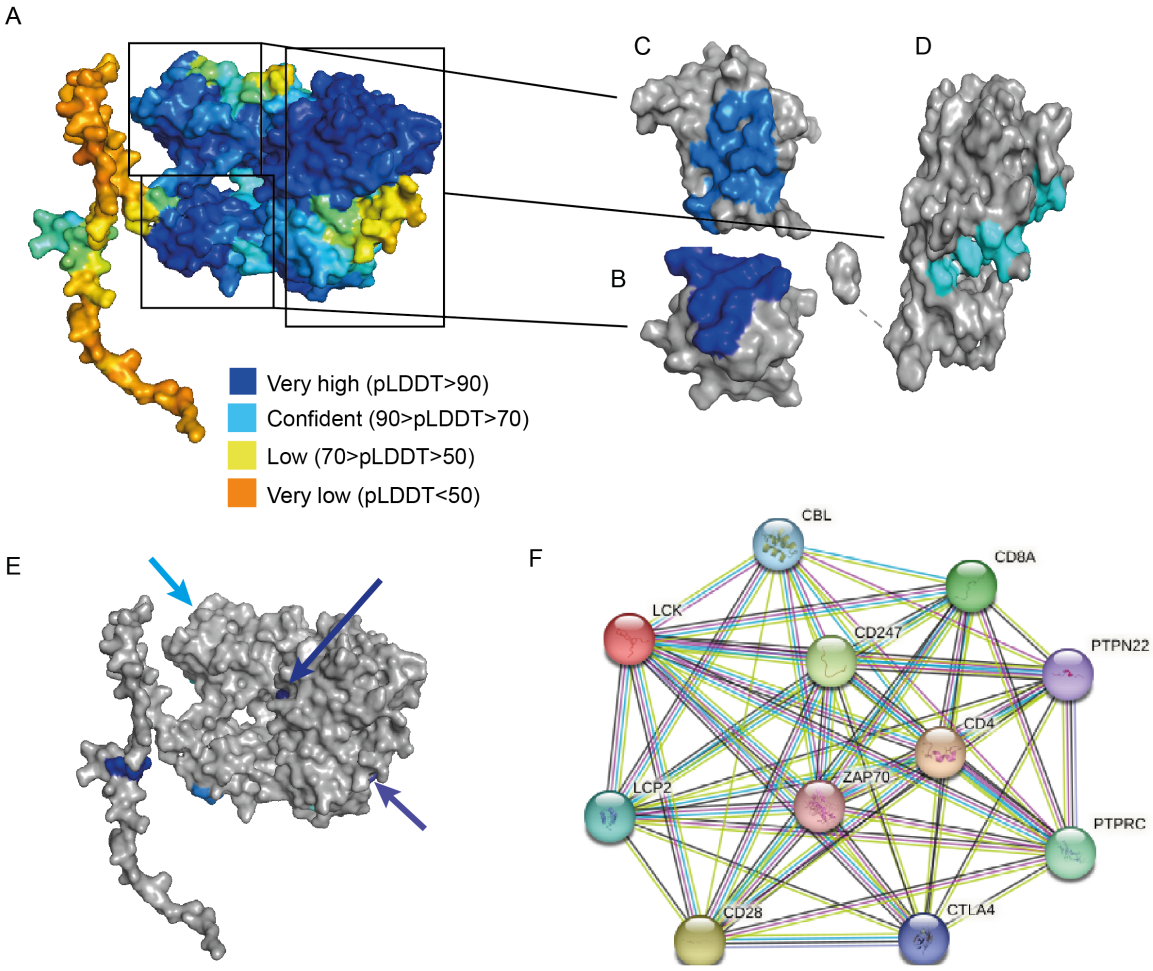

Figure 4.7: **LCK structure and RBDs.** **A.** Alphafold predicted structure of LCK with confidence indicated by colour (AF-P06239-F1). **B.** NMR structure of A chain amino acids 64-120 (PDB 1KIK), relation to whole protein indicated by black box in A. Blue residues identify positive patches predicted by BindUP software for RBD prediction. **C.** 1.00Å Xray structure of A chain amino acids 122-226 (PDB 1LKK), relation to whole protein indicated by black box in A. Blue residues identify positive patches predicted by BindUP software for RBD prediction. **D.** 1.81Å Xray structure of amino acids 225-509 (PDB 6PDJ), relation to whole protein indicated by black box in A. Blue residues identify positive patches predicted by BindUP software for RBD prediction. **E.** RBDetect prediction of RBDs based on amino acid sequence, shown in blue. Arrows indicate hidden predicted RBDs. **F.** STRING analysis of LCK interacting proteins. The colours of the connecting lines indicate the type of interaction: teal is a known interaction from a curated database; pink is an experimentally determined interaction; green is predicted based on gene neighbourhood; red is predicted on gene fusion; blue is predicted from gene co-occurence; yellow predicted from textmining; black predicted from co-expression and lilac predicted from protein homology.

with other Src kinase family members, Hck and Fgr, LCK was found to directly phosphorylate tank-binding kinase 1 (TBK1) to prevent its dimerisation and activation. Consequently, the transcription factor IRF3 is not phosphorylated and cannot translocatesignallingucleus to transcribe type I and III interferons. Blockage of the TBK1 signalling cascade dampens the cellular anti-viral innate immune response. Interestingly LCK, Hck and Fgr are all strongly induced during viral infection by TBK1-mediated IRF3 mobilisation, thus forming a negative feedback loop [203]. Unsurprisingly, viruses have hijacked this feedback loop to switch off expression of interferons. Virion protein 11/12 is a tegument protein of Herpes simplex virus type 1 (HSV-1) and has been shown to interact with LCK during HSV-1 infection and activate LCK through phosphorylation, thus preventing the cells from entering an anti-viral state [206]. A similar mechanism has also been observed in Herpesvirus saimiri infection [202].

In addition to the IRF3 signaling pathway, LCK is also involved in the nuclear factor (NF)-κB pathway, the activation of which is crucial for HIV-1 infection. The binding of HIV-1 gp120 to CD4 on the surface of T lymphocytes triggers NF-κB translocation to the nucleus and enhances the expression of pro-inflammatory cytokines, signallingsion molecules and chemokines. This increases the permissiveness of the cell to infection and promotes proviral integration. Interaction between CD4 and LCK was shown to be crucial for nuclear translocation of NF-κB and did not occur in cells expressing a mutant CD4 that prevented interaction between LCK and the cytoplasmic tail of CD4 [207].

HIV-1 replication is tightly coupled to T cell activation, which is governed by signalling through the T cell receptor (TCR) after interacting with antigen-presenting cells and forming an immunological synapse. After TCR engagement, LCK is recruited to the immunological synapse where it acts as a signaling master switch [208]. The activation of LCK at these sites is positively regulated by PTPRC/CD45 [209], which has also been identified as part of the HIV-1 particle proteome [57]. T cell activation is beneficial to HIV-1 infection because it facilitates the transcription of latent provirus and the progression of the viral life cycle. However, activation-induced cell death in T lymphocytes reduces the lifespan of infected cells. HIV-1 Nef alters the formation of the

immunological synapse and TCR stimulation to strike a balance between partial T cell activation, to promote HIV-1 spreading, but prevent activation-induced cell death. Part of the remodelling involves endosomal accumulation of LCK, inhibiting its recruitment to the immunological synapse [210]. Following endosomal accumulation, LCK is sequestered to the trans-golgi network by Nef, where it activates Ras-Erk signalling to promote the production of the T lymphocyte survival factor interleukin (IL)-2 and enhances viral spread [211, 212].

LCK has also been shown to function at later stages of the HIV-1 life cycle. Strasner et al revealed LCK knockout decreased the release of viral particles from HIV-1 infected cells. The effect was not due to reduced infection, transcription or translation. Instead, HIV-1 appeared to accumulate intracellularly. The plasma membrane-targeting palmitoylation domain of LCK was shown to be necessary for its effect on the targeting of HIV-1 Gag to the plasma membrane of T cells [213].

Whilst it is obvious LCK plays multiple roles in HIV-1 infection, it would be interesting to establish a role for LCK inside viral particles, especially given our lack of understanding of LCK RNA binding. It is not clear how being packaged inside the viral particle would be beneficial to any of the previously described roles of LCK in the literature. Therefore, this research may perhaps uncover a completely novel function of LCK in viral infection.

4.1.7 CIRBP

Cold-inducible RNA-binding protein (CIRBP) was originally described as a cold-shock domain containing cold-stress response protein that alters the translational efficiency of target mRNAs and has since been implicated in the cellular stress response to UV radiation and hypoxia [214]. It also harbours a highly conserved N-terminal RNA-recognition motif (RRM) through which it binds 5' or 3' UTRs of target mRNAs and influences their stability, splicing, nuclear export and translation, enabling the cell to respond rapidly to environmental stimuli [215]. Figure 4.8 highlights the RRM and its positive surface which interacts with RNA. The C-terminus of CIRBP is a glycine-rich region containing many RNA-binding RGG repeats which are thought to control

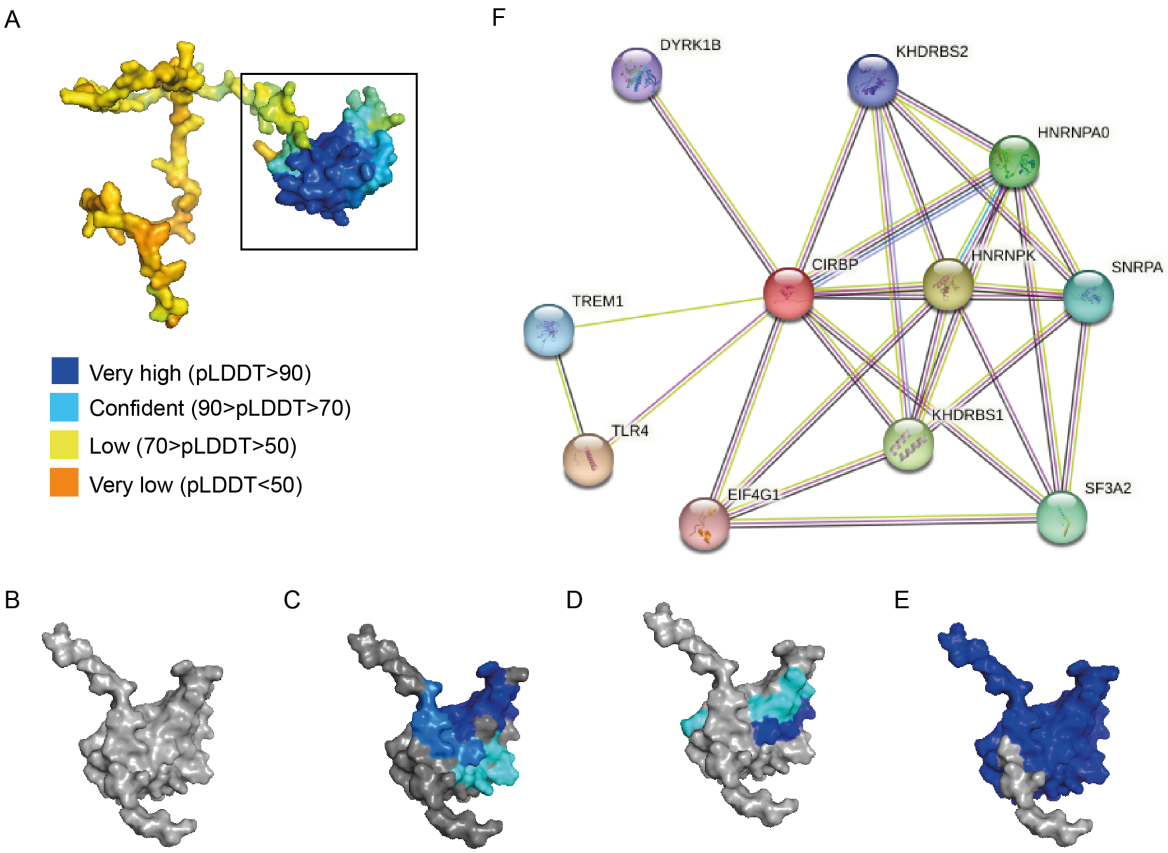

Figure 4.8: **CIRBP structure and RBDs. A.** Alphafold predicted structure of CIRBP with confidence indicated by colour (AF-Q14011-F1). **B.** 1.77Å resolution Xray structure of amino acids 1-91 (PDB 5TBX), relation to whole protein indicated by black box in A. **C.** Blue regions indicate the top three positive patches on the surface of CIRBP predicted by BindUP. **D.** RBDetect prediction of RBDs based on amino acid sequence, shown in blue. **E.** Canonical RBD, RRM motif, highlighted in blue. **F.** STRING analysis of CIRBP interacting proteins. The colours of the connecting lines indicate the type of interaction: teal is a known interaction from a curated database; pink is an experimentally determined interaction; green is predicted based on gene neighbourhood; red is predicted on gene fusion; blue is predicted from gene co-occurence; yellow predicted from textmining; black predicted from co-expression and lilac predicted from protein homology.

subcellular localisation and promote the migration of CIRBP from the nucleus to the cytoplasm. A large proportion of CIRBP is then further re-localised to stress granules where it acts as a translational repressor [216].

Many studies have revealed CIRBP is a damage-associated molecular pattern (DAMP) which promotes inflammation through the release of pro-inflammatory cytokines [217, 218]. Transcriptomics studies highlighted that CIRBP is upregulated during porcine reproductive and respiratory syndrome virus infection where it increased expression of the pro-inflammatory cytokines tumour necrosis factor-α (TNFα), IL-1β and IL-6. This effect was observed upon over-expression of CIRBP but was amplified further during viral infection [219].

Another transcriptome study, this one focusing on N^6-methyladenosine (m_6A) modification of mRNA in Flaviviridae infection, showed m_6A modification of CIRBP mRNA transcripts is decreased during infection. This promoted alternative splicing of CIRBP transcripts and consequently, reduced expression of the long isoform of CIRBP. Since the long isoform of CIRBP is not translated as efficiently as the shorter isoform, it was thought that this may be a mechanism to maintain abundant CIRBP protein levels during dengue virus (DENV), zika virus (ZIKV), WNV and HCV infection. Interestingly, both CIRBP isoforms were found to have a pro-viral function in DENV, ZIKV and HCV infection [220]. CIRBP expression was also increased in chickens infected with Newcastle disease virus but no investigation into functionality was conducted [221].

Despite being highlighted in several viral infections, the role of CIRBP during infection is poorly understood. Therefore, it would be of interest to elucidate the function of CIRBP in HIV-1 infection. The pro-inflammatory cytokines TNFα, IL-1β and IL-6 all stimulate HIV-1 replication in T cells and monocyte-derived macrophages [222]. Thus, HIV-1 may interact with CIRBP to promote expression of these cytokines. Alternatively, the mRNA stabilising activity of CIRBP may be required to help protect HIV-1 RNAs. CIRBP is known to interact with chaperone hnRNPK (Figure 4.8F), which we also find in our ivRBPome, suggesting that these RBPs could work together.

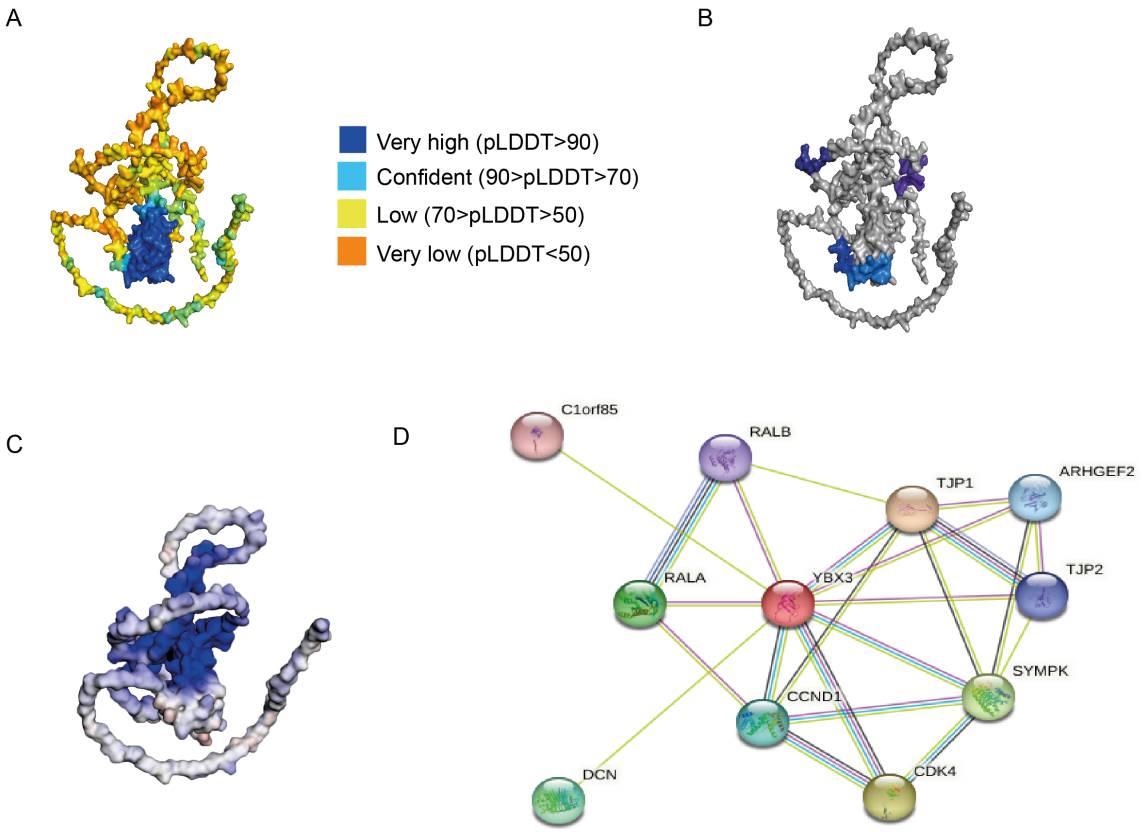

Figure 4.9: **YBX3 predicted structure and RBDs.** **A.** Alphafold predicted structure of YBX3 with confidence indicated by colour (AF-P16989-F1). **B.** RBDetect prediction of RBDs based on amino acid sequence, shown in blue. **C.** Electrostatic charge of surface, red for negative and blue for positive, modelled using PDB2PQR and APBS software [223]. **D.** STRING analysis of CIRBP interacting proteins. The colours of the connecting lines indicate the type of interaction: teal is a known interaction from a curated database; pink is an experimentally determined interaction; green is predicted based on gene neighbourhood; red is predicted on gene fusion; blue is predicted from gene co-occurence; yellow predicted from textmining; black predicted from co-expression and lilac predicted from protein homology.

4.1.8 YBX3

Y-box-binding protein 3 (YBX3) is another cold shock protein and harbours a cold shock domain (CSD) which enables YBX3 to bind both DNA and RNA. Its CSD sits in the core of the protein flanked by highly disordered N and C terminal tails (Figure 4.9). The Y-box protein family have incredibly diverse functions including the regulation of transcription, splicing, translation and mRNA stability. YBX3 itself has been identified in DNA repair and transcription of genes involved in cell survival, proliferation and differentiation [224–226]. Unsurprisingly, most of the literature explores the role of YBX3 in various cancers [227–230].

YBX3 was first discovered as a DNA-binding factor that regulated the transcription of genes containing a YBX promoter sequence [231]. However, roles in mRNA stability and the post-transcriptional control of translation have since been described [232,233]. Most recently, Cooke et al revealed YBX3 binds the 3' UTRs of the mRNAs encoding the amino acid transporters *SLC7A5* and *SLC3A2* which stabilises them. Depletion of YBX3 resulted in a depletion of the amino acids transported by these proteins [234].

Two links have been made between YBX3 and viral infection. First, YBX3 gene expression was found to be downregulated in SARS-CoV-2 infection [235]. Secondly, YBX3 was shown to restrict IAV infection through interaction with the viral RNP. Overexpression of YBX3 resulted in the inhibition of primary vRNA transcription and viral protein expression in infected cells. Conversely, YBX3 knockdown increased viral replication and led to accumulation of vRNPs [236]. No relationship between YBX3 and HIV-1 has been described and it was not identified in any of the other HIV-1 particle RBPomes. Additionally, none of the identified proteins that interact with YBX3 (Figure 4.9D) are found in any of the RBPomes either. Therefore, any potential relationship between HIV-1 and YBX3 is completely novel.

My main interest in YBX3 lies in the strong enrichment of this protein in crosslinked ivRIC eluates (Figure 4.2). YBX3 was consistently the RBP with the highest fold change over controls and that with the lowest false discovery rate (FDR). It is also enriched in viral particles 10-fold over the SupT1 whole cell proteome. The strength of this enrichment alone makes YBX3 a candidate worth investigating.

4.2 Results

4.2.1 Generation of MKRN1-KO SupT1 cell line

To determine whether the candidate ivRBPs had any effect on HIV-1 infection, viral replication was studied in the absence of ivRBP expression. This was achieved by generating knockout (KO) cell lines using CRISPR/Cas9 Nanoblades [237] in a CD4+ T lymphocytic cell line (SupT1). These cells were chosen as HIV-1 primarily infects CD4+ T lymphocytes and these transformed CD4+ T cell lines are easily maintained in tissue culture whilst being efficiently infected by HIV-1. KO SupT1 clones were already established for FAM120A, LSM14A, MSN, LCK, CIRBP and YBX3 in the Castello lab. For MKRN1-KO, I generated Nanoblades targeting two different regions of the MKRN1 gene. Nanoblades are viral vector delivery system for a Cas9-guide(g)RNA RNP. They incorporate gRNAs and a murine leukaemia virus (MLV) Gag polyprotein fused to Cas9, separated by a proteolytic site that can be cleaved by MLV protease to release the Cas9 (Figure 4.10A). Vesicular stomatitis virus glycoprotein (VSVG) and baboon endogenous retrovirus R-less glycoprotein (BaEVRless) form the envelope, enabling entry of the viral vector into a vast array of cell lines. Co-transfection of HEK293T cells with plasmids encoding each of the components results in the production of fusogenic viral like particles which are released into the culture medium [237]. The MKRN1 Nanoblades, TS576 and TS578, targeted two regions of the MKRN1 gene overlapping exon 4, a critical exon (Figure 4.10B). Cleavage at both of these sites induced the deletion of the DNA in between causing a frameshift resulting in a premature stop codon. A premature stop codon within a transcript will cause it to be degraded by nonsense-mediated decay. Beyond their high efficiency in difficult-to-transfect cell lines, Nanoblades present the advantage of containing Cas9 as protein (rather than encoded in DNA) enabling the expression of Cas9 for a limited timeframe. Short-term expression of the editing machinery and gRNAs reduces potential off target effects. After Nanoblade treatment, single cells were isolated and screened for successful KO.

Monoclonal cultures were first validated via PCR to identify any changes in the size of amplified DNA fragments. Visualisation of the PCR products via agarose gel

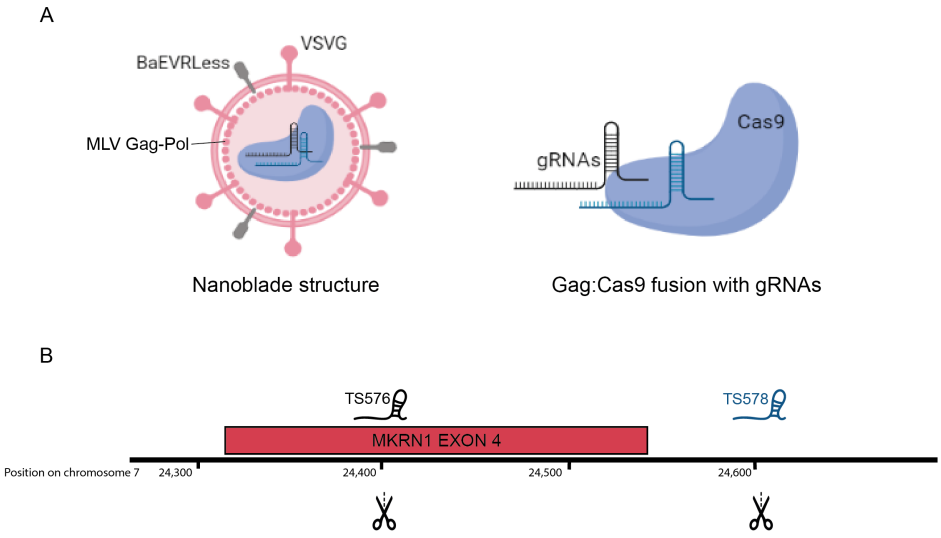

Figure 4.10: **CRISPR/Cas9 Nanoblades targeting MKRN1. A.** Structure of Nanoblade containing Cas9, gRNAs, and vesicular stomatitis virus glycoprotein (VSVG) and baboon endogenous retrovirus R-less glycoprotein (BaEVRless) envelope. **B.** Region of MKRN1 gene on chromosome 7 targeted by the gRNAs. The region between the two cleavage sites is deleted.

electrophoresis highlighted clones 15-17 as potential KOs as they produced bands larger than the WT positive control (400 bp) (Figure 4.11A). Two bands were detected for clone 18 which suggested one allele of MKRN1 had been successfully cut. Western blotting of whole cell lysate confirmed total loss of MKRN1 expression in clone 18 and reduced expression in clone 15 (Figure 4.11B). For all subsequent experiments using MKRN1-KO cells, clone 18 was used.

To establish whether the SupT1 KO cell lines were viable and grew at the same rate as WT SupT1 cells a viability assay was performed. Cells were seeded on day 0 at a density of $1x10^5$ cells/well and were counted every 24 h. A one-way ANOVA was used to determine whether the area under curve for each KO was significantly different to the WT. There was no difference in the growth rate of MKRN1-KO, FAM120A-KO and MSN-KO SupT1 cells compared to WT. However, LSM14A-KO, LCK-KO, CIRBP-KO and YBX3-KO cells grew slightly slower than the WT (Figure 4.12A). This difference appeared to be greatest from day 3 onward. These results were considered when designing future experiments. No difference in cell viability was observed for any of the KO clones (Figure 4.12B).

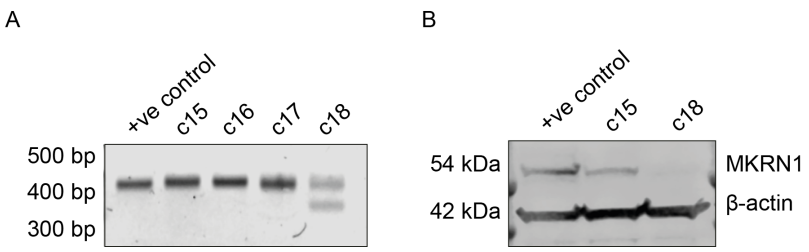

Figure 4.11: **CRISPR/Cas9 knockout of MKRN1 in SupT1 cells.** A. Clones screened by PCR and visualised on agarose gel. WT MKRN1 produces a band at 400 bp. B. MKRN1 protein levels in WCL were assessed by western blot in KO clones 15 and 18 and WT SupT1 cells. Western blot analysis was conducted using anti-MKRN1 and anti-βactin antibodies. MKRN1 expression produces a band at 54 kDa.

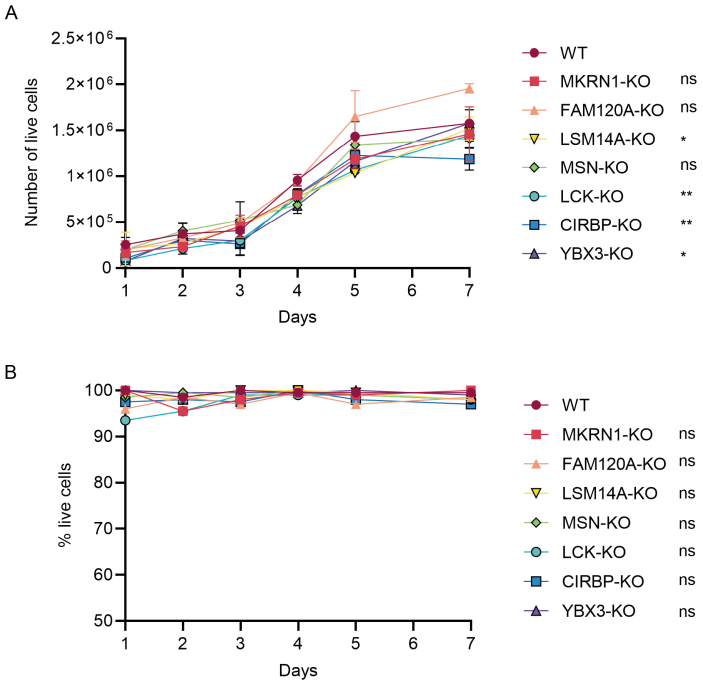

Figure 4.12: **Growth rates of SupT1-KO cells differ slightly.** Cells were seeded on day 0 at a density of 1x10^5 cells/well and were counted every 24 h. Number of live cells (**A**) and percentage of live cells **(B)** was determined by trypan blue staining. For both a one-way ANOVA was used to determine whether the area under curve for each KO was significantly different to the WT, n=2. For number of live cells MKRN1-KO p=0.0625, for FAM120A-KO p=0.0998, for LSM14A-KO p=0.0275, for MSN-KO p=0.3983, for LCK-KO p=0.0060, for CIRBP-KO p=0.0086 and for YBX3-KO p=0.0260. For percentage of live cells MKRN1-KO p=0.4617, for FAM120A-KO p=0.0690, for LSM14A-KO p=0.9783, for MSN-KO p=0.8771, for LCK-KO p=0.0632, for CIRBP-KO p=0.1227 and for YBX3-KO p=0.9997.

4.2.2 MKRN1-KO cells have increased expression of viral proteins and RNA during HIV-$1_{\text{-R-E-mCherry-Nef}}$ infection

To determine if MKRN1 regulated HIV-1 infection, MKRN1-KO CD4+ T lymphocytic cells were infected with a fluorescent reporter virus (HIV-$1_{\text{R-E-Nef-mCherry}}$) and vRNA and protein expression were measured. HIV-$1_{\text{R-E-Nef-mCherry}}$ is a 'single-round' virus that lacks HIV-1 envelope (Env) and Vpr and thus can infect once but cannot produce new infectious progeny in the infected cell. The pNL4-3-R-E-Nef-mCherry backbone can be co-transfected with vesicular stomatitis virus glycoprotein (VSVG) to generate pseudotyped single-round viruses. The advantage of using VSVG to pseudotype the viral particles is that it mediates viral-cell fusion through recognition of the low-density lipoprotein receptor, which is expressed almost ubiquitously [238]. Therefore, these pseudotyped single-round viruses can be used to infect a broad range of cell types. I exploited this advantage in later experiments. The particular virus used here expresses mCherry protein fused to the N-terminus of the viral protein Nef via a linker sequence. The linker harbours the 2A self-cleaving peptide from thosea asigna virus and enables self-cleavage of mCherry from Nef shortly after synthesis, rescuing the functional activity of Nef. These labelled single-round viruses enable the quantification of HIV-1 infection through measuring fluorescence. For each molecule of HIV-1 Nef produced, a molecule of mCherry is generated. Therefore, mCherry fluorescence is a proxy for Nef expression within the cell.

To investigate whether MKRN1-KO affected viral protein expression, MKRN1-KO and WT CD4 T lymphocytic cells (SupT1) were infected with HIV-$1_{\text{R-E-Nef-mCherry}}$ at an MOI of 1. The progress of infection was observed in an atmosphere-controlled plate reader which measured red fluorescence within the cells every 15 minutes over the course of 72 h. No fluorescence was detected in uninfected cells, demonstrating the robustness of the assay (Figure 4.13A). In both infected MKRN1-KO and WT cells, the fluorescent signal derived from HIV-1-R-E-Nef-mCherry progressively increased throughout the infection (72 h). However, levels of red fluorescence were substantially higher in MKRN1-KO than WT cells throughout (RMSRE=0.1247), indicating greater expression of Nef in absence of MKRN1 (Figure 4.14A). The decrease in fluorescence

observed between 0 and 6 h was due to incoming viral particles. Whilst Nef is not a component of HIV-1 particles, it is often passively incorporated into viral particles due to abundant expression of Nef at sites of viral particle assembly. Equally, this occurs with HIV-$1_{\text{R-E-Nef-mCherry}}$. mCherry is expressed as a fusion mCherry-T2A-Nef that is quickly processed into mCherry and Nef. We have shown that the bulge of red signal at the beginning of the infection corresponds to incoming viral particles because when adding RT inhibitors this early signal is still detectable while the late viral gene expression is not detected.

In order to determine whether MKRN1-KO affected the expression of other viral proteins and to validate the results by an orthogonal method, the expression of the HIV-1 Gag polyprotein was analysed by Western blotting. MKRN1-KO and WT SupT1 cells were infected with HIV-$1_{\text{R-E-Nef-mCherry}}$ at an MOI of 1 and lysed at 24 and 48 hpi. Lysates collected at 48 hpi were western blotted against β-actin and HIV-1 p24 which detects both p24 (CA) and p55 (Gag polyprotein). The signal intensity of detected bands was calculated in comparison to WT after normalisation to β-actin expression. Normalisation was conducted as it was observed that β-actin expression was consistently lower in KO cells despite controlling the number of cells lysed and the quantity of protein analysed. No signal was measured for either p24 or p55 in uninfected WT cells (Figure 4.13B, C and D) indicating the p24/55 antibody has high specificity. The detected signal intensity of p24 and p55 in MKRN1-KO cell lysates at 48 hpi was substantially higher in the absence of MKRN1 than in WT cells (Figure 4.14B, C and D). This suggests that ablation of MKRN1 enhances HIV-$1_{\text{R-E-Nef-mCherry}}$ viral fitness. Altogether, these results confirmed that HIV-1 protein expression was greater in MKRN1 KO cells, suggesting that MKRN1 may act as an anti-viral protein.

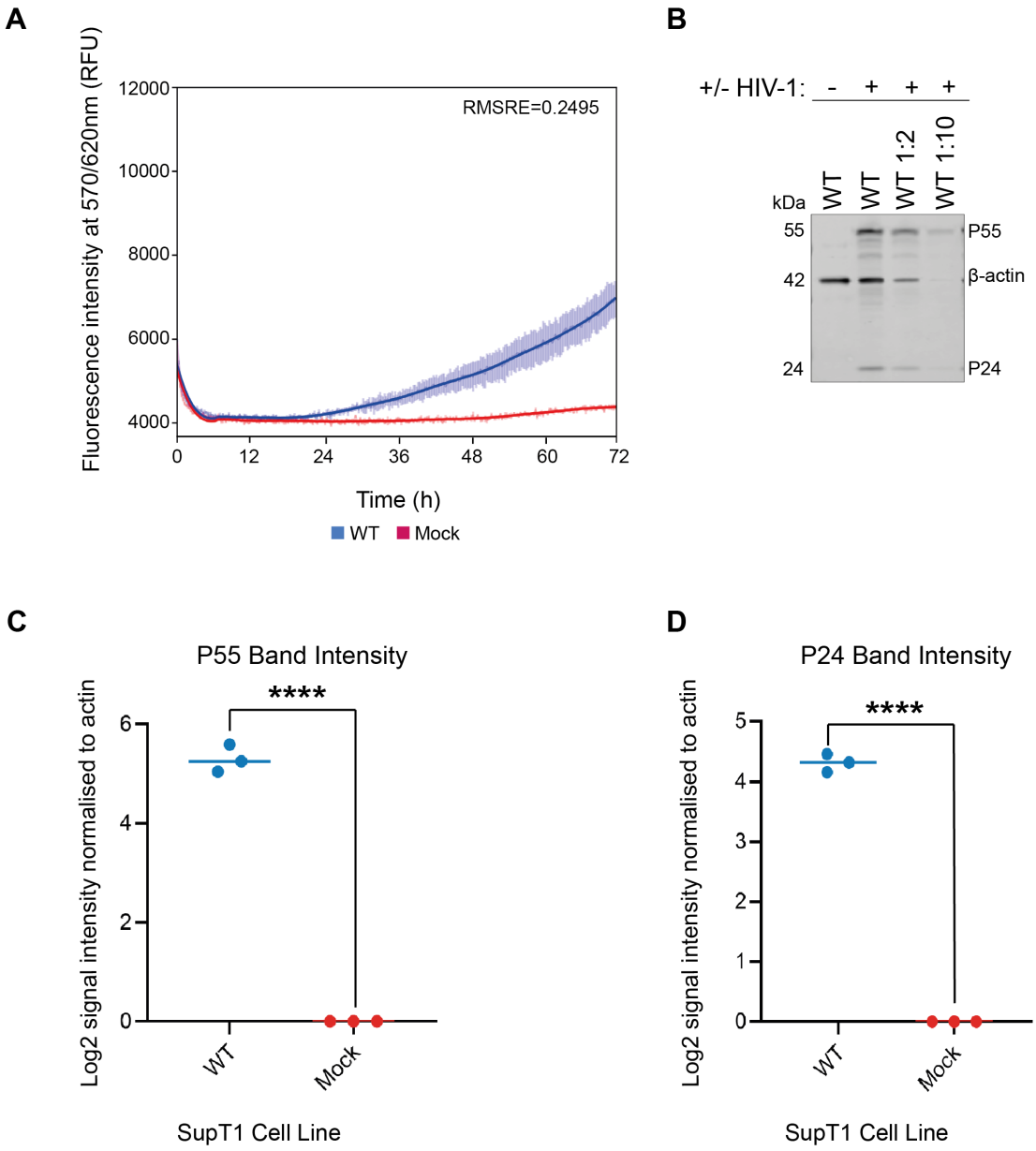

Figure 4.13: **HIV-1 Nef and Gag are not detected in uninfected cells A.** WT SupT1 cells (blue) infected with HIV-1$_{\text{-R-E-mCherry-Nef}}$ and uninfected 'mock' (red) SupT1 cells. Fluorescence intensity was measured at 570/62 nm every 15 minutes for 72 hpi. RMSRE was calculated to determine the difference between the two curves, n=4 **B.** Western blot analysis of WCL from WT cells infected with HIV-1$_{\text{-R-E-mCherry-Nef}}$ and uninfected cells collected at 48 hpi. Analysis was conducted using anti-βactin antibody and anti-HIV-1-P24 antibody, which recognises both p24 and p55. A dilution series of WCL for WT was used to plot a standard curve of signal intensity for anti-βactin and anti-HIV-1-p24. Signal intensity for mock was normalised to βactin signal in WT WCL. Quantification of the signal intensity for HIV-1 P55 and P24 is shown in **C.** and **D.** respectively. Individual points represent biological replicates and the bar represents the median. Statistical significance was determined using an unpaired T-test. P$\leq$0.0001 indicated by ****, n=3.

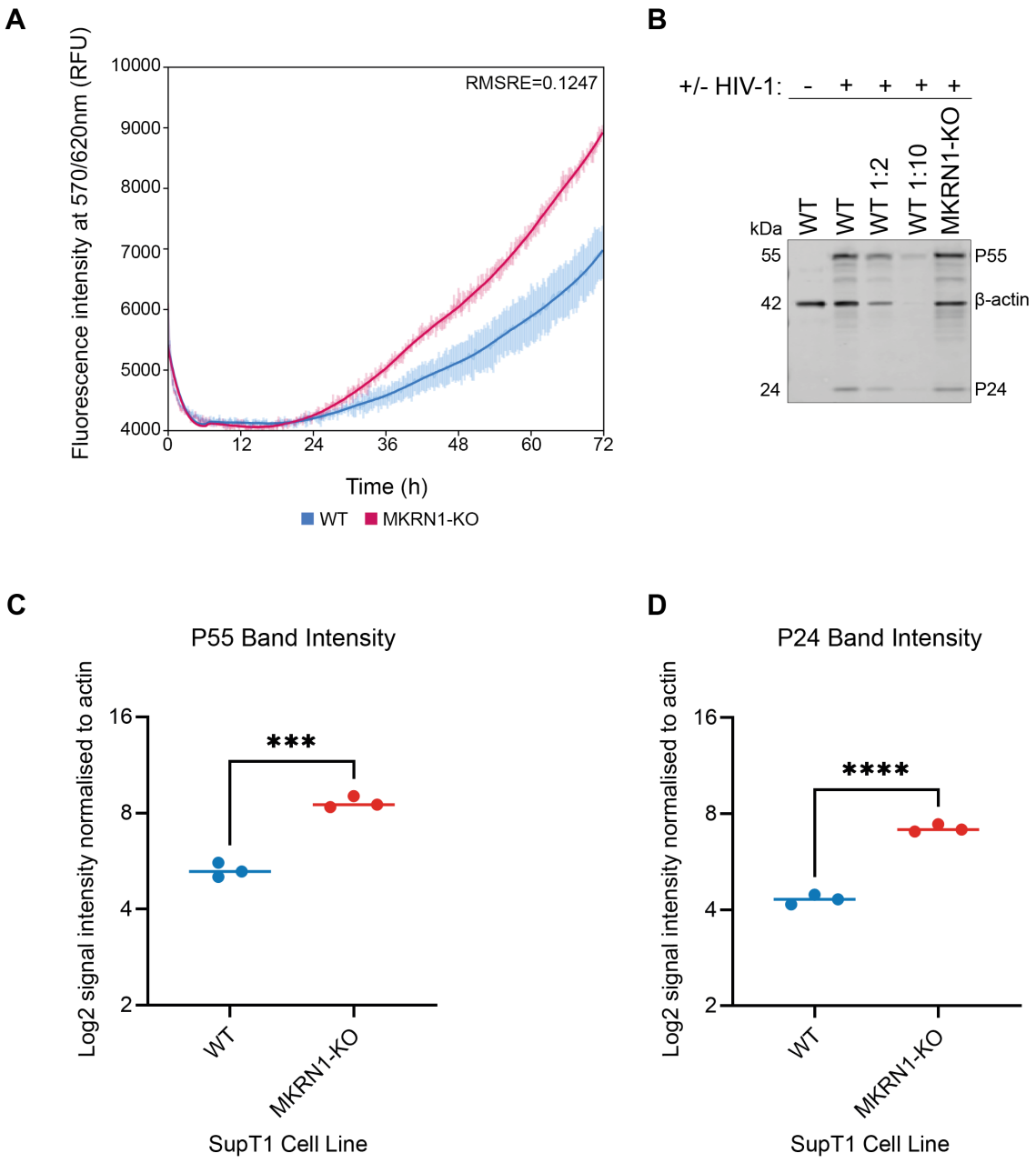

Figure 4.14: **MKRN1-KO increases expression of HIV-1 proteins. A.** WT (blue) and MKRN1-KO (red) SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$. Fluorescence intensity was measured at 570/62 nm every 15 minutes for 72 hpi. RMSRE was calculated to determine the difference between the two curves, n=4 **B.** Western blot analysis of WCL from WT and MKRN1-KO SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$ collected at 48 hpi. Analysis was conducted using anti-βactin antibody and anti-HIV-1-P24 antibody, which recognises both p24 and p55. A dilution series of WCL for WT was used to plot a standard curve of signal intensity for anti-βactin and anti-HIV-1-p24. Signal intensity for MKRN1-KO was normalised to βactin signal in WT WCL. Quantification of the signal intensity for HIV-1 P55 and P24 is shown in **C.** and **D.** respectively. Individual points represent biological replicates and the bar represents the median. Statistical significance was determined using an unpaired T-test. P$\leq$0.001 indicated by *** and P$\leq$0.0001 indicated by ****, n=3.

In addition to analysing viral protein expression within cell lysate, the expression of the three major splice variants of HIV-1 RNA was also observed. The lysates collected previously from HIV-$1_{\text{-R-E-mCherry-Nef}}$ infected cells at 24 and 48 hpi were analysed via RT-qPCR to measure the expression of unspliced (genomic), single-spliced and completely spliced HIV-1 RNA. The different splice variants of HIV-1 RNA were investigated to provide insight into key gene expression events in the HIV-1 life cycle. A $\Delta\Delta$CT analysis was performed to normalise vRNA expression in lysates to actin mRNA expression, then compare vRNA expression in MKRN1-KO lysate to WT lysate. None of the splice variants were detected in mock-infected cells at either 24 or 48 hpi (Figure 4.15) indicating that viral RNAs were specifically detected by RT-qPCR. At both 24 and 48 hpi expression of all HIV-1 splice variants was moderately increased in MKRN1-KO compared to WT T lymphocytes (Figure 4.16). Although, this increase was only significant in genomic RNA at 48 hpi. Together with protein expression, this data indicates viral replication was increased in MKRN1-KO cells and therefore suggests MKRN1 performs an inhibitory role in HIV-1 infection.

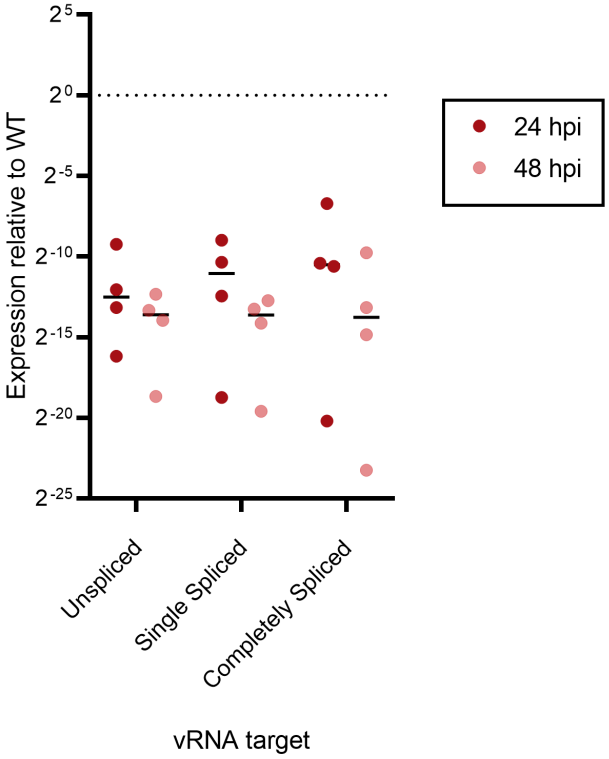

Figure 4.15: **No HIV-1 RNA is detected in uninfected cells.** RT-qPCR analysis of total cell RNA purified from WCL of uninfected 'mock' WT SupT1 cells, collected at 24 and 48 hpi. Expression of HIV-1 RNA in mock cells was calculated relative to WT cells using $\Delta\Delta CT$ normalised to actin mRNA expression at 24 and 48 hpi. Comparative expression of unspliced genomic, single-spliced and completely spliced vRNA is shown. Individual points represent the mean CT value calculated from technical replicates of 4 biological repeats of the experiment. The bar represents the median of these values. Statistical significance was determined using Welch's T test for populations with unequal standard deviations. P values were $\leq$0.0001 for all conditions.

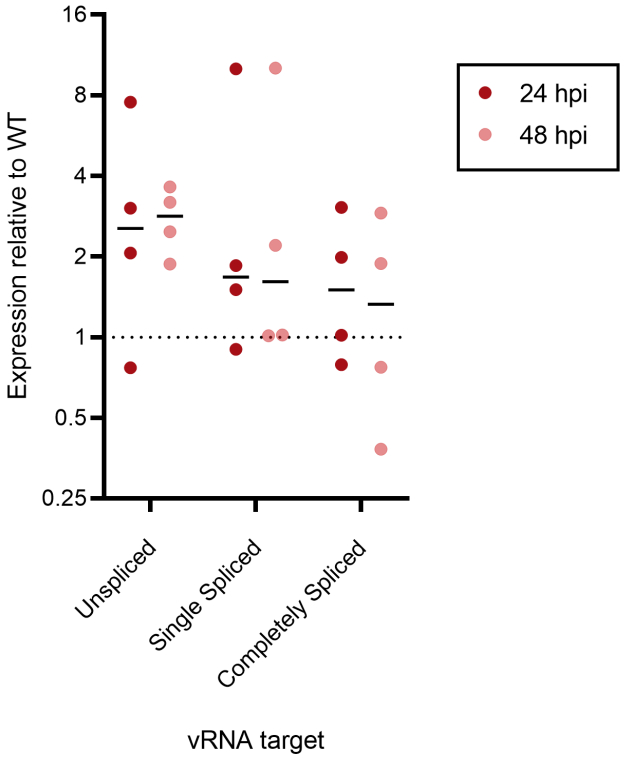

Figure 4.16: **MKRN1-KO increases expression of HIV-1 RNA splice variants.** RT-qPCR analysis of total cell RNA purified from WCL of WT and MKRN1-KO SupT1 cells infected with HIV-1$_{\text{-R-E-mCherry-Nef}}$, collected at 24 and 48 hpi. Expression of HIV-1 RNA in MKRN1-KO cells was calculated relative to WT cells using $\Delta\Delta CT$ normalised to actin mRNA expression at 24 and 48 hpi. Comparative expression of unspliced genomic, single-spliced and completely spliced vRNA is shown. Individual points represent the mean CT value calculated from technical replicates of 4 biological repeats of the experiment. The bar represents the median of these values. Statistical significance was determined using Welch's T test for populations with unequal standard deviations. For unspliced vRNA at 24 and 48 hpi the p values were 0.2082 and 0.0192 respectively. For single-spliced vRNA at 24 and 48 hpi the p values were 0.3198 and 0.3232 respectively. For completely spliced vRNA at 24 and 48 hpi the p values were 0.2624 and 0.4566 respectively.

4.2.3 FAM120A-KO cells have decreased expression of viral proteins and completely spliced RNA during HIV-$1_{\text{-R-E-mCherry-Nef}}$ infection

The same analysis of viral protein expression was also conducted in FAM120A-KO CD4+ T lymphocytic cells in parallel to MKRN1-KO. Cells were infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$ and the increase in red fluorescence was measured over the course of 72 h. Red fluorescence is substantially reduced in FAM120A-KO cells compared to WT, indicating less Nef-mCherry expression (Figure 4.17A). The same trend is observed when measuring HIV-1 Gag expression. Cell lysates collected from HIV-$1_{\text{-R-E-mCherry-Nef}}$ infected WT and FAM120A-KO cells at 48 hpi were western blotted against HIV-1 p24 and p55. Comparison of the detected signal intensity of p24 and p55 revealed there was significantly less expression of p24 and p55 in FAM120A-KO cells (Figure 4.17B, C and D). Comparative analysis of HIV-1 RNA expression in the same lysates by RT-qPCR indicated that there was also a slight reduction in the expression of genomic or single-spliced vRNAs at 24 and 48 hpi in FAM120A-KO cells compared to WT. However, the calculated p values were all $>$0.5 therefore this was not conclusive. The expression of completely spliced vRNA was affected to a greater extent, expression at 24 hpi was moderately reduced (p=0.0231) and at 48 hpi there was a slight reduction (Figure 4.18). Taken together, these results indicate that HIV-$1_{\text{-R-E-mCherry-Nef}}$ was able to integrate into the genome of FAM120A-KO cells as neither vRNA nor protein expression was completely abolished. However, the expression of HIV-1 completely spliced vRNA and proteins was reduced, suggesting FAM120A plays a pro-viral role during infection that affects viral gene expression.

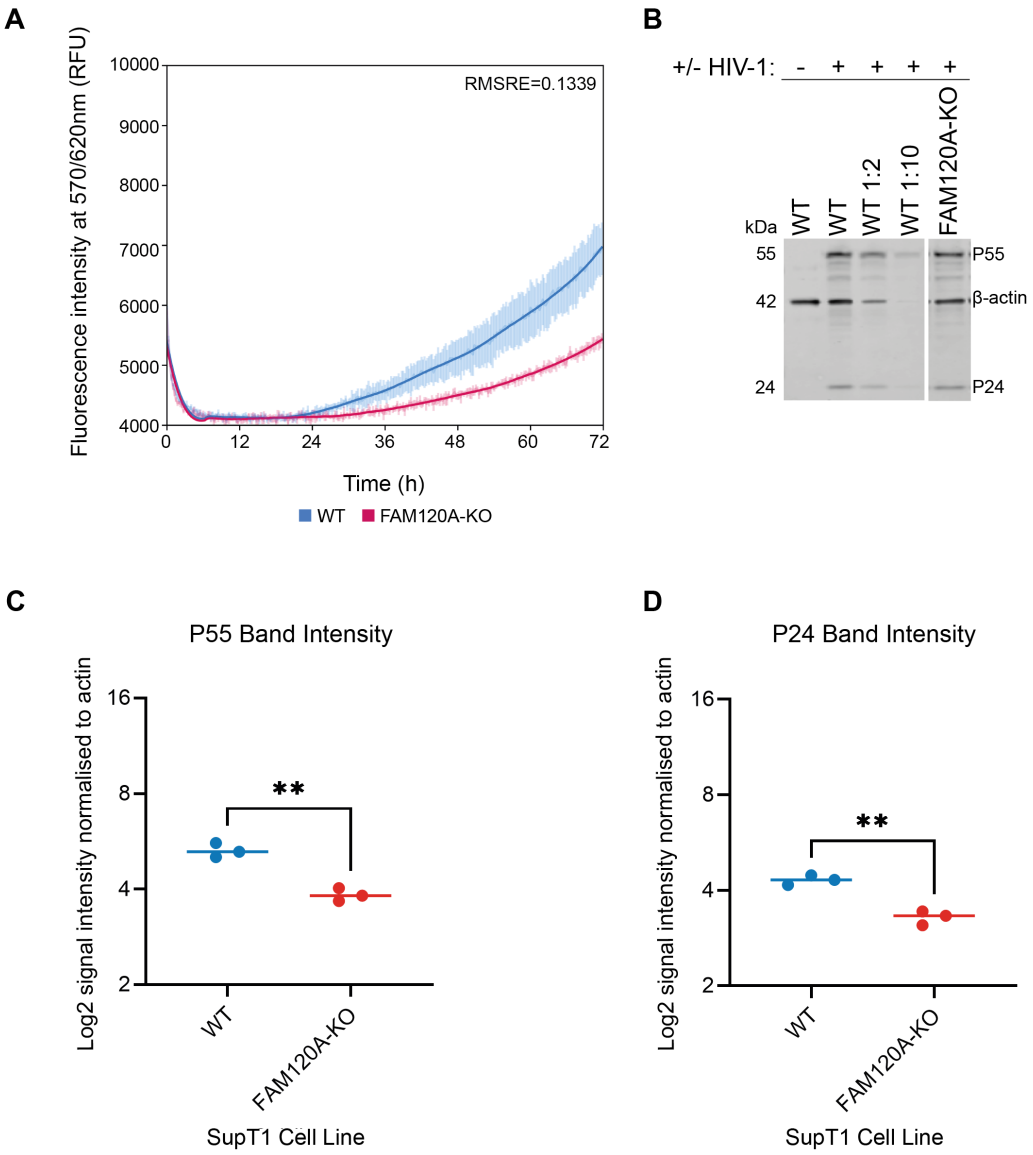

Figure 4.17: **FAM120A-KO decreases expression of HIV-1 proteins. A.** WT (blue) and FAM120A-KO (red) SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$. Fluorescence intensity was measured at 570/620 nm every 15 minutes for 72 hpi. RMSRE was calculated to determine the difference between the two curves, n=4. **B.** Western blot analysis of WCL from WT and FAM120A-KO SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$ collected at 48 hpi. Analysis was conducted using anti-βactin antibody and anti-HIV-1-P24 antibody, which recognises both p24 and p55. A dilution series of WCL for WT was used to plot a standard curve of signal intensity for anti-βactin and anti-HIV-1-p24. Signal intensity for FAM120A-KO was normalised to βactin signal in WT WCL. Quantification of the signal intensity for HIV-1 P55 and P24 is shown in **C.** and **D.** respectively. Individual points represent biological replicates and the bar represents the median. Statistical significance was determined using an unpaired T-test. $P \leq 0.01$ indicated by **, n=3.

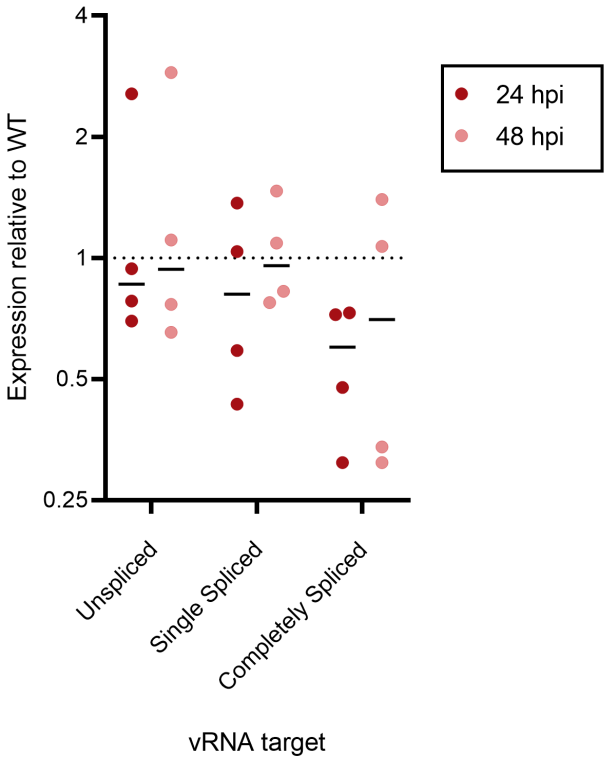

Figure 4.18: **FAM120A-KO decreases expression of HIV-1 completely spliced RNA.** RT-qPCR analysis of total cell RNA purified from WCL of WT and FAM120A-KO SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$, collected at 24 and 48 hpi. Expression of HIV-1 RNA in FAM120A-KO cells was calculated relative to WT cells using $\Delta\Delta CT$ normalised to actin mRNA expression at 24 and 48 hpi. Comparative expression of unspliced genomic, single-spliced and completely spliced vRNA is shown. Individual points represent the mean CT value calculated from technical replicates of 4 biological repeats of the experiment. The bar represents the median of these values. Statistical significance was determined using Welch's T test for populations with unequal standard deviations. For unspliced vRNA at 24 and 48 hpi the p values were 0.6179 and 0.5450 respectively. For single-spliced vRNA at 24 and 48 hpi the p values were 0.5511 and 0.8202 respectively. For completely spliced vRNA at 24 and 48 hpi the p values were 0.0231 and 0.4729 respectively.

4.2.4 LSM14A-KO has no overall effect on HIV-$1_{\text{-R-E-mCherry-Nef}}$ infection but may delay initial gene expression

Analysis conducted in parallel in LSM14A-KO CD4+ T lymphocytes suggests that, unlike MKRN1 and FAM120A-KO, LSM14A-KO had little impact on HIV-$1_{\text{-R-E-mCherry-Nef}}$ infection. Analysis of Nef expression by measuring mCherry fluorescence over the course of a 72 h infection revealed no difference in Nef expression between LSM14A-KO and WT cells (Figure 4.19A). Between 48 and 72 hpi the signal intensity increases in LSM14A-KO cells compared to WT. However, overall the root mean square relative error (RMSRE) calculated suggests very little difference between the two intensity curves (RMSRE=0.0545). Analysis of Gag expression by western blotting showed the same trend with no difference in p24 or p55 expression between LSM14A-KO and WT cells (Figure 4.19B, C and D). RT-qPCR analysis of vRNA expression in HIV-$1_{\text{-R-E-mCherry-Nef}}$ infected LSM14A-KO and WT SupT1s revealed that LSM14A-KO has a greater effect at the RNA level. The absence of LSM14A appears to delay viral gene expression as the expression of single-spliced and completely spliced vRNAs in reduced at 24 hpi. However, at 48 hpi the expression of all splice variants is substantially increased in LSM14A-KO cells compared to WT (Figure 4.20). These results suggest that LSM14A-KO initially impairs viral gene expression but this effect can be recovered and is likely why no effect is observed at the protein level.

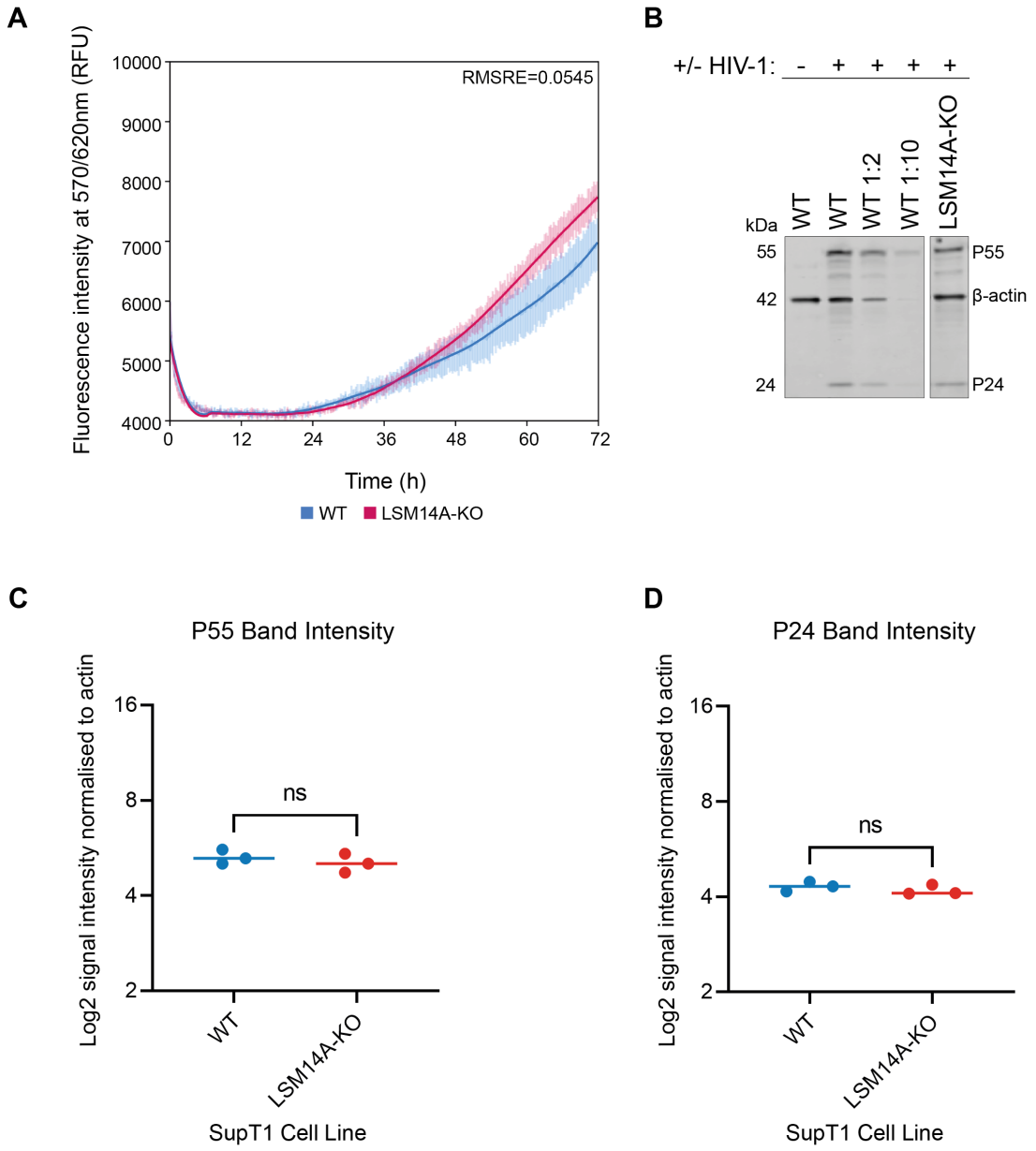

Figure 4.19: **LSM14A-KO does not affect expression of HIV-1 proteins.** **A.** WT (blue) and LSM14A-KO (red) SupT1 cells infected with HIV-1$_{\text{-R-E-mCherry-Nef}}$. Fluorescence intensity was measured at 570/620 nm every 15 minutes for 72 hpi. RMSRE was calculated to determine the difference between the two curves, n=4. **B.** Western blot analysis of WCL from WT and LSM14A-KO SupT1 cells infected with HIV-1$_{\text{-R-E-mCherry-Nef}}$ collected at 48 hpi. Analysis was conducted using anti-βactin antibody and anti-HIV-1-P24 antibody, which recognises both p24 and p55. A dilution series of WCL for WT was used to plot a standard curve of signal intensity for anti-βactin and anti-HIV-1-p24. Signal intensity for LSM14A-KO was normalised to βactin signal in WT WCL. Quantification of the signal intensity for HIV-1 P55 and P24 is shown in **C.** and **D.** respectively. Individual points represent biological replicates and the bar represents the median. Statistical significance was determined using an unpaired T-test, n=3.

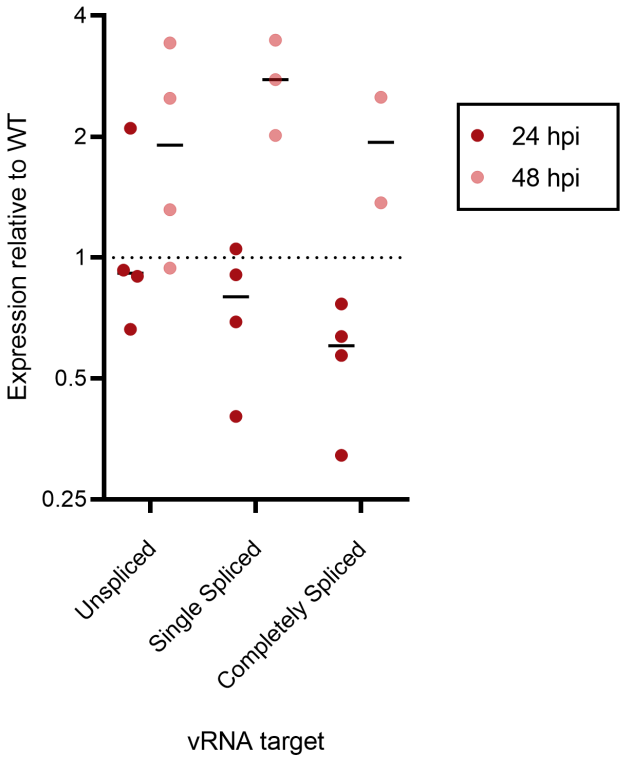

Figure 4.20: **LSM14A-KO delays expression of completely spliced vRNA.** RT-qPCR analysis of total cell RNA purified from WCL of WT and LSM14A-KO SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$, collected at 24 and 48 hpi. Expression of HIV-1 RNA in LSM14A-KO cells was calculated relative to WT cells using $\Delta\Delta CT$ normalised to actin mRNA expression at 24 and 48 hpi. Comparative expression of unspliced genomic, single-spliced and completely spliced vRNA is shown. Individual points represent the mean CT value calculated from technical replicates of 4 biological repeats of the experiment. The bar represents the median of these values. Statistical significance was determined using Welch's T test for populations with unequal standard deviations. For unspliced vRNA at 24 and 48 hpi the p values were 0.6775 and 0.1626 respectively. For single-spliced vRNA at 24 and 48 hpi the p values were 0.1920 and 0.0532 respectively. For completely spliced vRNA at 24 and 48 hpi the p values were 0.0197 and 0.3470 respectively.

4.2.5 MSN-KO increases the expression of Gag and unspliced vRNA

To determine whether MSN performs a critical function within HIV-1 infection, the same experiments were conducted in parallel in MSN-KO CD4+ T lymphocytic cells. Measurement of red fluorescence over 72 h in HIV-$1_{\text{-R-E-mCherry-Nef}}$ infected WT and MSN-KO SupT1 cells showed no difference in fluorescence between the two cell lines (RMSRE=0.0627). Therefore, the expression of HIV-1 Nef was not affected by MSN-KO (Figure 4.21A). Conversely, western blotting of cell lysate collected from HIV-$1_{\text{-R-E-mCherry-Nef}}$ infected MSN-KO cells suggested that P55 and P24 expression at 48 hpi was significantly increased in MSN-KO cells compared to WT (Figure 4.21B, C and D). Similar to the phenotype observed in LSM14A-KO cells, RT-qPCR analysis of infected cell lysate revealed delayed viral gene expression in MSN-KO cells. Completely spliced, single-spliced and unspliced vRNAs were downregulated at 24 hpi in MSN-KO compared to WT cell. However, expression of unspliced and completely spliced RNAs were then increased compared to WT at 48 hpi in MSN-KO cells (Figure 4.22). The greatest downregulation at 24 hpi was observed in unspliced vRNA whereas the greatest upregulation at 48 hpi was observed in completely spliced RNA. It appears that the increase in expression of completely spliced vRNA recovers the initial decrease so that at 48 hpi, Nef expression was not affected. Additionally, increased expression of unspliced vRNA reflects the increased expression of Gag protein. Overall, this suggests MSN plays an inhibitory role related to Gag RNA and protein in HIV-$1_{\text{-R-E-mCherry-Nef}}$ infection.

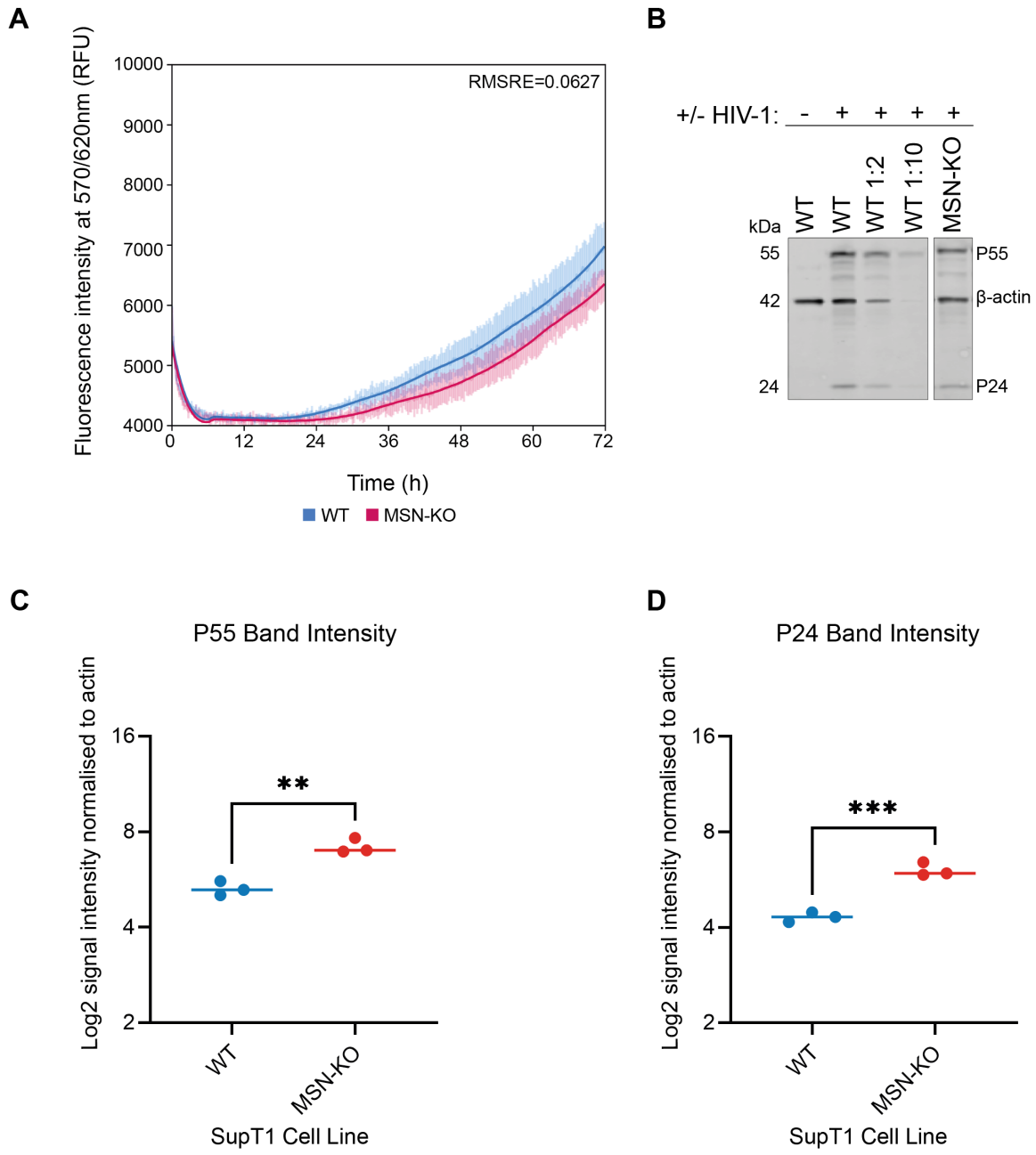

Figure 4.21: **MSN-KO increases expression of HIV-1 Gag but not Nef. A.** WT (blue) and MSN-KO (red) SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$. Fluorescence intensity was measured at 570/620 nm every 15 minutes for 72 hpi. RMSRE was calculated to determine the difference between the two curves, n=4. **B.** Western blot analysis of WCL from WT and MSN-KO SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$ collected at 48 hpi. Analysis was conducted using anti-βactin antibody and anti-HIV-1-P24 antibody, which recognises both p24 and p55. A dilution series of WCL for WT was used to plot a standard curve of signal intensity for anti-βactin and anti-HIV-1-p24. Signal intensity for MSN-KO was normalised to βactin signal in WT WCL. Quantification of the signal intensity for HIV-1 P55 and P24 is shown in **C.** and **D.** respectively. Individual points represent biological replicates and the bar represents the median. Statistical significance was determined using an unpaired T-test. P$\leq$0.01 indicated by ** and P$\leq$0.001 indicated by ***, n=3.

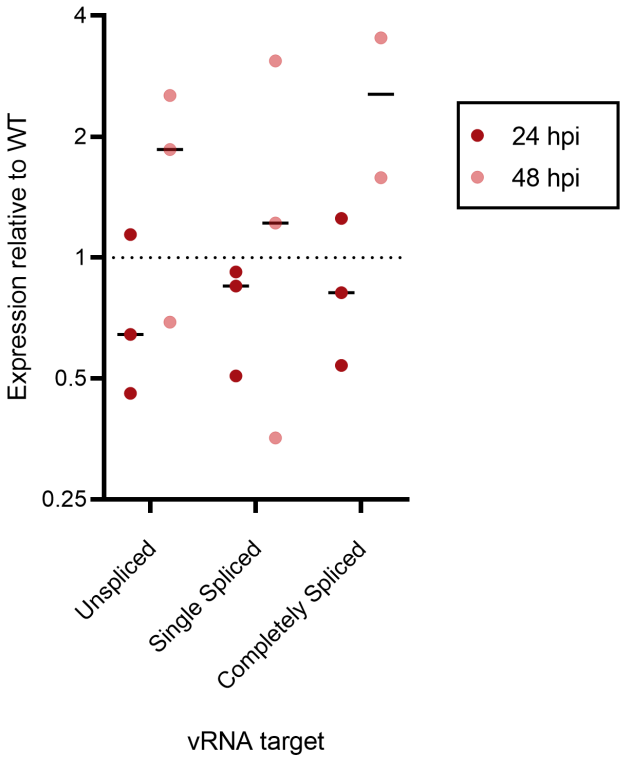

Figure 4.22: **MSN-KO initially decreases but then increases expression of HIV-1 RNA splice variants.** RT-qPCR analysis of total cell RNA purified from WCL of WT and MSN-KO SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$, collected at 24 and 48 hpi. Expression of HIV-1 RNA in MSN-KO cells was calculated relative to WT cells using $\Delta\Delta CT$ normalised to actin mRNA expression at 24 and 48 hpi. Comparative expression of unspliced genomic, single-spliced and completely spliced vRNA is shown. Individual points represent the mean CT value calculated from technical replicates of 4 biological repeats of the experiment. The bar represents the median of these values. Statistical significance was determined using Welch's T test for populations with unequal standard deviations. For unspliced vRNA at 24 and 48 hpi the p values were 0.3415 and 0.3265 respectively. For single-spliced vRNA at 24 and 48 hpi the p values were 0.1999 and 0.5631 respectively. For completely spliced vRNA at 24 and 48 hpi the p values were 0.5926 and 0.3559 respectively.

4.2.6 LCK-KO increases the initial expression of completely spliced vRNA and increases Nef protein expression

Again, the same analysis of viral RNA and protein expression in HIV-$1_{\text{-R-E-mCherry-Nef}}$ infection was conducted in parallel in LCK-KO cells. In contrast to MSN-KO, LCK-KO appears to substantially increase the expression of Nef and decrease the expression of Gag. Measuring the increase in fluorescence over a 72 hour infection with HIV-$1_{\text{-R-E-mCherry-Nef}}$ revealed a substantial increase in fluorescence over 72 hpi in LCK-KO cells (RMSRE=0.2455), corresponding to an increase in Nef expression (Figure 4.23A). However, western blotting of infected cell lysate against p24 and p55 suggested a very slight but significant decrease in the absence of LCK (Figure 4.23B, C and D). Curiously, RT-qPCR analysis of viral RNA expression in infected WT and LCK-KO SupT1 cells revealed a trend in RNA expression which contrasts with the trend observed at the protein level. Initially at 24 hpi, the expression of unspliced and completely spliced vRNA is slightly increased, whilst single spliced expression is unchanged. However, at 48 hpi, the expression of all splice variants is decreased (Figure 4.24). Although, none of these results were significant. The initial increase in completely spliced RNA may have been sufficient to drive the increase in Nef protein expression, which suggested that LCK may negatively regulate HIV-1 early gene expression but this result was not sustained at the RNA level.

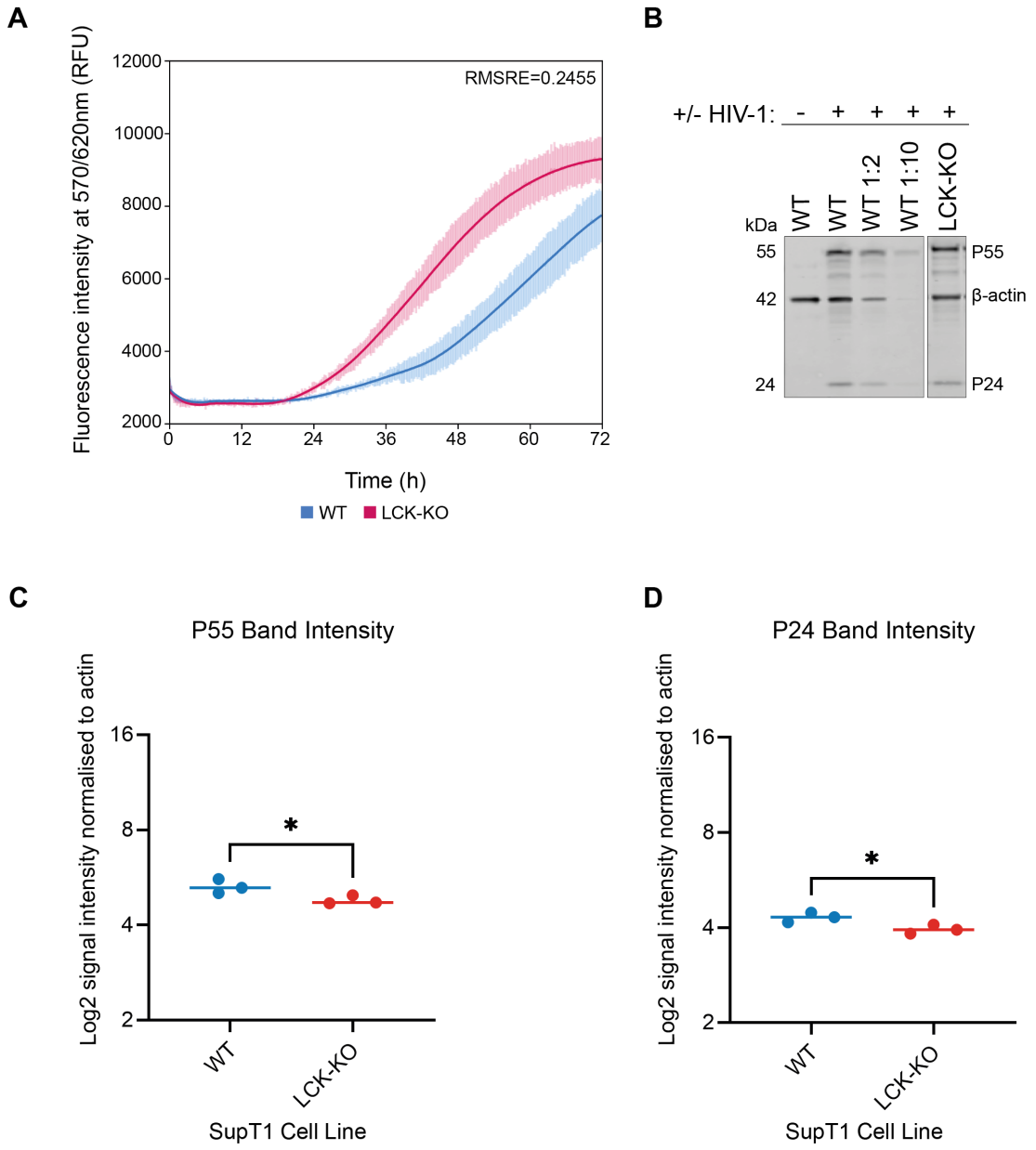

Figure 4.23: **LCK-KO increases expression of HIV-1 Nef. A.** WT (blue) and LCK-KO (red) SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$. Fluorescence intensity was measured at 570/620 nm every 15 minutes for 72 hpi. RMSRE was calculated to determine the difference between the two curves, n=4. **B.** Western blot analysis of WCL from WT and LCK-KO SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$ collected at 48 hpi. Analysis was conducted using anti-βactin antibody and anti-HIV-1-P24 antibody, which recognises both p24 and p55. A dilution series of WCL for WT was used to plot a standard curve of signal intensity for anti-βactin and anti-HIV-1-p24. Signal intensity for LCK-KO was normalised to βactin signal in WT WCL. Quantification of the signal intensity for HIV-1 P55 and P24 is shown in **C.** and **D.** respectively. Individual points represent biological replicates and the bar represents the median. Statistical significance was determined using an unpaired T-test. $P \leq 0.05$ indicated by *, n=3.

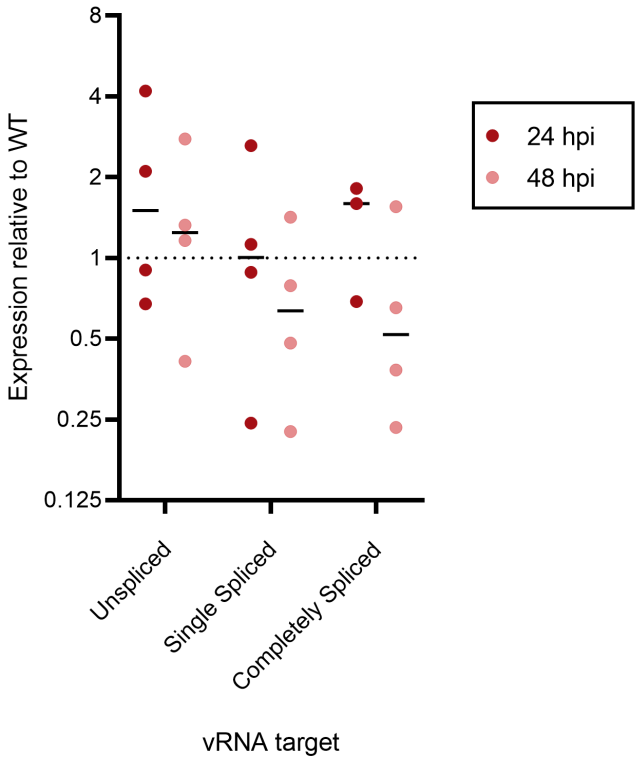

Figure 4.24: **LCK-KO initially upregulates vRNA expression but upregulation was not sustained.** RT-qPCR analysis of total cell RNA purified from WCL of WT and LCK-KO SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$, collected at 24 and 48 hpi. Expression of HIV-1 RNA in LCK-KO cells was calculated relative to WT cells using $\Delta\Delta CT$ normalised to actin mRNA expression at 24 and 48 hpi. Comparative expression of unspliced genomic, single-spliced and completely spliced vRNA is shown. Individual points represent the mean CT value calculated from technical replicates of 4 biological repeats of the experiment. The bar represents the median of these values. Statistical significance was determined using Welch's T test for populations with unequal standard deviations. For unspliced vRNA at 24 and 48 hpi the p values were 0.3154 and 0.4593 respectively. For single-spliced vRNA at 24 and 48 hpi the p values were 0.6940 and 0.3688 respectively. For completely spliced vRNA at 24 and 48 hpi the p values were 0.3997 and 0.3931 respectively.

4.2.7 CIRBP-KO increases expression of Nef and vRNAs at 48 hpi but has no effect on Gag expression

In order to determine the effect of CIRBP-KO on HIV-1 infection, CIRBP-KO and WT CD4+ T lymphocytes were infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$ and vRNA and protein expression was measured. Measuring mCherry fluorescence over 72 h revealed a greater increase in fluorescence occurred in CIRBP-KO, indicating more Nef was produced during infection of these cells (Figure 4.25A). The same increase in protein expression was not observed for the Gag polyprotein at 48 hpi. Western blot analysis shows no significant difference in P55 or P24 expression in the WCL of infected CIRBP-KO cells compared to WT (Figure 4.25B, C and D). Despite the increase in Nef observed at a protein level in CIRBP-KO cells, RT-qPCR analysis of the total cell RNA purified from infected CIRBP-KO cells at 24 and 48 hpi revealed completely spliced vRNA was slightly downregulated at 24 hpi, but then upregulated at 48 hpi, in comparison to WT (Figure 4.26). This is in contrast to the upregulation observed for Nef at a protein level which suggests a low level of completely spliced RNAs is sufficient to maintain Nef protein expression. In agreement with Gag protein expression, there was no significant change in expression of unspliced RNA in CIRBP-KO cells at 24 hpi but there was a slight increase in expression at 48 hpi. The greatest effect was seen in single-spliced RNA which was increased by 2 fold at 48 hpi. However, the protein expression was not measured for any of the proteins encoded by single-spliced transcripts to confirm if this effect was also observed at the protein level. The disconnect between the effect of CIRBP-KO on expression at the RNA and protein level may link to CIRBP's function as a translational regulator. Overall, in the absence of CIRBP, there is a greater expression of the viral accessory protein Nef and vRNAs at 48 hpi, which indicates CIRBP may inhibit these factors during infection.

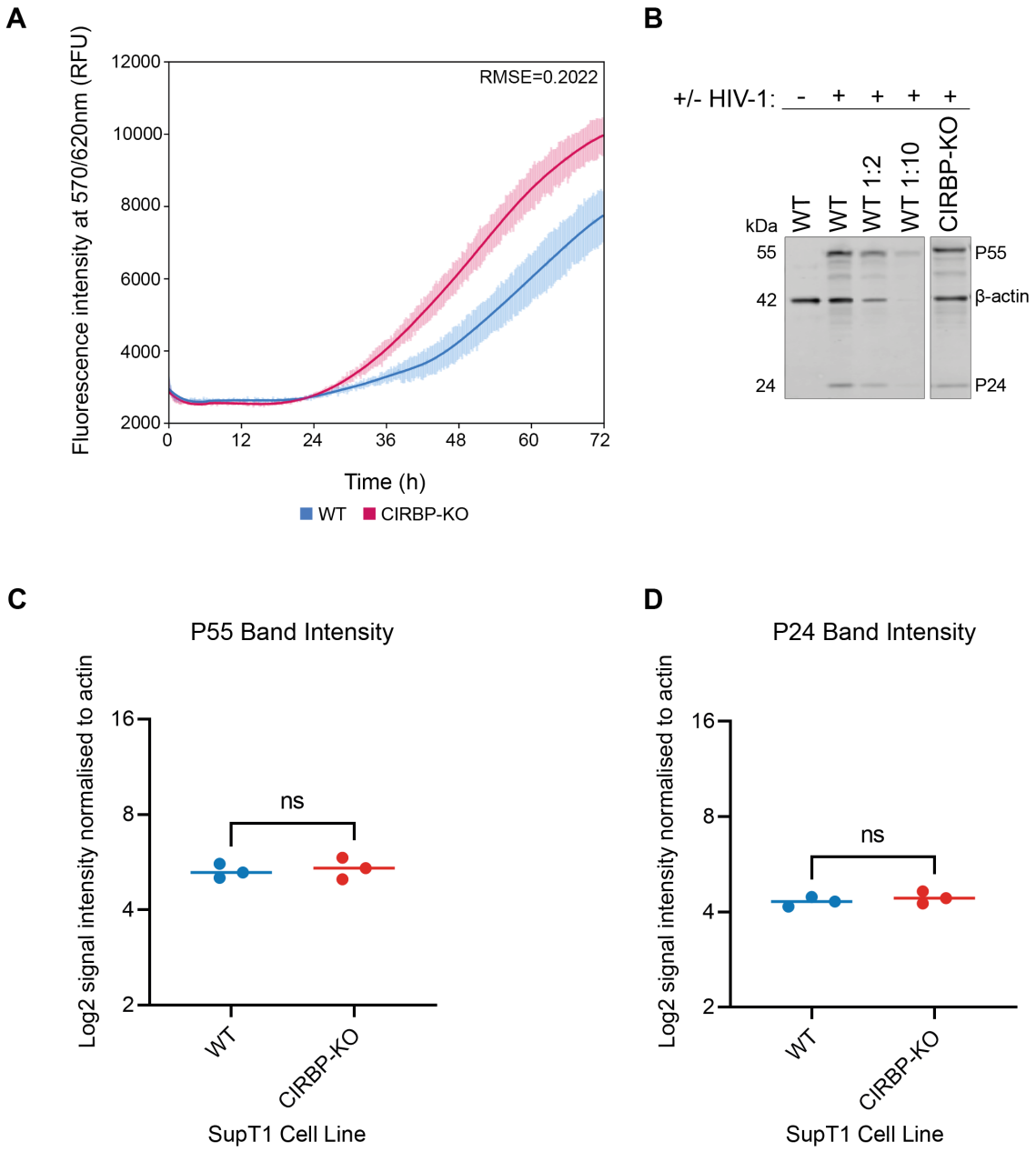

Figure 4.25: **CIRBP-KO increases expression of HIV-1 Nef but decreases expression of Gag.** **A.** WT (blue) and CIRBP-KO (red) SupT1 cells infected with HIV-1$_{\text{-R-E-mCherry-Nef}}$. Fluorescence intensity was measured at 570/620 nm every 15 minutes for 72 hpi. RMSRE was calculated to determine the difference between the two curves, n=4. **B.** Western blot analysis of WCL from WT and CIRBP-KO SupT1 cells infected with HIV-1$_{\text{-R-E-mCherry-Nef}}$ collected at 48 hpi. Analysis was conducted using anti-βactin antibody and anti-HIV-1-P24 antibody, which recognises both p24 and p55. A dilution series of WCL for WT was used to plot a standard curve of signal intensity for anti-βactin and anti-HIV-1-p24. Signal intensity for CIRBP-KO was normalised to βactin signal in WT WCL. Quantification of the signal intensity for HIV-1 P55 and P24 is shown in **C.** and **D.** respectively. Individual points represent biological replicates and the bar represents the median. Statistical significance was determined using an unpaired T-test, n=3.

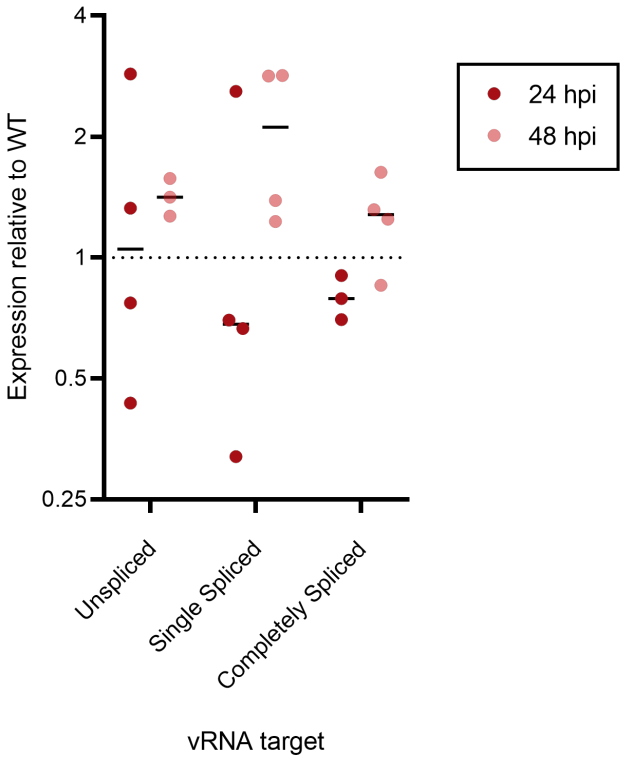

Figure 4.26: **CIRBP-KO downregulates completely spliced and single-spliced vRNA expression at 24 hpi and upregulates expression at 48 hpi.** RT-qPCR analysis of total cell RNA purified from WCL of WT and CIRBP-KO SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$, collected at 24 and 48 hpi. Expression of HIV-1 RNA in CIRBP-KO cells was calculated relative to WT cells using $\Delta\Delta CT$ normalised to actin mRNA expression at 24 and 48 hpi. Comparative expression of unspliced genomic, single-spliced and completely spliced vRNA is shown. Individual points represent the mean CT value calculated from technical replicates of 4 biological repeats of the experiment. The bar represents the median of these values. Statistical significance was determined using Welch's T test for populations with unequal standard deviations. For unspliced vRNA at 24 and 48 hpi the p values were 0.5625 and 0.2974 respectively. For single-spliced vRNA at 24 and 48 hpi the p values were 0.9011 and 0.0938 respectively. For completely spliced vRNA at 24 and 48 hpi the p values were 0.0740 and 0.2003 respectively.

4.2.8 YBX3-KO slightly increases Nef and Gag protein expression but delays vRNA expression

The same analysis of viral protein and RNA expression was conducted in YBX3-KO SupT1s to determine the effect of YBX3 on HIV-1 infection. Fluorescence analysis revealed a greater increase in fluorescence in HIV-$1_{\text{-R-E-mCherry-Nef}}$ infected YBX3-KO cells compared to WT, indicating greater Nef expression occurred in the absence of YBX3 (Figure 4.27A). Similarly, western blot analysis of HIV-$1_{\text{-R-E-mCherry-Nef}}$ infected cell lysate showed a slight increase in P55 and P24 expression ($p \leq 0.05$ and $p \leq 0.01$ respectively) in YBX3-KO cells at 48 hpi (Figure 4.27B, C and D). Despite viral protein expression increasing in YBX3-KO cells, RT-qPCR analysis of viral gene expression in HIV-$1_{\text{-R-E-mCherry-Nef}}$ WT and YBX3-KO cells revealed a delay in vRNA expression (Figure 4.28. Expression of completely spliced and single-spliced vRNA was substantially decreased by 65% in YBX3-KO cells at 24 hpi compared to WT cells ($p \leq 0.05$ and $p \leq 0.01$ respectively). Unspliced vRNA was also decreased at 24 hpi in YBX3-KO cells, although not significantly. Interestingly, the expression of all splice variants was recovered by 48 hpi. Together these results suggest two things: 1) YBX3 is required for initial gene expression and 2) the loss of YBX3 can be rescued by another protein to recover RNA and protein expression by 48 hpi.

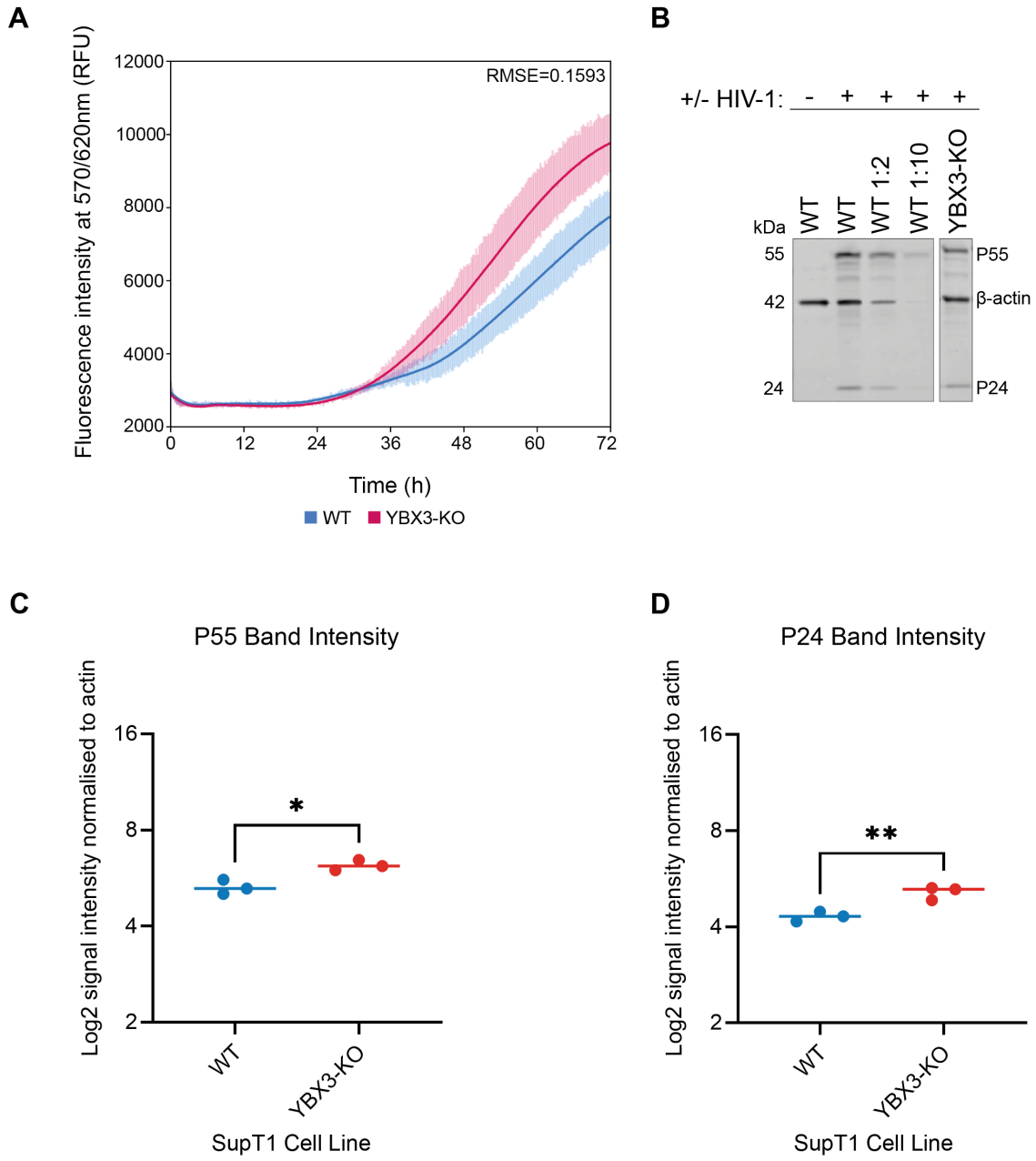

Figure 4.27: **YBX3-KO increases expression of HIV-1 Nef and Gag. A.** WT (blue) and YBX3-KO (red) SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$. Fluorescence intensity was measured at 570/620 nm every 15 minutes for 72 hpi. RMSRE was calculated to determine the difference between the two curves, n=4. **B.** Western blot analysis of WCL from WT and YBX3-KO SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$ collected at 48 hpi. Analysis was conducted using anti-βactin antibody and anti-HIV-1-P24 antibody, which recognises both p24 and p55. A dilution series of WCL for WT was used to plot a standard curve of signal intensity for anti-βactin and anti-HIV-1-p24. Signal intensity for YBX3-KO was normalised to βactin signal in WT WCL. Quantification of the signal intensity for HIV-1 P55 and P24 is shown in **C.** and **D.** respectively. Individual points represent biological replicates and the bar represents the median. Statistical significance was determined using an unpaired T-test. $P \leq 0.05$ indicated by * and $P \leq 0.01$ indicated by **, n=3.

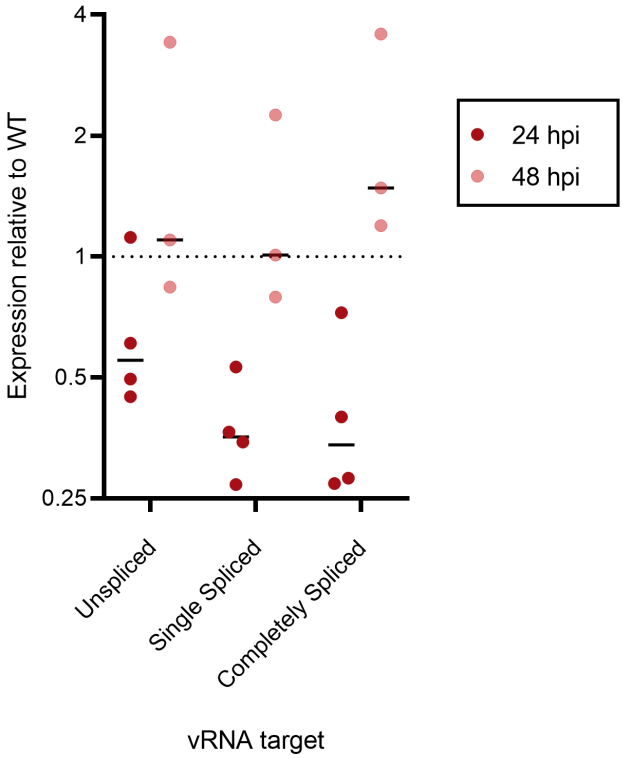

Figure 4.28: **YBX3-KO delays expression of vRNAs.** RT-qPCR analysis of total cell RNA purified from WCL of WT and YBX3-KO SupT1 cells infected with HIV-$1_{\text{-R-E-mCherry-Nef}}$, collected at 24 and 48 hpi. Expression of HIV-1 RNA in YBX3-KO cells was calculated relative to WT cells using $\Delta\Delta CT$ normalised to actin mRNA expression at 24 and 48 hpi. Comparative expression of unspliced genomic, single-spliced and completely spliced vRNA is shown. Individual points represent the mean CT value calculated from technical replicates of 4 biological repeats of the experiment. The bar represents the median of these values. Statistical significance was determined using Welch's T test for populations with unequal standard deviations. For unspliced vRNA at 24 and 48 hpi the p values were 0.1185 and 0.4388 respectively. For single-spliced vRNA at 24 and 48 hpi the p values were 0.0015 and 0.5201 respectively. For completely spliced vRNA at 24 and 48 hpi the p values were 0.0119 and 0.2857 respectively.

4.2.9 MKRN1-KO cells are less permissive to infection and produce fewer viral particles

After establishing the effect of the ivRBP-KOs on vRNA and protein expression, I determined whether KO of the ivRBPs affected viral particle production. Whilst HIV-$1_{\text{-R-E-mCherry-Nef}}$ enabled efficient study of viral replication, its use is limited to a single round of infection. To enable investigation of the later stages of the HIV-1 life cycle

and additional rounds of infection, the use of fully-competent HIV-1 is required. An assay was established (Figure 4.29) to measure the number of viral particles produced by infected ivRBP-KO cells. WT and ivRBP-KO CD4+ T lymphocytes (SupT1) were infected with HIV-$1_{mCherry\text{-}Nef}$ at an MOI of 0.1. As with the single-round virus, mCherry-Nef is cleaved shortly after synthesis to ensure Nef functionality is maintained. After overnight incubation, the supernatant was removed from the cells and replaced with fresh media to ensure only mCherry-Nef from successfully integrated pro-viruses was measured. At four days post-infection (dpi), half the supernatant was collected and replaced with fresh supplemented media. This was repeated at five, six and seven dpi. At seven dpi, the cells were collected, fixed, stained and analysed via flow cytometry to determine the percentage of live and infected cells. An enzyme-linked immunosorbent assay (ELISA) was conducted to determine the concentration of HIV-1 p24 in the supernatants. P24 concentration is indicative of viral particle concentration within the supernatant as simple stoichiometry suggests that 0.1pg of P24 corresponds to approximately 1000 viral particles (on the assumption that a viral particle contains 1500 copies of Gag) [239].

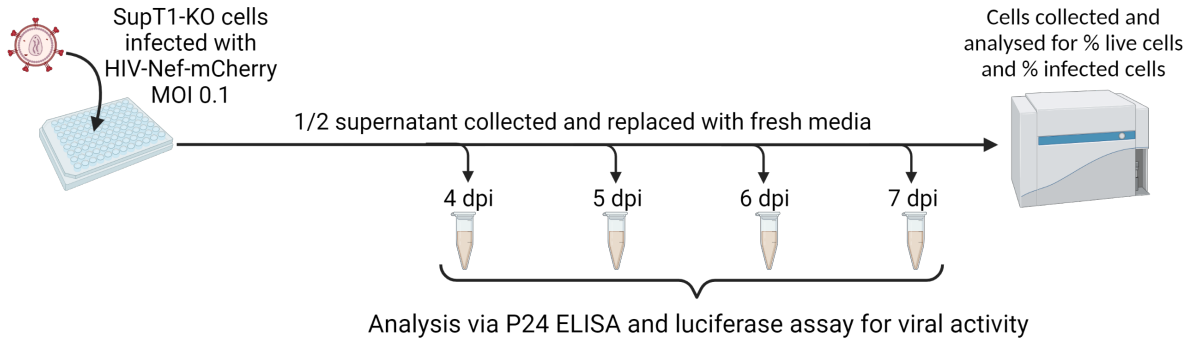

Figure 4.29: **An assay designed to determine the effect of ivRBP-KO on viral particle production.**

To determine the specificity of the assay an identical protocol was applied in parallel to uninfected CD4+ T lymphocytes. This analysis revealed that no p24 was detected in the supernatant of uninfected cells, indicating the specificity of the p24 ELISA (Figure 4.30A). Additionally, there was no difference in cell viability between HIV-$1_{mCherry\text{-}Nef}$ infected cells and uninfected cells (Figure 4.30B) and no mCherry-Nef expression was

detected in uninfected cells at 7 dpi (Figure 4.30C). Together these results indicate that the assay conditions enabled the specific detection of HIV-1 p24 and mCherry-Nef.

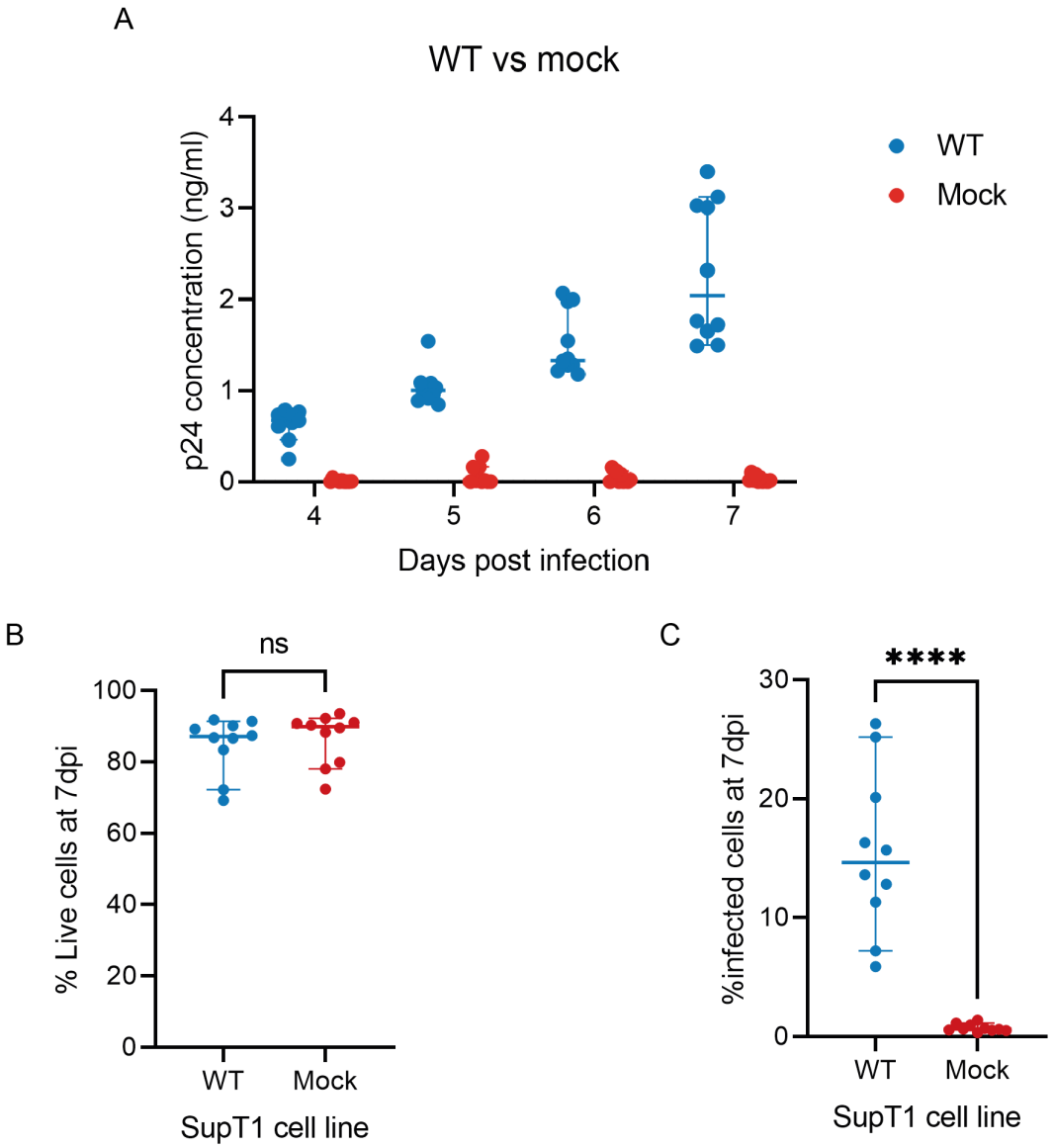

Figure 4.30: **No p24 can be detected in the supernatant of uninfected cells A.** A P24 ELISA was conducted to measure the concentration of P24 in supernatants collected from HIV-$1_{mCherry\text{-}Nef}$ infected (blue) and uninfected 'mock' (red) SupT1 WT cells. Supernatants were collected at 4, 5, 6 and 7 dpi. Individual points on the plot represent biological replicates and the bar shows the median. A 2-sample Z test was conducted between WT and mock populations, n=10 in each population. A P value was calculated for the Z test to determine the significance. RMSE gives a relative error indication of difference, higher RMSE is indicative of greater difference. **B.** Percentage of live SupT1 cells at 7 dpi determined through flow cytometry of fixed cells using a viability stain. Significance determined with an unpaired t test. **C.** Percentage of HIV-$1_{mCherry\text{-}Nef}$ infected SupT1 cells at 7 dpi, determined by flow cytometry of fixed cells for mCherry fluorescence. Statistical significance was calculated with a Welch's t test for populations with unequal variances, **** signifies $P \leq 0.0001$.

Surprisingly, when the assay was conducted in MKRN1-KO cells, P24 ELISA analysis of the supernatant revealed the concentration of P24 was significantly reduced (p=0.0004) in MKRN1-KO supernatant compared to WT supernatant (Figure 4.31A). At 7 dpi, the P24 concentration peaked at $0.91\,\mathrm{ng\,mL^{-1}}$ for MKNR1-KOs whereas for WT it was $2.00\,\mathrm{ng\,mL^{-1}}$. Therefore, approximately 20,000 viral particles/mL were present in WT supernatant and only 9100 viral particles/mL were present in MKRN1-KO supernatant at 7 dpi (55% decrease). The decrease in p24 concentration was not due to excessive cell death (Figure 4.31B). Additionally, the percentage of infected cells at 7 dpi was substantially reduced (p=0.0380) in MKRN1-KO cells compared to WT (Figure 4.31C) which indicates viral spread was impaired in MKRN1-KO lymphocytes.

4.2.10 FAM120A-KO does not effect the number of viral particles released from HIV-$1_{\text{mCherry-Nef}}$ infected cells

The same analysis was conducted in parallel for FAM120A-KO SupT1s. There was no significant difference (p=0.5467) in the p24 concentration of supernatant from infected FAM120A-KO cells and WT cells (Figure 4.32A). This was supported by no difference (p=0.1207) in the percentage of infected FAM120A-KO cells at 7 dpi (Figure 4.32C). There was, however, a significant decrease in the percentage of live FAM120A-KO cells at 7 dpi (Figure 4.32B). This indicated FAM120A-KO cells were not as viable during HIV-1 infection. Therefore, FAM120A has no role in viral particle release but may be required for cell viability during HIV-1 infection.

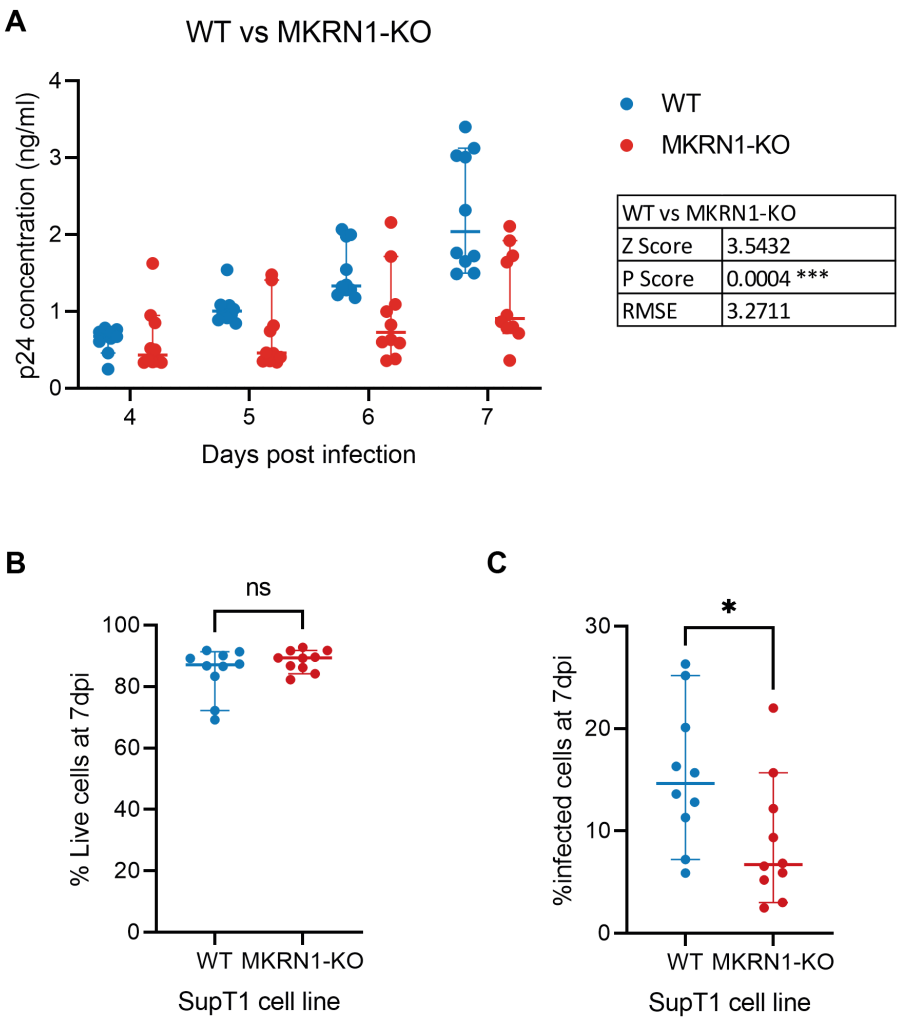

Figure 4.31: **MKRN1-KO cells release fewer HIV-1 viral particles. A.** A P24 ELISA was conducted to measure the concentration of P24 in supernatants collected from HIV-$1_{mCherry\text{-}Nef}$ infected SupT1 WT (blue) and MKRN1-KO (red) cells. Supernatants were collected at 4, 5, 6 and 7 dpi. Individual points on the plot represent biological replicates and the bar shows the median. A 2-sample Z test was conducted between WT and MKRN1-KO populations, n=10 in each population. A P value was calculated for the Z test to determine the significance. RMSE gives a relative error indication of difference, higher RMSE is indicative of greater difference. **B.** Percentage of live SupT1 cells at 7 dpi determined through flow cytometry of fixed cells using a viability stain. Significance determined with an unpaired t test. **C.** Percentage of HIV-$1_{mCherry\text{-}Nef}$ infected SupT1 cells at 7 dpi, determined by flow cytometry of fixed cells for mCherry fluorescence. Statistical significance was calculated with a Welch's t test for populations with unequal variances, * signifies $P \leq 0.05$.

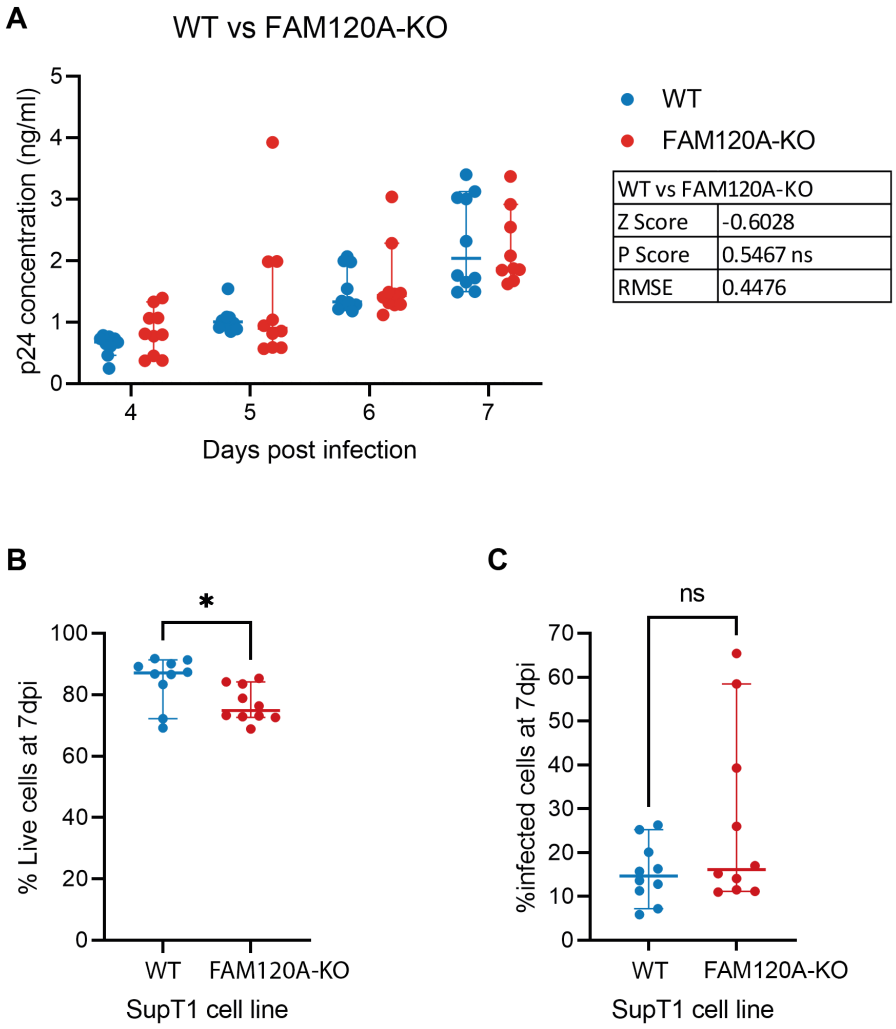

Figure 4.32: **FAM120A-KO does not affect the number of viral particles released from HIV-$1_{mCherry\text{-}Nef}$ infected cells. A.** A P24 ELISA was conducted to measure the concentration of P24 in supernatants collected from HIV-$1_{mCherry\text{-}Nef}$ infected SupT1 WT (blue) and FAM120A-KO (red) cells. Supernatants were collected at 4, 5, 6 and 7 dpi. Individual points on the plot represent biological replicates and the bar indicates the median. A 2-sample Z test was conducted between WT and FAM120A-KO populations, n=10 in each population. A P value was calculated for the Z test to determine the significance. RMSE gives a relative error indication of difference, higher RMSE is indicative of greater difference. **B.** Percentage of live SupT1 cells at 7 dpi determined through flow cytometry of fixed cells using a viability stain. Significance determined with an unpaired t test, * signifies $P \leq 0.05$. **C.** Percentage of HIV-$1_{mCherry\text{-}Nef}$ infected SupT1 cells at 7 dpi, determined by flow cytometry of fixed cells for mCherry fluorescence. Statistical significance was calculated with a Welch's t test for populations with unequal variances.

4.2.11 LSM14A-KO reduces the number of viral particles released from HIV-$1_{mCherry\text{-}Nef}$ infected cells

The same p24 ELISA analysis suggested that there was a reduced concentration of p24 in the supernatant of LSM14A-KO cells compared to WT cells (Figure 4.33A). At 7 dpi, the concentration of p24 in LSM14A-KO supernatant was 1.32 ng mL^{-1} which was a 34% decrease from 2.00 ng mL^{-1} in WT supernatant. This significant reduction was sustained throughout infection and suggests fewer viral particles were being released from infected LSM14A-KO cells. The decrease in P24 concentration may be contributed to by cell death as significantly fewer LSM14A-KO cells were alive at 7dpi (70% versus 86% in WT (Figure 4.33B). However, it could also indicate that LSM14A-KO cells were more susceptible to HIV-1 infection. Fewer infected cells were observed at 7 dpi in LSM14A-KOs but the 4% median decrease was not significant (Figure 4.33C), which suggests cell-cell transmission of infection may compensate for the reduction in free-virus.

4.2.12 MSN-KO reduces the release of viral particles from infected cells

The same assay conducted in MSN-KO cells revealed the absence of MSN resulted in a significant decrease (p=0.000001) of p24 in supernatant, indicating fewer viral particles were released from MSN-KO lymphocytes (4.34A). In some replicates MSN-KO induced a complete loss of p24 in the supernatant, although the median at 7 dpi was 1.31 ng mL^{-1} versus 2.00 ng mL^{-1} in WT supernatant. This decrease in viral particle release was not due to excessive cell death in MSN-KO cells (Figure 4.34B). The two replicates at 41.8% and 45.7% live cells did not correspond to the replicates producing 0 ng mL^{-1} p24. Supporting this, MSN-KO cells also had a significantly decreased (p=0.0027) percentage of infected cells at 7 dpi in comparison to WT cells (Figure 4.34C). Taken together, these results indicate MSN participates in viral particle release and potentially cell-cell transmission of HIV-1.

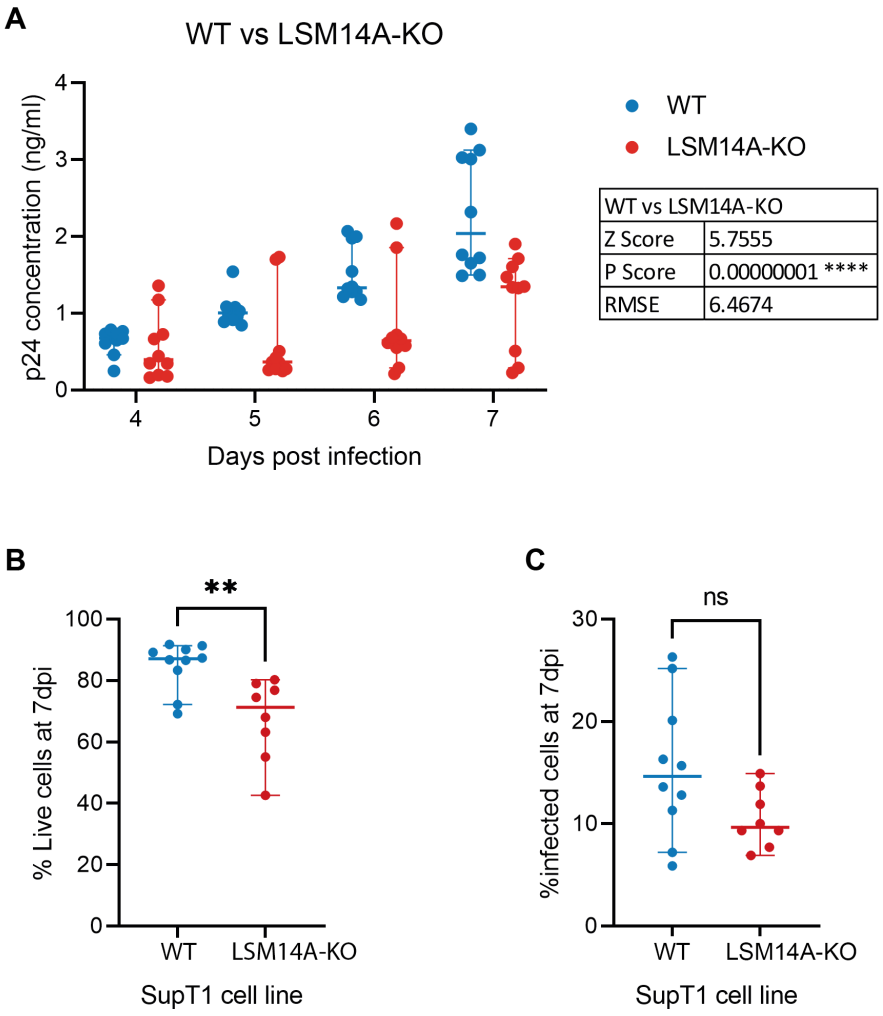

Figure 4.33: **LSM14A-KO decreases the release of viral particles from infected cells. A.** A P24 ELISA was conducted to measure the concentration of P24 in supernatants collected from HIV-$1_{mCherry\text{-}Nef}$ infected SupT1 WT (blue) and LSM14A-KO (red) cells. Supernatants were collected at 4, 5, 6 and 7 dpi. Individual points on the plot represent biological replicates and the bar indicates the median. A 2-sample Z test was conducted between WT and LSM14A-KO populations, n=10 in each population. A P value was calculated for the Z test to determine the significance. RMSE gives a relative error indication of difference, higher RMSE is indicative of greater difference. **B.** Percentage of live SupT1 cells at 7 dpi determined through flow cytometry of fixed cells using a viability stain. Significance determinedaffectan unpaired t test, ** signifies $P \leq 0.01$. **C.** Percentage of HIV-$1_{mCherry\text{-}Nef}$ infected SupT1 cells at 7 dpi, determined by flow cytometry of fixed cells for mCherry fluorescence. Statistical significance was calculated with a Welch's t test for populations with unequal variances.

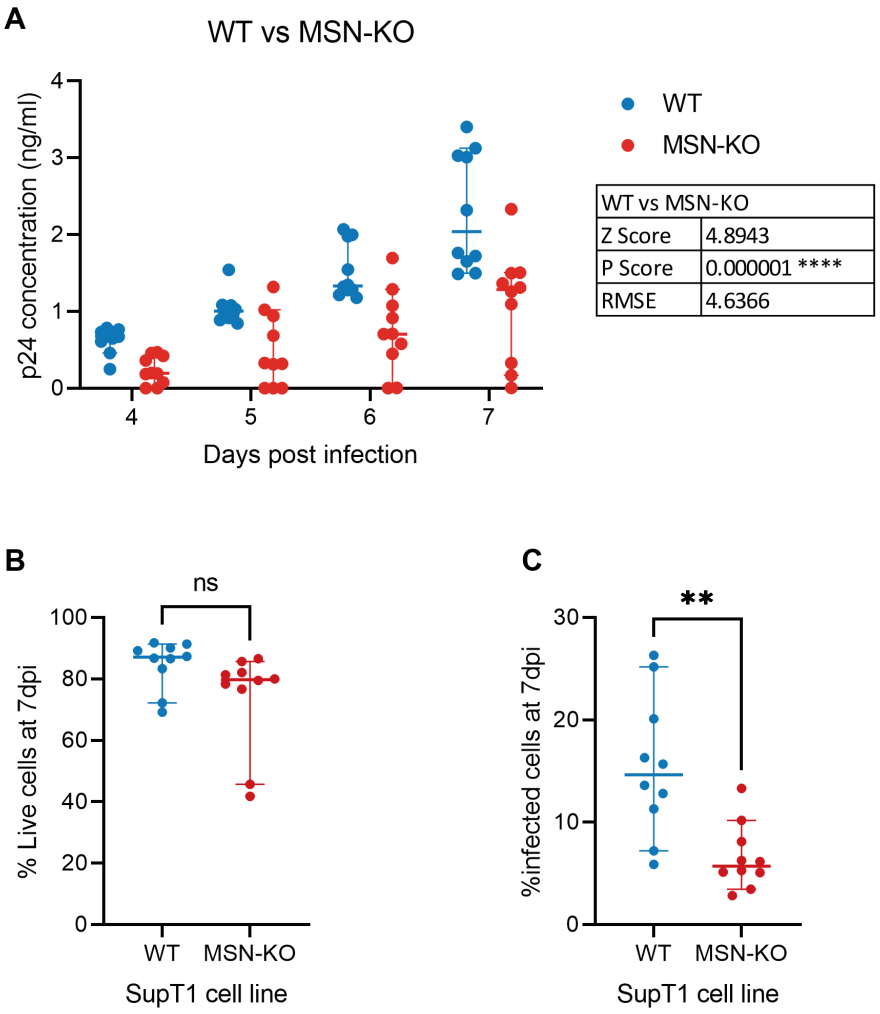

Figure 4.34: **MSN-KO decreases the release of viral particles from infected cells.** **A.** A P24 ELISA was conducted to measure the concentration of P24 in supernatants collected from HIV-$1_{mCherry\text{-}Nef}$ infected SupT1 WT (blue) and MSN-KO (red) cells. Supernatants were collected at 4, 5, 6 and 7 dpi. Individual points on the plot represent biological replicates and the bar indicates the median. A 2-sample Z test was conducted between WT and MSN-KO populations, n=10 in each population. A P value was calculated for the Z test to determine the significance. RMSE gives a relative error indication of difference, higher RMSE is indicative of greater difference. **B.** Percentage of live SupT1 cells at 7 dpi determined through flow cytometry of fixed cells using a viability stain. Significance determined with an unpaired t test. **C.** Percentage of HIV-$1_{mCherry\text{-}Nef}$ infected SupT1 cells at 7 dpi, determined by flow cytometry of fixed cells for mCherry fluorescence. Statistical significance was calculated with a Welch's t test for populations with unequal variances, ** signifies $P \leq 0.01$.

4.2.13 LCK-KO decreases the release of viral particles from infected cells

When conducted in LCK-KO cells, the P24 assay revealed LCK-KO results in a substantial decrease (p=0.00003) in the concentration of P24 in the supernatant of infected cells compared to WT (Figure 4.35A). In several replicates, LCK-KO supernatant had a P24 concentration of 0 ng mL^{-1}, although the median at 7 dpi was 0.40 ng mL^{-1}. It was possible that cell death contributed to a reduction in viral particle release as there was significantly more cell death (p=0.0053) in LCK-KO cells at 7 dpi (Figure 4.35B). However, if this were true it would be expected that the replicates with the highest percentage cell death would produce the lowest p24 concentration, which was not the case. There was also a substantial decrease (p=0.0037) in the percentage of infected LCK-KO cells at 7 dpi in comparison to WT cells (Figure 4.35C). This may be purely a result of the reduction in release viral particles or may suggest that cell-cell spread is also decreased in LCK-KO cells. Overall, this data suggests LCK may be involved in the assembly and/or release of HIV-1 viral particles from infected cells.

4.2.14 CIRBP-KO increases spread of Hlv-1 infection

In contrast to the other RBPs investigated, p24 ELISA analysis demonstrated there was a slight increase (p=0.1518) in p24 concentration in the supernatant of CIRBP-KO cells (Figure 4.36A). This was not a result of a change in cell viability as there was no significant difference in the percentage of live cells between CIRBP-KO and WT cells at 7 dpi (Figure 4.36B). The percentage of infected cells at 7 dpi was substantially higher (p=0.0431) in CIRBP-KO cells (Figure 4.36C) despite there being a small increase in viral particle release. This suggests that either CIRBP-KO cells are more permissive than WT cells to HIV-$1_{mCherry\text{-}Nef}$ infection, increasing the likelihood of productive infection, or CIRBP-KO cells have a higher level of cell-cell HIV-1 transmission than WT cells. Overall, these results suggest that the absence of CIRBP increases permissivity of cells to infection. Therefore, CIRBP may play an anti-viral role during HIV-1 infection.

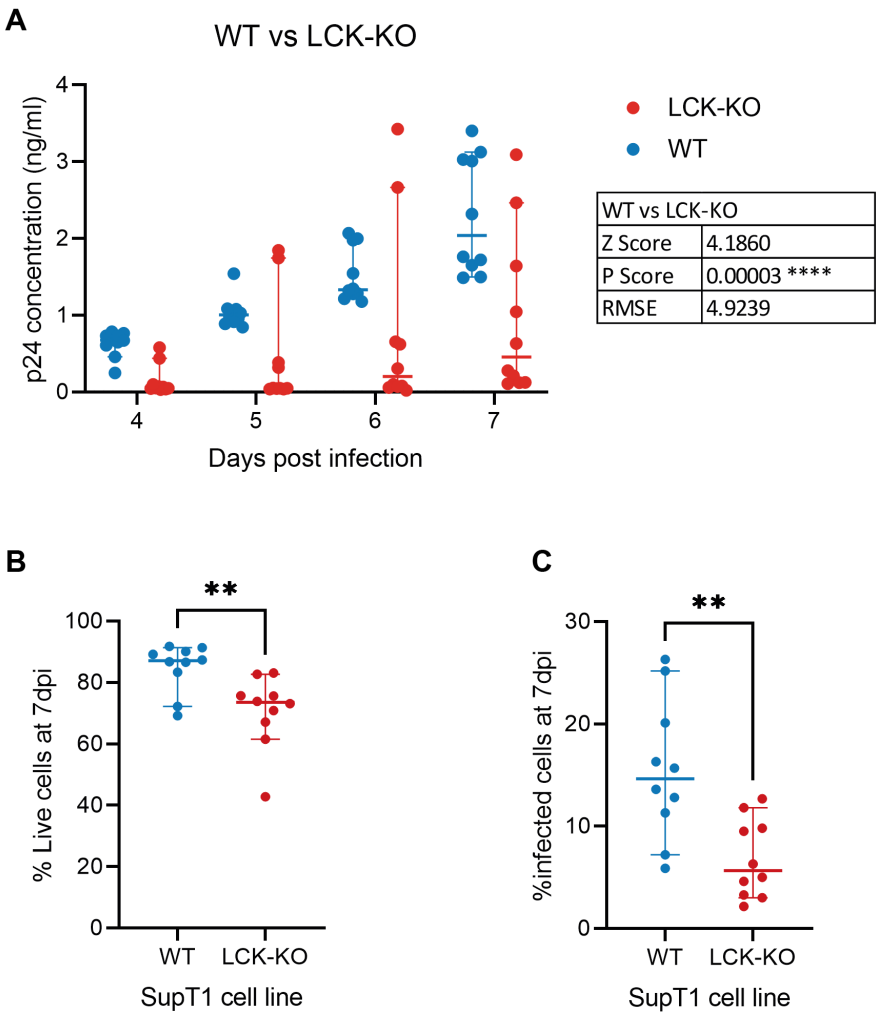

Figure 4.35: **LCK-KO decreases the release of viral particles from infected cells. A.** A P24 ELISA was conducted to measure the concentration of P24 in supernatants collected from HIV-$1_{mCherry\text{-}Nef}$ infected SupT1 WT (blue) and LCK-KO (red) cells. Supernatants were collected at 4, 5, 6 and 7 dpi. Individual points on the plot represent biological replicates and the bar indicates the median. A 2-sample Z test was conducted between WT and LCK-KO populations, n=10 in each population. A P value was calculated for the Z test to determine the significance. RMSE gives a relative error indication of difference, higher RMSE is indicative of greater difference. **B.** Percentage of live SupT1 cells at 7 dpi determined through flow cytometry of fixed cells using a viability stain. Significance determined with an unpaired t test. **C.** Percentage of HIV-$1_{mCherry\text{-}Nef}$ infected SupT1 cells at 7 dpi, determined by flow cytometry of fixed cells for mCherry fluorescence. Statistical significance was calculated with a Welch's t test for populations with unequal variances, ** signifies $P \leq 0.01$.

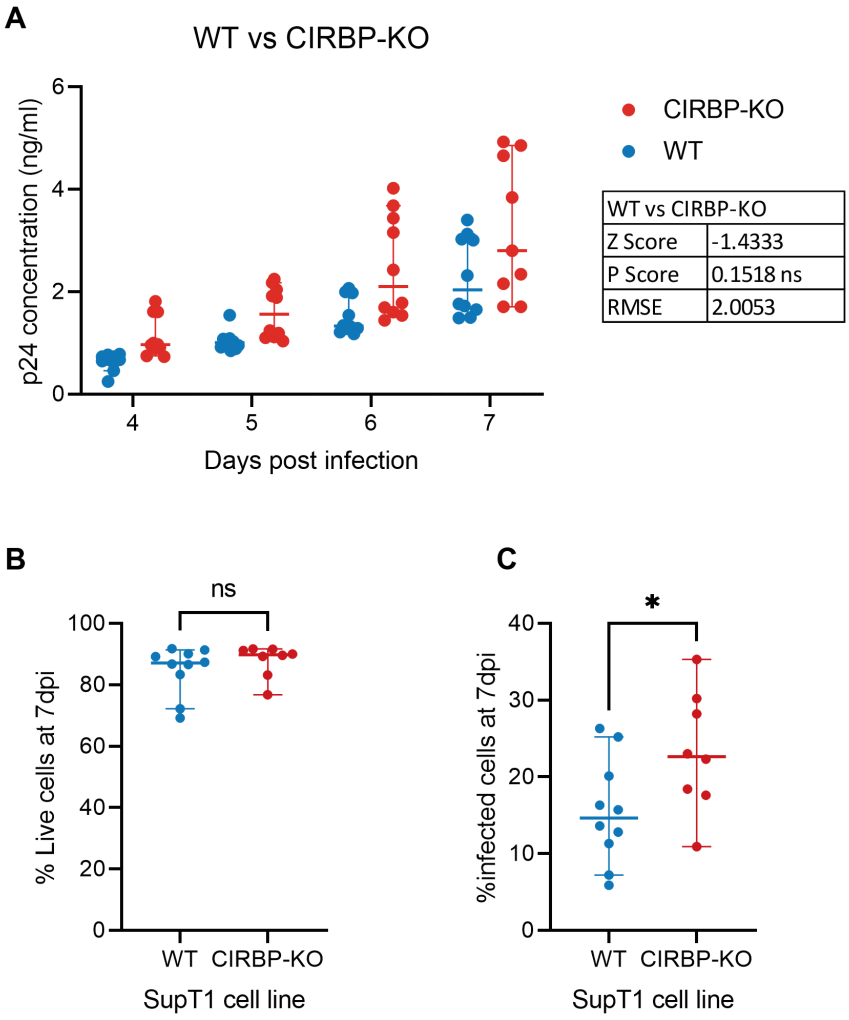

Figure 4.36: **CIRBP-KO does not effect the release of viral particles from infected cells but does increase percentage of infected cells. A.** A P24 ELISA was conducted to measure the concentration of P24 in supernatants collected from HIV-$1_{mCherry\text{-}Nef}$ infected SupT1 WT (blue) and CIRBP-KO (red) cells. Supernatants were collected at 4, 5, 6 and 7 dpi. Individual points on the plot represent biological replicates and the bar indicates the median. A 2-sample Z test was conducted between WT and CIRBP-KO populations, n=10 in each population. A P value was calculated for the Z test to determine the significance. RMSE gives a relative error indication of difference, higher RMSE is indicative of greater difference. **B.** Percentage of live SupT1 cells at 7 dpi determined through flow cytometry of fixed cells using a viability stain. Significance determined with an unpaired t test. **C.** Percentage of HIV-$1_{mCherry\text{-}Nef}$ infected SupT1 cells at 7 dpi, determined by flow cytometry of fixed cells for mCherry fluorescence. Statistical significance was calculated with a Welch's t test for populations with unequal variances, * signifies $P \leq 0.05$.

4.2.15 YBX3-KO

To determine the effect of YBX3-KO on viral particle release, the same assay was conducted in YBX3-KO cells. P24 analysis revealed HIV-$1_{mCherry\text{-}Nef}$ infected YBX3-KO cells released significantly fewer (p=0.00002) viral particles into the supernatant, as indicated by a decrease in p24 concentration (Figure 4.37A). This reduction was not caused by excessive cell death (Figure 4.37B). In agreement with reduced viral titre in the supernatant, there was a substantial decrease (p=0.0050) in the percentage of infected YBX3-KO cells at 7 dpi compared to WT SupT1s (Figure 4.37C). These results indicate viral particle release and possibly cell-cell spread is impaired in the absence of YBX3. Therefore, YBX3 positively impacts viral particle release in HIV-1 infection.

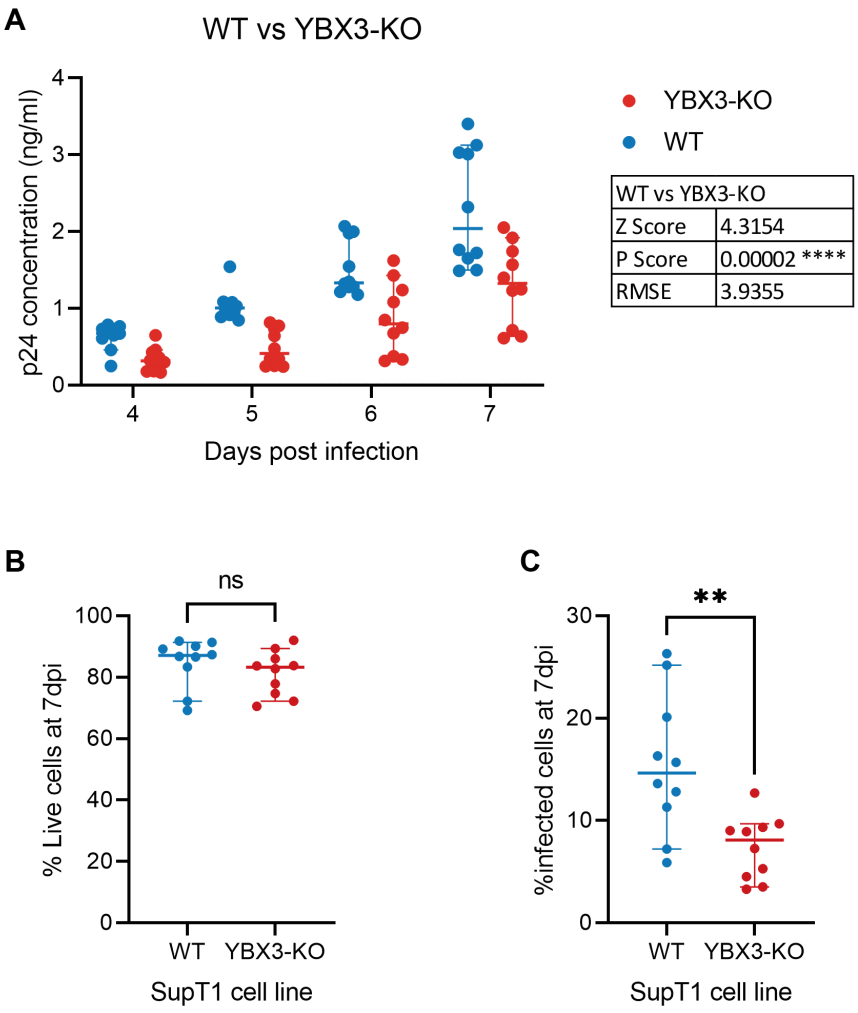

Figure 4.37: **YBX3-KO significantly decreases the release of viral particles from infected cells and rate of infection.** **A.** A P24 ELISA was conducted to measure the concentration of P24 in supernatants collected from HIV-$1_{mCherry\text{-}Nef}$ infected SupT1 WT (blue) and YBX3-KO (red) cells. Supernatants were collected at 4, 5, 6 and 7 dpi. Individual points on the plot represent biological replicates and the bar indicates the median. A 2-sample Z test was conducted between WT and YBX3-KO populations, n=10 in each population. A P value was calculated for the Z test to determine the significance. RMSE gives a relative error indication of difference, higher RMSE is indicative of greater difference. **B.** Percentage of live SupT1 cells at 7 dpi determined through flow cytometry of fixed cells using a viability stain. Significance determined with an unpaired t test. **C.** Percentage of HIV-$1_{mCherry\text{-}Nef}$ infected SupT1 cells at 7 dpi, determined by flow cytometry of fixed cells for mCherry fluorescence. Statistical significance was calculated with a Welch's t test for populations with unequal variances, ** signifies $P \leq 0.01$.

4.3 Discussion

Despite many studies showing the importance of cellular RBPs in the HIV-1 life cycle, we know relatively little about the functionality of RBPs within viral particles. This chapter provides new insights into the roles of these virion-incorporated RBPs in HIV-1 infection. I have shown through fluorescence, protein, RNA and flow cytometric analysis that each protein has a unique effect on HIV-1 replication and viral particle production (summarised in table 21). However, it was not fully established whether the viral particles produced were infectious and therefore, certain questions remain unanswered about the roles of these proteins within virions.

Before discussing each RBP in-depth, I will first address points relevant to all of the proteins. As mentioned, there is one outstanding question: are the viral particles produced in KO cells capable of productive infection? The limitation of the 7-day infection assay is that it does not provide sufficient detail on individual rounds of infection, only the overall infection over 7 days is observed. This indicates whether there is an altered level of infection. However, it cannot distinguish between an effect on viral particle release, relative infectivity of the viral particles or cell-cell transmission of virions. In cell-cell spread, which has been shown to account for 60% of viral transmission [240], infection could spread to a number of cells without releasing any viral particles into the supernatant, thus showing no effect in the p24 ELISA. Several approaches were trialled to determine the infectivity of KO viral particles. However, none were optimised within the time frame of this project. This was unfortunate as it would support (or oppose) the hypothesis that ivRBPs are incorporated into viral particles to aid early infection steps and is something that I would like to continue developing.

Secondly, whilst pseudotyped single-round viruses are an effective way to study HIV-1 infection, there are differences in their replication cycle which may affect the results. As pseudotyped viruses have VSVG embedded within their outer membrane, they do not mirror the binding of HIV-1 envelope proteins to the CD4 and chemokine co-receptors to initiate fusion with the cell membrane. VSVG-enveloped viruses instead enter the target cell following recognition of the LDL receptor, internalisation

into endosomes and subsequent escape from endosomes [241]. Whilst HIV-1 entry via endocytosis has been documented, the capsid cores are degraded and do not lead to productive infection [242]. Additionally, because VSVG-enveloped viruses deposit the capsid core wherever they escape from endosomes, their trafficking to the nucleus differs from that of plasma membrane-deposited capsids. These differences could establish an interaction with the pseudotyped virus that would not be relevant to HIV-1 infection. Equally, interactions with HIV-1 that do not occur with the psueodtyped virus may be overlooked. However, the experiments conducted in section 4.2, focus on downstream steps of HIV-1 replication post-integration so should not be strongly influenced by the differences in cell entry and trafficking. Whilst it has been validated that integration, transcription and translation occur with HIV-$1_{\text{-R-E-mCherry-Nef}}$, future work could improve upon this system by conducting all experiments with non-pseudotyped viral particles.

Another limitation of the approach taken to studying viral protein and RNA expression is that analysis was done on a population level rather than individual cells. Therefore, it is not possible to confirm whether an effect is due to altered viral replication within cells or altered susceptibility to infection. Given that virion-incorporated RBPs are hypothesised to affect the second round of infection, susceptibility to infection is likely to drive any phenotype observed in the first round of infection. However, it should not be ruled out that an RBP may have a function both within the cell and within the capsid during infection. Thus, it is critical that future investigation into any of the candidates should involve exploration of whether viral replication specifically is affected. One way this could be achieved is by using single-molecule fluorescence in situ hybridisation (smFISH) to uncover any effect on viral gene transcription.

The work in this chapter was also limited by the use of single clones. Poor KO efficiency was achieved for MKRN1 so pools of CRISPR-edited cells could not be used as unedited cells may have masked any effect from the KO. Therefore, single clones were isolated for each RBP-KO and only one clone was selected due to time pressures. Minute variations between single cells isolated from a cell line and off-

target effects from CRISPR/Cas9 gene editing mean that individual clones can behave very differently. Thus, a phenotype observed in one KO clone may not be observed in another clone of the same KO. To overcome this, multiple clones of each KO should be used to determine phenotype. Further investigation into any of the candidates studied in this chapter requires confirmation of the phenotype in additional clones.

Table 21 highlights the differences in phenotypes observed for each KO and demonstrates the complexity of interaction between HIV-1 and each RBP. It is plausible that some candidates have a multifactorial role within infection that is responsible for the differences in results between time points and stages of infection. Because of the complexity of the results, each candidate will be discussed individually below.

Table 21: **Summary of effect of ivRBP-KO on HIV-1 infection.** Colours indicate the effect on a particular stage of infection: red for increase, blue for decrease, white for no effect and orange for a mixed effect. Results that were significant are indicated by *.

ivRBP	**Effect on viral protein expression**	**Effect on vRNA expression**	**Effect on viral particle release**
MKRN1-KO	Increased*	Increased	Decreased*
FAM120A-KO	Decreased*	Decreased*	No effect
LSM14A-KO	No effect	Delayed for unspliced vRNA*	Decreased*
MSN-KO	Increased Gag*, not Nef	Delayed for all vRNAs	Decreased*
LCK-KO	Increased Nef, decreased Gag*	Increased completely spliced vRNA	Decreased*
CIRBP-KO	Increased Nef, decreased Gag	Delayed for all vRNAs	Increased
YBX3-KO	Increased*	Delayed for all vRNAs*	Decreased*

4.3.1 MKRN1 may be required in viral particle budding and release

The existing literature for the role of MKRN1 in viral infection suggests it plays an anti-viral role by ubiquitinating the capsid proteins of various viruses and causing premature degradation of the capsid [161,162,164]. Therefore, it was expected that HIV-1 protein and vRNA levels would be increased in MKRN1-KO SupT1s. Cell growth rate and viability was shown to be the same in WT and MKRN1-KO CD4+ T cells. Thus, these increases were not a result of faster growth of MKRN1-KO cells. It was, however, unexpected that fewer viral particles would be released from infected MKRN1-KO cells. The latter experiment suggests MKRN1 is required in some way for the release of viral particles. Inhibiting the release of viral particles would also explain the intracellular increase in viral protein and RNA. Rather than an increased ability of HIV-1 to replicate in MKRN1-KO cells driving the increased level of protein and RNA, it is the decreased ability of viral particles to leave MKRN1-KO cells that causes viral protein and RNA to accumulate in the cytoplasm. Therefore, I hypothesise that MKRN1 is required for HIV-1 viral particle release.

The function of MKRN1 as an E3 ubiquitin ligase makes this hypothesis particularly interesting. HIV-1 relies on the host ESCRT complex for release from cells and this process is dependent on the ubiquitination of Gag [243,244]. It is thought that Nedd4-like ubiquitin ligase is responsible for Gag ubiquitination as overexpression of Nedd4-2S rescues budding defects when Gag is de-ubiquitinated [245,246]. However, neither KO of MKRN1 nor Nedd4-2s completely abolishes viral particle release so perhaps both proteins are capable of regulating the process.

In addition to the ubiquitination of Gag, the ubiquitination of ALIX is also required for HIV-1 release. ALIX interacts with the $LYPX_nL$ motif in the P6 domain of Gag and the NC domain of Gag, and recruits Nedd4 to facilitate HIV-1 release via a mechanism involving ALIX ubiquitination [247]. Thus, another possibility for the involvement of MKRN1 in HIV-1 infection is that MKRN1 ubiquitinates ALIX or another protein necessary for HIV-1 budding.

A possible method of recruitment of MKRN1 to sites of viral assembly at the plasma membrane could be through interaction with vRNA. MKRN1 is known to bind to A-

rich sequences upstream of poly(A) tails in cellular mRNAs in a PABPC1-dependent manner [160]. HIV-1 RNA is polyadenylated and we also find PABPC1 bound to vRNA in our ivRBPome. Therefore, it is possible that MKRN1 interacts with the poly(A) tail of HIV-1 RNA to be localised to Gag and the ESCRT machinery.

4.3.2 FAM120A

Unlike MKRN1, there was no previously described role for FAM120A in viral infection. I hypothesised that FAM120A may be involved in controlling HIV-1 gene expression or trafficking of the vRNP based on the cellular function of FAM120A. Regarding viral gene expression, there was a significant decrease in completely spliced vRNA but not single-spliced or unspliced vRNA, which suggests FAM120A may be involved in nuclear export of completely spliced transcripts. These completely spliced RNAs are exported in a Rev-independent manner and instead use the NXF1 nuclear export pathway [85]. Whereas single-spliced and unspliced transcripts are exported in a Rev-dependent manner. No decrease in single-spliced and unspliced vRNAs is observed. However, there is a decrease in expression of both Nef and Gag at a protein level. Because RNA analysis is performed on whole cell lysate, there is no way to distinguish between RNA expression in the nucleus and the cytoplasm. It is possible that single-spliced and unspliced vRNAs are transcribed, yet their export from the nucleus is impaired, consequently, decreasing translation of these transcripts. However, no prior role in nuclear export has been described for FAM120A.

The decrease in Nef expression could further exacerbate the effect induced by FAM120A-KO as Nef regulates many aspects of HIV-1 infection. One particular role of Nef is in the recruitment and activation of Src family kinases [248]. Activation of Src kinases by Nef is an essential step in the downregulation of MHC class I molecules in infected cells, to evade the immune response [249]. FAM120A is required for the activation of many Src family kinases [173], therefore the absence of both FAM120A and Nef could increase the cellular immune response to HIV. To add to the increased immune surveillance within the cell, FAM120A is also involved in miRNA-mediated silencing pathways [174]. Cellular miRNAs are known to downregulate host anti-viral

proteins in HIV-1 infection [180, 181]. Thus, in the absence of FAM120A, miRNAs may not be directed to silence the anti-viral response. The expression of Gag polyprotein was also decreased in FAM120A-KO cells. Like the change in Nef expression, this could be due to altered immune surveillance in the cell. However, Gag expression may be inhibited by impaired trafficking to the plasma membrane. FAM120A has been shown to associate with motor proteins to modulate cell surface trafficking [172] and also interacts with proteins that are core components of the HIV-1 RNP [129] therefore trafficking may be disrupted by the absence of FAM120A.

Given the previous results, it was unexpected that there would be no effect on viral particle release. Whilst we assume the P24 concentration of the supernatant reflects the number of viral particles released, from this experiment alone, we have no way of knowing whether the viral particles are properly assembled or infectious. It is possible that FAM120A-KO causes defects in the packaging of the vRNP inside assembling virions. However, more data on infectivity of viral particles is required before any conclusions can be drawn from this.

4.3.3 LSM14A

Given its role in mRNA degradation in P-bodies and as an innate immune sensor for vRNA [189], I had predicted LSM14A would play an anti-viral role in HIV-1 infection by promoting the degradation of vRNAs and innate immune signalling. Therefore, it was unanticipated that LSM14A-KO delays initial expression of completely spliced RNAs. Although, there is no described role in the literature that suggests a mechanism by which LSM14A may promote early gene expression in HIV-1. It is possible that the slightly slower growth shown for LSM14A-KO cells causes this delay in initial gene expression. Further testing on metabolic rate of LSM14A-KO cells is required before conclusions can be drawn.

In contrast, the increased gene expression at 48 hpi observed in the absence of LSM14A is in agreement with the literature and suggests LSM14A may be involved in the degradation of vRNAs [250]. The increase in expression of vRNAs at 48 hpi may recover protein expression, hence why no difference is observed at the protein

level. This lack of effect on protein expression was also unexpected. However, it is possible that immune-evasion mechanisms utilised by HIV-1 circumvent LSM14A immune surveillance. HIV-1 recruits the cellular protein Staufen1 to viral RNPs to inhibit the formation of stress granules induced by the NC domain of Gag [251]. It may be that the action of Staufen1 protects the HIV-1 RNP from LSM14A in P-bodies and stress granules. Additionally, LSM14A-deficient mice were shown to have downregulation of the anti-viral innate immune signalling protein STING [176]. If this also occurs in human CD4+ T lymphocytes, the reduction in anti-viral signalling may maintain cell permissivity to HIV-1.

Whilst there is a significant decrease in p24 concentration of supernatants from infected LSm14-KO cells, there is no change in percentage of infected cells and a significant decrease in cell viability. Given the trend in the majority of the data, it is likely cell death contributed to the decrease in p24 release, rather than being a direct result of LSM14A-KO. It is clear that the interaction of LSM14A in HIV-1 infection is complex and possibly multi-factor. The results presented here are not sufficient to draw a conclusion and further investigation into LSM14A in infection is required.

4.3.4 MSN

Existing literature implies MSN has a multi-factorial role in HIV-1 infection. Most of the effect on HIV-1 infection stems from the regulation of the cell cytoskeleton and plasma membrane by MSN [191] and in turn, the clustering of immune receptors [194, 195]. Hence, it would be logical for MSN-KO to inhibit cell entry, trafficking of the capsid or viral particle release because of the dependence on the cytoskeleton, especially as disruption to the cytoskeleton has been shown to inhibit reverse transcription [252]. Therefore, it was surprising that MSN appears to differentially affect HIV-1 proteins instead of affecting the whole viral proteome. MSN has the strongest effect on Gag and unspliced RNA, with MSN-KO increasing the expression of both. It is possible that the use of pseudotyped virus for these experiments has affected the phenotype seen in MSN-KO cells as VSVG-mediated cell entry does not use the same cytoskeletal pathways as HIV-1 [241]. Change in expression of just one protein could suggest

differences in transcription, translation or nuclear export. However, MSN has no reported involvement in these processes as vRNA travels through the cytoplasm mainly by diffusion [253, 254].

It is, however, understandable that viral budding and the spread of infection is significantly decreased in MSN-KO cells. The actin cytoskeleton has been shown to move the vRNP and other proteins to cytoplasmic assembly sites [62]. Additionally, actin depolymerising agents have been shown inhibit HIV-1 assembly [255]. The interaction of HIV-1 Gag and actin in infected cells was also shown to enhance cell motility and promote dissemination of viral particles [256]. The absence of MSN could disrupt this interaction with actin and prevent proper assembly of the virions. Another ERM family member, ezrin, has been shown to be crucial for efficient virion release from HIV-1 infected cells [257]. However, no role in budding has been described for MSN. An additional potential mechanism for the decrease in viral spread is through the disruption of cell surface receptors required to form virological synapses. MSN is known to positively regulate CXCR4 expression and is required for HIV-1 induced CD4-CXCR4 clustering [195]. Knockdown of MSN was also shown to decrease transduction titres of X4-tropic HIV-1 and knockdown of all ERM proteins decreased efficiency of cell-cell fusion mediated by HIV-1 [200].

The results presented here support the current literature describing MSN involvement in HIV-1 infection. Furthermore, I have potentially uncovered a novel role for MSN in regulating the assembly and budding of HIV-1.

4.3.5 LCK

The Src family tyrosine kinase LCK has a well-established pro-viral role in HIV-1 infection. The majority of these roles involve activating various signalling pathways to prevent cells from entering an anti-viral state [206] and promoting the expression of pro-viral signalling factors [207, 210, 211]. It was unsurprising that LCK-KO resulted in a small but significant decrease in Gag protein expression and reduced expression of completely spliced and single-spliced vRNAs at 48 hpi. This suggests the absence of LCK makes the cell less permissive to viral replication which was also shown by

Liu et al [203]. In contrast to this, Nef protein expression was increased in LCK-KO cells and the expression of completely spliced vRNA was initially upregulated at 24 hpi compared to WT. Nef is known to trigger intracellular accumulation of LCK to inhibit TCR signal transduction [210]. Therefore it is interesting that the absence of LCK increases HIV-1 Nef. It is possible that the loss of LCK-mediated pro-viral signalling requires an increase in Nef to maintain cell permissiveness to HIV-1 infection.

Also in agreement with the literature is the effect of LCK-KO on viral particle release. Here I showed that the absence of LCK significantly reduces the release of HIV-1 particles over 7 days, which reinforces the findings of Strasner et al [213]. This is likely to be caused by a disruption in viral assembly as the palmitoylation domain of LCK is required for targeting of Gag to the plasma membrane [213]. However, it should be noted that LCK-KO cells grew slower than WT SupT1s and this may contribute to the reduction in early gene expression, Gag protein synthesis and viral particle release. Viability analysis in LCK-KO cells was not performed by Strasner et al for comparison.

Whilst no novel interactions were identified for LCK, the results presented in this chapter underpin the current literature surrounding LCK in HIV-1 infection. Although, further investigation could be conducted to uncover the mechanism behind the increase in Nef in the absence of LCK.

4.3.6 CIRBP

I had hypothesised that CIRBP would be a pro-viral protein based on its role in inducing the expression of pro-inflammatory cytokines which promote HIV-1 infection and its ability to stabilise mRNA. However, the results presented in this chapter do not support a pro-viral role for CIRBP in HIV-1 infection. The effect of CIRBP-KO on vRNA was delayed; expression of all variants was slightly low at 24 hpi but was increased compared to WT at 48 hpi. The greatest increase was seen in single spliced vRNAS. However, as protein expression for the viral proteins encoded by single-spliced RNA was not measured, it is unknown whether any effect would be seen at the protein level. The initial decrease in vRNA expression may be a result of the slightly slower growth of CIRBP-KO cells compared to WT SupT1s. Despite this, expression of HIV-1 Nef

was increased in CIRBP-KO cells. Additionally, the release of viral particles and the permissivity of cells to infection was also increased in the absence of CIRBP which suggests CIRBP negatively regulates these processes. Whilst no link between HIV-1 and CIRBP has been uncovered, this result is inconsistent with current literature reporting a pro-viral role for CIRBP in flavivirus infection. Gokhale et al revealed that knockdown of CIRBP inhibited DENV, ZIKV and HCV infection. They found that innate immune signalling and ER stress pathways triggered by flavivirus infection decreased M6-adenylation of *CIRBP*, leading to increased alternative splicing of CIRBP and promotion of pro-inflammatory signalling [220]. Whilst the pro-inflammatory signalling induced by CIRBP may be beneficial for flavivirus infection, the regulation of infection by these pathways is often highly complex and affects individual viruses differently. Some pro-infammatory cytokines are known to stimulate HIV-1 infection (eg TNFα, IL-1β, IL-6). However, interferons α, β and γ have all been shown to inhibit HIV-1 infection [258–260]. It is possible that in HIV-1 infection the pro-inflammatory cytokines promoted by CIRBP lead to the inhibition of infection. Therefore, the work presented here may have highlighted a novel role for CIRBP as a negative regulator of HIV-1 and warrants further investigation to uncover the mechanism behind this control.

4.3.7 YBX3

Despite YBX3 being the most enriched protein in ivRIC eluates, YBX3-KO only has a small, mixed effect on HIV-1 infection. At the protein level YBX3-KO significantly, but only slightly, increased the expression of Nef and the Gag polyprotein in HIV-$1_{\text{-R-E-mCherry-Nef}}$ infection. This would indicate YBX3 acts as an anti-viral protein, which is supported by the anti-viral action of YBX3 described in IAV infection [236]. RNA expression analysis in YBX3-KO cells, however, implies an absence of YBX3 delays expression of the three different splice variants of vRNA, which suggests YBX3 is required for the initial expression of these transcripts. eCLIP analysis revealed YBX3 binds to a broad range of cellular mRNAs without a specific binding motif [234] and RNA immunoprecipitation analysis has shown interaction with over 80% of the cell transcriptome [261]. Therefore, the promiscuous binding activity of YBX3 may enable

it to interact with HIV-1 RNA.

Interestingly, YBX3-KO significantly impaired the spread of HIV-1. Both the concentration of P24 in supernatant collected from YBX3-KO cells and the percentage of infected cells were decreased in YBX3-KO compared to WT. This was slightly surprising as YBX3 is known to regulate transcription of genes involved in cell survival, proliferation and differentiation [224–226]. Whilst dysregulation of these genes may not have resulted in cell death in these experiments, it is possible that YBX3 controls transcription of other genes involved in viral infection which decreases budding. Although, this has not been reported. There was no difference in percentage of cell death between WT and YBX3-KO cells during infection. However, the proliferation assay performed on uninfected SupT1 KOs suggested that YBX3-KO grows slower than WT cells and this may have contributed to the reduction in viral particle release.

4.3.8 Conclusions

The data presented in this chapter demonstrates that host RBPs that interact with HIV-1 RNA inside viral particles play roles in the first round of HIV-1 infection. From the pool of ivRBPs that were studied, I have identified both pro- and anti-viral regulators of HIV-1 infection. Interestingly, transmission of HIV-1, either by free virus or cell-cell spread, was significantly reduced in MKRN1-KO, MSN-KO, LCK-KO and YBX3-KO SupT1s. Conversely, CIRBP-KO significantly increased HIV-1 transmission, indicating that anti-viral regulators were also identified by ivRIC. However, the infectivity of viral particles produced in KO cells was not directly determined and therefore it cannot be concluded whether these proteins are packaged into virions for the next round of infection. Further work to determine the infectivity of KO-virions and the role of the candidates within the second round of infection is required.

Whilst many of the RBPs studied have interesting effects on HIV-1 infection that warrant further investigation, only one could be thoroughly investigated within the time frame of this DPhil. I chose to continue exploring the role of MKRN1 in HIV-1 infection as the phenotype observed in MKRN1-KO was robust and my hypothesis that MKRN1 regulates viral particle release from the plasma membrane would be a novel

and exciting contribution to the field. Chapter 5 continues exploring the relationship with MKRN1 and HIV-1. However, if given more time, there are a few ways in which I would improve and build upon the research in this chapter. These are discussed in the section below.

4.4 Future work

The experiments conducted in this chapter give a good indication of whether the ivRBPs investigated have any role in HIV-1 infection. However, simple improvements could be made which would reveal more about the function of each ivRBP. Firstly, all experiments were conducted in SupT1 clones. Whilst this is a physiologically relevant CD4+ T lymphocyte line, it is a cancerous cell line which can harbour a number of unknown mutations which may contribute to the observed phenotypes. Further investigation of these proteins in primary CD4+ T lymphocytes and macrophages would provide irrefutable evidence that these interactions truly occur in HIV-1 infection. As a minimum, experiments need to be repeated in additional SupT1 clones to account for clonal variation.

Additionally, the use of fully-competent HIV-1 would avoid any potential discrepancy in results between virus and VSVG pseudotyped single-round virus, due to the differences in their life cycles. It is possible that interactions may not have been observed due to the different method of cell entry used by VSVG enveloped viruses [241]. Whilst this would not have affected all the ivRBPs investigated, LCK, MSN and LSM14A are all linked to the endosomal network [190, 262, 263]. An Env-expression vector can be pseudotyped with an Env-defective backbone vector to produce single-round virions with Env glycoproteins [264, 265]. This would enable the study of infection with CD4-mediated cell entry under BSL2 conditions. However, challenges arise with low infection efficiency using the CD4 pathway. Additionally, the safety of Env-coated single-round viruses has also been questioned [266]. Therefore, they may be no better than using WT HIV-1. The use of pseudotyped viruses does not mean the research presented in this chapter is irrelevant, but if LCK, MSN or LSM14A were explored further, they would benefit from the use of WT HIV-1.

The fluorescence assay conducted with HIV-$1_{\text{-R-E-mCherry-Nef}}$ is a robust technique for modelling viral protein expression. Although, it was limited in this case to studying only one protein, Nef. The Castello lab has also established HIV-$1_{\text{-R-E-Gag-mCherry}}$ and HIV-$1_{\text{-R-E-IN-Scarlet}}$ (see chapter six). It would be interesting to determine whether the same effect is observed across different HIV-1 proteins and would provide further support for the role of the ivRBP within infection. This experiment would also benefit from additional controls to ensure any effect of the KO on metabolic rate does not influence the results. LSM14A-KO, LCK-KO, CIRBP-KO and YBX3-KO were shown to grow slightly slower than WT SupT1 cells. Therefore, expressing a second non-viral fluorescent protein throughout infection would enable the measurement of cellular protein expression in addition to viral replication. Additionally, western blot analysis could also be conducted for all of the HIV-1 proteins, not just the Gag polyprotein, to highlight whether expression of particular proteins is affected.

The RT-qPCR analysis performed also provided valuable insight into the effects of ivRBP-KO on completely spliced, single-spliced and unspliced vRNAs. However, because RNA was purified from whole cell lysate, it is unknown whether localisation of the vRNAs was affected. By performing subcellular fractionation to isolate nuclear, cytoplasmic and and purifying RNA from each fraction, localisation of particular RNAs could be determined. This would be of particular importance to RBPs that may have a role in nuclear export of vRNAs.

In addition to the P24 ELISA to measure viral particle release, I also wanted to analyse the infectivity of the viral particles produced in ivRBP-KO cells. This project was established from the hypothesis that RBPs required for early stages of infection are pre-assembled into viral particles to gain access to the vRNA. Whilst my research has shown the effects on later stages of HIV-1 infection, I have only been able to start uncovering the role these proteins play inside the capsid. To really understand the interaction we need to first determine whether viral particles produced in KO cells are capable of productive infection. Several attempts were made to study the infectivity of the KO viral particles. However, I was unsuccessful within the time frame of the project due to the poor efficiency of infection.

There are several different approaches that could be taken to assess infectivity. Firstly, the supernatant from HIV-1 infected KO cells could be collected and added to uninfected cells. The use of a fluorescently-tagged HIV-1 would enable infected cells to be quantified via flow cytometry. Similarly, the supernatant could be added to a HIV-1 reporter cell line such as TZM-bl. These are a HeLa-derived cell line engineered to express CD4 and CCR5 receptors with endogenous expression of CXCR4. They express luciferase under the control of a HIV-1 Tat promoter, thus luminescence can be used to quantify productive HIV-1 infection. In addition to studying free-virus infection, cell-cell transmission could be studied with co-culture experiments. Fluorescent HIV-1 infected KO cells could be co-cultured with uninfected WT cells and the spread of virus could be determined by measuring the proportion of fluorescent cells via flow cytometry. Alternatively, a more permissive cell line could be used to produce the viruses. Whilst SupT1 CD4+ T lymphocytes are a physiologically appropriate cell line in which to study HIV-1 infection, they produce poor viral titres. Generating the KO cells in a more permissive cell line such as HEK293 may enable the production of sufficient viral titres to study KO-viral particle infectivity.

I believe the addition of these experiments would further our understanding of the roles of the ivRBPs in HIV-1 infection. However, given the strong phenotype observed in MKRN1-KO cells, further experimentation on the other RBPs was not within the scope of this project and the focus was directed to MKRN1.

5 The one RING ligase to rule them all: characterisation of the role of MKRN1 in HIV-1 infection

5.1 Introduction

HIV-1 virion assembly and budding are tightly regulated processes driven by the Gag polyprotein but require multiple host cell factors to complete. Based on the results in Chapter 4, I hypothesised that MKRN1 may be a novel regulator in this stage of the viral life cycle. In this chapter, I aim to test this hypothesis by further exploring the effect of MKRN1-KO on viral budding with microscopic analysis and elucidating the interactome of MKRN1 in HIV-1 infection.

5.1.1 HIV-1 budding and the ESCRT pathway

HIV-1 Gag is responsible for targeting the vRNP to the plasma membrane, recruiting the viral and cellular proteins required and packaging the vRNA, all of which are thought to occur simultaneously. The different domains of Gag (MA, CA, NC and p6) mediate the different processes involved. MA^{Gag} is responsible for targeting the vRNP to specialised microdomains in the inner leaflet of the plasma membrane known as lipid rafts. These are rich in cholesterol, sphingolipids and phosphatidyl inositol (4,5) bisphosphate ($PI(4,5)P_2$) which are crucial for virus production [267]. Targeting lipid rafts is achieved through a bipartite signal comprising of an N-myristol fatty acid modification and a $PI(4,5)P_2$ binding site in MA. The N terminal MA^{Gag} is also responsible for the recruitment of the viral Env protein in the plasma membrane as the virions bud from the cell surface [268]. The CA^{Gag} polypeptide mediates the protein-protein interaction required to induce bending of the membrane to form the spherical virion. Polymerisation of Gag does not extensively occur before reaching the plasma membrane. Instead it arrives at the membrane as monomers or dimers which can fold in compact conformations enabling the MA-membrane, NC-vRNA and Gag-Gag interactions [269, 270]. As Gag polymerises at the membrane it forms a hexagonal

lattice with regular defects that allows the lattice to bend and gives it an outward curvature [271]. The NC^{Gag} domain contains two ZnF motifs which capture the viral gRNA and $tRNA^{Lys3}$ to prime reverse transcription in the future capsid. Finally, the C terminal $p6^{Gag}$ domain recruits the viral accessory protein Vpr and contains two 'late assembly (L-)domains' which interact with ESCRT cofactors [268].

The cellular ESCRT machinery is required for termination of Gag polymerisation and virion release. In non-infected cells, ESCRT facilitates membrane fission reactions that release vesicles and separate daughter cells during cytokinesis [272] and is therefore repurposed to conduct virion membrane fission. The human ESCRT machinery consists of four different ESCRT sub-complexes (ESCRT-0, ESCRT-I, ESCRT-II and ESCRT-III) which are made up of over 30 different proteins. These complexes include multiple cofactors like TSG101, ALIX, VPS4 and ubiquitination machinery [273]. However, ESCRT-0 and ESCRT-II are dispensable for HIV-1 budding [274, 275].

Recruitment of early-acting ESCRT factors to the budding virion is facilitated by $p6^{Gag}$. $P6^{Gag}$ has two L-domains: PTAP in the N terminus and $LYPX_nL$ in the C terminal α-helix [276]. Both L-domains substantially contribute to HIV-1 replication. However, $LYPX_nL$ is less critical [277]. The PTAP L-domain mimics cellular ESCRT-1 recruiting motifs and recognises the ESCRT-I complex member TSG101 [278], whilst the $LYPX_nL$ L-domain recruits the ESCRT cofactor ALIX [279]. In addition to $p6^{Gag}$ recruitment of ALIX, interaction between the Bro domain of ALIX and NC^{Gag} has also been reported [133].

After recruitment of ESCRT-I, the remaining ESCRT complexes assemble to mediate membrane fission and recycling of ESCRT factors [272, 280]. In normal ESCRT function, ESCRT-II bridges the ESCRT-I and ESCRT-III complexes. However, in HIV-1 budding, ESCRT-I and ALIX mediate the recruitment of ESCRT-III and the cofactor VPS4 to the viral budding site [274, 275, 281]. Humans have 12 different ESCRT-III proteins known as chromatin-modifying proteins/charged multivesicular body proteins (CHMP). However, only CHMP2 and CHMP4 are required for HIV-1 budding [97]. Binding of ALIX to the $LYPX_nL$ motif induces a conformational change

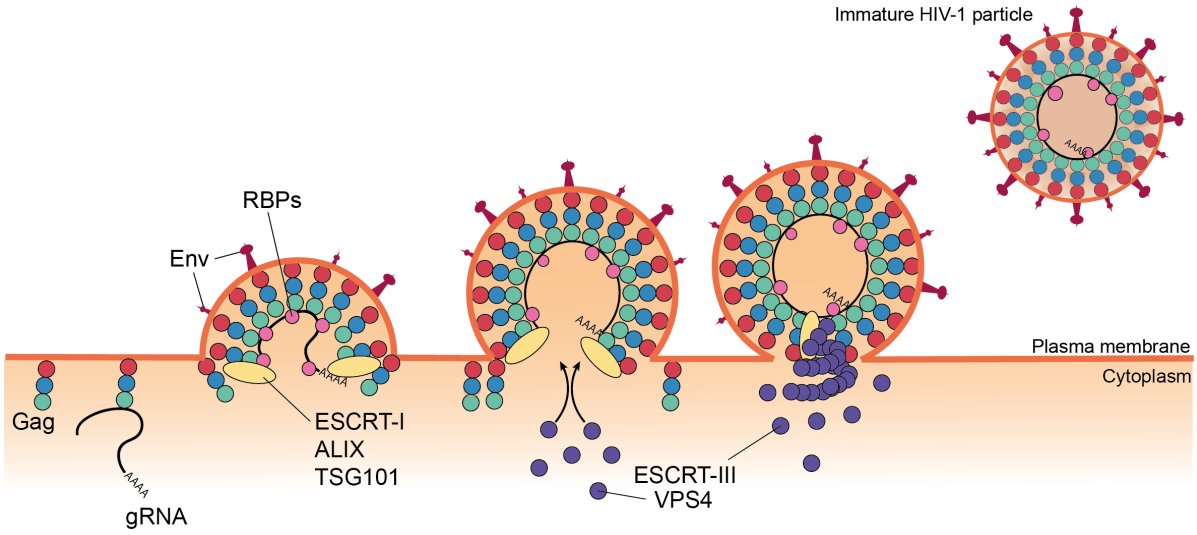

Figure 5.1: **Mechanism of HIV-1 budding.** Simplified schematic based on models of HIV-1 budding by Freed [92] and Sundquist et al [268]. HIV-1 Gag (MA red, CA blue, NC green) interacts with the ESCRT-I complex including ALIX and TSG101 (yellow), which recruit ESCRT-III (purple). Membrane scission is driven by ESCRT-III polymerisation with the ATPase VPS4. Shown here is the 'dome' model with ESCRT-III forming a spiral with consecutively narrower rings that pulls the membrane together.

that enables ALIX to dimerise and recruit CHMP2 and 4. Interaction with ALIX relieves autoinhibition of CHMP4, inducing polymerisation of CHMP4 filaments within the neck of the budding virion. The CHMP4 filaments recruit CHMP2, which in turn recruits the type I ATPase VPS4. VPS4 proteins form enzymatically active hexameric rings which stack on top of each other, with each budding virion requiring three to five VPS4 dodecamers [282]. The ESCRT assembly process occurs in multiple stages with the accumulation of Gag and ALIX at the budding site, taking approximately 10 minutes for the entire complex to assemble [282, 283].

There are two models for how ESCRT-III conducts membrane fission. The first is the 'dome' model [284] which proposes that the ESCRT-III complex forms a spiral with consecutively narrower rings to give the shape of a dome. Interactions between the membrane and the ESCRT dome pull the membrane with the curvature of the dome until the neck of the budding virion closes and scission can occur. The second, more recent, model is the 'buckling' model [285] which suggests the ESCRT-III proteins assemble in a cone-like structure which spirals outwards forming consecutively larger

rings. This spiral grows away from the budding neck and triggers an instability which flattens the cone in on itself and then scission occurs [286]. After membrane scission the immature viral particle is released and VPS4 catalyses ATP hydrolysis to disassemble the ESCRT complex [287].

5.1.2 The role of ubiquitin and ubiquitin ligases in HIV-1 budding

The first indication that ubiquitin ligases were involved in HIV-1 budding came from the observation that unconjugated ubiquitin is enriched in HIV-1 particles [171]. Further investigation confirmed that ubiquitin played a role in HIV-1 budding and that depletion of pools of cellular ubiquitin inhibited viral particle release [243,288,289]. Over the past two decades, the mechanism by which ubiquitin facilitates HIV-1 budding and release has slowly been uncovered.

HIV-1 Gag is ubiquitinated near both NC^{Gag} and $p6^{Gag}$ domains and occurs in a manner dependent on the association of Gag with the plasma membrane [290]. Blocking this ubiquitin conjugation or mutating the lysine residues of Gag inhibits HIV-1 release [291]. Furthermore, overexpression or targeting of the Nedd4-like ubiquitin ligase, Nedd4-2, to sites of virion assembly overcomes budding defects induced by Gag mutation [245, 292, 293]. The Nedd4 proteins are a family of HECT ubiquitin ligases with conserved C2 domains that mediate plasma membrane localisation and function in the budding of many retroviruses [294–297]. The ubiquitination of Gag is also required for the recruitment of both TSG101 [298,299] and ALIX [300–302]. During IFN-mediated inhibition of HIV-1 replication, the IFN-induced ubiquitin-like protein ISG15 inhibited the ubiquitination of Gag and consequently, the recruitment of TSG101 and viral budding [303]. Both TSG101 and ALIX are also themselves ubiquitinated [247, 302, 304]. However, in the absence of TSG101 and ALIX ubiquitination, there were no observed budding defects, provided Gag was ubiquitinated [244].

Most retroviruses recruit the host ESCRT components through an additional L-domain motif, PPXY, which interacts with Nedd4-like ubiquitin ligases. Although HIV-1 Gag lacks a PPXY motif, it engages Nedd4 family members Nedd4-2 and Nedd4-

1 to facilitate virus release [247]. This recruitment mechanism has not been fully uncovered but Nedd4-1 has been shown to interact with ALIX through the $LYPX_nL$ L-domain in a ubiquitin-dependent manner [247]. Other members of the Nedd4-like family have also been implicated in Gag ubiquitination, including WWP1, WWP2 and Itch. However, overexpression of these other family members did not rescue HIV-1 budding defects [246].

In summary, it is clear that ubiquitination of Gag and ESCRT proteins regulates HIV-1 budding. However, there are still many questions unanswered regarding the specific mechanisms, such as how Nedd4-1 is recruited by ALIX and whether ubiquitin ligases other than Nedd4-2 are responsible for HIV-1 ubiquitination. Further investigation of proteins involved in these processes would help to fully understand this process.

5.1.3 MKRN1 has no known connection to the ESCRT pathway

Given the strong reduction in release of viral particles from MKRN1-KO cells and the requirement for ubiquitin ligases in the HIV-1 budding process, I hypothesised that MKRN1 may ubiquitinate viral proteins or any of the known regulators of viral budding. However, there is no documented role for MKRN1 in the regulation of viral budding or in the other functions of the ESCRT complex in the multivesicular body pathway. Concentration of MKRN1 in the vRNA RNP might facilitate the ubiquitination of targets that does not occur in uninfected conditions. Alternatively, the effect of MKRN1 may be due to other factors with yet unknown roles in budding.

Whilst exploring the role of MKRN1 in ribosome-associated quality control, Hildebrandt et al performed affinity purification coupled with stable isotope labelling with amino acids in cell culture (SILAC)-based quantitative mass spectrometry to identify the protein interactors of GFP-MKRN1 [160]. They identified 53 proteins that were significantly enriched in GFP-MKRN1 affinity purification compared to controls. Almost all of the identified proteins had been previously found in association with polyadenylated RNA transcripts. I conducted STRING analysis of the proteins and compared with KEGG pathway terms which highlighted an enrichment for ribosomal proteins and members of mRNA surveillance pathways (Figure 5.2A). In agreement

with this, Gene Ontology (GO) terms for biological processes revealed the majority of MKRN1 interactors are involved in the regulation of gene expression (Figure 5.2B). Further investigation by Hildebrandt et al revealed the ubiquitination targets of MKRN1. They performed ubiquitin remnant profiling to compare the relative abundance of di-glycine-modified lysines in MKRN1 knockdown and control cells. Overall they identified 1264 proteins as targets of MKRN1 ubiquitination but noted depletion of MKRN1 significantly decreased the abundance of ubiquitination on 21 proteins (Figure 5.2C). Importantly, they observed no change in the relative abundance of these proteins in MKRN1 knockdown cells, suggesting that ubiquitination of these proteins does not target them for proteasomal degradation [160].

None of the interactors or ubiquitination targets of MKRN1 are part of the ESCRT machinery or known to regulate viral budding. However, we identify several interactors and targets of MKRN1 in our HIV-1 ivRBPome such as PABPC1, YBX3 and FAM120A (Figure 5.2A and C). The role of MKRN1 may thus be exerted through this network of ivRBPs. This could be in addition to or instead of a role in viral budding. We can be confident that MKRN1 and these interacting proteins were incorporated into HIV-1 particles through interaction with the vRNA rather than as part of a protein complex due to irradiation at 254 nm during ivRIC only inducing crosslinking between proteins and RNA. Additionally, stringent washing following oligo d(T) capture ensured any protein-protein interactions were disrupted. Hence, only proteins directly bound to the HIV-1 RNA are identified in the ivRBPome.

Before investigating the role of MKRN1 in viral budding or regulation of the RNP, I first validated that the results observed in Chapter 4 were not limited to the single CD4+ T lymphocyte MKRN1-KO clone studied. Three HEK293 Flp/In MKRN1-KO clones were generated to confirm the restriction of HIV-1 budding was observed in different cell lines and different clones. Additionally, the phenotype was also confirmed in a HEK293T MKRN1-KO clone. The infectivity of HIV-1 particles produced in each MKRN1-KO cell line was also determined and the absence of MKRN1 in viral particles was found to have no effect on infectivity, further supporting a role for MKRN1 in budding. I then took a bi-directional approach to study viral budding. First, I utilised

single-molecule fluorescence in situ hybridisation (smFISH) to visualise vRNA and Gag in HIV-1 infected MKRN1-KO cells. The aims of this was to a) validate the accumulation of vRNA and Gag in MKRN1-KO cells by an orthogonal approach; b) demonstrate whether the observed effect is due to an increase in transcription and c) to visualise the localisation of the accumulating viral components. Secondly, I conducted a co-immunoprecipitation of MKRN1 from HIV-1 infected cell lysate and used quantitative mass spectrometry to reveal the interactome of MKRN1 in HIV-1 infection. The goal of this analysis was to determine if MKRN1 interacted with any HIV-1 proteins or members of ESCRT and highlight any changes to the MKRN1 interactome upon HIV-1 infection.

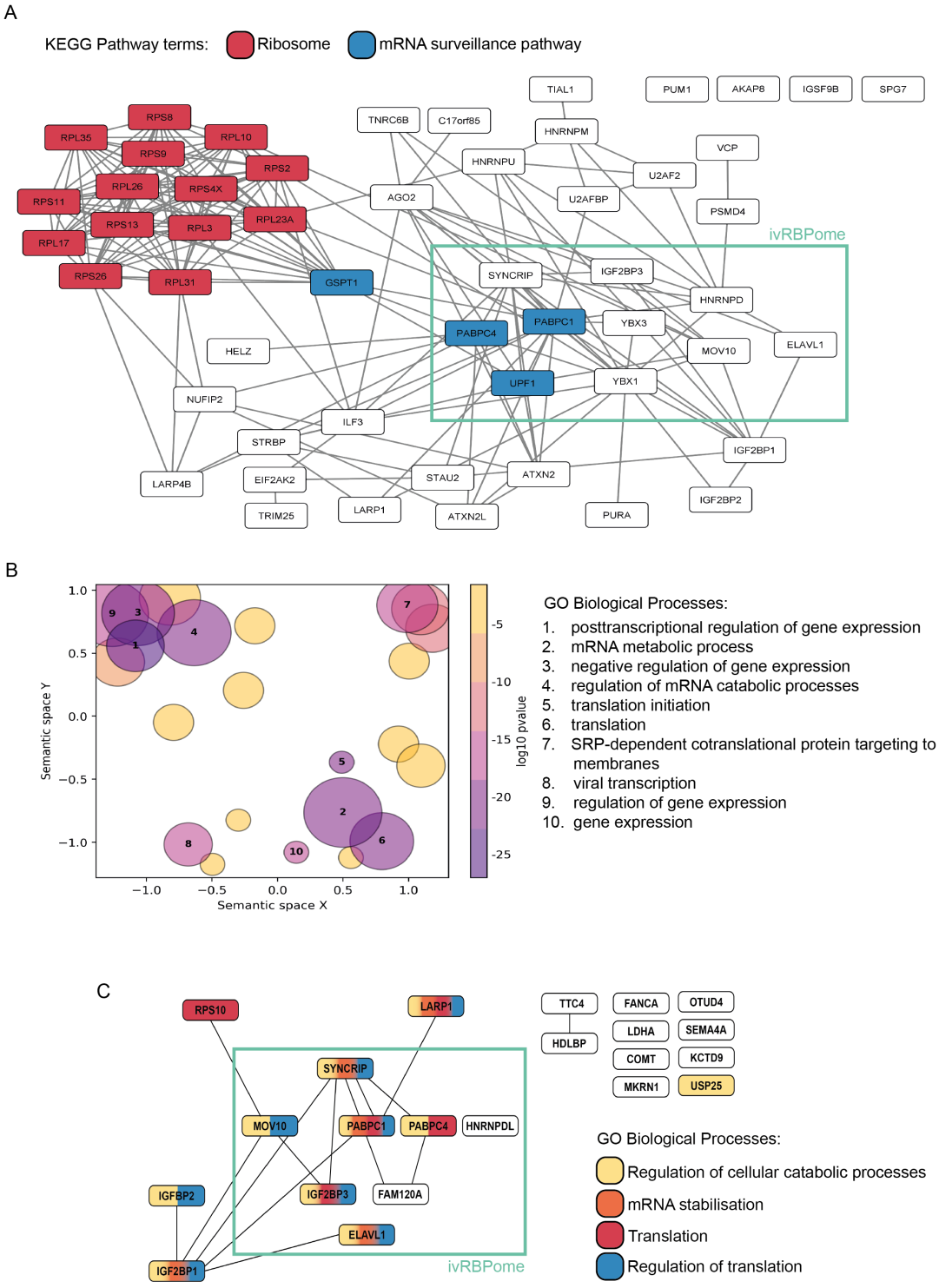

Figure 5.2: **Interactors and targets of MKRN1. A.** A Network map of MKRN1 interactors identified through affinity purification and quantitative mass spectrometry [160]. The map was generated using STRING [142] and Cytoscape [138] with KEGG pathway terms labelled in red and blue. Proteins also identified in our ivRBPome are highlighted in the green box. **B.** The top 10 gene ontology (GO) descriptions for the biological processes of the proteins in **A**. **C.** A network map of ubiquitination targets of MKRN1 [160] with GO biological processes labelled by colour. Proteins also identified in our ivRBPome are highlighted in the green box.

5.2 Results

5.2.1 Generation of MKRN1-KO HEK293 Flp/In T-rex cells

Whilst the results presented in Chapter 4 uncovered an interesting potential role for MKRN1 in HIV-1 infection, all experiments were conducted in a single CD4+ T lymphocyte clone. To validate these results I needed to test the MKRN1-KO phenotype in additional clones. MKRN1-KO cells were generated human embryonic kidney cell line (HEK293) engineered to host Flp recombination target (FRT) sites and a Tat repressor gene within the cell genome (HEK293 Flp/In T-rex). Whilst not a physiological target of HIV-1 infection, HEK293 cells support HIV transcription and gene expression. HEK293 Flp/In T-rex cells were specifically chosen to generate the KO in due to their potential for future rescue experiments. The FRT sites inserted in the genome enable the integration of a gene of interest via Flp recombinase-mediated DNA recombination at the FRT site [305]. This permits reintroduction of MKRN1 and mutants of it in future investigation. To generate the KO line, the same CRISPR Nanoblades approach was used as in Chapter 4 [237]. After Nanoblade treatment, single cells were isolated and screened for successful KO. Western blotting of whole cell lysate confirmed total loss of MKRN1 expression in clones 2B, 8 and 9 (Figure 5.3). There are four isoforms of MKRN1 produced by alternative splicing of sizes 53 kDa, 48 kDa, 45 kDa and 35 kDa, but the largest is accepted as the canonical/dominant isoform. Expression of all four isoforms was lost in clones 2B, 8 and 9. The identity of the protein producing a band at 70 kDa was unknown but was likely to be a cross-reaction of the antibody with a cellular protein.

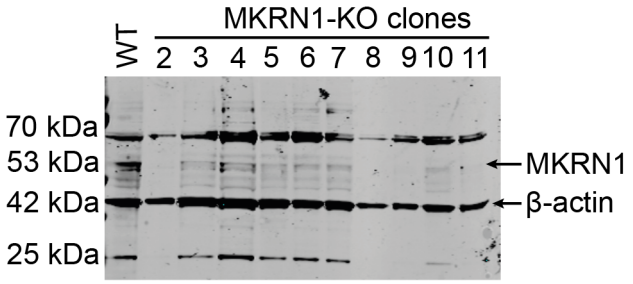

Figure 5.3: **CRISPR/Cas9 KO of MKRN1 in HEK293 Flp/In cells** HEK293 Flp/In cells were transduced with CRISPR/Cas9 nanoblades to generate a MKRN1-KO. MKRN1 expression levels in WCL collected from individual clones was analysed by western blot using anti-MKRN1 and anti-βactin antibodies. The canonical isoform of MKRN1 produces a band at 53 kDa.

5.2.2 MKRN1-KO in HEK293 cells inhibits HIV-1 particle release

To determine whether HIV-1 viral particle release was inhibited in HEK293 MKRN1-KO clones a virus release assay was established. To prevent bias due to clonal variation, HEK293 Flp/In MKRN1-KO clones C2B, C8 and C9 were selected for further investigation. A fully validated and characterised HEK293T MKRN1-KO clone [160] kindly gifted by Julian König (IMB Mainz, Germany) was also included in the analysis. HEK293T are a derivative of HEK293 that express a temperature-sensitive simian virus 40 (SV40) T-antigen mutant, enabling plasmids carrying the SV40 origin of replication to efficiently replicate when transfected into the cell line [306]. HEK293T are commonly used for the production of viral vectors due to this efficient replication and reduced antiviral activity [307] and therefore provide another experimental setting to test the MKRN1-KO phenotype. Additionally, the HEK293T MKRN1-KO clone was generated using an alternative CRISPR gene editing approach [160] and therefore provides a control against any common off-target effect caused by the gRNAs used in our Nanoblade approach.

To test if intracellular and extracellular levels of P24 were regulated by MKRN1, WT and KO HEK293T and HEK293 Flp/In T-rex cells were transfected with plasmids encoding HIV-$1_{\text{-R-E-Nef-mCherry}}$ and VSVG, or GFP as a transfection control. After 48 h the transfected cells were visualised for GFP and mCherry fluorescence to determine

transfection efficiency before supernatants and whole cell lysates were collected. No difference in percentage of cells expressing GFP or mCherry at 48 hpt was observed for any of the MKRN1-KO clones (Figure 5.4). This suggested transfection efficiency of both the GFP plasmid and the HIV-$1_{\text{-R-E-Nef-mCherry}}$ was the same in WT and MKRN1-KO cells. Therefore, any alteration observed in viral protein expression were unlikely to be a result of differences in transfection.

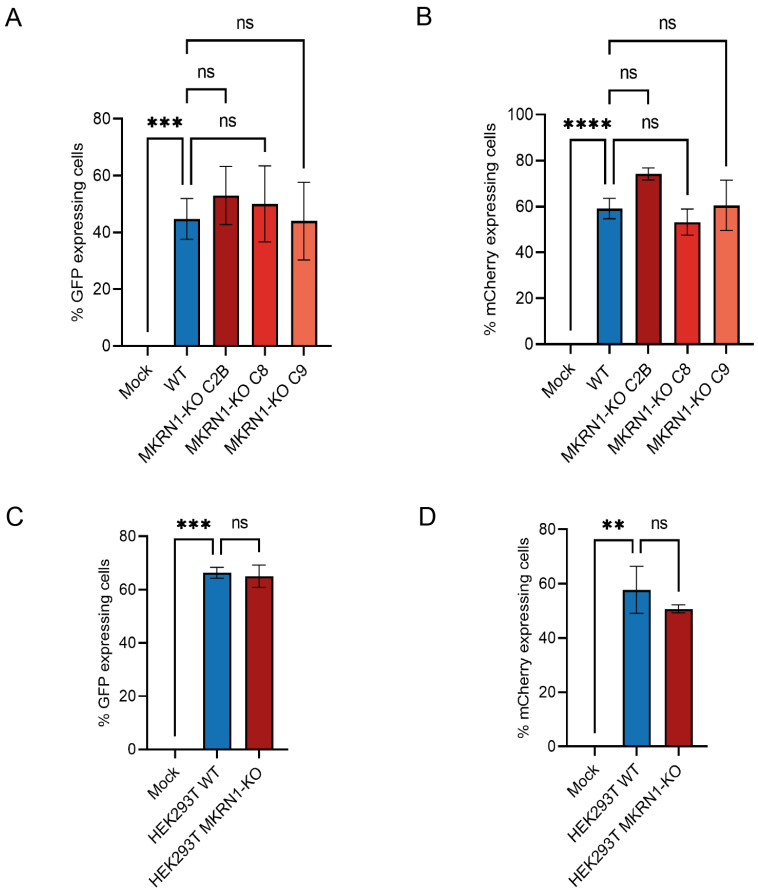

Figure 5.4: **Viral plasmid transfection is equally efficient in WT and MKRN1-KO HEK293 cell lines** Transfection efficiency of pNL4.3-R-E-T2A-Nef-mCherry and GFP plasmids determined by counting number of fluorescent cells at 48 h post transfection visualised on an EVOS M7000 microscope with excitation at 488 nm and 561 nm. Statistical significance was determined by one-way ANOVA where ns=non-significant, ** represents $p<0.01$, *** represents $p<0.001$ and **** represents $p<0.0001$.**A.** Percentage of HEK293 Flp/In cells expressing GFP. **B.** Percentage of HEK293 Flp/In cells expressing mCherry. **C.** Percentage of HEK293T expressing GFP. **B.** Percentage of HEK293T cells expressing mCherry.

Whole cell lysates and supernatants were collected from the transfected cells at 48 hpt. Samples were analysed by Western blotting using the HIV-1 P24 antibody that also detected the P55 Gag polyprotein. In contrast to CD4 T cells, I did not detect an increase in P55 and P24 in the KO lines. No significant change in P24 or P55 expression was detected in MKRN1-KO HEK293 Flp/In cells when compared to WT cells (Figure 5.5A and B). Conversely, a decrease in P24 was observed in MKRN1-KO HEK293T when compared to WT (Figure 5.5E and F). These results partially conflicted with those previously observed in MKRN1-KO CD4+ T lymphocytes. Still, Gag and P24 accumulation at the plasma membrane (see below) was observed when using HIV-1Gag-mCherry, still supporting a defect in morphogenesis. Differences could be due to technical (e.g. transfection versus infection) or biological (HEK293T versus CD4+ T cells) reasons.

Importantly, P24 expression was strongly reduced in the supernatant of all MKRN1-KO clones (Figure 5.5C, D, G and H). No P55 was detected in the supernatant of HEK293 WT or MKRN1-KO clones which indicated that most viral particles purified had successfully matured. By contrast p55 was still detectable in viruses purified from HEK293T suggesting an slower maturation. Slightly higher levels of P55 was detected in viruses isolated from MKRN1 KO HEK293T cells when compared to WT. In addition, P24 was strongly reduced in MKRN1-KO HEK293T cells (Figure 5.5E and F). These results confirmed a reduction in viral particle production in MKRN1-KO cells that was consistent with the results observed in CD4+ T cells. Moreover, the reduction in P24 but not P55 observed in HEK293T suggested potential effects in viral particle maturation.

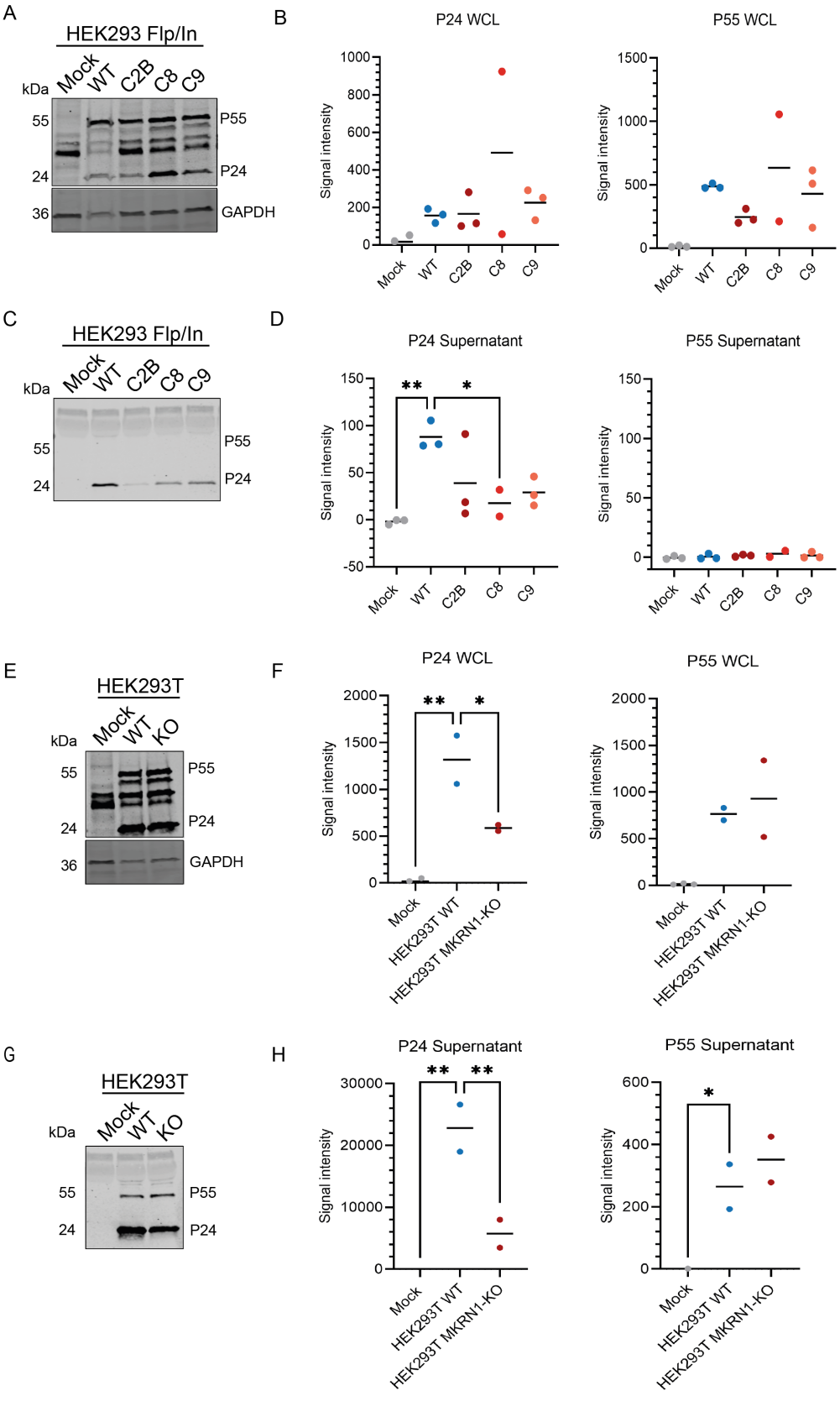

Figure 5.5: **P24 was reduced in supernatant from HEK293 Flp/In and HEK293T MKRN1-KO cells.**

Caption on next page.

Figure 5.5: **P24 was reduced in supernatant from HEK293 Flp/In and HEK293T MKRN1-KO cells. A.** WCL from HEK293 Flp/In Trex WT and MKRN1-KO cells transfected with VSVG and HIV-$1_{\text{-R-E-Nef-mCherry}}$ plasmids. Lysates were collected at 48 hpi and normalised to protein concentration before analysis via western blot with anti-GAPDH and a HIV-1 P24 antibody which detected both P24 and P55. A representative blot is shown, but a minimum of two replicates were analysed per clone. **B.** Signal intensity of P24 and P55 bands across replicates was plotted and statistical significance was determined by one-way ANOVA, for all $p>0.05$. **C.** Supernatant from HEK293 Flp/In Trex WT and MKRN1-KO cells transfected with VSVG and HIV-$1_{\text{-R-E-Nef-mCherry}}$ plasmids were collected at 48 hpi and analysed via western blot with a HIV-1 P24 antibody. One representative blot is shown, but supernatants from a minimum of two replicates were analysed for each clone. **D.** Signal intensity of P24 and P55 bands across replicates was represented. Statistical significance was determined by one-way ANOVA. * represents $p<0.05$ and ** represents $p<0.01$. **E.** WCL from HEK293T WT and MKRN1-KO cells transfected with VSVG and HIV-$1_{\text{-R-E-Nef-mCherry}}$ plasmids. Lysates were collected at 48 hpi and normalised to protein concentration before analysis via western blot with anti-GAPDH and a HIV-1 P24 antibody which detected both P24 and P55. A representative blot is shown, but a minimum of two replicates were analysed per cell line. **F.** Signal intensity of P24 and P55 bands across replicates was plotted and statistical significance was determined by one-way ANOVA, * represents $p<0.05$ and ** represents $p<0.01$. **G.** Supernatant from HEK293T WT and MKRN1-KO cells transfected with VSVG and HIV-$1_{\text{-R-E-Nef-mCherry}}$ plasmids were collected at 48 hpi and analysed via western blot with a HIV-1 P24 antibody. One representative blot is shown, but supernatants from a minimum of two replicates were analysed for each cell line. **H.** Signal intensity of P24 and P55 bands across replicates was plotted and statistical significance was determined by one-way ANOVA, * represents $p<0.05$ and ** represents $p<0.01$.

To exclude that the observed effects are due to differences in free P24, the VSVG-pseudotyped HIV-$1_{\text{-R-E-Nef-mCherry}}$ viral particles were purified from the supernatants of transfected WT and MKRN1-KO cells using a 10% sucrose cushion. The purified particles were then titrated by P24 ELISA. All the virus preparations showed lower levels of P24 in the supernatant, and in all clones except one this difference was statistically significant ($p=0.0166$, $p=0.0002$ and $p=0.0052$ for C2B, C8 and HEK293T MKRN1-KO, respectively) (Figure 5.6). HEK293 Flp/In MKRN1-KO C9 also had reduced P24 concentration in comparison to WT, but with higher variability rendering non-significant results. Altogether, these results strongly support that MKRN1 regulates HIV-1 viral particle production as observed in MKRN1-KO CD4+

T lymphocytes in Chapter 4. Moreover, the consistency of these results across cell clones and lines suggests the fundamental nature of the MKRN1-driven regulatory role.

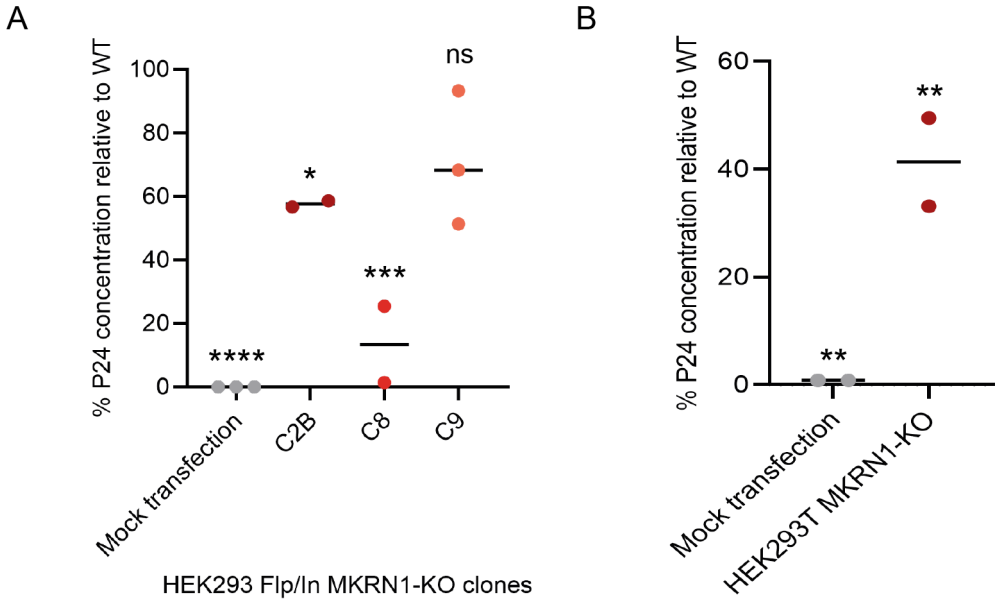

Figure 5.6: **Fewer viral particles were released from MKRN1-KO HEK293 Flp/In and HEK293T clones. A.** Concentration of P24 in supernatants collected from MKRN1-KO HEK293 Flp/In clones relative to supernatant from WT, determined by P24 ELISA. Statistical significance was determined by one-way ANOVA where ns=non-significant, * represents $p<0.05$, *** represents $p<0.001$ and **** represents $p<0.0001$.**B.** Concentration of P24 in supernatants collected from MKRN1-KO HEK293 Flp/In clones relative to supernatant from WT, determined by P24 ELISA. Statistical significance was determined by one-way ANOVA where ** represents $p<0.01$.

5.2.3 MKRN1-KO viral particles have the same infectivity as WT particles.

RBPs identified inside HIV-1 viral particles were hypothesised to regulate early stages of HIV-1 infection from inside the viral capsid. An outstanding question from Chapter 4 was whether MKRN1 had a function within HIV-1 particles and whether this affected infectivity. To investigate whether viral particles produced in MKRN1-KO cells were infectious, the purified HIV-$1_{\text{-R-E-Nef-mCherry}}$ particles from HEK293 Flp/In MKRN1-KO and HEK293T MKRN1-KO supernatants, and the corresponding WT supernatants were normalised according to P24 concentration measured by the ELISA. Viral titre

was calculated assuming 1 pg of P24 corresponded to $1x10^4$ viral particles [308]. WT HEK293 Flp/In cells were infected with WT and KO virus at an MOI of 10 to achieve a high percentage of infection. Viral particle infectivity was determined by percentage of cells expressing mCherry at 48 hpi, measured by flow cytometry. No difference was observed when comparing viruses produced in WT HEK293 Flp/IN WT and MKRN1-KO cells (Figure 5.7A and C). This suggests that MKRN1 does not have a function regulating reverse transcription or integration. Infection with virus generated in HEK293T MKRN1-KO cells was reduced by 22% compared to WT (Figure 5.7). However, only one replicate could be conducted thus the statistical significance of this result could not be determined. Further repeats will be carried out to test if this slight effect in virus infectivity is reproducible or just due to experimental variability.

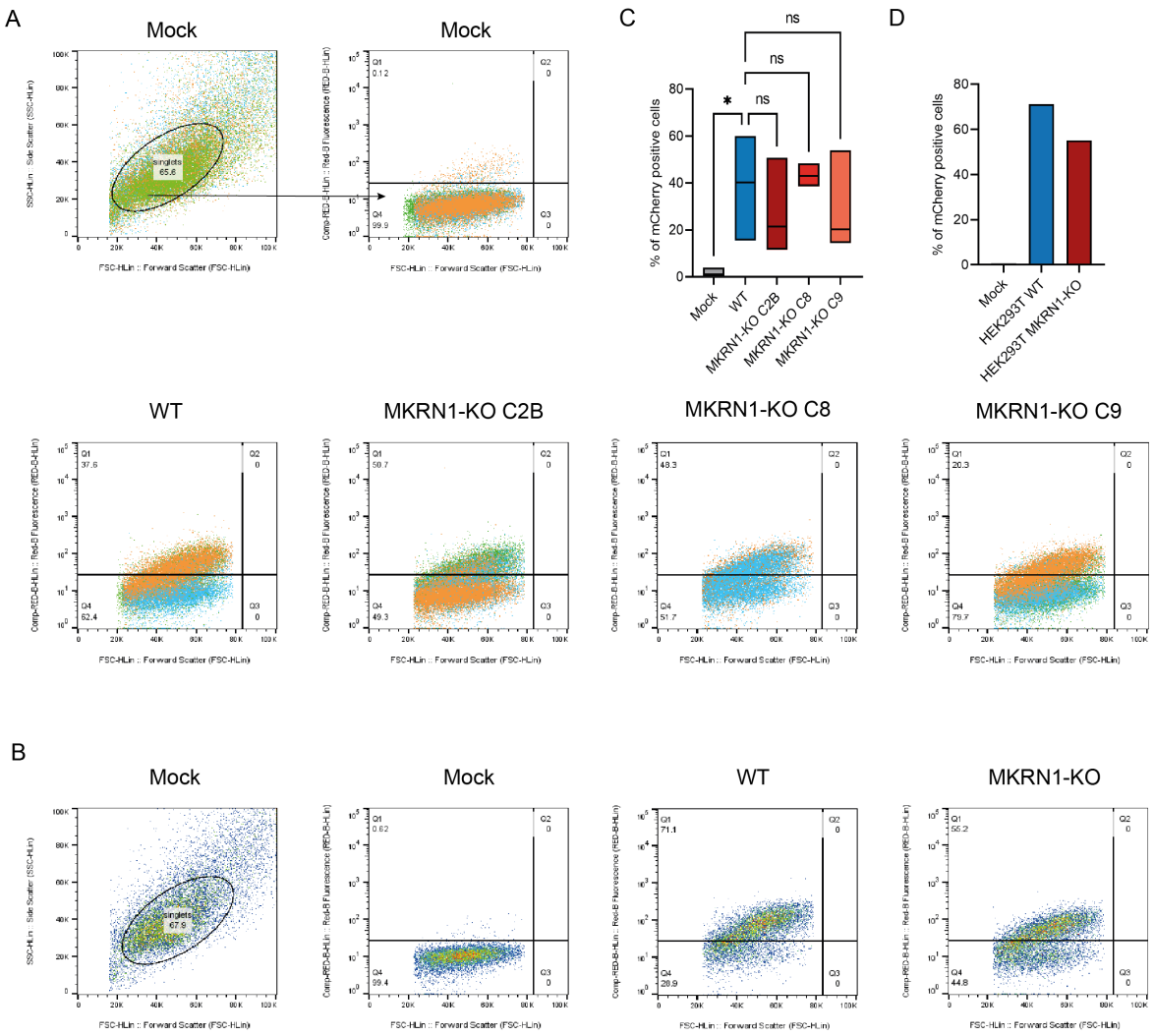

Figure 5.7: **MKRN1-KO viral particles have the same infectivity as WT viral particles A.** WT HEK293 Flp/In cells were infected with HIV-1$_{\text{-R-E-Nef-mCherry}}$ produced in WT and MKRN1-KO HEK293 Flp/In clones C2B, C8 and C9 at an MOI of 10. Percentage of infection was determined by flow cytometry with excitation at 561 nm. Virus produced in 3 different transfections of each cell line were used, these are indicated in green, orange and blue. The gating strategy is shown for mock only but was applied to all plots. **C.** Corresponding plot showing percentage of cells expressing mCherry. Statistical significance was calculated using Kruskal-Wallis test where ns=non-significant and * represents $p<0.05$. **B.** As in **A** but for WT HEK293 Flp/In cells infected with HIV-1$_{\text{-R-E-Nef-mCherry}}$. Gating strategy is shown for mock only but was applied to all plots. **D.** Corresponding plot showing percentage of cells expressing mCherry. No statistical analysis was conducted as n=1.

5.2.4 Quantification of HIV-1 RNAs in MKRN1-KO cells

To explore the spatial and temporal aspects of HIV-1 replication at a single-molecule and cellular level, I carried out smFISH with fluorescently labelled probes complementary to the HIV-1 RNA. 32 short antisense DNA oligonucleotide probes were designed to target the 3' region between the A7 acceptor splice site, Nef and the poly(A) tail and labelled with a single fluorescent dye to detect all splice variants of HIV-1 RNA (3'-670). WT and MKRN1-KO HEK293 Flp/In cells seeded on glass coverslips were infected with HIV-$1_{\text{-R-E-Gag-mCherry}}$ at an MOI of 1. Cells were fixed at 24, 48 and 72 hpi and permeabilised before hybridization with 3'-670 probes. Cells were also stained with DAPI to visualise nuclei, and either phalloidin or Cellmask Green, to highlight the cell outer edges. In addition to visualising HIV-1 RNAs, HIV-1 Gag-mCherry fluorescence was also observed for each time point. At 24 hpi, dye-labelled vRNAs accumulate at defined foci in the nucleus, which is compatible with high transcript density occurring at transcription foci [309, 310]. Single dots, compatible with single RNA molecules [309, 310], were also observed diffuse throughout the nucleoplasm and cytoplasm in both WT and MKRN1-KO cells (Figure 5.8). The number of HIV-1 RNA single molecules and transcription foci increased in both cell types as infection progressed (24 hpi $<$48 hpi $<$72 hpi) (Figure 5.8 and 5.8A-C). Moreover, single HIV-1 RNA molecules became more heavily localised in the cytoplasm over nuclei, which agrees with their transit to the cytoplasm for translation and viral particle assembly. Analysis of single-molecules of vRNA with the help of Dr Jeff Lee (Davis lab, Department of Biochemistry, University of Oxford) observed an increase in vRNA copies throughout the infection. However, there was no significant difference in the median distribution of vRNAs when comparing MKRN1-KO to WT cells (Figure 5.8B and C). That said, the distribution of vRNAs within the cell population differed at 72 hpi. Most WT cells expressed over 1000 HIV RNAs. Whereas, MKRN1-KO cells showed a bi-modal distribution with one group of cells expressing over 1000 vRNAs/cell and the other group showing 100 vRNAs/cell (Figure 5.8B). Together, these analyses show there was no difference in the median number of vRNAs per cell although differences can be observed on a cell-to-cell level. Therefore, increased transcription was not the

reason for the accumulation of vRNA and proteins observed in MKRN1-KO CD4+ T lymphocytic cells.

Figure 5.8: **Immunofluorescence and smFISH of HIV-1 RNA and Gag in HIV-1 infected MKRN1-KO cells.** Figure shown on next page. WT and MKRN1-KO HEK293 Flp/In cells were infected with HIV-$1_{\text{-R-E-Gag-mCherry}}$ at an MOI of 1 and fixed at 24, 48 and 72 hpi. Cells were hybridised with dye-coupled oligonucleotide probes complementary to the 3' of HIV-1 RNA, this revealed foci corresponding to the main splice variants of HIV-1 RNA. Transcription foci are seen in the nuclei of infected cells and the vRNAs are diffuse throughout the cytoplasm. Signal from Gag-mCherry reveals an enrichment of Gag at 72 hpi in MKRN1-KO cells. Cells were imaged by fluorescence microscopy on a 100x lens.

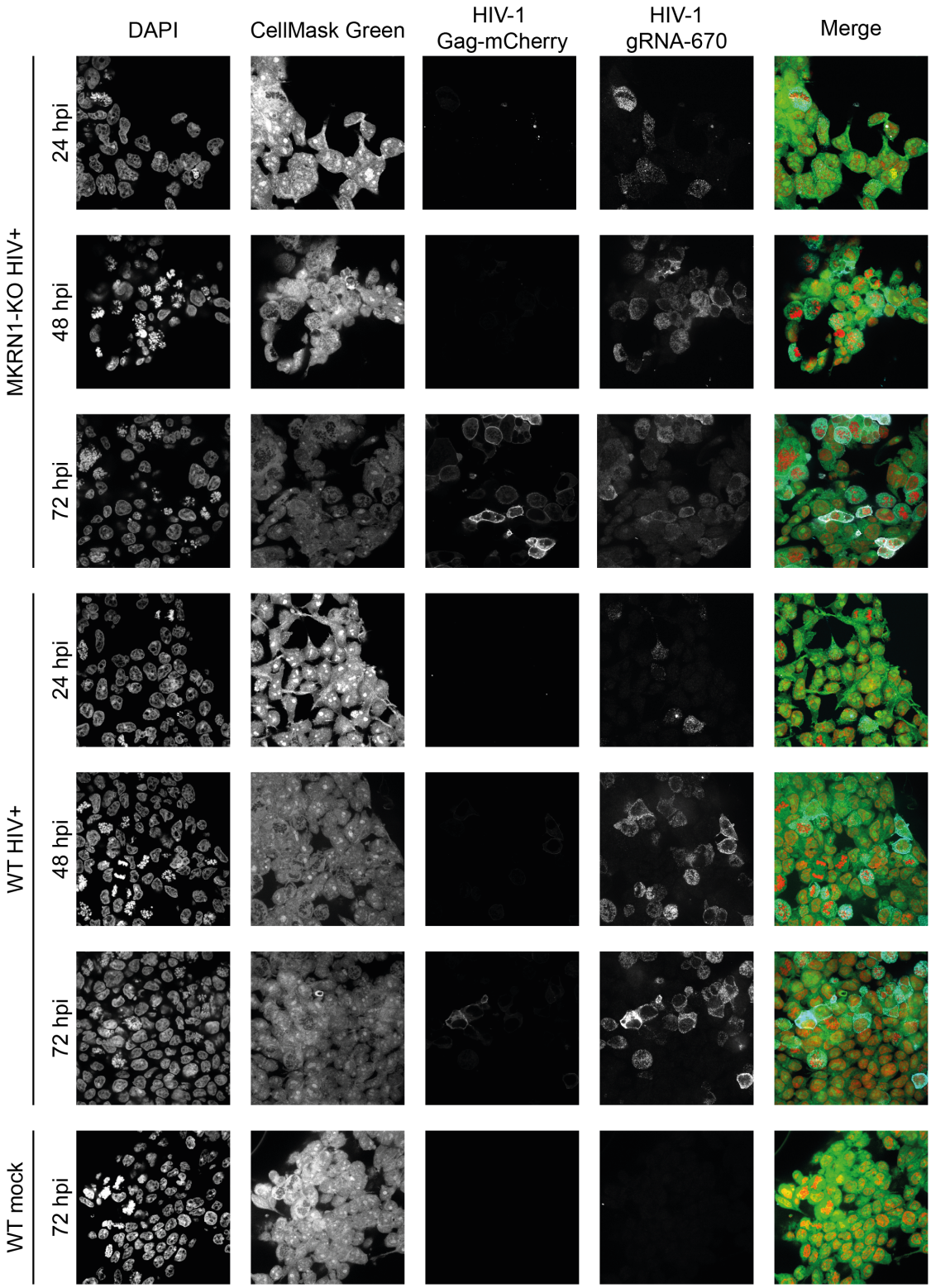

Figure 5.8: **Immunofluorescence and smFISH of HIV-1 RNA and Gag in HIV-1 infected MKRN1-KO cells.** Caption on previous page.

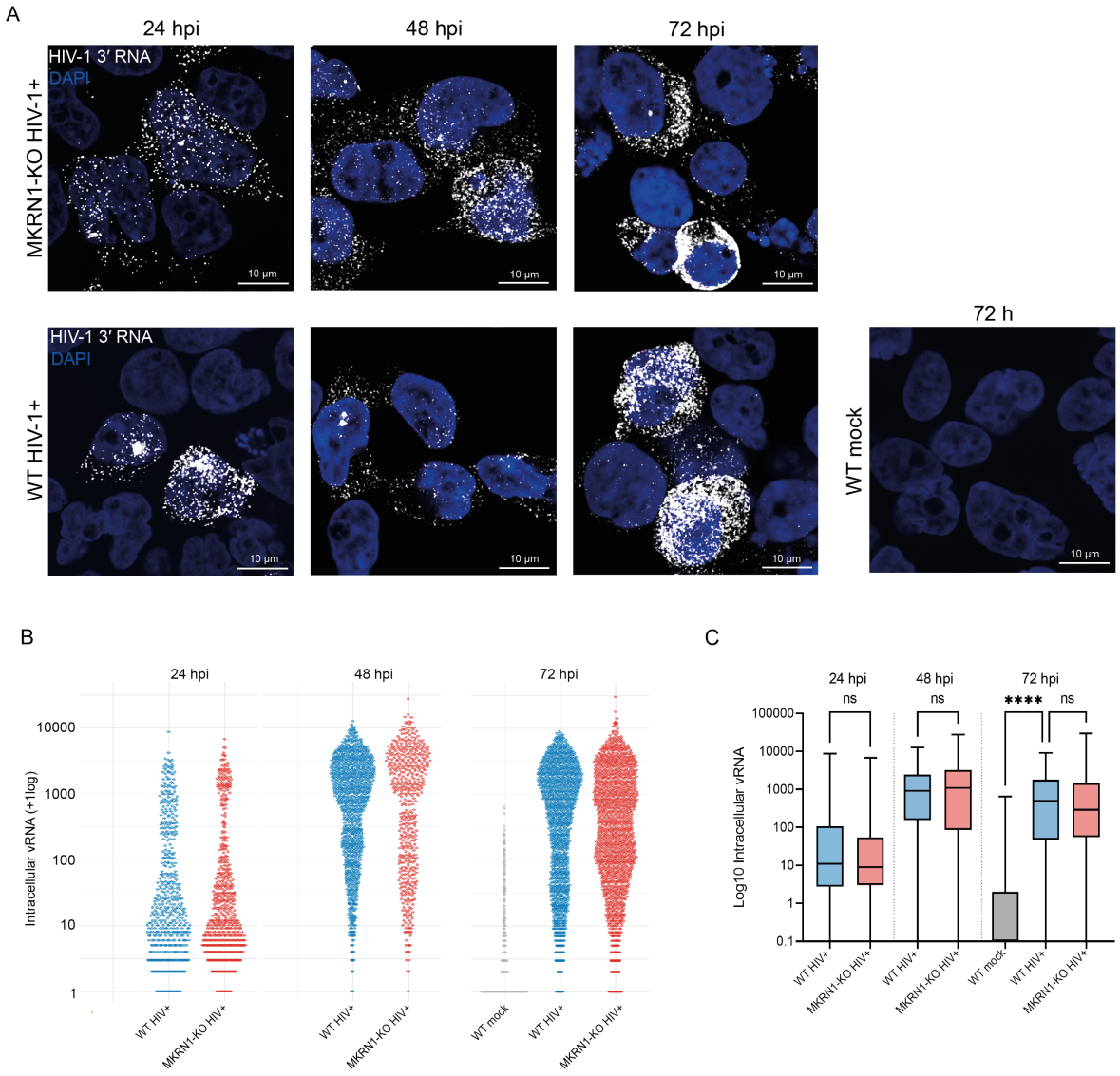

Figure 5.8: **Single-molecule detection of HIV-1 RNAs in infected cells. A.** Representative x100 magnified field of view of HIV-1$_{\text{-R-E-Gag-mCherry}}$ infected WT and MKRN1-KO HEK293 Flp/In cells fixed at 24, 48 and 72 hpi, MOI of 1. smFISH against 3' HIV-1 RNA was used to quantify vRNA counts. Scale bar shows 10 µm. **B.** Discrete separation of vRNA count distribution amongst all imaged infected WT and MKRN1-KO cells at 24, 48 and 72 hpi, where each point represents 1 cell. Shapiro-Wilk normality testing revealed the data was not normally distributed. Therefore, a Kruskal-Wallis test was used to determine the statistical significance between the medians of each group. This was statistically significant and a Dunn's multiple comparisons test was used to determine only the mock condition was significantly different ($p<0.0001$). MKRN1-KO vs WT 24 hpi ($p>0.999$), MKRN1-KO vs WT 48 hpi ($p>0.9999$) and MKRN1-KO vs WT 72 hpi ($p=0.2694$).

In addition to quantifying HIV-1 RNA during infection, Gag-mCherry fluorescence was visualised and quantified at 24, 48 and 72 hpi. Gag-mCherry was detected in distinct foci at 24 hpi for both WT and MKRN1-KO cells. It became more diffuse throughout the cytosol at 48 hpi with a noticeable accumulation at the plasma membrane (Figure 5.8). Gag-mCherry accumulation at the plasma membrane became even stronger at 72 hpi, likely reflecting strong commitment to viral particle formation. Interestingly, MKRN1-KO cells exhibited a substantial increase in Gag-mCherry signal compared to WT cells at 72 hpi and this 'extra' signal appears to accumulate in the plasma membrane. A magnified view of the cells highlighted the enrichment of Gag-mCherry at the plasma membrane in MKRN1-KO compared to WT (Figure 5.9). This outstanding result suggests that Gag was aberrantly accumulated in the plasma membrane of MKRN1-KO cells and supports that MKRN1 is required for efficient budding of virions.

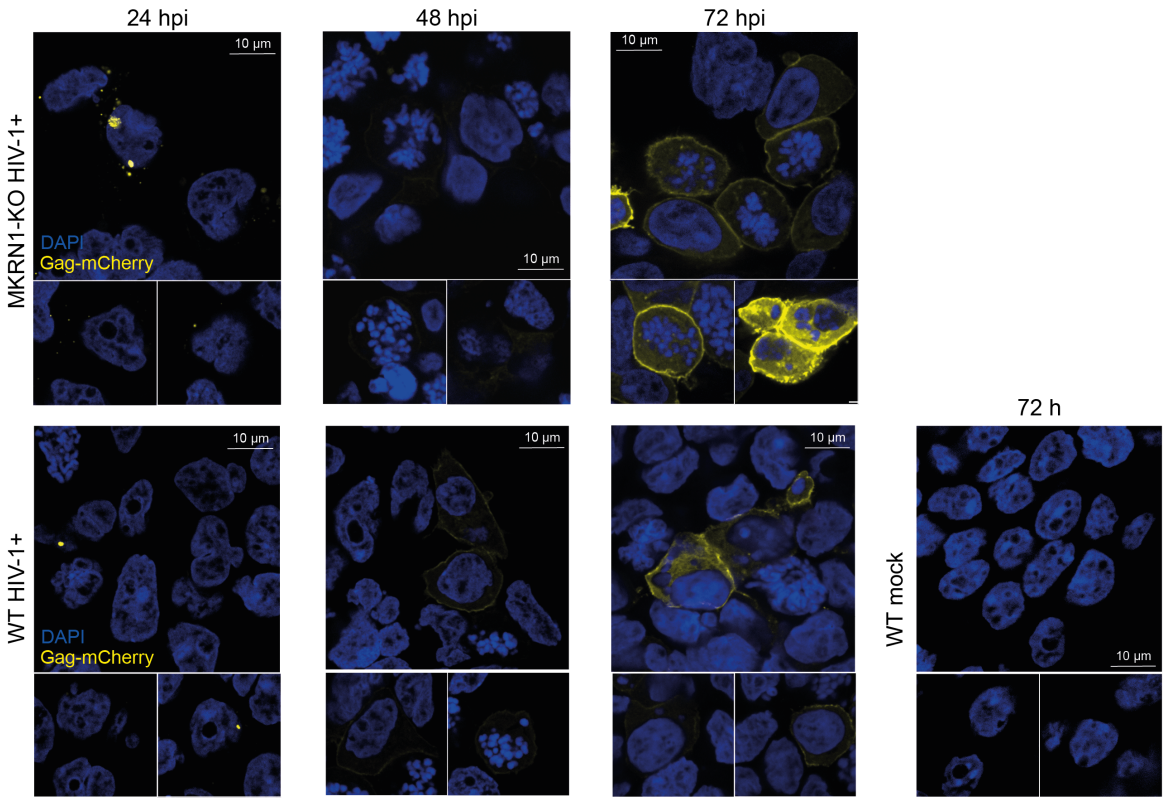

Figure 5.9: **Gag is enriched at the plasma membrane in MKRN1-KOs** Representative x100 magnified field of view of HIV-1$_{\text{-R-E-Gag-mCherry}}$ infected WT and MKRN1-KO HEK293 Flp/In cells fixed at 24, 48 and 72 hpi, MOI of 1. Scale bar = 10 µm. Gag-mCherry (yellow) is strongly expressed in the plasma membrane of MKRN1-KO cells at 72 hpi.

To quantify this phenomenon, linear cross-sections were drawn across infected cells. Infected cells were selected at random and cross-sections were drawn based on cytoplasmic staining to avoid bias. Fluorescence intensity of Gag-mCherry was measured along the cross-section and the intensity was normalised across the three biological repeats (n=30 cells/condition). As observed in the representative images, there was no overall difference in Gag-mCherry intensity between the two cells lines at 24 (p=0.794) or 48 hpi (p=0.632) (Figure 5.10A and B). However, Gag-mCherry became more enriched at the edges of the cell cross-section, i.e. the plasma membrane. This effect was particularly strong at 72 hpi where Gag-mCherry fluorescence was substantially greater (p=0.0301) in MKRN1-KO compared to WT (Figure 5.10C). In comparison, cross-sections were also measured for fluorescence intensity of vRNAs and do not show an increase in MKRN1-KO cells (Figure 5.11). HIV-

1 RNAs were diffused evenly throughout the cell at 24 hpi then became more enriched towards the edges of the cell at 48 and 72 hpi. However, there was no difference between intensity in MKRN1-KO and WT cells (p=0.871 for 24 hpi, p=0.153 for 48 hpi and p=0.471 for 72 hpi). This difference may be due to the fact that only 2 copies of vRNA and over 2000 copies of Gag assemble to form the viral particles. Alternatively, it may suggest that vRNAs are not accumulating at the plasma membrane because they are not being efficiently packaged into assembling virions. Taken together, these results support my hypothesis that MKRN1 regulates HIV-1 budding and that in the absence of MKRN1, assembling particles accumulate in the plasma membrane.

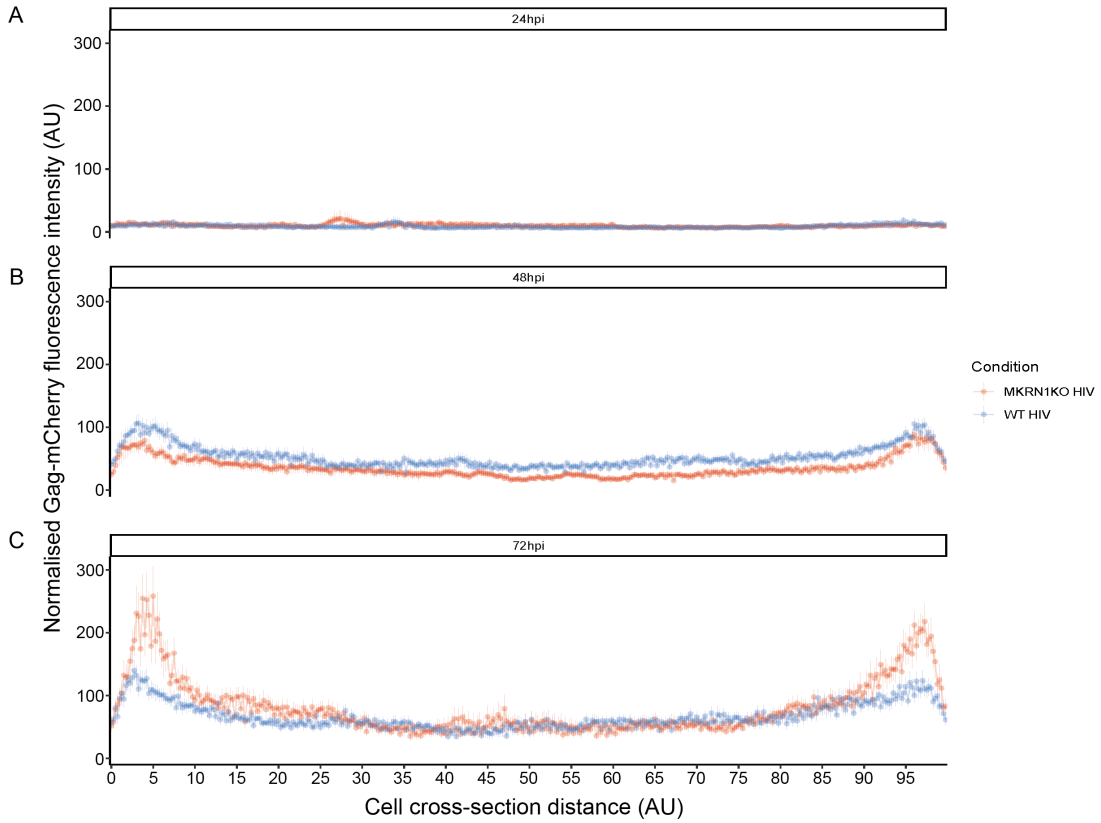

Figure 5.10: **MKRN1-KO cells have increased localisation of HIV-1 Gag to the plasma membrane** Fluorescence signal intensity of Gag-mCherry measured using a linear cross-section of infected WT (blue) and MKRN1-KO (red) HEK293 Flp/In cells. Cells were chosen randomly and linear cross-sections were drawn based on cytoplasmic stain whilst Gag-mCherry signal was hidden. Fluorescence intensity was normalised across three biological replicates. Gag-mCherry intensity at 24, 48 and 72 hpi is shown in **A, B** and **C**, respectively. The area under the curve was compared between WT and MKRN1-KO cells at each time point and a T-test was conducted for each (n=30 cells/condition). For 24, 48 and 72 hpi, p=0.794, 0.632 and 0.030, respectively.

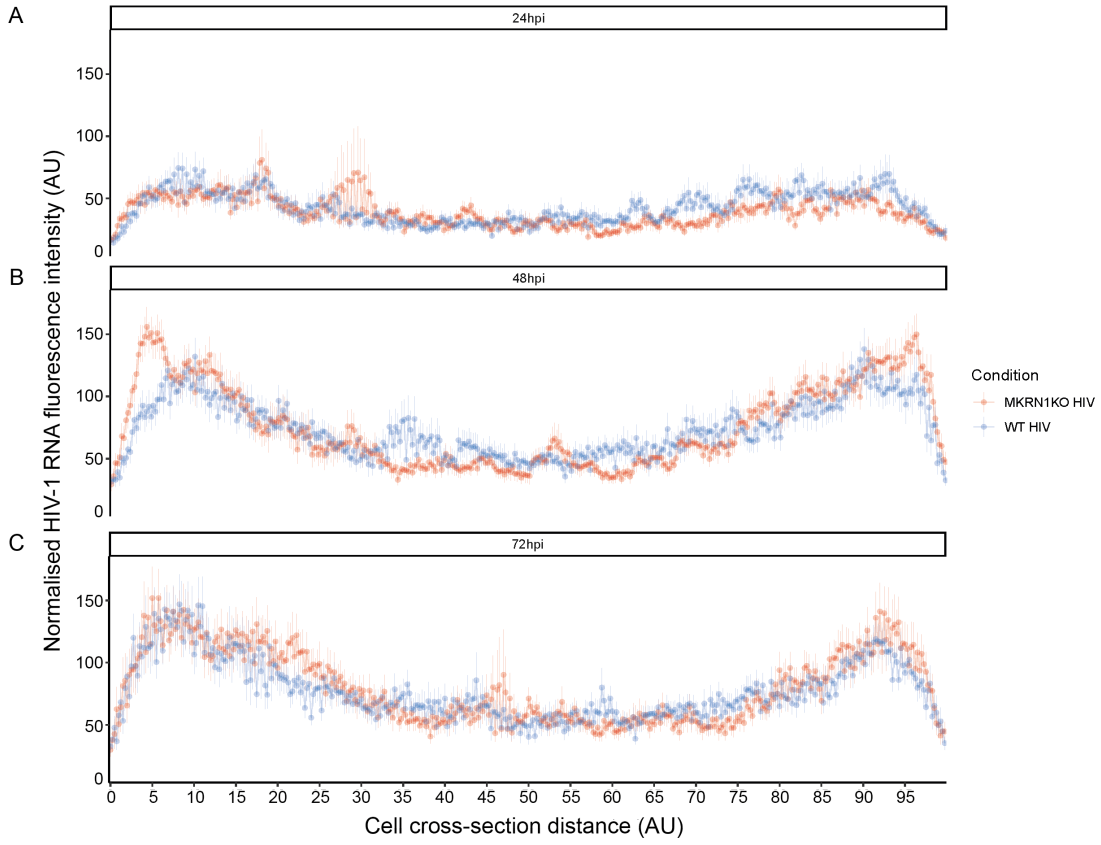

Figure 5.11: **No difference in vRNA localisation in MKRN1-KO cells** Fluorescence signal intensity of dye-coupled probe hybridised to vRNA measured using a linear cross-section of infected WT (blue) and MKRN1-KO (red) HEK293 Flp/In cells. Cells were chosen randomly and linear cross-sections were drawn based on cytoplasmic stain whilst vRNA signal was hidden. Fluorescence intensity was normalised across three biological replicates. vRNA intensity at 24, 48 and 72 hpi is shown in **A, B** and **C**, respectively. The area under the curve was compared between WT and MKRN1-KO cells at each time point and a T-test was conducted for each (n=30 cells/condition). For 24, 48 and 72 hpi, p=0.871, 0.153 and 0.471, respectively.

5.2.5 Generation of HEK293 Flp/In GFP-MKRN1 cell line

To further explore the potential role of MKRN1 in viral budding, I wanted to determine the interactome of MKRN1 in HIV-1 infected cells. I planned to do so by co-immunoprecipitating MKRN1 from HIV-1 infected cells and identifying the proteins precipitated with MKRN1 by mass spectrometry. However, I first needed to generate a system to enable the capture of MKRN1 from cell lysates using a high specificity and affinity antibody, in this case the GFP-Trap nanobody which is well-established within

the lab. To achieve this, the gene for MKRN1 was cloned by restriction digestion into a pcDNA5-FRT-TO-GFP vector, fusing an N-terminal GFP tag to MKRN1. Transfection of pcDNA5-FRT-TO-GFP-MKRN1 into HEK293 Flp/In T-rex cells, enables integration of GFP-MKRN1 into the FRT sites within the cell genome. This provides inducible expression of GFP-MKRN1 under the control of a tetracycline-induced promoter. The assembled pcDNA5-FRT-TO-GFP-MKRN1 plasmid was transformed into *E. coli* and the resulting colonies were screened for expression of the plasmid. N-terminal GFP-tagged MKRN1 has been shown to be enzymatically active within the literature [160]. However, correct function of GFP-MKRN1 was not determined here. All colonies screened showed MKRN1 was successfully ligated into the pcDNA5-FRT-TO-GFP vector at the correct site (Figure 5.12A). The pcDNA5-FRT-TO-GFP-MKRN1 plasmid was purified and confirmed via sequencing (data not shown), then transfected into HEK293 Flp/In T-rex cells. Cell lysates were immunoblotted against MKRN1 and GFP (Figure 5.12B). GFP-MKRN1 was successfully expressed in HEK293 Flp/In T-rex cells and detected by both anti-MKRN1 and anti-GFP antibodies at the expected size of 79 kDa. Expression of endogenous MKRN1 is lost in HEK293 Flp/In T-rex GFP-MKRN1, probably because of compensation from functional gene expression. No expression of free GFP was observed, although a band is detected at 37 kDa with anti-GFP which implies potential proteolytic cleavage at a site within the N terminus of MKRN1. To further validate the expression of GFP-MKRN1, the cells were induced with doxycycline for 24 h and fluorescence was observed by flow cytometry. A similar percentage of cells expressed GFP-MKRN1 and the GFP control. However, GFP-MKRN1 was expressed to a much lower level than unfused GFP in HEK293 Flp/In GFP cells (Figure 5.12C and D), which has also been observed in our laboratory for many other proteins with key regulatory roles in cell biology. No fluorescence was observed in uninduced controls for GFP-MKRN1, demonstrating tight control of gene expression.

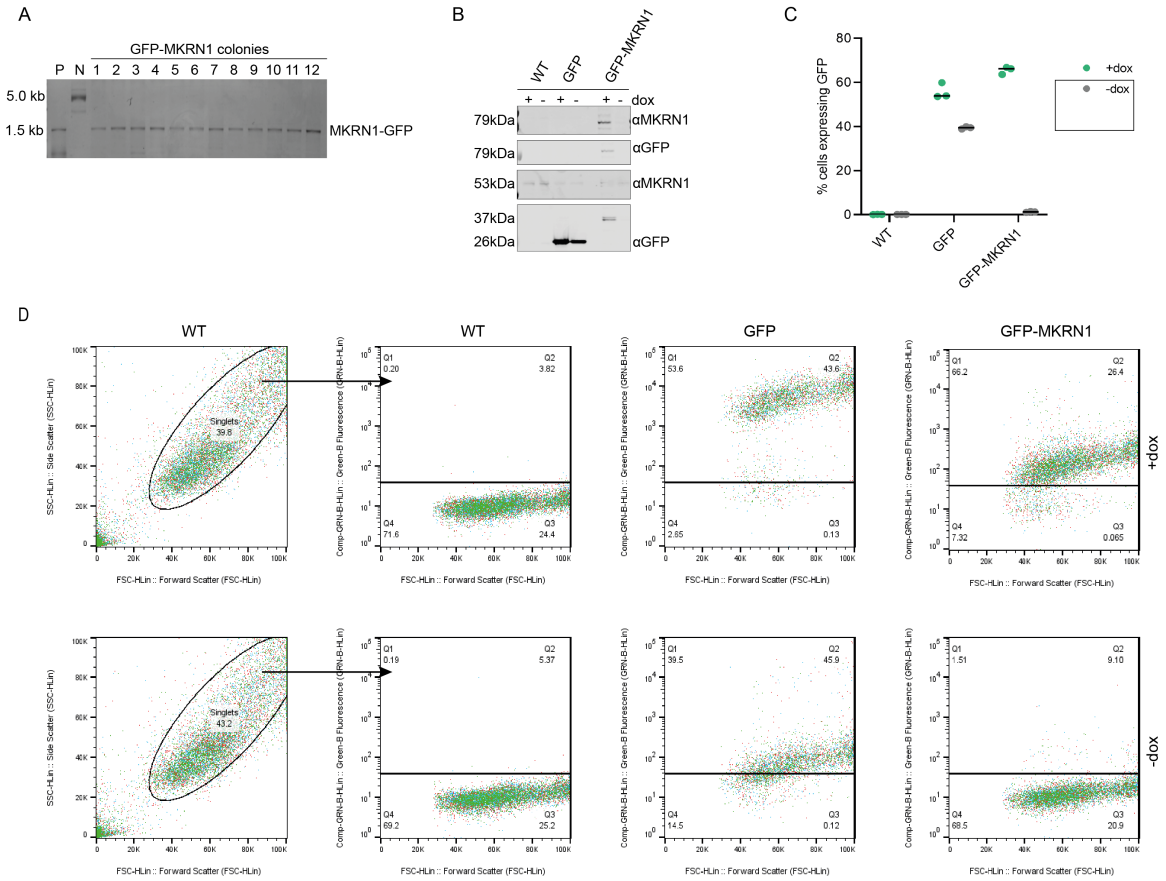

Figure 5.12: **Generation of HEK293 Flp/In GFP-MKRN1 cell line A.** DH5α *E. coli* colonies transformed with pcDNA5-FRT-TO-GFP-MKRN1 plasmid were screened for successful assembly of MKRN1 into the pcDNA5-FRT-TO-GFP vector. PCR amplification products were visualised on a 0.8% agarose gel. Correctly assembled GFP-MKRN1 is detected at 1.5 kb. Amplification of the MKRN1 gene from a pcDNA3-FLAG-MKRN1 plasmid was used as a positive control (P) and the empty pcDNA5-FRT-TO-GFP vector was used as a negative control (N). **B.** HEK293 Flp/In WT, GFP and GFP-MKRN1 cells were induced with 1 mg mL^{-1} doxycycline (dox) for 24 h before lysis. Lysates were immunoblotted against MKRN1 and GFP. Expression of GFP-MKRN1 is detected with both anti-GFP and anti-MKRN1 at 79 kDa. **C.** Expression of GFP was determined using flow cytometry at 510 nm in dox induced and uninduced cells, the fluorescence intensity of each cell is shown in **D.** Gating strategy is shown for WT but applied to all samples. The three colours represent three separate samples.

5.2.6 Optimisation of GFP-MKRN1 immunoprecipitation

To identify the subset of proteins interacting with MKRN1 during HIV-1 infection, I performed co-immunoprecipitation of GFP-MKRN1. In brief, HEK293 Flp/In GFP-MKRN1, GFP and parental cells were infected with HIV-1$_{\text{-R-E-Nef-mCherry}}$ at an MOI of

1. Both parental and GFP HEK293 Flp/In cells were used as a binding control as it was not known how the difference in expression between GFP and GFP-MKRN1 would affect the IP. HEK293 Flp/In GFP-MKRN1 cells were induced 24 h prior to infection and then again at the point of infection (48 h total). Whereas HEK293 Flp/In GFP were induced for only 4 h to reduce GFP expression to similar level as GFP-MKRN1. A mock infection was conducted in parallel to provide a baseline for MKRN1 interactors within the cell. Immediately before lysis, cells were imaged to confirm GFP expression and successful infection with HIV-$1_{\text{-R-E-Nef-mCherry}}$ (Figure 5.13). Despite a shorter induction time, GFP expression was far greater than GFP-MKRN1. Cells were lysed with 1x RIPA buffer at 24 hpi as this allows enough time for a full gene expression cycle of HIV-1 infection but limits the loss of GFP-MKRN1 through the release of viral-like particles. Lysates were then subjected to IP with GFP-Trap agarose beads following the manufacturer's protocol. Briefly, beads were incubated with lysates for 1 h at 4°C and washed eight times to remove contaminating proteins. Composition of wash buffers included physiological levels of salt (150 mM NaCl) and gradually reduced detergent concentration from 1% to 0% with subsequent washes. GFP-MKRN1 and interacting proteins were eluted from the beads using a low pH solution.

To determine whether GFP-MKRN1 was successfully isolated, I analysed input lysates and eluates from each IP by silver staining and western blot (Figure 5.14). The silver staining revealed proteins of a wide range of molecular weights in input samples, which reflected the whole cell proteome as expected. Eluate lanes had far more bands than expected, indicating that contaminating proteins were not successfully removed in the washing steps. (Figure 5.14A). Although, the western blot does confirm that GFP-MKRN1 was successfully captured and eluted (Figure 5.14B).

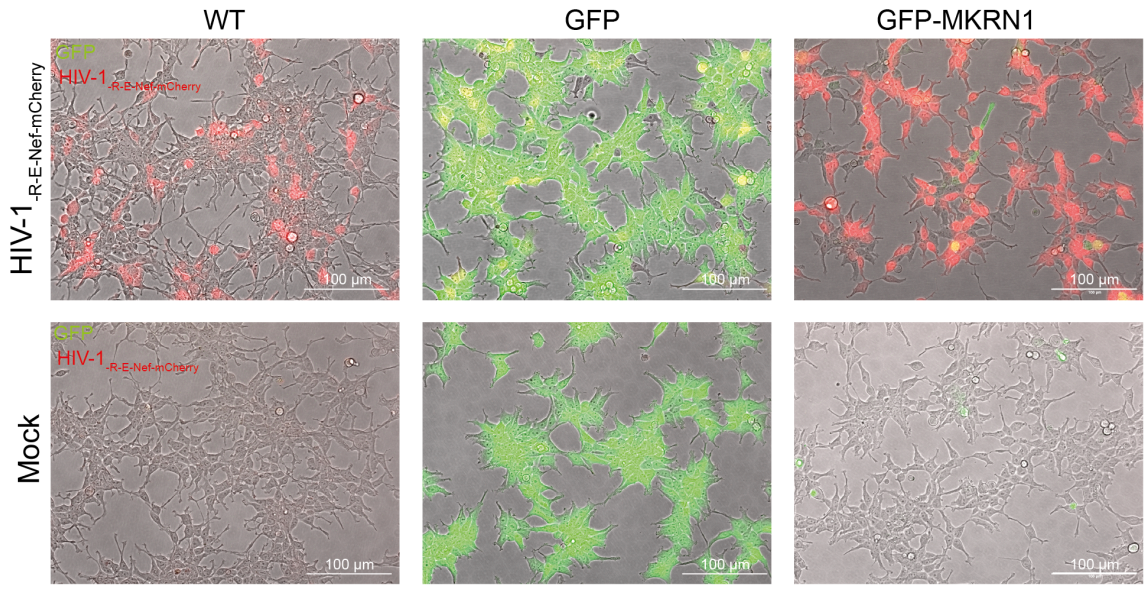

Figure 5.13: **Infection of HEK293 Flp/In GFP-MKRN1 for immunoprecipitation** HEK293 Flp/In WT, GFP and GFP-MKRN1 cells were infected with HIV-1$_{\text{-R-E-Nef-mCherry}}$ at an MOI of 1. At 24 hpi mCherry and GFP fluorescence was measured using a fluorescent cell scanner. HEK293 Flp/In GFP-MKRN1 cells were induced with 1 mg mL^{-1} doxycycline 24 h prior to infection and again at the point of infection. HEK293 Flp/In GFP cells were induced with 1 mg mL^{-1} doxycycline 4 h before imaging.

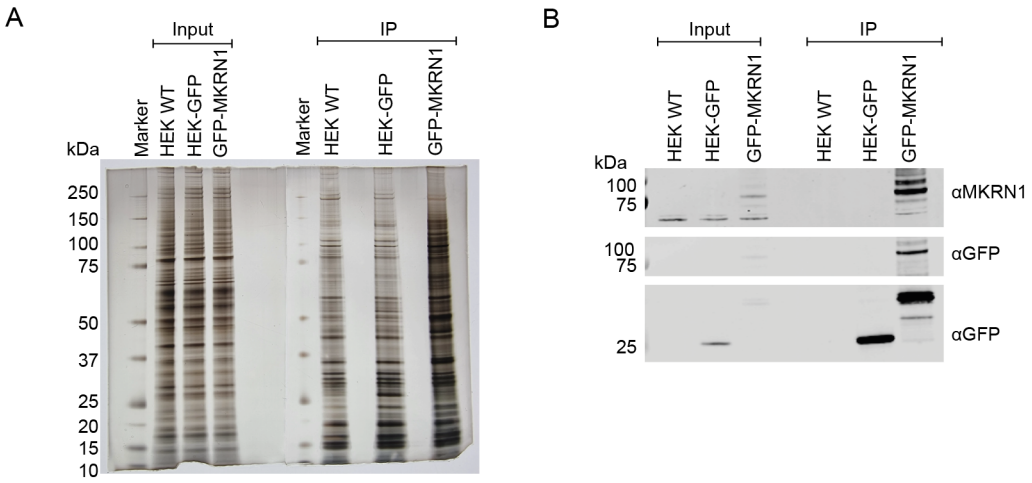

Figure 5.14: **GFP-MKRN1 is poorly isolated using GFP co-immunoprecipitation.** Doxycycline-induced HEK293 Flp/In cells infected with HIV-1$_{\text{-R-E-Nef-mCherry}}$ at an MOI of 1 were lysed at 24 hpi. Whole cell lysate was subjected to immunoprecipitation with GFP-Trap agarose beads. Inputs, flow through and eluates generated in the immunoprecipitation were analysed by silver staining (**A**) and western blotted against MKRN1 and GFP (**B.**).

In order to reduce contamination in the GFP-MKRN1 and control IP eluates, the IP was repeated with isolated cytoplasmic lysates. This was done as MKRN1 is a cytoplasmic protein and the whole cell lysis conducted in the previous experiment likely introduced non-physiologically relevant interactions with nuclear proteins which generated so much contamination in the IP. HEK293 Flp/In T-rex WT, GFP and GFP-MKRN1 were induced and infected as previously. However, at 24 hpi cells were subject to lysis and subcellular fractionation using 1% NP40 lysis buffer, this is a weaker detergent that disrupts the plasma membrane but is less efficient at breaking mitochondrial and nuclear envelopes. Lysates were centrifuged to remove nuclear and mitochondrial fractions, isolating cytosolic lysate. The IP was conducted in the same way and the inputs, flow through and eluates were analysed by silver staining and western blotting. Again, the silver staining revealed proteins of all molecular weights in input samples, which reflects the cytoplasmic proteome (Figure 5.15A). The majority of these proteins were removed in the washing steps and could be seen in the flow through. However, there was still a large amount of contamination seen in the IP eluates, as indicated by several bands in the HEK293 Flp/In T-rex WT eluate that were also present in GFP-MKRN1 eluates. Western blotting again confirmed that GFP-MKRN1 was successfully captured by GFP-Trap agarose beads (Figure 5.15B).

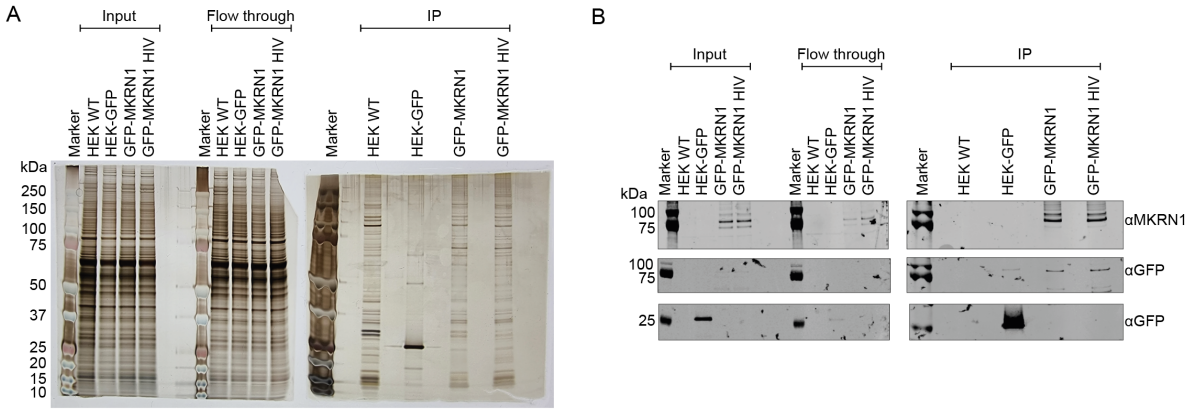

Figure 5.15: **More efficient immunoprecipitation of GFP-MKRN1 is achieved when cytoplasmic fraction is isolated.** Doxycycline-induced HEK293 Flp/In cells infected with HIV-$1_{\text{-R-E-Nef-mCherry}}$ at an MOI of 1 were lysed at 24 hpi. The cytoplasmic fraction was subjected to immunoprecipitation with GFP-Trap agarose beads. Inputs, flow through and eluates generated in the immunoprecipitation were analysed by silver staining (**A**) and western blotted against MKRN1 and GFP (**B.**).

To further optimise IP of GFP-MKRN1, I added an additional step in the preparation of lysates. HEK293 Flp/In WT, GFP and GFP-MKRN1 were induced and infected as previously and the cytoplasmic lysates were isolated at 24 hpi by subcellular fractionation. Prior to binding GFP-Trap agarose beads, the input lysates were 'pre-cleared' with binding control agarose beads for 10 mins. The aim of pre-clearing is to remove any background interaction between proteins in the lysate with the agarose beads to reduce contamination in the eluates. However, pre-clearing risks reducing the efficiency of target protein capture, which could be problematic given the relatively low abundance of GFP-MKRN1. After this step, the pre-cleared lysates were subjected to IP with GFP-Trap agarose and stringently washed eight times before elution in low pH buffer. Input lysate, pre-cleared lysate, flow through and eluates from the IP were analysed via silver staining and western blotting. Again, the silver staining revealed proteins of all molecular weights in input samples (Figure 5.16A). The intensity of these bands was reduced in the pre-cleared lysate, showing the binding control beads effectively removed contaminating interactions. The remaining proteins were successfully removed in the washing steps, indicated by the bands in the flow through lanes. Importantly, the HEK293 WT eluate was completely clear of contaminating proteins and only a small number of faint bands remain in HEK293 GFP eluate. This suggests the bands present in GFP-MKRN1 eluates were bona fide protein-protein interactions with GFP-MKRN1. Western blotting confirms that GFP-MKRN1 was still efficiently captured despite the addition of the pre-clearing step (Figure 5.16B). These results indicate that the addition of subcellular fractionation and pre-clearing to the IP protocol optimised the technique and enabled efficient isolation of GFP-MKRN1 and its interactors.

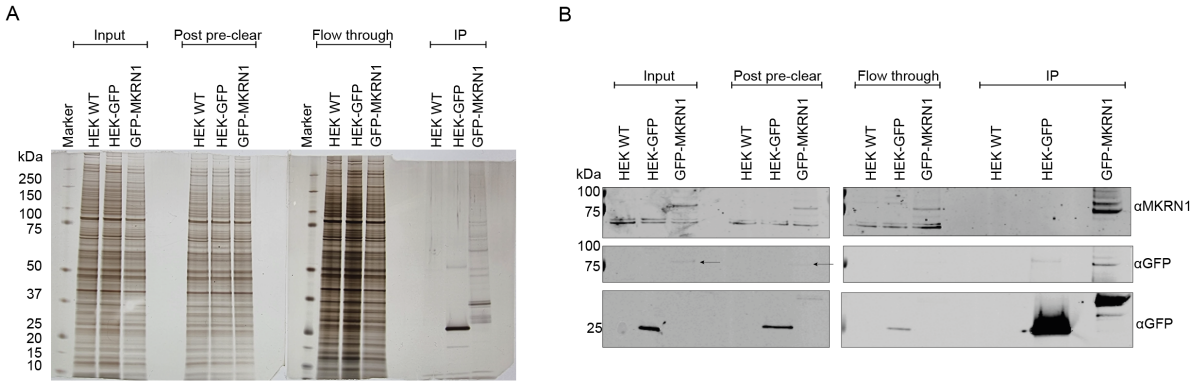

Figure 5.16: **GFP-MKRN1 is efficiently captured using an optimised GFP co-immunoprecipitation.** Doxycycline-induced HEK293 Flp/In cells infected with HIV-$1_{\text{-R-E-Nef-mCherry}}$ at an MOI of 1 were lysed at 24 hpi and the cytoplasmic fraction was subjected to immunoprecipitation with GFP-Trap agarose beads. Prior to the addition of GFP-Trap agarose beads, the lysates were pre-cleared using binding control agarose beads. Inputs, pre-cleared inputs, flow through and eluates generated in the immunoprecipitation were analysed by silver staining (**A**) and western blotted against MKRN1 and GFP (**B.**).

5.2.7 GFP-MKRN1 interactome in HIV-1 infection comprises 87 host cell proteins

Once the optimum protocol for GFP-MKRN1 IP was established, a single replicate of the IP was performed to be analysed by label-free mass spectrometry. HEK293 Flp/In WT, GFP and GFP-MKRN1 cells were induced and infected as previously described. Immediately prior to collecting cell lysates for IP analysis, the cells were imaged on a fluorescence cell scanner to confirm expression of GFP and infection with HIV-$1_{\text{-R-E-Nef-mCherry}}$ (Figure 5.17). The IP was performed as before, including subcellular fractionation and pre-clearing steps. The IP eluates were prepared for mass spectrometry by trypsin digestion followed by desalting using C18 resin. Liquid-chromatography and mass spectrometry (LC-MS) analysis and quantitative proteomics identified 522 proteins with $<$1% FDR (raw data attached in Additional File 1). Of these proteins 383 are enriched in GFP-MKRN1 HIV-1 infected (HIV+) samples over GFP HIV+ samples (Figure 5.18A). Comparison of GFP-MKRN1 HIV+ samples with WT HIV+ samples revealed 177 proteins are enriched in GFP-MKRN1 HIV+ (Figure 5.18B, full list of proteins in Appendix 2, Table 1). It was unknown

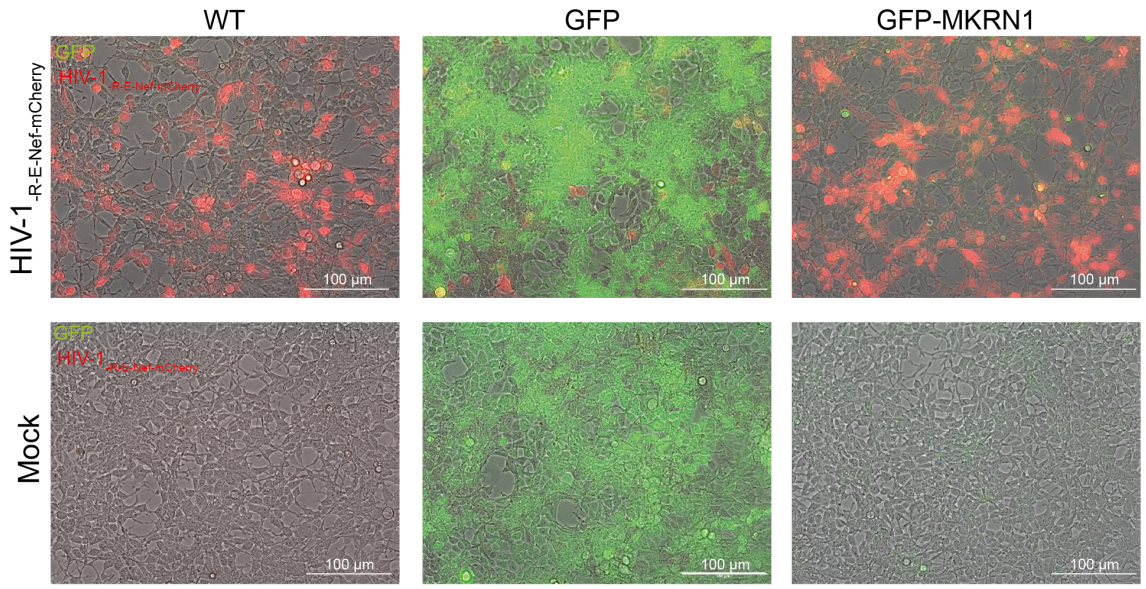

Figure 5.17: **Infection of HEK293 Flp/In GFP-MKRN1 for immunoprecipitation and proteomic analysis** HEK293 Flp/In WT, GFP and GFP-MKRN1 cells were infected with HIV-1$_{\text{-R-E-Nef-mCherry}}$ at an MOI of 1. At 24 hpi mCherry and GFP fluorescence was measured using a fluorescent cell scanner. HEK293 Flp/In GFP-MKRN1 cells were induced with 1 mg mL^{-1} doxycycline 24 h prior to infection and again at the point of infection. HEK293 Flp/In GFP cells were induced with 1 mg mL^{-1} doxycycline 3 h before imaging.

whether GFP or WT cells would provide a better control for background interactions due to the disproportionately high expression of GFP which would saturate GFP-Trap agarose beads and prevent background interactions. WT HIV+ appears to be the more stringent control. However, there were some proteins that were more abundant in GFP HIV+ samples that were not detected in WT HIV+ samples. Therefore, comparison with both WT HIV+ and GFP HIV+ samples was performed to determine which proteins were enriched in GFP-MKRN1 HIV+. These proteins were then subject to further comparison with GFP-MKRN1 mock to establish which proteins are most enriched during HIV-1 infection (Figure 5.18C). Of the 177 proteins enriched in MKRN1 samples, 87 of these were enriched in HIV-1 infection (protein and gene IDs listed in Table 23 at the end of the results section). The interactome of MKRN1 in HIV-1 infection was highly interconnected and features proteins associated with the ribosome, spliceosome, mRNA surveillance pathway and proteasome. Additionally, many of these proteins such as PABPC1 and PABPC4 are known interactors of MKRN1. Unfortunately, no

HIV-1 proteins were detected in any of the samples. However, this may be a result of the time point chosen to collect lysates; 24 hpi may be too early for interactions with MKRN1 to accumulate. In addition to this, no proteins involved in the ESCRT pathway were found to interact with MKRN1 during infection, or in uninfected cells. This suggests that MKRN1 may be affecting the budding of viral particles through a different mechanism other than the ESCRT pathway. Alternatively, the conditions of the IP may need further optimisation to capture interaction with HIV-1 proteins, such as increasing the length of infection to enable the formation of more viral particles. However, because only a single replicate was performed, no statistical analysis has been conducted on these results. Therefore, these results are not conclusive and only provide an indication of proteins in the MKRN1 interactome. I am currently conducting the repeats of this experiment.

Figure 5.17: **177 proteins are identified in the MKRN1 interactome, 87 of which are more abundant in HIV-1 infection.** Figure is on the next page. Single-replicate IP eluates of HEK293 WT, GFP and GFP-MKRN1 were analysed by mass spectrometry. Spectral counts are summed for each identified protein and then divided by protein length, generating a spectral abundance factor (SAF) for each protein. The SAF value of each identified protein is then normalized against the sum of all SAFs within an individual biological sample, resulting in the normalized SAF (NSAF) value. These are then multiplied by 100 (NSAF100) for ease of analysis. **A.** NSAF100 values for proteins identified in the HIV-1 infected GFP-MKRN1 IP are plotted against NSAF100 values for proteins identified in the HIV-1 infected GFP IP. Whether a protein is enriched in a specific IP is determined by its position above or below the x=y gradient. Proteins below this line are more enriched in GFP-MKRN1 HIV+ IP (red), whereas proteins above are more enriched in GFP HIV+ IP (blue). Of the 522 proteins identified, 383 are more enriched in GFP-MKRN1 HIV+. The point for GFP-MKRN1 fell outside of the axes range of this plot at (4.855, 14.712). The enrichment of GFP-MKRN1 in GFP HIV+ samples was due to mapping of GFP to the GFP sequence in GFP-MKRN1, no reads mapped to the MKRN1 sequence. **B.** As in **A** but for GFP-MKRN1 HIV+ IP compared to WT HIV+ IP. 177 proteins are enriched in GFP-MKRN1 HIV+ over WT HIV+ IP. **C.** As comparison to WT HIV+ IP provided a more stringent control, the 177 proteins enriched in GFP-MKRN1 HIV+ IP over WT were further analysed to determine the enrichment of proteins in HIV-1 infection. NSAF100 values for GFP-MKRN1 HIV+ IP were compared to GFP-MKRN1 mock IP. Again, an x=y gradient was used to determine the enrichment in HIV+ conditions (purple) vs mock conditions (orange). 87 proteins were enriched in GFP-MKRN1 IP in HIV-1 infection, the 10 most abundant proteins are labelled.

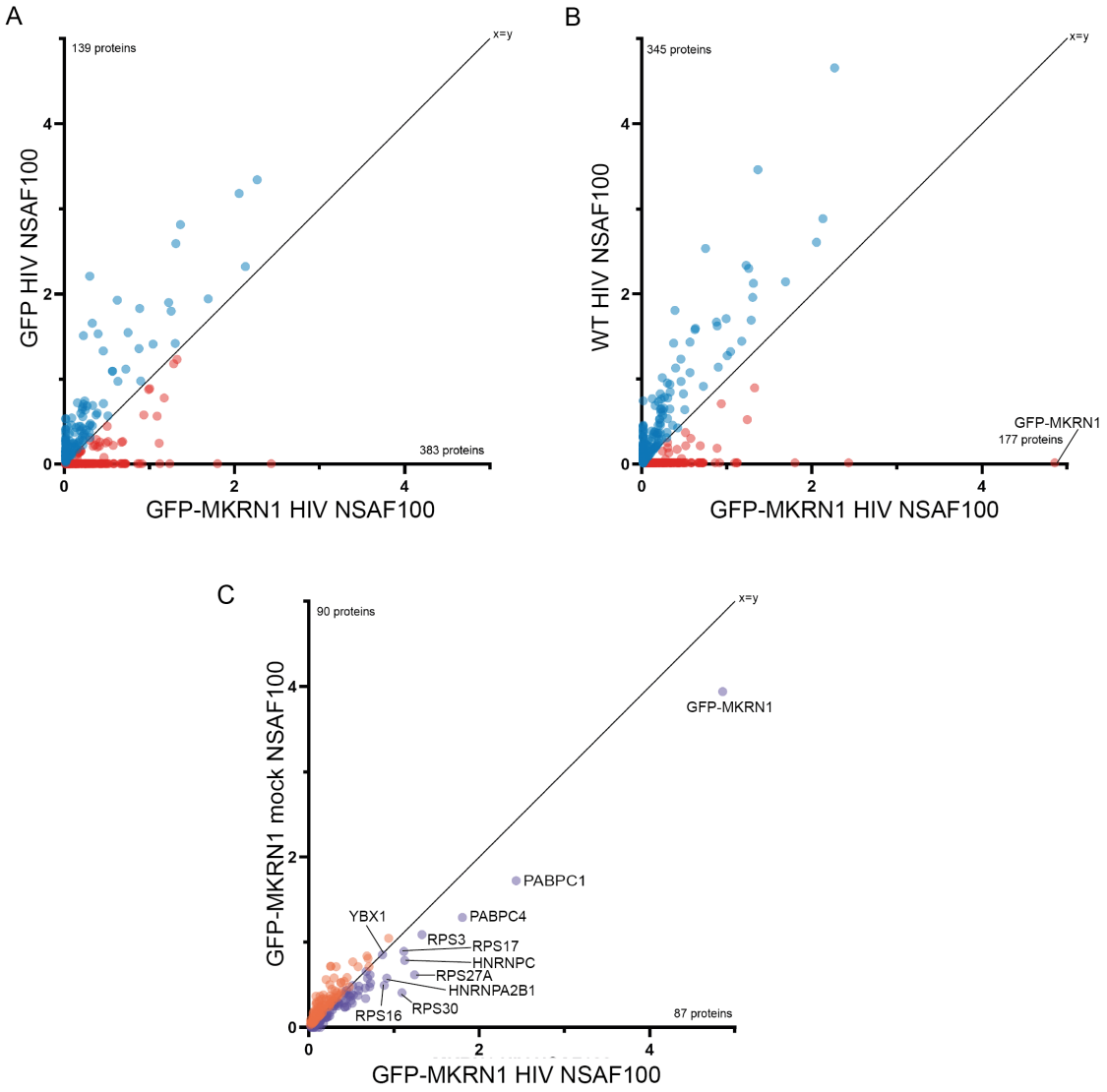

Figure 5.18: **177 proteins are identified in the MKRN1 interactome, 87 of which are more abundant in HIV-1 infection.** Caption on previous page.

To validate the likelihood of the identified MKRN1 interactors being true interactions, the GFP-MKRN1 IP dataset was compared to a previous affinity purification of MKRN1 conducted by Hildebrandt et al [160]. Of the 90 proteins identified in the interactome of MKRN1 in uninfected cells, 15 had been previously reported as interactors of MKRN1 (Figure 5.19B). This suggests that my experiment has expanded the MKRN1 RNA interactome. 16 proteins overlapped between the MKRN1 HIV-1 interactome and the Hildebrandt interactome (Figure 5.19A) suggesting that some interactors of MKRN1 are conserved during HIV-1 infection but many new interactions occur. A similar

analysis was conducted to compare the interactome of MKRN1 in HIV-1 infection to the ivRBPome identified in HIV-1 particles [135]. Interestingly, 31 of the 87 proteins identified in the MKRN1 HIV-1 interactome (36%), were also present inside HIV-1 viral particles (Figure 5.19C), which supported the interaction of these proteins with MKRN1 in HIV-1 infection. However, 22 of the ivRIC proteins were also identified as MKRN1 interactors in the absence of HIV-1 infection (Figure 5.19D). This suggested that despite interactors being present inside viral particles, not all interactions of MKRN1 are important for infection.

Again, the proteins identified in both the MKRN1 HIV-1 interactome and the ivRBPome, were highly interconnected and mainly associated with the ribosome and spliceosome (Figure 5.19E). Additionally, GO term analysis indicated that all of these proteins are involved in regulation of gene expression and many of them have functions within RNA metabolism, stability, transport and splicing, protein targeting to the plasma membrane and viral transcription (Table 22). Notably, 13 proteins identified in both the MKRN1 HIV-1 interactome and the HIV-1 ivRBPome were also confirmed interactors of MKRN1 by Hildebrandt et al: PABPC1, PABPC4, HNRNPD, HNRNPU, SYNCRIP, ILF3, RPL31, RPL35, RPS2, RPS8, RPS9, IGF2BP3 and YBX1. MKRN1 forms a complex with PABPC1 upstream of poly(A) tails of cellular mRNAs, and it is likely that this interaction also occurs at the poly(A) region of HIV-1 gRNA. Taken together, these results indicate that MKRN1 is incorporated into virions as part of a larger complex of proteins. Whilst none of these proteins were ESCRT-related, it is possible that this complex instead regulates the stability of the vRNP and/or ensures its plasma membrane trafficking and incorporation into virions.

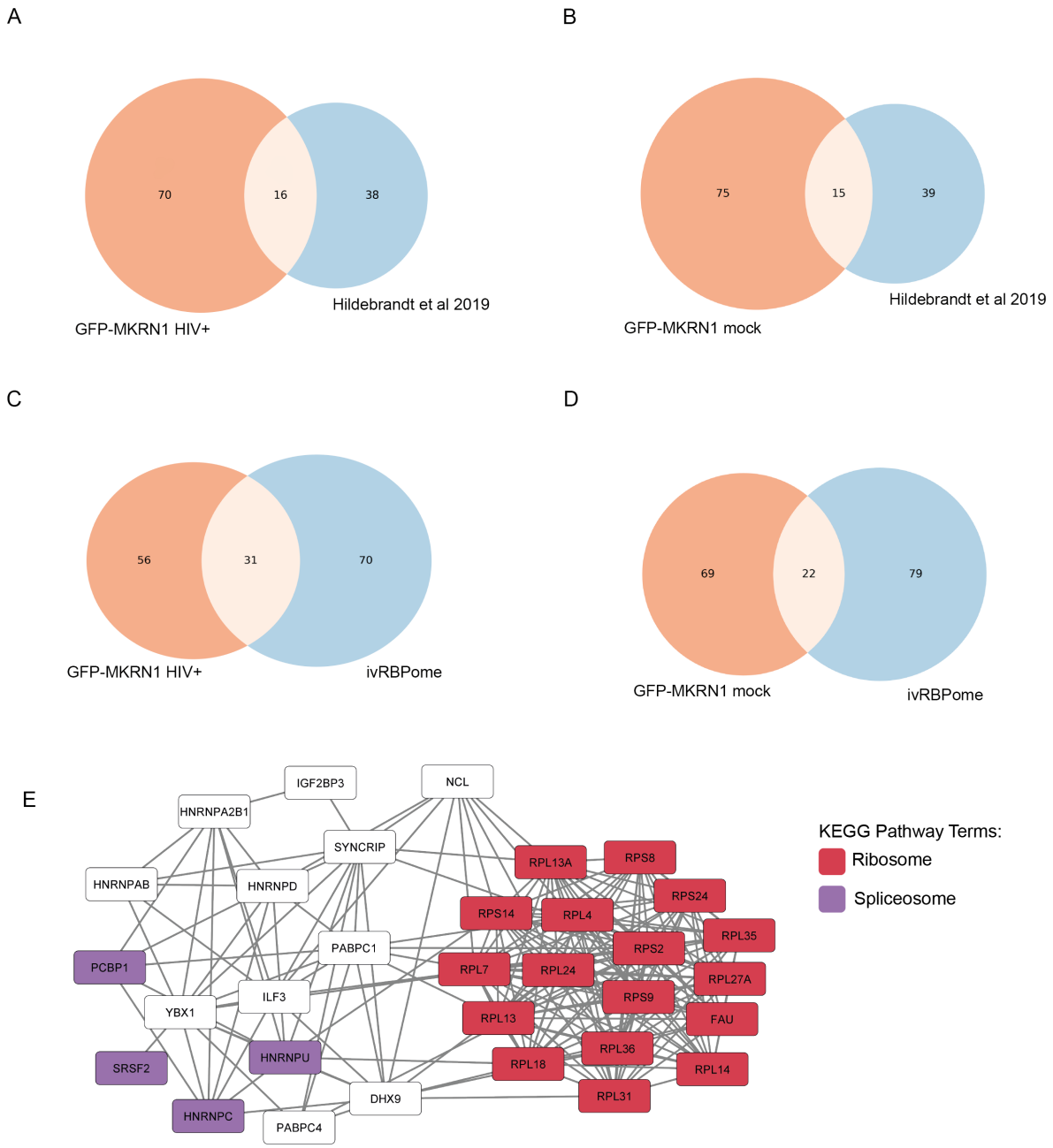

Figure 5.19: **GFP-MKRN1 interactome overlaps with previous MKRN1 interactomes and ivRBPome datasets. A-B** Comparison of proteins identified in the IP of GFP-MKRN1 in HIV-1 **(A)** and mock **(B)** infected cells with the proteins identified in MKRN1 affinity purification in Hildebrandt et al [160]. **C-D.** Comparison of proteins identified in GFP-MKRN1 IP in HIV-1 **(C)** and mock infected **(D)** cells with proteins identified in the HIV-1 ivRBPome [135]. **E.** STRING network of proteins identified in both GFP-MKRN1 IP in HIV-1 and mock infected samples and in the ivRBPome. Data show a high interconnection between MKRN1-interacting proteins. KEGG pathway analysis reveals the majority of the proteins identified are associated with the ribosome (red), spliceosome (purple) and other RNA related processes (white).

Table 22. GO terms of proteins in MKRN1 interactome in HIV-1 infection. GO terms were generated through STRING analysis. Observed gene count indicates the number of proteins in the network the GO term corresponds to. Strength describes how large the enrichment effect is and is calculated by number of proteins observed divided by the number of proteins expected to be annotated with that term in a random network of the same size. False discovery rate measures how significant the enrichment is, p values are shown.

Table 22:

GO term	term description	observed gene count	strength	FDR
GO:0010468	Regulation of gene expression	31	0.61	4.80E-17
GO:0016070	RNA metabolic process	27	1.03	6.02E-23
GO:0033036	Macromolecule localization	23	0.77	1.33E-12
GO:0034613	Cellular protein localization	22	0.94	3.81E-15
GO:0016032	Viral process	21	1.23	6.68E-20
GO:0006401	RNA catabolic process	20	1.71	2.98E-27
GO:0006412	Translation	20	1.54	2.88E-24
GO:0033365	Protein localization to organelle	20	1.23	1.11E-18
GO:0006886	Intracellular protein transport	20	1.1	2.34E-16
GO:0019083	Viral transcription	18	1.99	3.48E-28
GO:0000184	Nuclear-transcribed mrna catabolic process, nonsense-mediated decay	18	1.98	4.09E-28
GO:0006413	Translational initiation	18	1.91	2.98E-27
GO:0006396	RNA processing	18	1.12	1.14E-14
GO:0006614	SRP-dependent cotranslational protein targeting to membrane	17	2.05	1.73E-27
GO:0043488	Regulation of mrna stability	9	1.49	1.81E-09
GO:0042254	Ribosome biogenesis	9	1.29	8.84E-08
GO:0000398	mRNA splicing, via spliceosome	9	1.29	9.17E-08
GO:0006364	rRNA processing	8	1.38	1.82E-07
GO:0048255	mRNA stabilization	7	1.99	2.54E-10
GO:0070934	CRD-mediated mRNA stabilization	4	2.7	8.87E-08
GO:0002181	Cytoplasmic translation	4	1.54	0.00061
GO:0048024	Regulation of mrna splicing, via spliceosome	4	1.36	0.0028
GO:0051028	mRNA transport	4	1.23	0.0081
GO:1990428	miRNA transport	2	2.62	0.0022

5.2.8 Cloning of MKRN1 catalytic mutants for further study

To determine whether the observed effects on HIV-1 infection in MKRN1-KO cells were due to its function as an E3 ubiquitin ligase, catalytically inactive mutants of MKRN1 were designed and cloned. These mutants harbour point mutations at three sites within the RING domain of the protein. The first, H307E, has been previously described in the literature to completely inhibit MKRN1 ubiquitin ligase catalytic function [311, 312]. I designed two further point mutations that disrupt the structure of the catalytic RING domain: C281A and M285E. C281 is a conserved cysteine residue required for coordination of a zinc atom within the RING domain, disruption of which inhibits mahogunin ring finger 1 (MGRN1) ubiquitin ligase activity. M285E compromises the stability of an adjacent cysteine which disrupts the coordination of the zinc atom [313]. All three were cloned to ensure a successful catalytic mutant was generated. The previously cloned pcDNA5-FRT-TO-GFP-MKRN1 plasmid was subject to site-directed mutagenesis to individually introduce the three point mutations. The success of mutagenesis was confirmed with sequencing (data not shown) before the plasmids were transfected into HEK293 Flp/In T-rex MKRN1-KO cells. The Flp/In system enables integration of the mutant MKRN1 genes within FRT sites in the cell genome. MKRN1-KO cells were specifically chosen to ensure no expression of WT MKRN1 masked effects generated by the mutants. HEK293 Flp/In T-rex GFP-MKRN1-mutant cells were induced with 1 mg mL doxycycline for 24 h. After 24 h induction, lysates were collected and analysed for expression of the mutants. Western blotting shows that all three mutants were strongly expressed, more so than MKRN1-GFP, and were detected by both anti-MKRN1 and anti-GFP (Figure 5.20A). Some cleavage of the fusion occurred after the N-terminal GFP. Flow cytometry of doxycyline induced cells revealed that a similar percentage of cells expressed the GFP MKRN1-mutants compared to unfused GFP or GFP-MKRN1 (Figure 5.20B). When focusing on the green fluorescent signal in individual cells, I could appreciate that unfused GFP is expressed at higher levels than GFP-MKRN1 mutants (Figure 5.20C). Together, these results suggest successful expression of MKRN1 mutant cells. However, further analysis is required to determine if the mutations successfully inhibit catalytic activity.

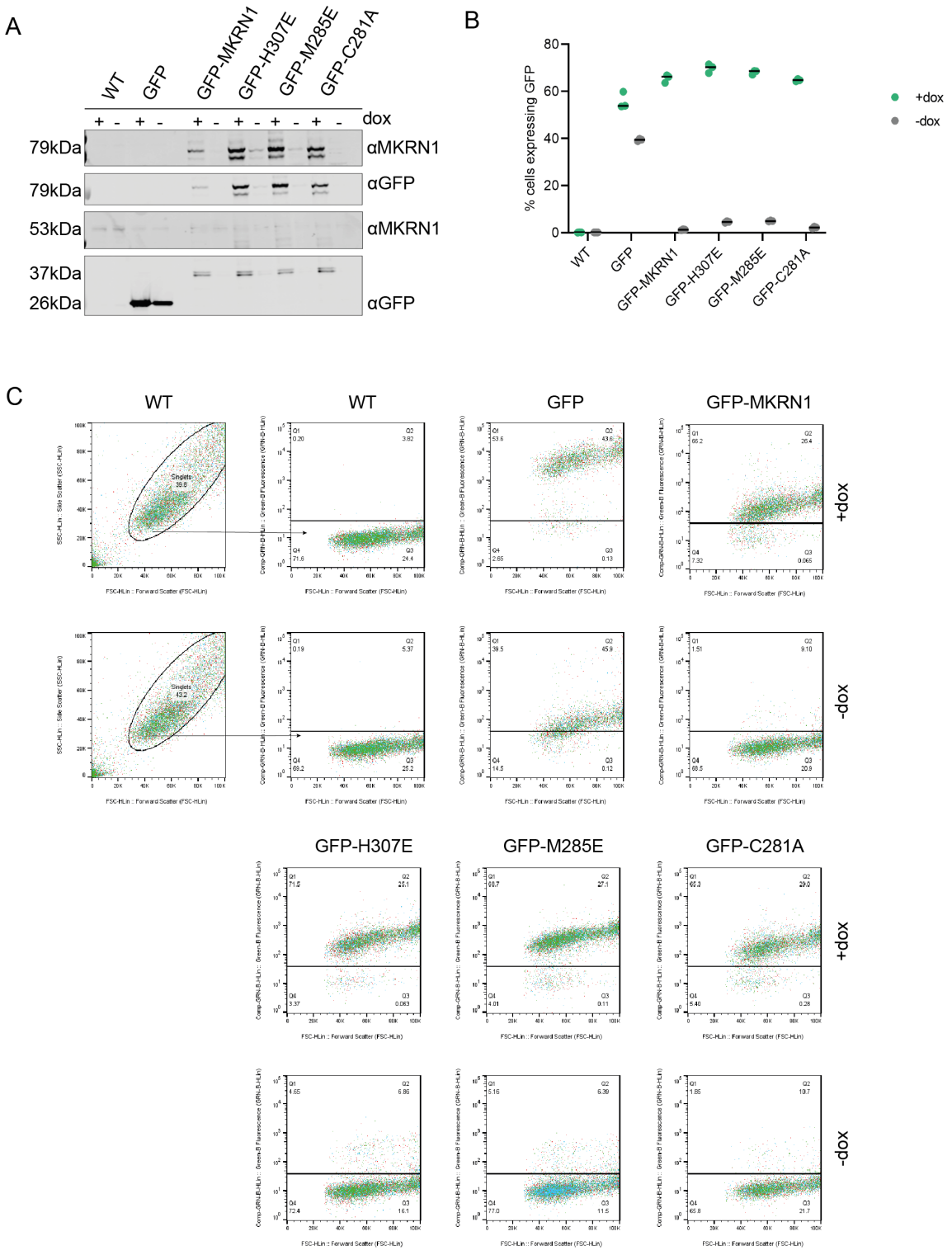

Figure 5.20: **Generation of HEK293 Flp/In GFP-MKRN1 catalytic mutant cell line. A.** HEK293 Flp/In WT, GFP, GFP-MKRN1, GFP-H307E, GFP-M285E and GFP-C281A cells were induced with 1 mg mL^{-1} doxycycline (dox) for 24 h before lysis. Lysates were immunoblotted against MKRN1 and GFP. Expression of GFP-MKRN1 is detected with both anti-GFP and anti-MKRN1 at 79 kDa. **B.** Percentage cells expressing GFP with and without 24 h doxycyline induction (+/-dox) determined by flow cytometry. Flow cytometry shown in **C.** Gating strategy is shown for WT but was applied to all plots.

Table 23. Proteins identified in MKRN1 interactome during HIV-1 infection compared to our ivRBPome and the affinity purification of MKRN1 conducted by Hildebrandt et al, 2019.

Table 23:

Accession	Gene Name	Description	Enriched in HIV-1 infection	Present in ivRBPome	Present in Hildebrandt AP
Q12904	AIMP1	Aminoacyl tRNA synthase complex-interacting multifunctional protein 1 OS=Homo sapiens OX=9606 GN=AIMP1 PE=1 SV=2 - [AIMP1_HUMAN]	Y	N	N
Q9UKV5	AMFR	E3 ubiquitin-protein ligase AMFR OS=Homo sapiens OX=9606 GN=AMFR PE=1 SV=2 - [AMFR_HUMAN]	Y	N	N
Q17R89	ARHGAP44	Rho GTPase-activating protein 44 OS=Homo sapiens OX=9606 GN=ARHGAP44 PE=1 SV=1 - [RHG44_HUMAN]	Y	N	N
Q494R4	CCDC153	Coiled-coil domain-containing protein 153 OS=Homo sapiens OX=9606 GN=CCDC153 PE=1 SV=2 - [CC153_HUMAN]	Y	N	N
Q5W186	CST9	Cystatin-9 OS=Homo sapiens OX=9606 GN=CST9 PE=2 SV=1 - [CST9_HUMAN]	Y	N	N
Q08211	DHX9	ATP-dependent RNA helicase A OS=Homo sapiens OX=9606 GN=DHX9 PE=1 SV=4 - [DHX9_HUMAN]	Y	Y	N
P60981	DSTN	Destrin OS=Homo sapiens OX=9606 GN=DSTN PE=1 SV=3 - [DEST_HUMAN]	Y	N	N
P26641	EEF1G	Elongation factor 1-gamma OS=Homo sapiens OX=9606 GN=EEF1G PE=1 SV=3 - [EF1G_HUMAN]	Y	N	N
P62861	FAU	40S ribosomal protein S30 OS=Homo sapiens OX=9606 GN=FAU PE=1 SV=1 - [RS30_HUMAN]	Y	Y	N
Q92522	H1-10	Histone H1.10 OS=Homo sapiens OX=9606 GN=H1-10 PE=1 SV=1 - [H1X_HUMAN]	Y	N	N
P16403	H1-2	Histone H1.2 OS=Homo sapiens OX=9606 GN=H1-2 PE=1 SV=2 - [H12_HUMAN]	Y	N	N
Q9P035	HACD3	Very-long-chain (3R)-3-hydroxyacyl-CoA dehydratase 3 OS=Homo sapiens OX=9606 GN=HACD3 PE=1 SV=2 - [HACD3_HUMAN]	Y	N	N
Q13151	HNRNPA0	Heterogeneous nuclear ribonucleoprotein A0 OS=Homo sapiens OX=9606 GN=HNRNPA0 PE=1 SV=1 - [ROA0_HUMAN]	Y	N	N
P22626	HNRNPA2B1	Heterogeneous nuclear ribonucleoproteins A2/B1 OS=Homo sapiens OX=9606 GN=HNRNPA2B1 PE=1 SV=2 - [ROA2_HUMAN]	Y	Y	N
Q99729	HNRNPAB	Heterogeneous nuclear ribonucleoprotein A/B OS=Homo sapiens OX=9606 GN=HNRNPAB PE=1 SV=2 - [ROAA_HUMAN]	Y	Y	N
P07910	HNRNPC	Heterogeneous nuclear ribonucleoproteins C1/C2 OS=Homo sapiens OX=9606 GN=HNRNPC PE=1 SV=4 - [HNRPC_HUMAN]	Y	Y	N
Q14103	HNRNPD	Heterogeneous nuclear ribonucleoprotein D0 OS=Homo sapiens OX=9606 GN=HNRNPD PE=1 SV=1 - [HNRPD_HUMAN]	Y	Y	Y
O14979	HNRNPDL	Heterogeneous nuclear ribonucleoprotein D-like OS=Homo sapiens OX=9606 GN=HNRNPDL PE=1 SV=3 - [HNRDL_HUMAN]	Y	N	N
P52597	HNRNPF	Heterogeneous nuclear ribonucleoprotein F OS=Homo sapiens OX=9606 GN=HNRNPF PE=1 SV=3 - [HNRPF_HUMAN]	Y	N	N
P31943	HNRNPH1	Heterogeneous nuclear ribonucleoprotein H OS=Homo sapiens OX=9606 GN=HNRNPH1 PE=1 SV=4 - [HNRH1_HUMAN]	Y	N	N
Q00839	HNRNPU	Heterogeneous nuclear ribonucleoprotein U OS=Homo sapiens OX=9606 GN=HNRNPU PE=1 SV=6 - [HNRPU_HUMAN]	Y	Y	Y
Q9BUJ2	HNRNPUL1	Heterogeneous nuclear ribonucleoprotein U-like protein 1 OS=Homo sapiens OX=9606 GN=HNRNPUL1 PE=1 SV=2 - [HNRL1_HUMAN]	Y	N	N
Q9Y6M1	IGF2BP2	Insulin-like growth factor 2 mRNA-binding protein 2 OS=Homo sapiens OX=9606 GN=IGF2BP2 PE=1 SV=2 - [IF2B2_HUMAN]	Y	N	Y
O00425	IGF2BP3	Insulin-like growth factor 2 mRNA-binding protein 3 OS=Homo sapiens OX=9606 GN=IGF2BP3 PE=1 SV=2 - [IF2B3_HUMAN]	Y	Y	Y
Q12905	ILF2	Interleukin enhancer-binding factor 2 OS=Homo sapiens OX=9606 GN=ILF2 PE=1 SV=2 - [ILF2_HUMAN]	Y	N	N
Q12906	ILF3	Interleukin enhancer-binding factor 3 OS=Homo sapiens OX=9606 GN=ILF3 PE=1 SV=3 - [ILF3_HUMAN]	Y	Y	Y
P52292	KPNA2	Importin subunit alpha-1 OS=Homo sapiens OX=9606 GN=KPNA2 PE=1 SV=1 - [IMA1_HUMAN]	Y	N	N
Q92615	LARP4B	La-related protein 4B OS=Homo sapiens OX=9606 GN=LARP4B PE=1 SV=3 - [LAR4B_HUMAN]	Y	N	Y
P61326	MAGOH	Protein mago nashi homolog OS=Homo sapiens OX=9606 GN=MAGOH PE=1 SV=1 - [MGN_HUMAN]	Y	N	N

Table 23:

Accession	Gene Name	Description	Enriched in HIV-1 infection	Present in ivRBPome	Present in Hildebrandt AP
P43243	MATR3	Matrin-3 OS=Homo sapiens OX=9606 GN=MATR3 PE=1 SV=2 - [MATR3_HUMAN]	Y	N	N
C9JLW8	MCRIP1	Mapk-regulated corepressor-interacting protein 1 OS=Homo sapiens OX=9606 GN=MCRIP1 PE=1 SV=1 - [MCRI1_HUMAN]	Y	N	N
Q09161	NCBP1	Nuclear cap-binding protein subunit 1 OS=Homo sapiens OX=9606 GN=NCBP1 PE=1 SV=1 - [NCBP1_HUMAN]	Y	N	N
P19338	NCL	Nucleolin OS=Homo sapiens OX=9606 GN=NCL PE=1 SV=3 - [NUCL_HUMAN]	Y	Y	N
P11940	PABPC1	Polyadenylate-binding protein 1 OS=Homo sapiens OX=9606 GN=PABPC1 PE=1 SV=2 - [PABP1_HUMAN]	Y	Y	Y
Q13310	PABPC4	Polyadenylate-binding protein 4 OS=Homo sapiens OX=9606 GN=PABPC4 PE=1 SV=1 - [PABP4_HUMAN]	Y	Y	Y
Q86U42	PABPN1	Polyadenylate-binding protein 2 OS=Homo sapiens OX=9606 GN=PABPN1 PE=1 SV=3 - [PABP2_HUMAN]	Y	N	N
P25787	PSMA2	Proteasome subunit alpha type-2 OS=Homo sapiens OX=9606 GN=PSMA2 PE=1 SV=2 - [PSA2_HUMAN]	Y	N	N
P25788	PSMA3	Proteasome subunit alpha type-3 OS=Homo sapiens OX=9606 GN=PSMA3 PE=1 SV=2 - [PSA3_HUMAN]	Y	N	N
P20618	PSMB1	Proteasome subunit beta type-1 OS=Homo sapiens OX=9606 GN=PSMB1 PE=1 SV=2 - [PSB1_HUMAN]	Y	N	N
Q99460	PSMD1	26S proteasome non-ATPase regulatory subunit 1 OS=Homo sapiens OX=9606 GN=PSMD1 PE=1 SV=2 - [PSMD1_HUMAN]	Y	N	N
O00232	PSMD12	26S proteasome non-ATPase regulatory subunit 12 OS=Homo sapiens OX=9606 GN=PSMD12 PE=1 SV=3 - [PSD12_HUMAN]	Y	N	N
O43242	PSMD3	26S proteasome non-ATPase regulatory subunit 3 OS=Homo sapiens OX=9606 GN=PSMD3 PE=1 SV=2 - [PSMD3_HUMAN]	Y	N	N
P26599	PTBP1	Polypyrimidine tract-binding protein 1 OS=Homo sapiens OX=9606 GN=PTBP1 PE=1 SV=1 - [PTBP1_HUMAN]	Y	N	N
P62913	RPL11	60S ribosomal protein L11 OS=Homo sapiens OX=9606 GN=RPL11 PE=1 SV=2 - [RL11_HUMAN]	Y	N	N
P26373	RPL13	60S ribosomal protein L13 OS=Homo sapiens OX=9606 GN=RPL13 PE=1 SV=4 - [RL13_HUMAN]	Y	Y	N
P40429	RPL13A	60S ribosomal protein L13a OS=Homo sapiens OX=9606 GN=RPL13A PE=1 SV=2 - [RL13A_HUMAN]	Y	Y	N
P50914	RPL14	60S ribosomal protein L14 OS=Homo sapiens OX=9606 GN=RPL14 PE=1 SV=4 - [RL14_HUMAN]	Y	Y	N
Q07020	RPL18	60S ribosomal protein L18 OS=Homo sapiens OX=9606 GN=RPL18 PE=1 SV=2 - [RL18_HUMAN]	Y	Y	N
P46778	RPL21	60S ribosomal protein L21 OS=Homo sapiens OX=9606 GN=RPL21 PE=1 SV=2 - [RL21_HUMAN]	Y	N	N
P35268	RPL22	60S ribosomal protein L22 OS=Homo sapiens OX=9606 GN=RPL22 PE=1 SV=2 - [RL22_HUMAN]	Y	N	N
P83731	RPL24	60S ribosomal protein L24 OS=Homo sapiens OX=9606 GN=RPL24 PE=1 SV=1 - [RL24_HUMAN]	Y	Y	N
P46776	RPL27A	60S ribosomal protein L27a OS=Homo sapiens OX=9606 GN=RPL27A PE=1 SV=2 - [RL27A_HUMAN]	Y	Y	N
P62899	RPL31	60S ribosomal protein L31 OS=Homo sapiens OX=9606 GN=RPL31 PE=1 SV=1 - [RL31_HUMAN]	Y	Y	Y
P42766	RPL35	60S ribosomal protein L35 OS=Homo sapiens OX=9606 GN=RPL35 PE=1 SV=2 - [RL35_HUMAN]	Y	Y	Y
Q9Y3U8	RPL36	60S ribosomal protein L36 OS=Homo sapiens OX=9606 GN=RPL36 PE=1 SV=3 - [RL36_HUMAN]	Y	Y	N
P83881	RPL36A	60S ribosomal protein L36a OS=Homo sapiens OX=9606 GN=RPL36A PE=1 SV=2 - [RL36A_HUMAN]	Y	N	N
P62891	RPL39	60S ribosomal protein L39 OS=Homo sapiens OX=9606 GN=RPL39 PE=1 SV=2 - [RL39_HUMAN]	Y	N	N
P36578	RPL4	60S ribosomal protein L4 OS=Homo sapiens OX=9606 GN=RPL4 PE=1 SV=5 - [RL4_HUMAN]	Y	Y	N
P18124	RPL7	60S ribosomal protein L7 OS=Homo sapiens OX=9606 GN=RPL7 PE=1 SV=1 - [RL7_HUMAN]	Y	Y	N

Table 23:

Accession	Gene Name	Description	Enriched in HIV-1 infection	Present in ivRBPome	Present in Hildebrandt AP
P46783	RPS10	40S ribosomal protein S10 OS=Homo sapiens OX=9606 GN=RPS10 PE=1 SV=1 - [RS10_HUMAN]	Y	N	N
P62277	RPS13	40S ribosomal protein S13 OS=Homo sapiens OX=9606 GN=RPS13 PE=1 SV=2 - [RS13_HUMAN]	Y	N	Y
P62263	RPS14	40S ribosomal protein S14 OS=Homo sapiens OX=9606 GN=RPS14 PE=1 SV=3 - [RS14_HUMAN]	Y	Y	N
P62244	RPS15A	40S ribosomal protein S15a OS=Homo sapiens OX=9606 GN=RPS15A PE=1 SV=2 - [RS15A_HUMAN]	Y	N	N
P62249	RPS16	40S ribosomal protein S16 OS=Homo sapiens OX=9606 GN=RPS16 PE=1 SV=2 - [RS16_HUMAN]	Y	N	N
P08708	RPS17	40S ribosomal protein S17 OS=Homo sapiens OX=9606 GN=RPS17 PE=1 SV=2 - [RS17_HUMAN]	Y	N	N
P15880	RPS2	40S ribosomal protein S2 OS=Homo sapiens OX=9606 GN=RPS2 PE=1 SV=2 - [RS2_HUMAN]	Y	Y	Y
P62266	RPS23	40S ribosomal protein S23 OS=Homo sapiens OX=9606 GN=RPS23 PE=1 SV=3 - [RS23_HUMAN]	Y	N	N
P62847	RPS24	40S ribosomal protein S24 OS=Homo sapiens OX=9606 GN=RPS24 PE=1 SV=1 - [RS24_HUMAN]	Y	Y	N
P62979	RPS27A	Ubiquitin-40S ribosomal protein S27a OS=Homo sapiens OX=9606 GN=RPS27A PE=1 SV=2 - [RS27A_HUMAN]	Y	N	N
P23396	RPS3	40S ribosomal protein S3 OS=Homo sapiens OX=9606 GN=RPS3 PE=1 SV=2 - [RS3_HUMAN]	Y	N	N
P62081	RPS7	40S ribosomal protein S7 OS=Homo sapiens OX=9606 GN=RPS7 PE=1 SV=1 - [RS7_HUMAN]	Y	N	N
P62241	RPS8	40S ribosomal protein S8 OS=Homo sapiens OX=9606 GN=RPS8 PE=1 SV=2 - [RS8_HUMAN]	Y	Y	Y
P46781	RPS9	40S ribosomal protein S9 OS=Homo sapiens OX=9606 GN=RPS9 PE=1 SV=3 - [RS9_HUMAN]	Y	Y	Y
Q9Y467	SALL2	Sal-like protein 2 OS=Homo sapiens OX=9606 GN=SALL2 PE=1 SV=4 - [SALL2_HUMAN]	Y	N	N
P53007	SLC25A1	Tricarboxylate transport protein, mitochondrial OS=Homo sapiens OX=9606 GN=SLC25A1 PE=1 SV=2 - [TXTP_HUMAN]	Y	N	N
Q9UBX3	SLC25A10	Mitochondrial dicarboxylate carrier OS=Homo sapiens OX=9606 GN=SLC25A10 PE=1 SV=2 - [DIC_HUMAN]	Y	N	N
P08579	SNRPB2	U2 small nuclear ribonucleoprotein B” OS=Homo sapiens OX=9606 GN=SNRPB2 PE=1 SV=1 - [RU2B_HUMAN]	Y	N	N
Q01130	SRSF2	Serine/arginine-rich splicing factor 2 OS=Homo sapiens OX=9606 GN=SRSF2 PE=1 SV=4 - [SRSF2_HUMAN]	Y	Y	N
Q16629	SRSF7	Serine/arginine-rich splicing factor 7 OS=Homo sapiens OX=9606 GN=SRSF7 PE=1 SV=1 - [SRSF7_HUMAN]	Y	N	N
Q13242	SRSF9	Serine/arginine-rich splicing factor 9 OS=Homo sapiens OX=9606 GN=SRSF9 PE=1 SV=1 - [SRSF9_HUMAN]	Y	N	N
O60506	SYNCRIP	Heterogeneous nuclear ribonucleoprotein Q OS=Homo sapiens OX=9606 GN=SYNCRIP PE=1 SV=2 - [HNRPQ_HUMAN]	Y	Y	Y
Q3ZCQ8	TIMM50	Mitochondrial import inner membrane translocase subunit TIM50 OS=Homo sapiens OX=9606 GN=TIMM50 PE=1 SV=2 - [TIM50_HUMAN]	Y	N	N
Q13595	TRA2A	Transformer-2 protein homolog alpha OS=Homo sapiens OX=9606 GN=TRA2A PE=1 SV=1 - [TRA2A_HUMAN]	Y	N	N
P62995	TRA2B	Transformer-2 protein homolog beta OS=Homo sapiens OX=9606 GN=TRA2B PE=1 SV=1 - [TRA2B_HUMAN]	Y	N	N
Q13263	TRIM28	Transcription intermediary factor 1-beta OS=Homo sapiens OX=9606 GN=TRIM28 PE=1 SV=5 - [TIF1B_HUMAN]	Y	N	N
P67809	YBX1	Y-box-binding protein 1 OS=Homo sapiens OX=9606 GN=YBX1 PE=1 SV=3 - [YBOX1_HUMAN]	Y	Y	Y

5.3 Discussion

Previous experiments assessing HIV-1 infection in MKRN1-KO cells lead me to hypothesise that MKRN1 regulates the production of virions, possibly through the ubiquitination of HIV-1 Gag, cellular ESCRT proteins or unknown cellular proteins with roles in assembly. This chapter aimed to 1) confirm these observations using other cell lines and clones, 2) determine if viral molecules accumulate at the plasma membrane using fluorescence microscopy and 3) identify the interactome of MKRN1 in HIV-1 infection.

A virus release assay was conducted in three different HEK293 Flp/In MKRN1-KO clones and a HEK293T MKRN1-KO clone. These cell lines were transfected with HIV-$1_{\text{-R-E-Nef-mCherry}}$ and VSVG for 48 h then analysed for P24 and P55 expression within cell lysates and supernatant. Unlike in CD4+ T lymphocytes, MKRN1-KO HEK293 cell lysates did not show increased levels of P24 and P55. This could be a result of the experimental approach as P24 and P55 expression in MKRN1-KO CD4+ T cells was measured in HIV-$1_{\text{-R-E-Nef-mCherry}}$ infected cells whereas in MKRN1-KO HEK293s, P24 and P55 expression was determined in transfected cells. Moreover, differences could also be biological as HIV-1 gene expression is slower in CD4+ T cells than in HEK293T/Flp/In cells and the transcriptome and proteome of both cell types strongly differ. Transfection efficiency did not differ between MKRN1-KO and WT cells. Therefore, alterations in P24 expression were not caused by differential transfection. Whole cell lysates of WT and MKRN1-KO cells transfected with GFP could be western blotted as a control to determine if cellular gene expression was different between WT and MKRN1-KO cells. Whilst the growth rate and viability of MKRN1-KO CD4+ T cells did not differ from WT CD4+ T cells, the same assays were not conducted for the various HEK293 KO clones. While no visual difference in shape and proliferation were observed in these experiments, growth and viability will be formally tested in HEK293 Flp/In and HEK293T MKRN1-KO. WT and KO HEK293T/ Flp/In cells will also be transfected with HIV-$1_{\text{-R-E-Nef-mCherry}}$ and HIV-$1_{\text{-R-E-Gag-mCherry}}$ and assessed in our plate reader assay to test with higher sensitivity and resolution if there are differences in viral gene expression in both the first and second round of infection. This approach is more

sensitive and more comprehensive than western blot as it provides kinetic information, enabling me to determine if MKRN1-KO affects gene expression in a temporal manner. I could not confirm that P24 or P55 is increased in MKRN1-KO HEK293T/Flp/In cells. However, the smFISH data (discussed below) shows Gag accumulation at the plasma membrane in MKRN1-KO HEK293 Flp/In cells. These results suggest that even if the overall levels of Gag could be similar in WT and KO HEK293/T cells, their distribution in the cell differs. The accumulation of Gag at the plasma membrane in KO cells further reinforce the idea of defects in assembly in MKRN1-KO cells. The lack of concentration of vRNA molecules at the assembly sites indicate that the effect of MKRN1 could be associated to the recruitment of the vRNA to the plasma membrane. This could explain the reduction in viral particle production.

A clear decrease in P24 concentration in the supernatant from HEK293T/Flp/In MKRN1-KO cells was shown by western blot. This effect was further confirmed by the substantial decrease in P24 concentration in the purified virus preparations. These investigations strongly supported the notion of a reduction in viral particle formation and validated the effect observed in MKRN1-KO CD4+ T lymphocytes.

An unanswered question from the previous chapter was whether MKRN1 had any role inside the viral particle. We hypothesised that RBPs identified within HIV-1 particles were deliberately packaged to enable the initial steps of infection in the newly infected cell. Therefore, the infectivity of viral particles produced in MKRN1-KO HEK293T/Flp/In cells was measured after infecting WT cells using flow cytometry. No difference in infectivity was observed for HEK293 Flp/In cells, while a slight effect was found in HEK293T but requires confirmation. These results suggest that MKRN1 does not affect viral particle infectivity by regulating reverse transcription or integration. Conversely, Dr Manuel Garcia-Moreno showed that KO of the ivRBPs PURA and PURB substantially reduces viral infectivity [135]. The opposing results between MKRN1 and PURA/B support the notion that ivRBPs could play roles prior to virus assembly, during viral particle production or in the early steps upon infection of a new cell. The roles of each individual ivRBP should thus be assessed in a case-by-case manner. Regulation of virus assembly by MKRN1 may lead to its packaging into viral

particles because of its localisation to the genomic RNA.

Quantification of vRNAs by smFISH revealed that there was no increase in the number of transcription foci in MKRN1-KO cells compared to WT HEK293s. This supported my hypothesis that the increased levels of vRNA observed previously in MKRN1-KO cells is not a byproduct of increased viral transcription but the result of an intracellular accumulation of vRNA due to defective viral assembly. Interestingly, whilst the median copies of vRNAs in MKRN1-KO vs WT cells at 24, 48 and 72 hpi did not differ, a subpopulation of cells emerged in MKRN1-KOs that had fewer copies of vRNA per cell. Cells were infected at an MOI of 1 which means, in theory, one virus infects each cell. However, in practice, there will be a variation in the number of viruses infecting each cell, meaning some cells are not infected whereas others receive a high viral load. The subpopulation of 'low-copy' cells seen in MKRN1-KOs are likely cells that were infected at a very low level. However, if this were to be the case, the same effect would have been observed in WT cells. Alternatively, this effect could be caused by later activation of provirus in MKRN1-KO cells leading to a lower level of transcription. Regardless, it is possible that this low-copy subpopulation is not observed in WT cells because such a low level of vRNA is efficiently removed from the cell via the budding of viral particles. Whereas in MKRN1-KO cells, these vRNAs remain in the cytoplasm, which supports the idea that viral budding is affected by the absence of MKRN1. It is also possible that whilst the MKRN1-KO was produced from a single cell, more than one cell was isolated and the population of cells was not entirely clonal. This could be causing the observed bi-modal distribution and should be ruled out by sequencing the MKRN1-KO clone.

The localisation of vRNAs also did not differ between MKRN1-KO and WT cells. In both cell types vRNAs were diffuse throughout the cell and became more enriched at the plasma membrane over 72 h. This indicates the vRNAs are still able to interact with Gag at the plasma membrane as when the plasma membrane targeting or RNA binding activity of Gag is disrupted, HIV-1 RNAs are highly dynamic and move in and out of the proximity of the plasma membrane [314]. This also implies Gag-Gag interactions are not disrupted as in the presence of a Gag-mutant that has RNA- and

plasma membrane-binding ability but cannot assemble into viral particles, the vRNA is observed initially docked at the membrane but then drifts and dissociates from the membrane [314]. This suggests that any disruption to viral budding driven by MKRN1-KO, does not mechanistically involve Gag targeting to the plasma membrane, vRNA or Gag-Gag interaction. However, the higher concentration of Gag at the plasma membrane does not promote a more efficient recruitment of viral genomic RNA. This suggests that MKRN1 may play a role in the capture or assembly of the vRNA into viral particles.

Analysis of Gag localisation in HIV-1 infected MKRN1-KO cells provided further support for the defective assembly or release of viral particles. Gag accumulates at the plasma membrane in MKRN1-KO cells, confirming plasma-membrane targeting is not affected. However, a substantially higher signal intensity at 72 hpi is recorded in the membrane of MKRN1-KO cells in comparison to WT cells. This indicates that viral particles are able to bud from WT cells, removing Gag with them, whereas viral particle budding and release are impaired in MKRN1-KO cells. Hence, Gag intensity continues to increase in the plasma membrane. This theory is in agreement with experiments in which retroviral budding was inhibited with the proteasome inhibitor MG-132; Gag was visualised by confocal microscopy to accumulate on the plasma membrane [243].

Other studies describing HIV-1 budding defects utilise electron microscopy to visualise HIV-1 budding structures and quantify viral particle release by counting the number of structures attached to the plasma membrane [244, 291, 315]. Whilst this literature does not explicitly measure the intensity of Gag expression in the membrane, it follows that an increase in viral particle retention would increase the intensity of Gag at the plasma membrane. Similar experiments could be carried out here to test whether MKRN1 is affecting the recruitment of vRNA and its assembly into viral particles or the release of budding virions.

Another method by which a defect in viral particle assembly or release was quantified was through western blot analysis of infected cell lysate vs supernatant. Gag increased in cell lysate and decreased in the supernatant when viral budding was inhibited [167,246,247,292]. Therefore, observing the accumulation of Gag in MKRN1-

KO cells is sufficient to conclude that MKRN1-KO impairs HIV-1 budding. Further investigation to validate this would be to quantify the viral particles in the supernatant of WT and MKRN1-KO cells by RT-qPCR for viral gRNA or a P24 ELISA. This would undeniably show that the accumulation of Gag in MKRN1-KO cells is because viruses are not efficiently formed or released. Supernatants from the cells in the smFISH experiments were collected for this purpose but are yet to be analysed.

To investigate how MKRN1 regulates HIV-1 budding, the interactome of MKRN1 in HIV-1 infected cells was determined using IP in combination with LC-MS analysis. A HEK293 Flp/In GFP-MKRN1 cell line was generated to enable specific immunoprecipitation of the fusion protein using the highly selective nanobodies against GFP. Whilst MKRN1-GFP was expressed, I noticed replcaement of the endogenous protein by th erecombinant protein. This suggested that MKRN1 is a key regulatory protein, the expression of which is tightly regulated in cells. This observation was reinforced by the fact that GFP-tagged catalytically inactive mutants of MKRN1 were expressed to a greater extent than WT GFP-MKRN1. Although, expression of WT MKRN1 was still downregulated in the GFP-MKRN1 mutant cell lines. This agrees with previous studies showing that N-terminus fusion of MKRN1 does not affect its activity [152, 159, 160]. Future work must though functionally test the activity of GFP-MKRN1 with in vitro ubiquitination assays.

Surprisingly, the interactome of MKRN1 in infected cells did not include any HIV-1 proteins or members of the cellular ESCRT pathway. However, this data was only from one biological repeat and is therefore not conclusive. A subset of proteins identified as MKRN1-interactors were also shown as MKRN1-interactors by Hildebrandt et al [160], thus we can assume that these are likely to be true interactions. Additionally, many of these proteins were also identified as part of our ivRBPome which further supports their role in HIV-1 infection. It is possible that ESCRT proteins and HIV-1 proteins were not identified because of the experimental set up and this is something I want to improve further. Whilst the conditions for IP of GFP-MKRN1 were optimised, the MOI and duration of infection were not developed and could be limiting the results of the IP. I deliberately chose to infect with HIV-$1_{\text{-R-E-Nef-mCherry}}$ so that I could confirm infection

of cells with mCherry fluorescence but maintain WT expression of Gag in case the mCherry tag interfered with binding. However, Nef is expressed at an earlier stage of infection compared to Gag so it is possible that Gag was not abundant enough to see any interaction with MKRN1, especially considering other results in this thesis were observed at 48 or 72 hpi.

Whilst there are limitations to the interactome data, we can take the overlap with the Hildebrandt dataset and the ivRBPome as an indication that these overlapping proteins, at least, are likely true interactions. A total of 31 proteins in the MKRN1 HIV+ IP were also found in the ivRBPome and are all heavily involved in RNA regulation. GO term analysis of these proteins revealed all 31 proteins to be involved in regulation of gene expression as well as subsets being involved in RNA metabolism, processing, stability, splicing and transport (RPL4, RPL7, RPL13, RPL13A, RPL14, RPL18, RPL24, RPL27A, RPL31, RPL35, RPL36, RPS2, RPS8, RPS9, RPS14, RPS24, FAU, SYNCRIP, PABPC1, PABPC4, YBX1, DHX9, HNRNPA2B1, HNRNPC, HNRNPD, HNRNPU, SRSF2, IGF2BP3). There was also an enrichment for proteins involved in protein localisation, protein transport and protein targeting to the plasma membrane (RPL4, RPL7, RPL13, RPL13A, RPL14, RPL18, RPL24, RPL27A, RPL31, RPL35, RPL36, RPS2, RPS8, RPS9, RPS14, RPS24, FAU, HNRNPA2B1, HNRNPU, SRSF2, DHX9 and YBX1). These proteins are all highly interconnected (protein-protein interaction enrichment p value $<1x10^{-16}$), suggesting they are likely to be incorporated into HIV-1 particles as a complex. In fact, all 31 of these proteins have been identified bound to HIV-1 vRNAs inside cells, supporting the idea they may form part of the HIV-1 RNP [116, 117]. Interestingly, IGF2BP3, YBX1, hnRNPA2B1 and SYNCRIP were all enriched on HIV-1 gRNA specifically [117]. Although, outside of our RBPome, MKRN1 has never been identified in association with HIV-1 RNA which suggests MKRN1 may interact with the vRNA indirectly, at least for cytoplasmic stages of infection. However, MKRN1 has such low cellular abundance that this may prevent it from being detected; we find approximately one copy of MKRN1 bound to each vRNA in the HIV-1 particle.

In addition to being heavily enriched in RNA regulatory roles, many of the identified

proteins are also known regulators of HIV-1 infection. Some have been shown to be anti-viral, such as SRSF2 which alters HIV-1 splicing to promote Tat production [316, 317] but also inhibits Rev function and nuclear export of vRNAs [318]. However, most have been described as pro-viral proteins. HNRNPA2B1, enhances HIV-1 transcription by unfolding LTR promoter G-quadruplexes [319], although depletion of HNRNPA2B1 increased HIV-1 Gag expression [320]. Knockdown of HNRNPU was found to inhibit HIV-1 replication [110]. Similarly, depletion of HNRNPD reduced Gag synthesis and HNRNPD has been shown to relocate to the cytoplasm during infection to associate with Gag [320]. YBX1, SYNCRIP and IGF2BP3 are all known to stabilise HIV vRNAs [100, 117, 321] and knockdown of YBX1 and IGF2BP3 decreases HIV-1 late gene expression [117]. Additionally, SYNCRIP increases protein expression from HIV-1 rev-dependent transcripts [322]. DHX9 is incorporated into HIV-1 viral particles by binding to the primer binding site segment of gRNA and increases infectivity of virions by enhancing processivity of HIV-1 reverse transcriptase on gRNA [101, 323]. RPL7, major ribosomal protein involved in the regulation of translation, binds HIV-1 Gag via the NC domain and has been suggested to act as a chaperoning factor contributing to the start of Gag assembly [324]. Given the roles of these proteins, they likely form a complex on the vRNA to protect it from degradation and ensure its incorporation into assembling virions.

Another protein identified in the IP and also found interacting with MKRN1 by Hildebrandt et al and in HIV-1 viral particles in our ivRBPome is PABPC1. The interaction between MKRN1 and PABPC1 is not surprising as they are known to regulate the ribosome quality control complex. However, the enrichment of PABPC1 interaction with MKRN1 in HIV-1 infection is interesting as PABPC1 is cleaved by HIV-1 protease. The cellular function of PABPC1 is to promote the translation of mRNAs through interaction with the 3' poly(A) tails and protect the poly(A) tails from degradation. However, HIV-1 does not use poly(A)-dependent translation and instead contains an internal ribosome entry site (IRES) in the 5' UTR of the vRNA which undergoes IRES-dependent translation initiation [325]. HIV-1 protease cleaves PABPC1 to inhibit cellular poly(A)-dependent translation and reduce cellular protein

synthesis in favour of viral protein synthesis [326]. Therefore, it is surprising to find it enriched in HIV-1 infection. However, it was identified bound to HIV-1 gRNA in our ivRBPome and was found bound to vRNAs twice by Knoener et al [116, 117], thus PABPC1 must have another function in HIV-1 infection. Perhaps MKRN1 and PABPC1 act together to prevent degradation of the HIV-1 3' poly(A) tail.

The highly connected network of RNA regulatory proteins and lack of viral proteins or ESCRT proteins in the interactome of MKRN1 suggests that MKRN1 indirectly regulates viral budding. The stability or transport of the vRNA could be part of this mechanism. However, HIV-1 gRNA is not required for the budding of viral particles [327] therefore the accumulation of Gag would not be observed. Alternatively, MKRN1 may ubiquitinate some of these proteins to aid assembly of the vRNP as PABPC1, PABPC4, IGF2BP3 and SYNCRIP are all ubiquitination targets of MKRN1 [160]. Additional investigation into the interactome of MKRN1 is required before any further conclusions are drawn as it is believed that the experimental conditions were not optimal for capture of MKRN1 interactions during HIV-1 infection. I also aim to investigate the RNA-binding and ubiquitination activity of MKRN1. This will be partly achieved using the MKRN1 catalytic mutants designed and cloned in this chapter and will be discussed in detail in Chapter 7. Currently, it is unclear exactly what the mechanism of action is. However, the data presented in this chapter provides further evidence that MKRN1 regulates viral particle release and may control the correct assembly of virions.

6 Establishing a system to visualise HIV-1 infection from viral particle entry to integration

6.1 Introduction

The life cycle of HIV-1 is a complex process involving many cellular interactors. Recent scrutiny of the early steps of HIV-1 infection have revealed the viral capsid to be essential for protecting the vRNA and cDNA from cytoplasmic innate immune factors. It is now understood that the capsid remains intact throughout its journey to the nucleus. However, exactly when and how the core loses integrity is still under debate. It is also well-established that cellular proteins can support or restrict HIV-1 infection. These proteins play essential roles in infection and therefore must be packaged inside the capsid to access the viral RNA. Although, little is known about how these proteins are packaged and act inside the capsid and whether productive infection can occur in their absence. Powerful microscopic techniques such as super-resolution microscopy and live-cell single virus particle tracking have helped us to elucidate the dynamics of the early stages of the HIV-1 life cycle [314, 328, 329]. Here I establish a new system to visualise HIV-1 infection from viral particle entry to integration using super-resolution microscopy of fluorescently labelled HIV-1.

Fluorescent labelling of viral proteins has enabled the tracking of individual viral particles during HIV-1 infection. Different HIV-1 proteins have been fluorescently labelled to explore different stages of the viral life cycle, commonly Vpr and Gag. The limitation of fluorescent labelling of Vpr is that it often inhibits viral transcription, thus labelled Vpr must be co-expressed with WT Vpr [328, 330]. Similar issues also arise when labelling Gag as each domain conducts a crucial role within in infection. The first replication-competent labelled Gag was achieved by inserting GFP between MA and CA domains [329]. However, it is insufficient for studying the early stages of infection as fluorescent signal is lost upon fusion of virions with the plasma membrane. To achieve labelling of the capsid, CA or another protein within the capsid needs to be labelled which has proven difficult due to the structural complexity of the HIV-1

capsid. HIV-1 CA is a 24 kDa protein, a similar size to fluorescent proteins such as GFP or RFP. CA hexamers are the core structural element of immature viral particles, the building blocks of the mature capsid and they mediate protein-protein and RNA-protein interactions throughout infection, all of which requires extreme flexibility and specificity from CA. Accordingly, mutagenic analyses of HIV-1 CA found it to be highly genetically fragile with up to 89% of single amino acid mutations tested resulting in abolishment of viral replication [331]. It is therefore unsurprising that the addition of a fluorescent protein to CA disrupts functionality and often results in a reduction or complete loss of infectivity. Usually, complementation with WT virus is required to restore infectivity [76, 332]. More recent techniques have enabled the direct labelling of capsid through genetic code expansion and click labelling. This process introduces a noncanonical amino acid carrying a highly reactive bio-orthogonal functional group into CA which can then be directly coupled to a photostable chemical dye without the need for complementation with WT [333]. Alternatively, proteins that interact with CA can be targeted. One method is to co-express a plasmid for the CA interactor PPIA (cyclophillin A)-dsRed with WT HIV-1 [74, 334]. PPIA promotes stability of the HIV-1 capsid and PPIA-dsRed binds to CA via an exposed loop in the N-terminus of CA that does not disrupt assembly of the hexameric lattice [335]. However, this requires efficient co-transfection of WT HIV-1 and PPIA-dsRed to achieve capsid labelling. Additionally, the marker is not a viral protein and therefore cannot unambiguously visualise the capsid.

Another approach to visualising the early stages of HIV-1 infection is to label the proteins within the capsid. Whilst still technically difficult due to the small size of the proteins, Vpr [328, 336] and IN have been successfully tagged with fluorophores. Attempts have been made to label RT but function and consequently viral replication was lost [337]. The benefit of tagging IN is that infection can be followed up to the point of integration of the viral cDNA into the host genome. HIV-1 IN is a 32 kDa protein that catalyses the integration of viral DNA into the host genome. It contains three canonical domains connected by flexible linkers: an N-terminal zinc-binding domain, a catalytic core and a C-terminal nonspecific DNA-binding domain [338, 339]. IN activity is

regulated by interactions with viral and cellular proteins which form the pre-integration complex (PIC) on the viral cDNA. The PIC protects IN from degradation, enhances the catalytic activity of IN and was thought to target IN to the nucleus, although this is now debatable given more recent discoveries regarding capsid uncoating and nuclear entry. Particularly, IN forms stable tetramers with the cellular protein LEDGF (lens epithelium–derived growth factor p75 splice variant), a stress-induced transcriptional co-activator, which binds the catalytic core of IN and tethers the viral cDNA to transcriptionally active regions of the host genome [340, 341].

The first labelling of IN was achieved by Arhel et al using FlAsH labelling to image intracytoplasmic and intranuclear HIV-1 complexes [62]. Since then several groups have generated different fluorescent-IN fusion proteins. It appears the C terminus of IN cannot be labelled whilst maintaining IN function [342]. This may be due to a structural rearrangement that IN oligomers undergo upon interaction with LEDGF which masks the C terminus [343]. However, fusion at the N terminus behaved normally. Three groups have exploited the transincorporation of Vpr by the p6 domain of Gag into HIV-1 particles and generated a Vpr-IN-fluorescent protein which had a protease cleavage site inserted between Vpr and IN [75, 344, 345]. However, the viral particles obtained had very low infectivity. Francis et al built on this method by introducing two mutations to IN (E11K and K186E) which are known to stabilise IN tetramers and improve infectivity of transcomplemented viruses [346]. These mutations did successfully increase infectivity of HIV-IN-eGFP and infection could be visualised up to integration [345]. Another approach generated Gag-IN-eGFP fusion proteins with a protease cleavage site inserted between Gag and IN, leading to release of IN-eGFP upon maturation of viral particles [347]. A similar approach was used by Mamede et al who developed two Gag-IN-fluorescent protein constructs in pCMV and psPAX2 lentiviral backbones. Impressively, they achieved a fully labelled population of particles without any non-fluorescent particles being present and could maintain high levels of infectivity with mRuby3, sfGFP and mNeonGreen labelling [342]. Currently, nobody has reported successful generation and incorporation of fluorescently labelled IN without fusing to either Gag or Vpr.

Under physiological conditions, IN is not fused to Gag or Vpr. Instead, HIV-1 IN is incorporated into assembling viral particles as part of the Gag-Pol polyprotein. During virion maturation Gag-Pol undergoes proteolytic cleavage which produces MA, CA, NC, p6 and two spacer peptides from Gag and PR, RT and IN from Pol. These processed proteins then rearrange to give the structure of the mature virion and IN associates with the vRNA [348]. In addition to the role of IN in integration, IN also plays an essential but poorly understood role in maturation of viral particles. It was first shown that mutation of IN resulted in a curious phenotype whereby empty capsids are formed and the vRNA is instead found outside of the capsid in the viral particle [349]. Since then many specific mutations of IN and the impact they have on viral particle maturation have been characterised. These mutations disrupt the binding of IN to vRNA by either preventing packaging of IN, directly blocking RNA binding or impairing the oligomerisation of IN, all of which render the virions defective [350, 351]. Therefore, it is imperative that any IN fusion proteins maintain proper incorporation of IN into viral particles and interaction with the vRNA. Whilst the fluorescent IN fusion proteins described do produce infectious virions, infectivity is often decreased [342,344,347] which suggests the interaction of IN during viral particle maturation may be affected. Given that I wish to use fluorescent IN to investigate the incorporation of cellular proteins into HIV-1 particles and uncover their role inside the viral particle and during the early stages of infection, it is crucial that the viral assembly process occurs unaltered and that mature capsid structure is maintained. It is possible that the non-physiological incorporation of IN into viral particles via fusion to Gag or Vpr disrupts normal assembly and may affect interaction with host proteins. Therefore, I have designed a fluorescent IN construct in which the fluorescent protein mScarlet is fused directly to IN (INscarlet) and is expressed and incorporated into assembling virions endogenously as part of Gag-Pol. The mScarlet fluorophore was chosen as its fluorescence is brighter than that of mCherry and other red fluorescent proteins [352].

In this chapter, I have designed and validated novel INscarlet HIV-1 constructs. This construct has been assembled in both HIV-1 replicons (-R-E) and fully infectious HIV-1 pNL4-3 backbone. All of the currently reported fluorescent IN fusion proteins

are expressed in HIV-1 replicons and require pseudotyping with VSV-G envelope to produce lentiviral particles [75, 342, 344, 345, 347]. VSV-G-enveloped virions enter cells through fusion with the LDL receptor followed by endocytosis [241]. Whilst HIV-1 entry via endocytosis has been documented, the capsid cores are degraded and do not lead to productive infection [242]. Therefore, to truly reveal the host-virus interactions occurring during the early stages of infection, the same CD4-mediated entry pathways must be used. However, to increase the flexibility of the tool, I also designed an INscarlet replicon that can be pseudotyped with VSV-G or HIV-1 Env. This enables the infection of a broad range of cell types as LDL receptors are ubiquitously expressed. The introduction of the fluorescent protein mScarlet into the N terminus of HIV-1 IN is expected to allow visualisation of the early stages of HIV-1 infection up to integration whilst maintaining high infectivity and functionality of IN. Therefore, this tool has the potential to be used in studies of the early stages of HIV-1 infection and the cellular proteins involved.

6.2 Results

To enable direct fluorescent labelling of the HIV-1 capsid for both fixed and live-cell imaging, I engineered an mScarlet-labelled HIV-1 IN construct which I introduced into both the pNL4-3 and pNL4-3-R-E backbones, generating HIV-$1_{\text{INscarlet}}$ and HIV-$1_{\text{-R-E-INscarlet}}$ respectively. The latter, when co-transfected with a plasmid expressing VSV-G, enables pseudotyping of single-round HIV-1 with a VSV-G envelope. Generating both fully infectious and single-round pseudotyped viruses broadens the scope of experiments that can be performed with INscarlet. Single-round HIV-$1_{\text{-R-E-INscarlet}}$ enables the study of HIV-1 infection in non-target cells in a biosafety level 2 (BSL2) environment, whilst HIV-$1_{\text{INscarlet}}$ enables the study of multiple rounds of infection under physiological conditions in CD4+ T lymphocytic cells or cells expressing CD4 and CXCR4. Attempts by other groups suggest tagging the C terminus of IN produces defective viruses whilst more success was found by tagging the N terminus [342]. Therefore, I chose to label the N terminus of IN. Previous attempts within the Castello lab to fluorescently label the N terminus of IN resulted in cleavage

of the fluorescent protein. Here, I inserted mScarlet within the N terminus IN sequence to minimise cleavage. The designed construct starts with the first five N-terminal amino acids of IN to maintain the protease cleavage site, followed by a short, flexible glycine-rich linker region, mScarlet, followed by another glycine-rich linker then the remaining IN sequence (Figure 6.1A). The glycine-rich linkers allow the independent and correct folding of IN and Scarlet preventing intramolecular disturbances. Restriction digest and ligation were used to introduce a *de novo* assembled INscarlet DNA fragment in both pNL4-3 and pNL4-3-R-E vectors (Figure 6.1B and C). Assembled plasmids were transformed into *E. coli* and the resulting colonies were screened for presence of the pNL4-3-INscarlet and pNL-4-3-R-E-INscarlet plasmids. The pNL4-3-INscarlet plasmid was correctly assembled in all colonies tested (Figure 6.1D) and pNL4-3-R-E-INscarlet was correctly assembled in all but one colony (Figure 6.1E). Colonies 1, 2 and 3 were confirmed by sequencing for both plasmids.

Figure 6.1: **Novel pNL4-3 HIV-1 and single round viruses were engineered to express mScarlet-tagged IN. A** INscarlet construct design with mScarlet inserted in the N-terminus of IN. **B.** Detailed view of the N-terminal insertion of mScarlet. **C.** HIV-$1_{INscarlet}$ and **D.** HIV-$1_{-R-E-INscarlet}$ comprise the coding region of pNL4-3 HIV-1 gRNA showing the insertion of mScarlet within the genome. In **D.** stars represent the frameshift mutations nullifying *vpr* and *env*. **E** Colony screen of Stbl3 *E. coli* containing pNL4-3-INscarlet and pNL4-3-R-E-INscarlet in **F.** Proper assembly of the plasmids is indicated by bands at 10.5 and 4.9 kb.

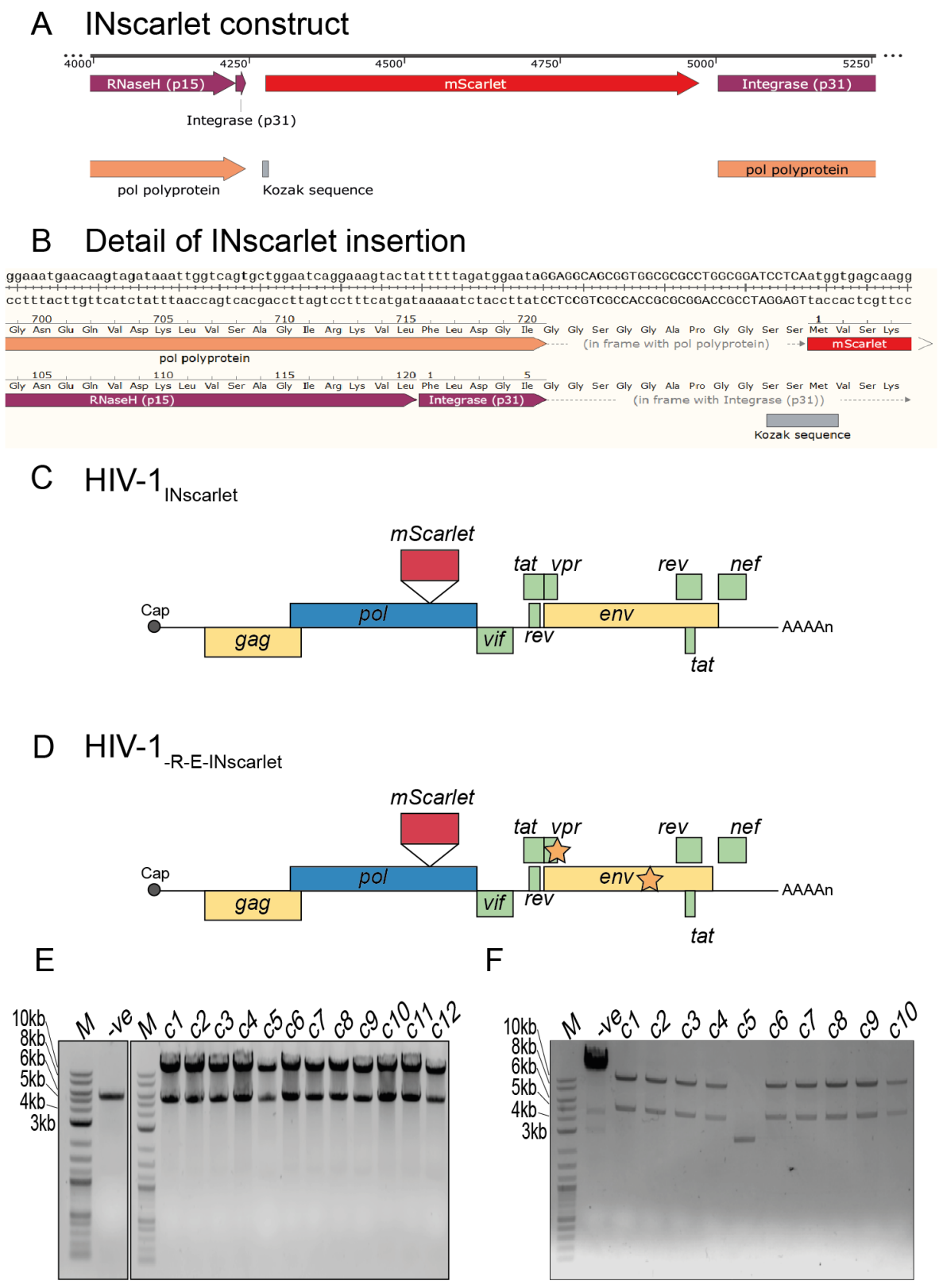

Figure 6.1: **Novel pNL4-3 HIV-1 and single round viruses were engineered to express mScarlet-tagged IN.** Caption on previous page.

To determine whether INscarlet was successfully expressed inside viral particles as a fusion protein or if any cleavage occurred, viral proteins were detected with a western blot. HIV-$1_{INscarlet}$, HIV-$1_{Nef\text{-}mCherry}$ and mock viral particles were produced in HEK293T cells and purified from cell supernatant. The viruses were then lysed and western blotted against RFP, which detects both mCherry and mScarlet, and the HIV proteins gp120, IN, Nef and p24/55. The same analysis was repeated with VSV-G pseudotyped HIV-$1_{\text{-R-E-INscarlet}}$, HIV-$1_{\text{-R-E-Nef-mCherry}}$ and mock viral particles. Both HIV-$1_{INscarlet}$ and HIV-$1_{\text{-R-E-INscarlet}}$ expressed the INscarlet fusion protein which is detected at 56 kDa by both anti-IN and anti-RFP antibodies (Figure 6.2). In both western blots a very faint band was observed at 32 kDa for IN which suggests in most cases IN is not cleaved from the fusion protein. However, an additional control using a WT NL4.3 virus should be included to compare the expression of the fusion protein to that of the native construct. Interestingly, Nef (and Nef-mCherry in HIV-$1_{\text{-R-E-Nef-mCherry}}$) was detected in all viruses. Whilst Nef is not actively incorporated as a viral particle protein, it is often passively included due to its localisation at the plasma membrane [353]. Detection with anti-Nef was conducted to confirm that the detected protein at 56 kDa in HIV-$1_{Nef\text{-}mCherry}$ and HIV-$1_{\text{-R-E-Nef-mCherry}}$ lanes was Nef-mCherry and not INscarlet.

Whilst INscarlet was detected well on a western blot, the fluorescence intensity of mScarlet may not be enough to visualise the virion as there are approximately only 125 copies of IN per virion [239]. To determine whether HIV-$1_{INscarlet}$ virions can be visualised by mScarlet fluorescence, viral particles were produced in HEK293T cells, then purified and concentrated from cell supernatant. Before imaging, the particles were immobilised on glass coverslips and immunolabelled with anti-gp120. The same preparation was performed on the supernatant of mock infected cells treated in the same way as a negative control, HIV-$1_{Gag\text{-}mCherry}$ virions were used as a positive control. This is a well-established virus within the Castello lab containing 1500-2000 copies of Gag per virion [239] that produces a strong Gag-mCherry signal. Super-resolution microscopy revealed that INscarlet signal was detected in HIV-$1_{INscarlet}$ and that viral particles were fluorescent (Figure 6.3). Detection of individual particles with anti-gp120 suggests that not every virion contained INscarlet (at a detectable level) as only 8.6%

of the gp120 signal overlapped with the INscarlet signal. However, the signal from INscarlet was sufficient to visualise individual viral particles. Comparatively, only 17% of gp120 signal overlapped with mCherry signal in HIV-$1_{Gag\text{-}mCherry}$ particles which was lower than expected. These results suggest that detection of gp120 and IN is inefficient leading to partial overlap. Further experiments should be performed to improve the sensitivity of the microscopy analysis.

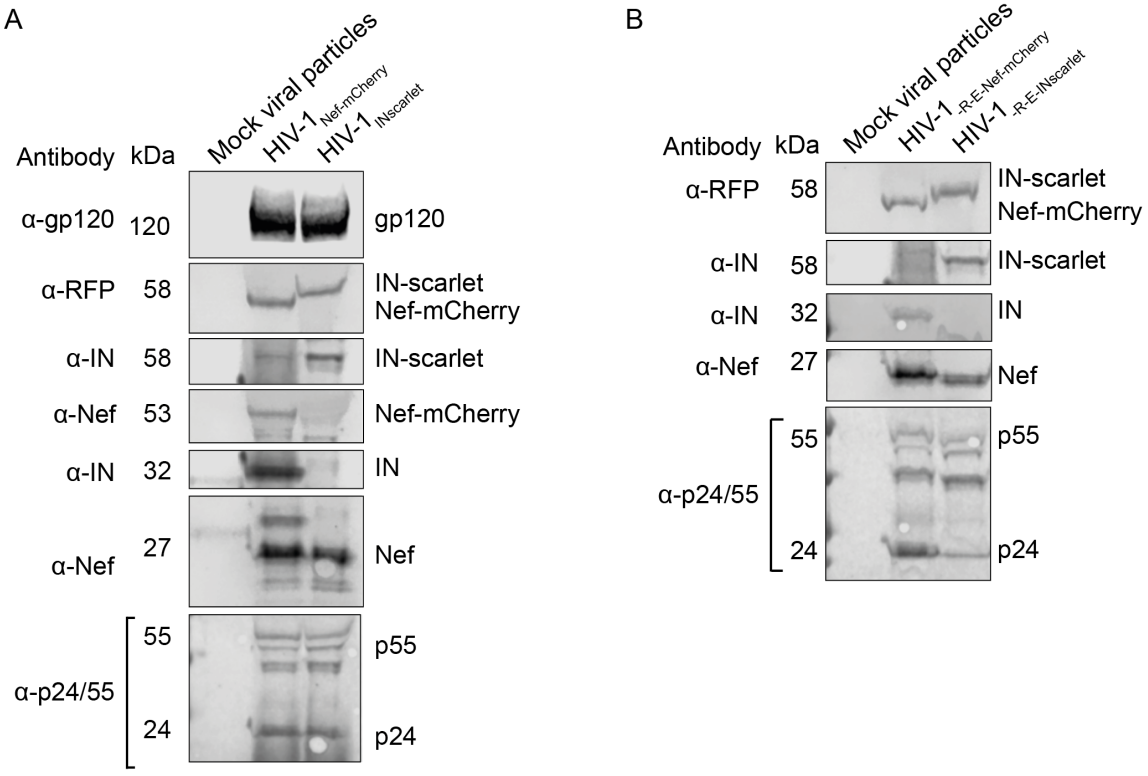

Figure 6.2: **INscarlet is expressed in HIV-1-INscarlet viral particles and single-round viruses.** **A.** HIV-$1_{INscarlet}$ viral particles were lysed and western blotted against gp120, IN, Nef and p24/55. Expression of INscarlet is detected at 56 kDa. HIV-$1_{Nef\text{-}mCherry}$ is used as a comparison for viral particle protein expression and mock viral particles collected from supernatants of uninfected cells were included as a control. **B.** Pseudotyped HIV-$1_{\text{-}R\text{-}E\text{-}INscarlet}$ viral particles were lysed and western blotted against IN, Nef and p24/55. Expression of INscarlet is detected at 56 kDa. HIV-$1_{\text{-}R\text{-}E\text{-}Nef\text{-}mCherry}$ is used as a comparison for viral particle protein expression and mock viral particles collected from supernatants of uninfected cells were included as a control.

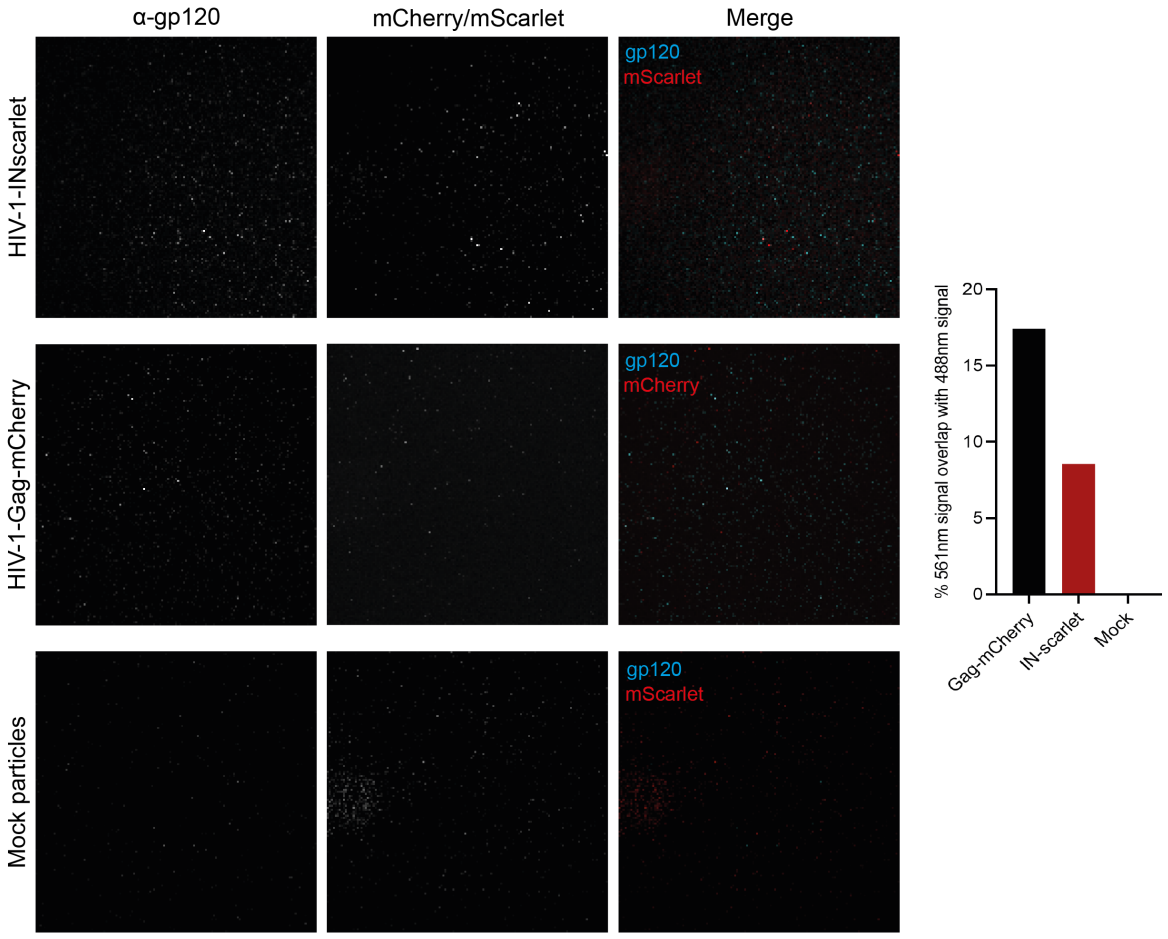

Figure 6.3: **Individual HIV-$1_{INscarlet}$ viral particles are fluorescent.** Viruses were produced in HEK293T cells, collected and fixed on glass coverslips. HIV-1 gp120 was detected using α-gp120 at 488 nm and INscarlet fluorescence was observed at 560 nm. HIV-$1_{Gag\text{-}mCherry}$ was used as a positive control. Images were captured on an Olympus SoRa spinning disc microscope at 60x magnification. Each imaged channel is shown individually (488 nm for gp120, 560 nm for mCherry/mScarlet) and then merged (gp120=blue, mScarlet=red). Percentage signal overlap was calculated by generating area masks for 488 nm and 561 nm and measuring the common areas.

After confirming that INscarlet was successfully expressed as a fusion protein and that mScarlet fluorescence could be observed, I next determined that the function of HIV-1 IN was not disrupted by the addition of mScarlet. To establish this, HIV-$1_{INscarlet}$ infection was measured in CD4+ T lymphocytes (SUpT1) cells in comparison to HIV-$1_{Gag\text{-}mCherry}$ and HIV-$1_{Nef\text{-}mCherry}$, which exhibit an infection fitness that is identical to WT HIV-1. First, SupT1 cells were spinoculated with HIV-$1_{INscarlet}$ and HIV-$1_{Gag\text{-}mCherry}$ at MOI 1. The culture media was replaced with fresh media after 4 h and the cells

were lysed at 48 hpi. Mock and infected cell lysates were western blotted against IN, p24/55 and RFP. A weak signal for INscarlet was detected in the cell lysate with a stronger band for IN, suggesting that more cleavage occurred in the second round of virus production or viruses with cleaved INscarlet were more successful in infection (Figure 6.4A). INscarlet was not detected in HIV-$1_{Gag\text{-}mCherry}$ or mock-infected lysates. Importantly, detection of p24 and p55 in HIV-$1_{INscarlet}$ infected lysates was comparable to HIV-$1_{Gag\text{-}mCherry}$, suggesting IN function was not disrupted and that the viral DNA was successfully integrated to enable viral gene expression. The detection of p24/55 in HIV-$1_{Gag\text{-}mCherry}$ lysates was shifted upwards due to the insertion of the mCherry protein within the MA domain.

Second, IN function was assessed in the infection of SupT1 cells by flow cytometry. SupT1s were spinoculated with HIV-$1_{INscarlet}$ and HIV-$1_{Nef\text{-}mCherry}$ at MOI 1. Additionally, cells were infected with a WT-complemented virus (HIV-$1_{INscarlet/WT}$) that was produced by co-transfecting HEK293Ts with plasmids for HIV-$1_{INscarlet}$ and HIV-1, resulting in viral particles that contain both WT IN and INscarlet. This was done to determine if the efficiency of infection could be increased whilst still expressing INscarlet. At 48 hpi the supernatant was collected for separate analyses and the cells were fixed and intracellularly immunolabelled with anti-p24. Anti-p24 and mScarlet fluorescence were measured in each cell via flow cytometry. The raw flow cytometry data for Figure 6.4 B and C is shown in Appendix 3 Figure 10.1. A similar percentage of cells expressed p24 in both HIV-$1_{INscarlet}$ and HIV-$1_{Nef\text{-}mCherry}$ infection, suggesting HIV-$1_{INscarlet}$ could infect cells with the same efficiency as HIV-$1_{Nef\text{-}mCherry}$. HIV-$1_{INscarlet/WT}$ had substantially reduced infection efficiency (Figure 6.4B). In the same experiment, mScarlet fluorescence was also measured. Whilst a similar percentage of cells expressed p24 and Nef-mCherry in HIV-$1_{Nef\text{-}mCherry}$ infected cells, only 12% of cells infected with HIV-$1_{INscarlet}$ expressed INscarlet, compared with 48% expressing p24 (Figure 6.4C). This supported that INscarlet was not incorporated into all virions at a detectable level (Figure 6.3) and was reflected in the signal intensity at 561 nm for HIV-$1_{INscarlet}$ infected cells being lower than for HIV-$1_{Nef\text{-}mCherry}$ infected cells. Whereas the signal intensity at 488 nm for P24 expression was the same in both infections

(Figure 10.1). The signal intensity of mScarlet is reported to be much higher than mCherry [352]. Therefore, the reduced red fluorescence intensity of HIV-$1_{INscarlet}$ infected cells was likely due to a combination of lower incorporation of INscarlet into viral particles and lower expression of INscarlet during infection compared to Nef-mCherry, which was expected. Predictably, even fewer cells expressed INscarlet when infected with HIV-$1_{INscarlet/WT}$. From this, it was concluded that IN function was not inhibited by fusion with INscarlet and complementing HIV-$1_{INscarlet}$ with WT HIV-1 was not necessary as it reduced infection efficiency. Therefore no further analysis was conducted with HIV-$1_{INscarlet/WT}$.

Finally, to further test the functionality of INscarlet and confirm whether INscarlet expression was maintained over multiple rounds of infection, HIV-$1_{INscarlet}$ and HIV-$1_{Nef\text{-}mCherry}$ particles were harvested from the supernatants of infected CD4+ T lymphocytic cells described previously. These particles were then used to infect fresh SupT1 cells at MOIs of 0.1, 1 and 10. At 48 hpi the cells were analysed by flow cytometry to measure the percentage of cells expressing mScarlet. At an MOI of 0.1 or 1 there was no difference in the percentage of cells expressing mScarlet between HIV-$1_{INscarlet}$ and HIV-$1_{Nef\text{-}mCherry}$ infected cells (Figure 6.4D, raw data shown in Appendix 3 Figure 10.2). However, at MOI 10, substantially more cells expressed Nef-mCherry. In agreement with this there was no difference between the fluorescence intensity of HIV-$1_{INscarlet}$ and HIV-$1_{Nef\text{-}mCherry}$ cells infected with MOI of 0.1 or 1. Although, when infected with an MOI of 10 HIV-$1_{Nef\text{-}mCherry}$ infected cells had a slightly higher intensity of red fluorescence. The fluorescence signal intensity increased as MOI increased for both viruses, reflecting greater infection (Figure 10.2). Taken together, these results confirm that the fusion of an mScarlet tag onto the N-terminus of HIV-1 IN did not ablate the function of IN and that these viruses were capable of multiple rounds of infection whilst maintaining INscarlet fluorescence.

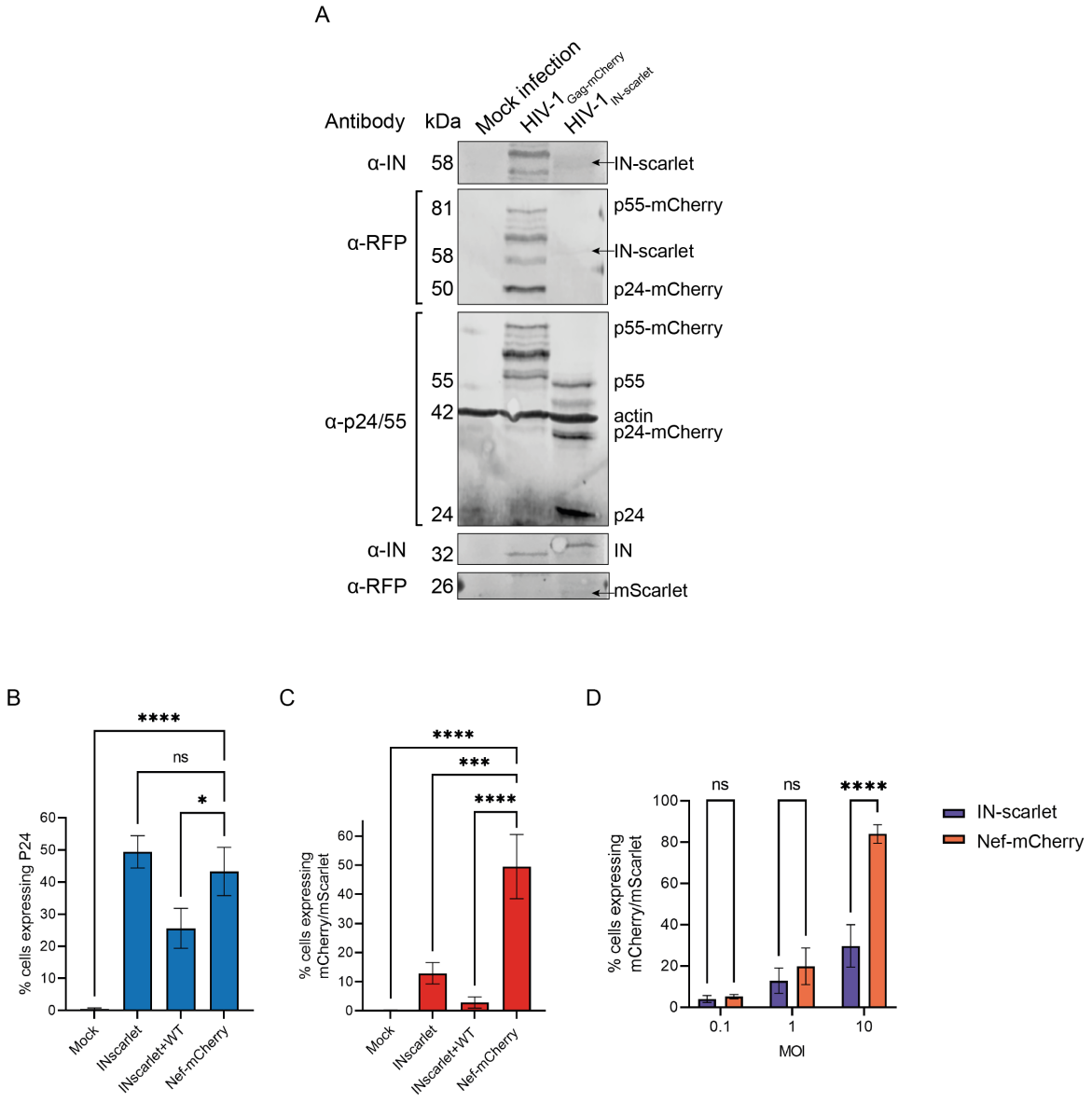

Figure 6.4: **Addition of mScarlet does not affect the function of IN. A.** SupT1 cells were infected with HIV-$1_{INscarlet}$ and HIV-$1_{Gag\text{-}mCherry}$ and lysed at 48 hpi. Mock-infected cells were lysed in parallel. Lysates were western blotted against IN, RFP (which detects mCherry and mScarlet) and p24/55. **B-C.** SupT1 cells were infected with HIV-$1_{INscarlet}$, HIV-$1_{Nef\text{-}mCherry}$ and HIV-$1_{INscarlet/WT}$ (a WT complemented virus produced by co-transfecting cells with WT HIV-1 and HIV-$1_{INscarlet}$). At 48 hpi cells were fixed and intracellularly stained with anti-p24. The percentage of cells expressing HIV-1 p24 **(B)** and mScarlet/mCherry **(C)** was determined by flow cytometry. A one-way ANOVA was conducted to determine the statistical significance of the difference between means of each group compared to HIV-$1_{Nef\text{-}mCherry}$ (ns=not significant, *= $p<0.05$, **= $p<0.01$, ***= $p<0.001$, ****= $p<0.0001$). Caption continued on next page.

Figure 6.4: **Addition of mScarlet does not affect the function of IN. D.** HIV-$1_{INscarlet}$ and HIV-$1_{Nef\text{-}mCherry}$ viral particles were collected from the supernatant of infected SupT1 cells, purified and used to infect fresh SupT1 cells at MOI of 0.1, 1 and 10. At 48 hpi the cells were fixed and the percentage of cells expressing mCherry/mScarlet was determined by flow cytometry. Two-way ANOVA was used to determine the statistical significance of the difference between means of each population (ns=not significant, ****= $p<0.0001$)

After confirming that HIV-$1_{INscarlet}$ was correctly expressed and maintained function of IN, I then determined whether INscarlet-labelled capsids could be visualised in the cytoplasm of infected cells. Whilst physiologically relevant for HIV-1 infection, the CD4+ T lymphocytic cell line SupT1 is unsuited for microscopic analysis due to them being suspension cells and having a distinct morphology with large nuclei and little cytoplasm. Instead, I visualised INscarlet-labelled capsids in HeLa cells that express CXCR4 and were engineered to express CD4-GFP under the control of a doxycyline-induced promoter. Using these cells had the added benefit of visualising the plasma membrane with CD4-GFP expression. HeLa-CD4-GFP were seeded on glass coverslips and induced 24 h prior to infection to allow plasma membrane expression of CD4-GFP. The cells were then infected with HIV-$1_{INscarlet}$, HIV-$1_{Nef\text{-}mCherry}$ and mock viral particles at an MOI of 10 and fixed at 6 hpi. This time point was chosen to enable fusion and entry of viral particles and trafficking to the nucleus while integration will not have occurred in most cases [354]. In this pilot experiment, a high MOI was used to account for the fact that not all viral particles expressed detectable levels of INscarlet. After immunolabelling with anti-p24 and staining with DAPI, infection was visualised using a super-resolution microscope. Individual capsids were seen as dots throughout the cytoplasm of infected cells detected by both mScarlet and anti-p24 fluorescence (Figure 6.5). The merge showed mScarlet and anti-p24 signal frequently overlapped which supports that each spot was a viral capsid. Magnification of HIV-$1_{INscarlet}$ showed this in more detail and revealed IN-scarlet inside the nucleus, at which point the p24 signal was lost (Figure 6.6). This further supported that the spots were capsids and revealed that INscarlet signal was also observed in the nucleus. Taken together, these results suggest that the labelling of HIV-1 capsids with INscarlet is an effective way to visualise the early stages of HIV-1 infection.

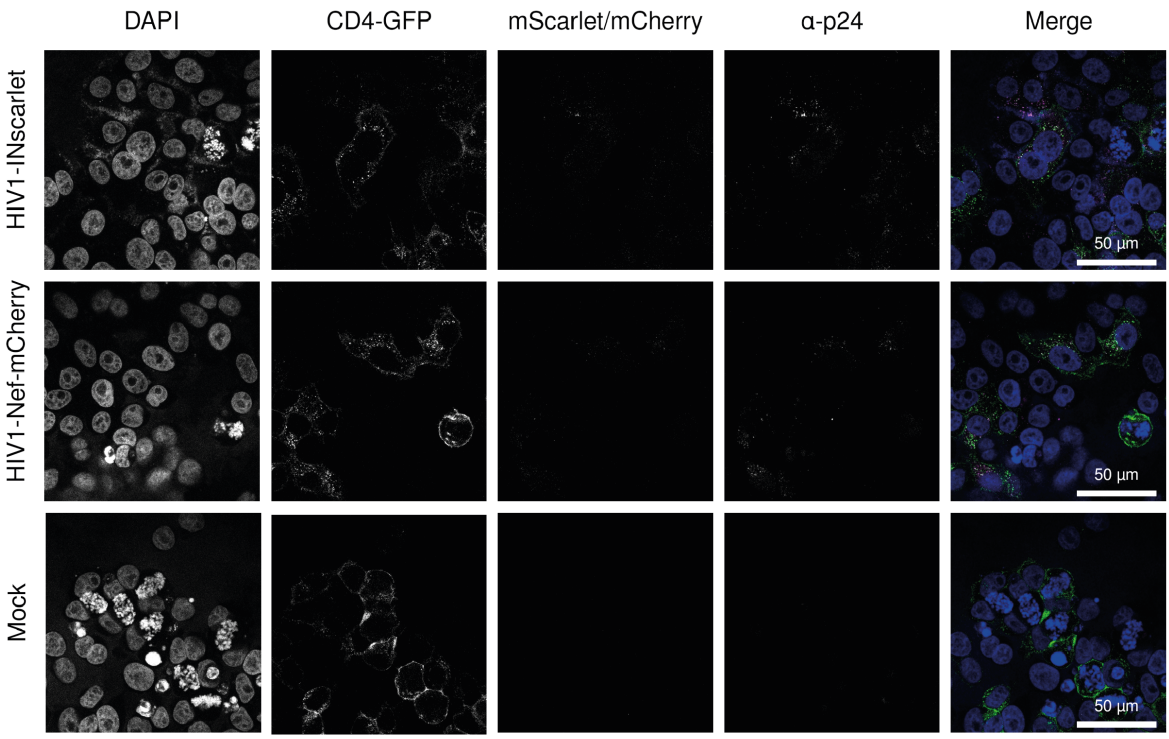

Figure 6.5: **INscarlet labelled capsids can be seen in the cytoplasm of infected cells.** HeLa-CD4-GFP cells were seeded on glass coverslips and CD4-GFP expression was induced 24 h prior to infection. Cells were then infected at MOI 10 with HIV-$1_{\text{INscarlet}}$, HIV-$1_{\text{Nef-mCherry}}$ and mock viral particles for 6 h before fixing and staining with DAPI. Cells were imaged on a Zeiss 880 Airyscan super-resolution confocal microscope at 100x magnification. Individual panels for CD4-GFP (488 nm), DAPI (408 nm), mCherry/Scarlet (560nm) and p24 (647 nm) are shown followed by the merge in which CD4 is green, DAPI is blue, mCherry/mScarlet is white and p24 is magenta.

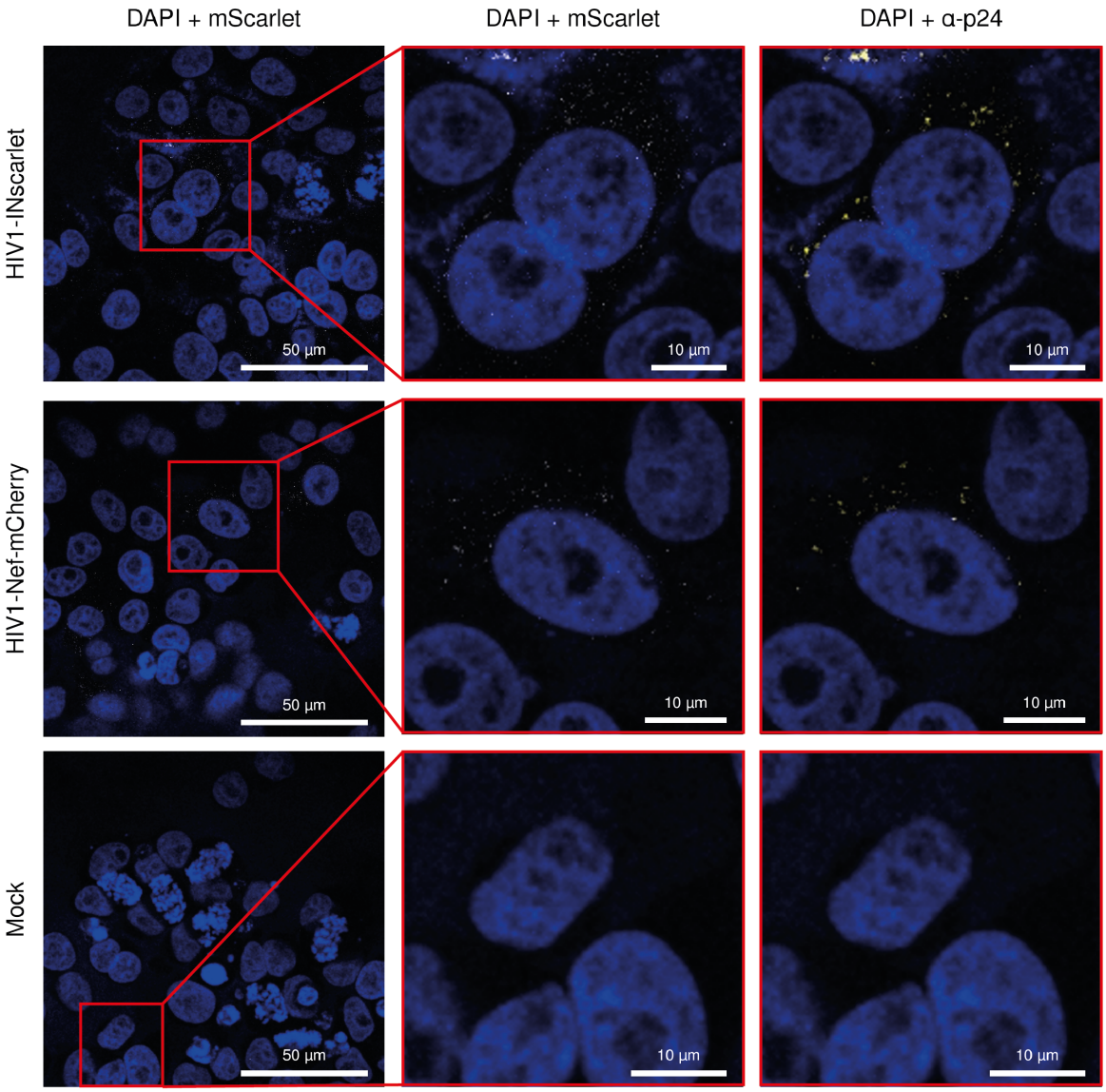

Figure 6.6: **INscarlet labelled capsids can be seen in the cytoplasm of infected cells at 6 hpi.** As in Figure 6.5 but only the DAPI (blue), mScarlet (white) and α-p24 (yellow) channels are shown. The right-hand panels contain magnified views of the left-hand images highlighting individual capsids labelled by mScarlet fluorescence in the cytoplasm of HeLa-CD4-GFP cells at 6 hpi.

6.3 Discussion

Many questions regarding the early stages of the HIV-1 life cycle remain unanswered or have only been partially resolved. Understanding the early stages of HIV-1 infection and revealing the cellular interactors during these stages is pertinent to the development of new anti-retroviral therapies. As demonstrated by numerous

discoveries achieved with fluorescently-labelled HIV-1 proteins, imaging techniques offer powerful tools to study the viral life cycle [75, 76, 355, 356]. Most of the systems used so far exploit the direct or indirect labelling of HIV-1 CA. Whilst these tools have extensively increased our understanding of capsid structure and uncoating, they can lead to disruption of CA interactions and are limited to visualising infection up to the point of capsid disassembly. By labelling HIV-1 IN, a protein within the capsid that is a key member of the PIC, we can visualise HIV-1 infection up to integration of viral DNA into the host genome. This enables the study of host proteins that are regulating key stages of the life cycle such as reverse transcription and integration.

In this chapter I have designed and validated two novel fluorescently labelled viruses to enable the visualisation of HIV-1 infection from cell entry to integration, HIV-$1_{\text{-R-E-INscarlet}}$ and HIV-$1_{\text{INscarlet}}$. The former can be produced by pseudotyping with VSV-G or another envelope glycoprotein and enables easier and safer visualisation of HIV-1 infection under BSL2 conditions. Crucially, the latter enables the study of early stages of infection under physiological conditions and permits multiple rounds of infection. This important for an accurate representation of HIV-1 infection and also for permitting the study of the incorporation of cellular proteins into assembling viral particles and following them in the infection of a new cell.

For HIV-$1_{\text{-R-E-INscarlet}}$ I validated expression of INscarlet within pseudotyped viral particles by western blotting. Whilst I did not directly confirm IN function was maintained, it can be assumed that it would function in the same way as HIV-$1_{\text{INscarlet}}$ because the construct design is identical. The formal demonstration of this, however, will be approached after defense of this thesis. This chapter focused on HIV-$1_{\text{INscarlet}}$ due to it being a closer to physiological infection. The other IN-labelled viruses reported in the literature are all expressed in replication or packaging deficient backbones and require pseudotyping with VSV-G to generate single-round viruses [75, 342, 344, 345]. Pseudotyped lentiviruses have benefits in terms of flexibility and safety of use. However, the mode of entry between VSV-G enveloped viruses and Env enveloped viruses is very different. Instead of CD4-mediated membrane fusion, VSV-G mediates cell entry through recognition of the

LDL receptor followed by internalisation into endosomes [241]. Whilst HIV-1 entry via endocytosis has been documented, the capsid cores are degraded and do not lead to productive infection [242]. Additionally, because VSV-G-mediated entry occurs through endosomes, processes mediated by the capsid in the cytosol, such as engagement with the cytoskeleton, may not take place. Therefore, the kinetics and efficiency of the trafficking of these capsids to the nucleus differs from that of plasma membrane-deposited capsids. Therefore, to study the early stages of HIV-1 infection under physiological conditions, Env-enveloped viruses need to be used. This is why the emphasis in this chapter was on the production and validation of HIV-$1_{\text{INscarlet}}$ and not the pseudotyped HIV-$1_{\text{-R-E-INscarlet}}$. However, before use of HIV-$1_{\text{-R-E-INscarlet}}$, function of IN and infectivity should be validated.

HIV-$1_{\text{INscarlet}}$ was shown to correctly express INscarlet and the function of IN was conserved. This was shown by the viral protein expression in infected cells. Maintaining protein function after the addition of a fluorophore has historically proven difficult due to the large size of the fluorophores in comparison to viral proteins [357]. However, here I achieved 48% infection of CD4+ T lymphocytes which is an improvement on older HIV-IN-fluorophore viruses [62, 344] and comparable to newer ones [342, 345]. Additionally, viruses could be collected after infection and used in a subsequent infection, something which is not possible with any of the other HIV-IN-fluorophore viruses. This enables visualisation of multiple rounds of infection if desired.

Whilst infectivity and IN function was a success, I could not prove labelling of 100% of viral particles. INscarlet was only detected in 12% of infected CD4+ T lymphocytes, compared to p24 being detected in 48% of infected cells. Whilst this can be down to differences between the two infections, these numbers suggest that at least a quarter of viral particles are labelled with enough INscarlet to be detected. Often fluorescent proteins are introduced to viral particles through co-transfection which results in competition between WT and labelled proteins during viral assembly and leads to a significant amount of WT particles. To avoid this INscarlet was introduced here as part of the HIV-1 backbone and therefore should be incorporated into every viral particle. This approach was also taken by Mamede et al, through which they

achieved 100% efficiency of particle labelling [342]. The reduced efficiency observed here suggests either INscarlet is not being incorporated into assembling virions or a subpopulation of labelled but defective virions is produced. Unlabelled or viral particles with low signal generate noise in microscopy experiments as they cannot be observed by imaging and will skew the data to suggest less infection has occurred than the infectious titre would suggest. This requires further work to optimise the incorporation of INscarlet to achieve full labelling of viral particles or to improve sensitivity of the analysis. One possibility would be the use of fluorescently labelled endogenous nanobodies (available for mScarlet), the use of HaLo tag and associated reagents, and more sensitive microscopy instruments.

Despite only a fraction of particles being labelled, INscarlet labelled capsids can be seen in the cytoplasm and nuclei of infected cells. This suggests labelling with INscarlet is a valid method to visualise the early stages of HIV-1 infection. The results presented here scratch the surface of possible experiments that can be conducted with HIV-$1_{\text{INscarlet}}$. To truly demonstrate its potential, I will utilise HIV-$1_{\text{INscarlet}}$ in future live-cell imaging experiments where I aim to use single molecule tracking to follow the journey of INscarlet labelled capsids from cell entry to integration. Most importantly, HIV-$1_{\text{INscarlet}}$ will be used to explore the role of cellular RBPs within the HIV-1 capsid, particularly MKRN1, which is discussed in more detail in Chapter 7.

7 General discussion and future directions

Many cellular RBPs have well-described roles in HIV-1 infection, both promoting and inhibiting viral replication. To provide just a few examples: UPF1 [99], YBX1 [100], DHX9 [101] and MOV10 [106] regulate reverse transcription; PPIA stabilises CA hexamers [358]; Crm1 facilitates Rev-mediated nuclear export of HIV-1 RNAs [359]; STAU1, HSP60, HSP70 and HSC70 chaperone Gag-vRNA complexes to the plasma membrane [129–132], and APOBEC3G binds to HIV-1 RNA to package into virions and restrict infection through deamination of viral cDNA [360]. Whilst these interactions are well described within the cell, relatively little is known about cellular RBPs within the viral particle. Except for APOBECs, it is not known how or why certain RBPs are packaged and whether they have specific roles in the viral particle or are required to be in the capsid for the next round of infection. Garcia-Moreno et al have set out to identify the RBPome of HIV-1 particles [135]. In this work, I have validated that these RBPs positively and negatively impact HIV-1 infection. The results presented in Chapter 4 reveal CIRBP-KO increases viral replication, implying CIRBP negatively regulates HIV-1 infection as an anti-viral protein. In HIV-1-infected MKRN1-KO cells, a substantial increase in viral replication was observed. However, viral particle release was remarkably reduced. Due to the function of MKRN1 as an E3 ubiquitin ligase and the heavy reliance of HIV-1 budding on ubiquitination, I hypothesised that MKRN1 positively regulated HIV-1 particle release. This regulation was thought to be achieved through the ubiquitination of Gag or a member of the cellular ESCRT pathway. Whilst several of the candidates tested in Chapter 4 appeared to have interesting functions within HIV-1 infection, the most striking result was observed for MKRN1. Therefore, MKRN1 was chosen for further investigation to elucidate the mechanism by which it regulates HIV-1 infection.

To test my hypothesis that MKRN1 regulated HIV-1 particle release I used smFISH combined with immunofluorescence to visualise the inhibition of viral particle release. Imaging showed that MKRN1-KO promoted a striking accumulation of Gag at the plasma membrane of HIV-1 infected cells. I then investigated the interactors of MKRN1 during HIV-1 infection to determine whether the regulation of viral particle release was

achieved through interaction with Gag or a member of the ESCRT pathway. This was achieved by co-immunoprecipitating GFP-MKRN1 from HIV-1 infected cells and identifying interacting proteins via quantitative mass spectrometry. This highlighted a pool of proteins interacting with MKRN1 during infection which were enriched in ribosomal and spliceosomal processes. Unexpectedly, no HIV-1 or ESCRT proteins were identified as interactors of MKRN1 suggesting MKRN1 regulates HIV-1 budding and release via an indirect pathway, not through the ubiquitination of HIV-1 Gag or ESCRT proteins. However, these results were not conclusive as only one repeat was conducted. The co-immunoprecipitation was performed at 24 hpi which was possibly too early to capture interactions involved with HIV-1 Gag and viral particle release. Further optimisation of this experiment is required before any conclusions can be made regarding the interactors of MKRN1 during HIV-1 infection. These repeats are the priority for further work and will either confirm the results we have already seen or alternatively, may highlight new interactions that were not identified in the first group of samples.

Alongside repeating co-immunoprecipitation and mass spectrometry analysis of GFP-MKRN1 interactors, there are a set of experiments which will enable better understanding of the role of MKRN1 in HIV-1 infection. First, I will investigate whether RNA plays an important role in the function of MKRN1. MKRN1 is a cytosolic RBP with low expression, yet we identify MKRN1 specifically incorporated into HIV-1 particles interacting with vRNA [135]. It is possible that the RNA-binding activity of MKRN1 enables its recruitment to sites of viral assembly through interaction with HIV-1 gRNA. To investigate the dependence of MKRN1 interactions on RNA I will conduct a separate co-immunoprecipitation on cytoplasmic lysate collected from HIV-1 infected GFP-MKRN1 and WT HEK293 Flp/In T-rex cells. However, the cells will be lysed in the presence of benzonase, a nuclease which degrades all forms of DNA and RNA [361]. This will degrade all RNAs, including HIV-1 RNAs, present in the lysates. GFP-MKRN1 will then be co-immunoprecipitated from the lysate and mass spectrometry will be conducted to identify the proteins that interact with MKRN1 in the absence of RNA. By comparing these proteins to the previous interactome of MKRN1

in the presence of RNA, it can be determined which interactions are dependent on the presence of RNA.

If RNA did prove to be required for the interaction of MKRN1 with certain proteins, it would be interesting to determine which RNAs MKRN1 interacts with during infection. This would support whether or not MKRN1 is recruited to sites of viral assembly through interaction with HIV-1 gRNA. Additionally, identifying the sequences within the RNA that MKRN1 recognises may reveal more about its function. Crosslinking immunoprecipitation coupled with next-generation sequencing (CLIP) refers to a group of techniques that enable the global identification of RNA targets of RBPs in physiological settings. The basic methodology involves 1) covalent crosslinking of protein-RNA complexes in live cells or viral particles; 2) cells lysis then treatment with RNases leaving small fragments of RNA (approximately 20-50 nucleotides) bound to the protein of interest; 3) immunoprecipitation of protein-RNA complexes followed by removal of non-specific RNAs and proteins. 4) Radioactive labelling of the purified protein-RNA complexes then separation by SDS-PAGE. 5) Isolation of bound RNA from the SDS-PAGE gel. 6) Ligation of adapters to the RNA and reverse transcription into cDNA which is amplified by PCR. 7) Sequencing of PCR products and mapping to reference genomes [362]. Application of CLIP to study Gag-RNA interactions has revealed the highly selective genome packaging process of HIV-1 [120] and the targeting of HIV-1 assembly to the plasma membrane through MA interactions with tRNAs [363]. Additionally, two CLIP variations were used to uncover how APOBEC3G is specifically incorporated into HIV-1 virions in the presence of a vast excess of cellular RNA [364, 365]. Therefore, a CLIP method would be an effective way to determine if and where MKRN1 is interacting with HIV-1 RNA and whether binding vRNA is how MKRN1 is incorporated into viral particles. Individual nucleotide (i)CLIP previously conducted on MKRN1 revealed MKRN1 predominantly binds A-rich stretches in the 3' UTR of protein-coding genes upstream of poly(A) sites [160]. HIV-1 RNAs are polyadenylated and therefore it is likely MKRN1 also recognises these sites on the vRNA.

In addition to RNA-binding activity, I also wish to further investigate the ubiquitin

ligase activity of MKRN1. Within this thesis, I generated three HEK293 Flp/In T-rex cell lines expressing catalytic mutants of MKRN1: GFP-MKRN1-C281A, GFP-MKRN1-M285E and GFP-MKRN1-H307E. These mutants were designed in collaboration with Dr Paul Elliot (Department of Biochemistry, University of Oxford), building on his experience on E3-ubiquitin ligases. While mutants have been purposely designed based on the current knowledge in MKR1, their catalytic activity should be confirmed via ubiquitination assays with purified GFP-MKRN1 (WT and mutants) and His-tagged ubiquitin, mixed with cellular extracts by assessing known targets. Activity can be assessed by looking at the ubiquitin state of known MKRN1 targets such as PABPC1, ubiquitin, ATP and E1 and E2 ubiquitin ligases, using anti-His antibodies in western blotting. The in vivo E1 and E2 partners of MKRN1 have not been identified but UBE1 and both UBE2D3 and UBCH5C have been used successfully for in vitro assays [163,312]. After confirming the inactivity of the mutants, HIV-1 infection can be assessed in a MKRN1-mutant cell line to determine whether loss of MKRN1 catalytic activity has the same effect as MKRN1 KO. This can be achieved by conducting the same analysis of viral protein expression, viral RNA expression, and viral particle release as performed in Chapter 4. This investigation will reveal whether the regulation of HIV-1 budding by MKRN1 is driven by its function as an E3 ubiquitin ligase. Whilst this may not be achieved as directly as ubiquitinating HIV-1 Gag or a member of the cellular ESCRT pathway, it is possible that MKRN1 ubiquitinates another protein that indirectly affects viral particle release. Hildebrandt et al determined the network of proteins ubiquitinated by MKRN1 and found that the majority of MKRN1-ubiquitination did not lead to proteasomal degradation of targets [160]. Additionally, MKRN1 has been shown to promote lysine-63 (K63)-linked ubiquitination of target proteins [366], the type of linkage known to regulate proteasome-independent pathways such as signal transduction [367,368]. Therefore, MKRN1-mediated ubiquitination may drive a signalling pathway necessary for viral particle assembly and release.

It would also be pertinent to uncover which proteins MKRN1 specifically ubiquitinates during HIV-1 infection. Whilst the co-immunoprecipitation and mass spectrometry results will reveal interactors of MKRN1, these proteins are not all

necessarily ubiquitination targets of MKRN1. This was demonstrated by Hildebrandt et al who identified 53 proteins interacting with MKRN1 but only 21 were ubiquitination targets [160]. Identifying the ubiquitination targets of MKRN1 during HIV-1 infection would be the key to proving and understanding the role of MKRN1 in HIV-1 budding. For this, a variety of methods could be used. A peptide-centric ubiquitinomics approach such as diglycine remnant profiling could be used, which was the method performed by Hildebrandt et al. In this method trypsin digestion of ubiquitinated proteins yields peptides containing characteristic diglycine-modified lysine residues which can be identified by mass spectrometry. However, due to the low abundance of ubiquitinated peptides compared to non-modified peptides in cell lysates, this approach is difficult without enrichment steps [369]. Alternatively, a protein-centric method such as tandem ubiquitin-binding entities (TUBEs) could be used to isolate and purify ubiquitinated proteins from HIV-1 infected cell lysates [370]. TUBEs consist of ubiquitin-binding domains with nanomolar affinities for ubiquitin and can be designed to recognise specific ubiquitin linkages or be pan-selective. Isolated ubiquitinated proteins can then be identified by mass spectrometry. However, a limitation of this method is that purification is performed under non-denaturing conditions and therefore ubiquitin-binding proteins are co-purified [371]. Either method described here would enable the identification of proteins ubiquitinated by MKRN1 during HIV-1 infection, underpinning the hypothesis that MKRN1 regulates HIV-1 particle release and revealing the pathways by which this regulation is achieved.

The final question to explore is why is MKRN1 incorporated into HIV-1 capsids? The results presented in Chapters 4 and 5 highlight the cellular role of MKRN1 in HIV-1 infection after integration has occurred. Additionally, they show that the absence of MKRN1 in HIV-1 particles does not affect the infectivity of particles in the second round of infection. Therefore, I hypothesised that MKRN1 binds to the vRNA to be recruited to sites of viral assembly and consequently, is incorporated into the viral particles, rather than MKRN1 having a specific role inside the capsid. The idea of recruitment via vRNA can be explored using the iCLIP experiments suggested above. However, to substantiate that MKRN1 has no role inside the capsid, cells should be infected with

WT and MKRN1-KO viral particles and analysed for vRNA expression at early and late time points throughout infection. This would enable the observation of any effect of MKRN1-KO on reverse transcription or integration. In addition, the HIV-$1_{INscarlet}$ virus that I designed and validated in Chapter 6 could be used to visualise MKRN1 inside the viral capsid during infection. One of the experiments considered when designing HIV-$1_{INscarlet}$ was a dual-fluorescence experiment in which HIV-$1_{INscarlet}$ is generated in GFP-MKRN1 expressing cells. GFP-MKRN1 would be incorporated into INscarlet-labelled HIV-1 particles, which would have dual-fluorescence. These dual-fluorescent particles would then be used to infect new cells and the dual-fluorescent capsids could be tracked, using single-molecule tracking, as they trafficked to the cytoplasm. The point at which the green GFP-MKRN1 signal dissociates from the red INscarlet signal would suggest whether MKRN1 regulated any stages within the second round of infection (Figure 7.1). For example, if GFP-MKRN1 signal is lost upon nuclear entry, it suggests MKRN1 may be involved in reverse transcription but is not involved in integration. Alternatively, if the dual-fluorescent signal was maintained in the nucleus and could be observed at sites of integration then that would provide evidence for MKRN1 regulating integration. However, based on the results collected so far, MKRN1 does not regulate the second round of infection.

Together, the experiments discussed here would enable the full elucidation of how MKRN1 regulates HIV-1 infection. This would be an entirely novel contribution to the field of HIV-1 biology as MKRN1 has no prior connection with HIV-1. Additionally, the mechanism by which MKRN1 regulates HIV-1 infection is unlike any previously described interaction of MKRN1 in viral infection [161, 162, 372]. This thesis has not only uncovered a new cellular regulator of HIV-1 infection but has also validated that cellular RBPs associated with vRNA in HIV-1 particles play vital roles in infection. This highlights a whole new proteome to explore in the regulation of early HIV-1 infection.

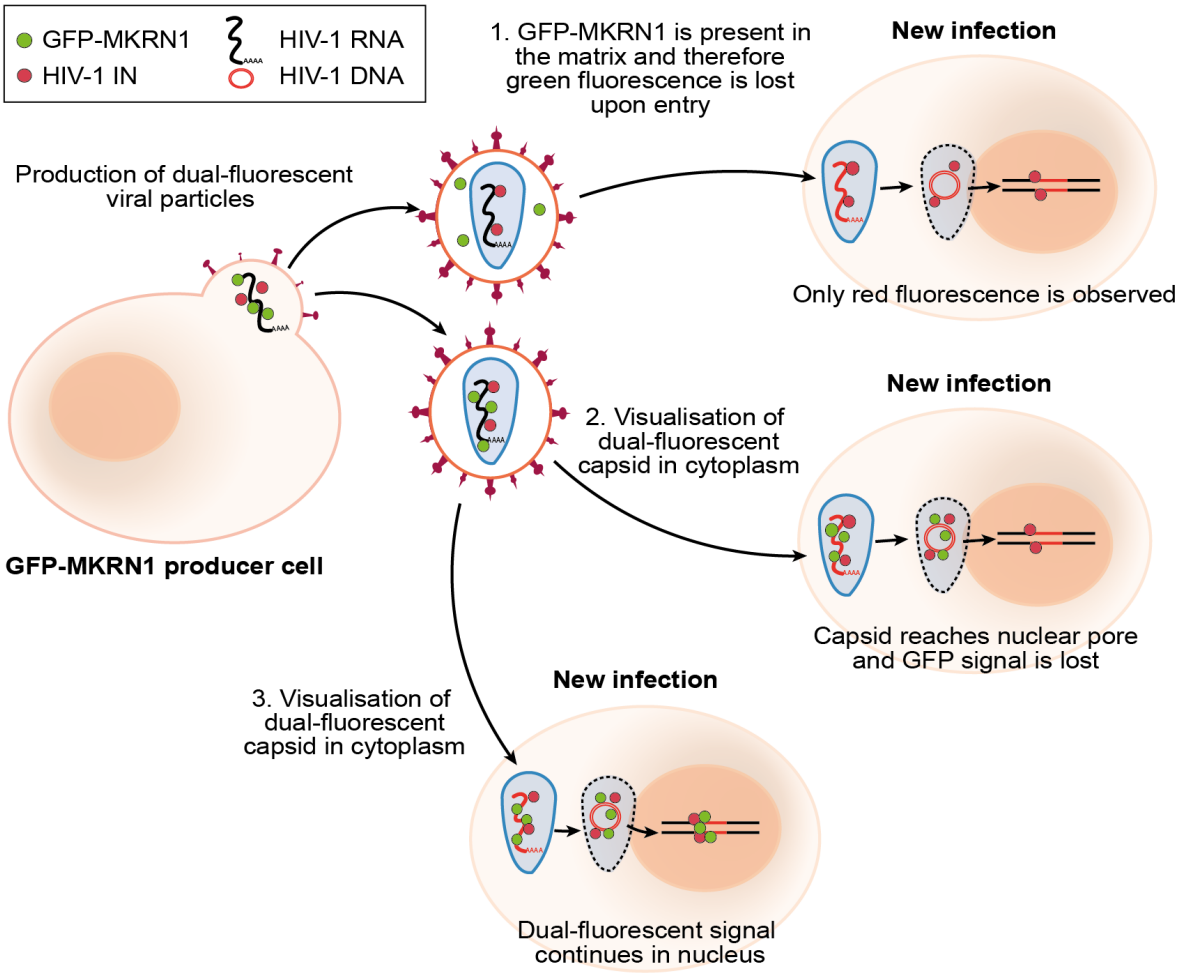

Figure 7.1: **Schematic of GFP-MKRN1 HIV-$1_{INscarlet}$ dual-fluorescent viral particles.** HIV-$1_{INscarlet}$ viruses produced in GFP-MKRN1 expressing cells will be dually labelled with GFP and mScarlet fluorescence as MKRN1 is incorporated into viral particles. Upon new infection there are three possible outcomes that can be observed by single-particle tracking and super-resolution microscopy. 1) GFP-MKRN1 is present in the matrix and green fluorescence is lost upon cell entry. 2) GFP-MKRN1 is present within the capsid and dual-fluorescent signal is observed up to nuclear import or 3) continues into the nucleus to sites of integration.

8 Appendix 1

Figure 8.0 ivRBPs identified by ivRIC enriched in crosslinked and HIV-1 infected eluates Fully labelled versions of Figure 4.2. **A.** Volcano plot showing the log2 fold change and the significance (p value) of each RBP between crosslinked viral particles and non-crosslinked viral particles using data from three biological replicates. False discovery rate (FDR) $<$1% shown in red, $<$10% shown in yellow. **B.** Same as in A but for 72hpi HIV-1 infected and uninfected conditions.

Figure 8.1 ivRBPs identified by ivRIC enriched in ivRIC eluates over inputs and the whole cell lysate. Fully labelled versions of Figure 4.2.**C.** Volcano plot showing the log2 fold change and thesignificance (p value) of each RBP between crosslinked ivRIC eluates and crosslinked ivRIC inputs. FDR $<$1% shown in blue and $<$10% shown in cyan for inputs. **D.** Scatterplot comparing the intensity of each RBP in the ivRIC eluate and whole cell lysate of lymphocytes.

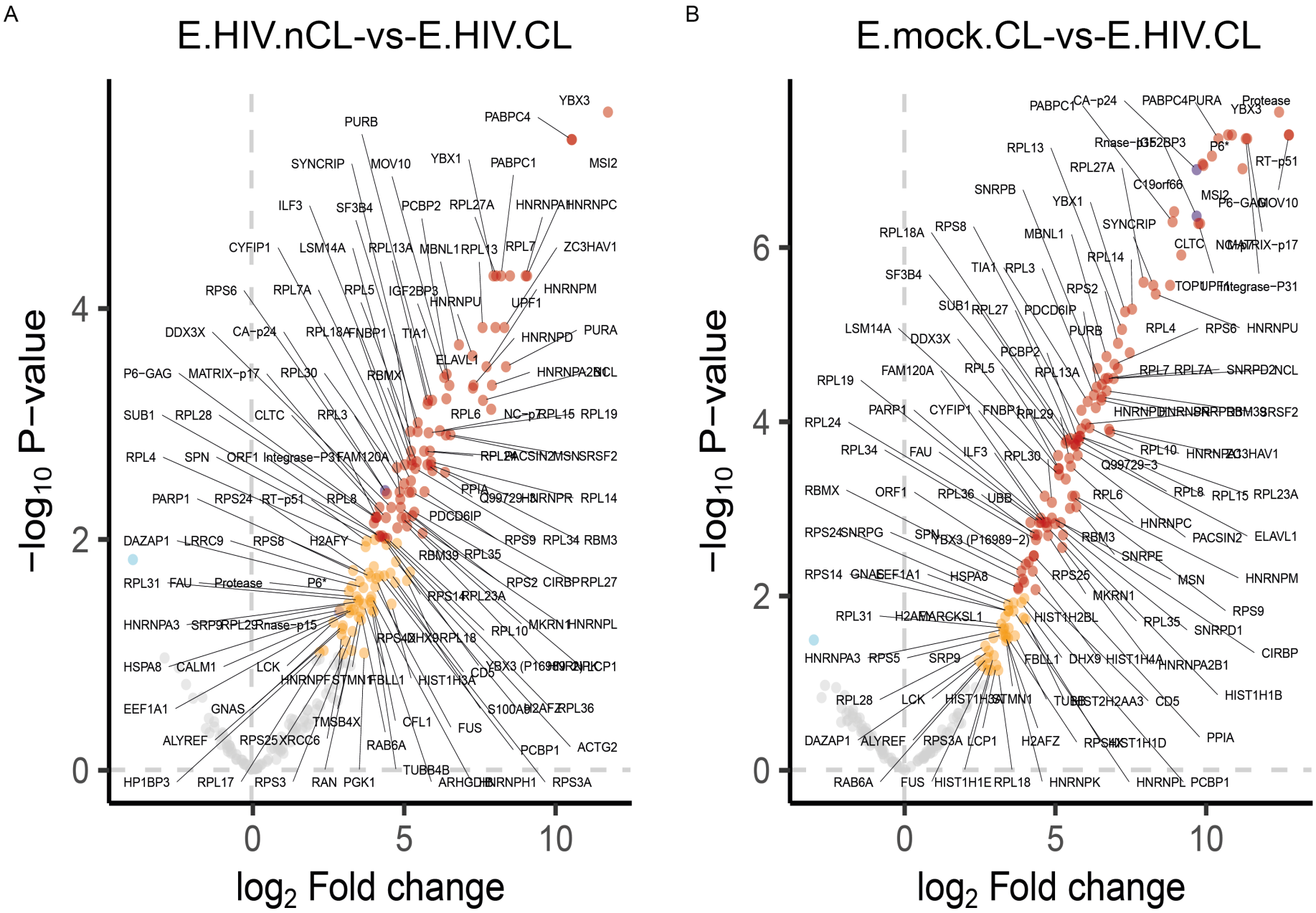

Figure 8.1: **ivRBPs identified by ivRIC enriched in crosslinked and HIV-1 infected eluates**

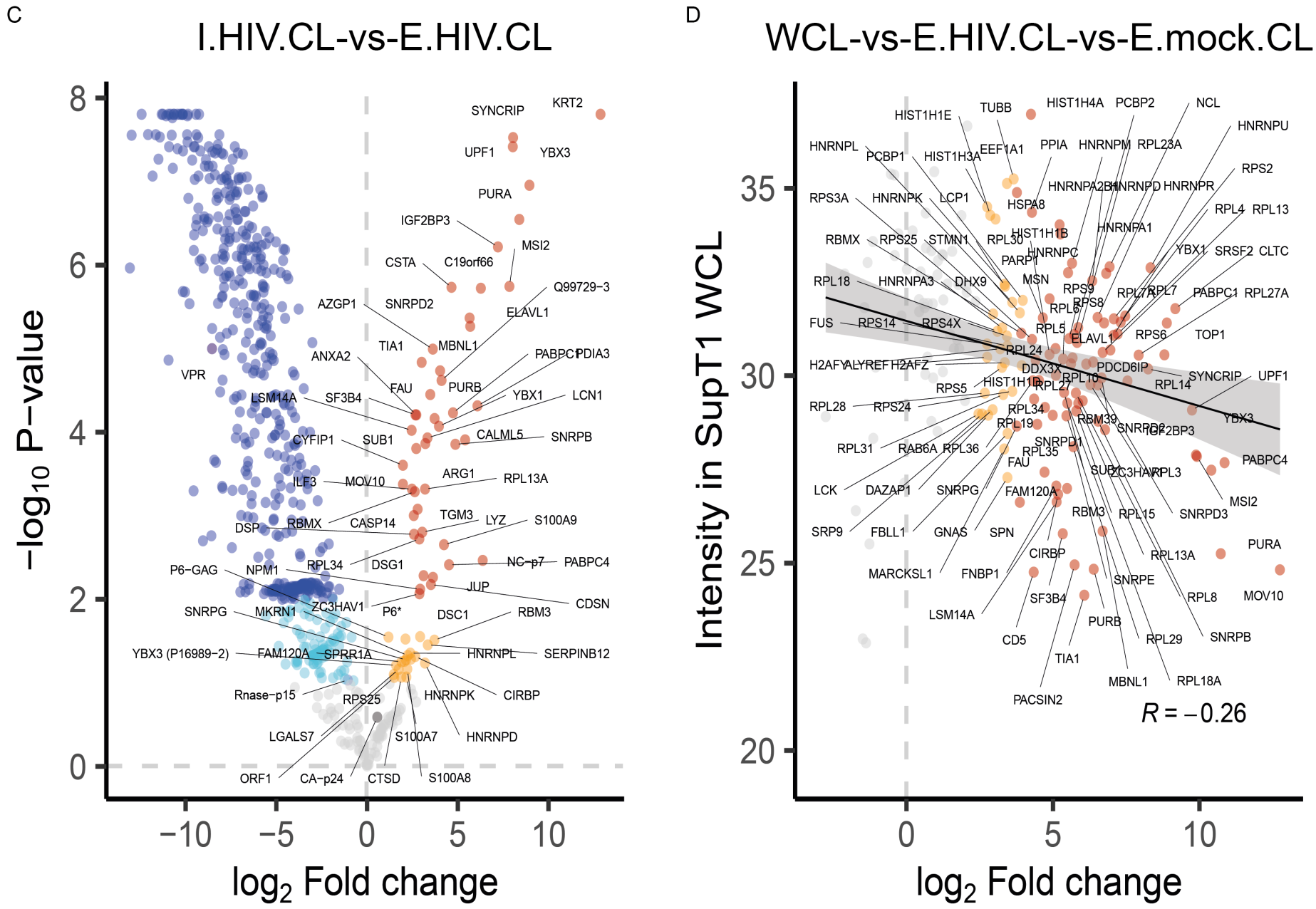

Figure 8.2: **ivRBPs identified by ivRIC enriched in ivRIC eluates over inputs and the whole cell lysate.**

9 Appendix 2

Table 1. Proteins identified in GFP-MKRN1 IP compared to our ivRBPome dataset and the Hildebrandt et al 2019 dataset of proteins affinity purified (AP) with MKRN1.

Accession	Gene Name	Description	Enriched in HIV-1 infection	Present in ivRBPome	Present in Hildebrandt AP
Q12904	AIMP1	Aminoacyl tRNA synthase complex-interacting multifunctional protein 1 OS=Homo sapiens OX=9606 GN=AIMP1 PE=1 SV=2 - [AIMP1_HUMAN]	Y	N	N
Q86V81	ALYREF	THO complex subunit 4 OS=Homo sapiens OX=9606 GN=ALYREF PE=1 SV=3 - [THOC4_HUMAN]	N	N	N
Q9UKV5	AMFR	E3 ubiquitin-protein ligase AMFR OS=Homo sapiens OX=9606 GN=AMFR PE=1 SV=2 - [AMFR_HUMAN]	Y	N	N
Q17R89	ARHGAP44	Rho GTPase-activating protein 44 OS=Homo sapiens OX=9606 GN=ARHGAP44 PE=1 SV=1 - [RHG44_HUMAN]	Y	N	N
Q9NVI7	ATAD3A	ATPase family AAA domain-containing protein 3A OS=Homo sapiens OX=9606 GN=ATAD3A PE=1 SV=2 - [ATD3A_HUMAN]	N	N	N
P36542	ATP5F1C	ATP synthase subunit gamma, mitochondrial OS=Homo sapiens OX=9606 GN=ATP5F1C PE=1 SV=1 - [ATPG_HUMAN]	N	N	N
Q07021	C1QBP	Complement component 1 Q subcomponent-binding protein, mitochondrial OS=Homo sapiens OX=9606 GN=C1QBP PE=1 SV=1 - [C1QBP_HUMAN]	N	N	N
O15234	CASC3	Protein CASC3 OS=Homo sapiens OX=9606 GN=CASC3 PE=1 SV=2 - [CASC3_HUMAN]	N	N	N
Q8N163	CCAR2	Cell cycle and apoptosis regulator protein 2 OS=Homo sapiens OX=9606 GN=CCAR2 PE=1 SV=2 - [CCAR2_HUMAN]	N	N	N
Q494R4	CCDC153	Coiled-coil domain-containing protein 153 OS=Homo sapiens OX=9606 GN=CCDC153 PE=1 SV=2 - [CC153_HUMAN]	Y	N	N
Q16630	CPSF6	Cleavage and polyadenylation specificity factor subunit 6 OS=Homo sapiens OX=9606 GN=CPSF6 PE=1 SV=2 - [CPSF6_HUMAN]	N	N	N
Q5W186	CST9	Cystatin-9 OS=Homo sapiens OX=9606 GN=CST9 PE=2 SV=1 - [CST9_HUMAN]	Y	N	N
P51398	DAP3	28S ribosomal protein S29, mitochondrial OS=Homo sapiens OX=9606 GN=DAP3 PE=1 SV=1 - [RT29_HUMAN]	N	N	N
O00571	DDX3X	ATP-dependent RNA helicase DDX3X OS=Homo sapiens OX=9606 GN=DDX3X PE=1 SV=3 - [DDX3X_HUMAN]	N	Y	N
Q08211	DHX9	ATP-dependent RNA helicase A OS=Homo sapiens OX=9606 GN=DHX9 PE=1 SV=4 - [DHX9_HUMAN]	Y	Y	N
P60981	DSTN	Destrin OS=Homo sapiens OX=9606 GN=DSTN PE=1 SV=3 - [DEST_HUMAN]	Y	N	N
P26641	EEF1G	Elongation factor 1-gamma OS=Homo sapiens OX=9606 GN=EEF1G PE=1 SV=3 - [EF1G_HUMAN]	Y	N	N
P38919	EIF4A3	Eukaryotic initiation factor 4A-III OS=Homo sapiens OX=9606 GN=EIF4A3 PE=1 SV=4 - [IF4A3_HUMAN]	N	N	N
Q15717	ELAVL1	ELAV-like protein 1 OS=Homo sapiens OX=9606 GN=ELAVL1 PE=1 SV=2 - [ELAV1_HUMAN]	N	Y	Y
P62861	FAU	40S ribosomal protein S30 OS=Homo sapiens OX=9606 GN=FAU PE=1 SV=1 - [RS30_HUMAN]	Y	Y	N
Q13283	G3BP1	Ras GTPase-activating protein-binding protein 1 OS=Homo sapiens OX=9606 GN=G3BP1 PE=1 SV=1 - [G3BP1_HUMAN]	N	N	N
Q92522	H1-10	Histone H1.10 OS=Homo sapiens OX=9606 GN=H1-10 PE=1 SV=1 - [H1X_HUMAN]	Y	N	N
P16403	H1-2	Histone H1.2 OS=Homo sapiens OX=9606 GN=H1-2 PE=1 SV=2 - [H12_HUMAN]	Y	N	N
Q9P035	HACD3	Very-long-chain (3R)-3-hydroxyacyl-CoA dehydratase 3 OS=Homo sapiens OX=9606 GN=HACD3 PE=1 SV=2 - [HACD3_HUMAN]	Y	N	N
Q13151	HNRNPA0	Heterogeneous nuclear ribonucleoprotein A0 OS=Homo sapiens OX=9606 GN=HNRNPA0 PE=1 SV=1 - [ROA0_HUMAN]	Y	N	N
P09651	HNRNPA1	Heterogeneous nuclear ribonucleoprotein A1 OS=Homo sapiens OX=9606 GN=HNRNPA1 PE=1 SV=5 - [ROA1_HUMAN]	N	Y	N
P22626	HNRNPA2B1	Heterogeneous nuclear ribonucleoproteins A2/B1 OS=Homo sapiens OX=9606 GN=HNRNPA2B1 PE=1 SV=2 - [ROA2_HUMAN]	Y	Y	N
P51991	HNRNPA3	Heterogeneous nuclear ribonucleoprotein A3 OS=Homo sapiens OX=9606 GN=HNRNPA3 PE=1 SV=2 - [ROA3_HUMAN]	N	Y	N

Accession	Gene Name	Description	Enriched in HIV-1 infection	Present in ivRBPome	Present in Hildebrandt AP
Q99729	HNRNPAB	Heterogeneous nuclear ribonucleoprotein A/B OS=Homo sapiens OX=9606 GN=HNRNPAB PE=1 SV=2 - [ROAA_HUMAN]	Y	Y	N
P07910	HNRNPC	Heterogeneous nuclear ribonucleoproteins C1/C2 OS=Homo sapiens OX=9606 GN=HNRNPC PE=1 SV=4 - [HNRPC_HUMAN]	Y	Y	N
Q14103	HNRNPD	Heterogeneous nuclear ribonucleoprotein D0 OS=Homo sapiens OX=9606 GN=HNRNPD PE=1 SV=1 - [HNRPD_HUMAN]	Y	Y	Y
O14979	HNRNPDL	Heterogeneous nuclear ribonucleoprotein D-like OS=Homo sapiens OX=9606 GN=HNRNPDL PE=1 SV=3 - [HNRDL_HUMAN]	Y	N	N
P52597	HNRNPF	Heterogeneous nuclear ribonucleoprotein F OS=Homo sapiens OX=9606 GN=HNRNPF PE=1 SV=3 - [HNRPF_HUMAN]	Y	N	N
P31943	HNRNPH1	Heterogeneous nuclear ribonucleoprotein H OS=Homo sapiens OX=9606 GN=HNRNPH1 PE=1 SV=4 - [HNRH1_HUMAN]	Y	N	N
P61978	HNRNPK	Heterogeneous nuclear ribonucleoprotein K OS=Homo sapiens OX=9606 GN=HNRNPK PE=1 SV=1 - [HNRPK_HUMAN]	N	Y	N
P14866	HNRNPL	Heterogeneous nuclear ribonucleoprotein L OS=Homo sapiens OX=9606 GN=HNRNPL PE=1 SV=2 - [HNRPL_HUMAN]	N	Y	N
P52272	HNRNPM	Heterogeneous nuclear ribonucleoprotein M OS=Homo sapiens OX=9606 GN=HNRNPM PE=1 SV=3 - [HNRPM_HUMAN]	N	Y	Y
O43390	HNRNPR	Heterogeneous nuclear ribonucleoprotein R OS=Homo sapiens OX=9606 GN=HNRNPR PE=1 SV=1 - [HNRPR_HUMAN]	N	Y	N
Q00839	HNRNPU	Heterogeneous nuclear ribonucleoprotein U OS=Homo sapiens OX=9606 GN=HNRNPU PE=1 SV=6 - [HNRPU_HUMAN]	Y	Y	Y
Q9BUJ2	HNRNPUL1	Heterogeneous nuclear ribonucleoprotein U-like protein 1 OS=Homo sapiens OX=9606 GN=HNRNPUL1 PE=1 SV=2 - [HNRL1_HUMAN]	Y	N	N
Q7Z6Z7	HUWE1	E3 ubiquitin-protein ligase HUWE1 OS=Homo sapiens OX=9606 GN=HUWE1 PE=1 SV=3 - [HUWE1_HUMAN]	N	N	N
Q9NZI8	IGF2BP1	Insulin-like growth factor 2 mRNA-binding protein 1 OS=Homo sapiens OX=9606 GN=IGF2BP1 PE=1 SV=2 - [IF2B1_HUMAN]	N	N	Y
Q9Y6M1	IGF2BP2	Insulin-like growth factor 2 mRNA-binding protein 2 OS=Homo sapiens OX=9606 GN=IGF2BP2 PE=1 SV=2 - [IF2B2_HUMAN]	Y	N	Y
O00425	IGF2BP3	Insulin-like growth factor 2 mRNA-binding protein 3 OS=Homo sapiens OX=9606 GN=IGF2BP3 PE=1 SV=2 - [IF2B3_HUMAN]	Y	Y	Y
Q12905	ILF2	Interleukin enhancer-binding factor 2 OS=Homo sapiens OX=9606 GN=ILF2 PE=1 SV=2 - [ILF2_HUMAN]	Y	N	N
Q12906	ILF3	Interleukin enhancer-binding factor 3 OS=Homo sapiens OX=9606 GN=ILF3 PE=1 SV=3 - [ILF3_HUMAN]	Y	Y	Y
O14654	IRS4	Insulin receptor substrate 4 OS=Homo sapiens OX=9606 GN=IRS4 PE=1 SV=1 - [IRS4_HUMAN]	N	N	N
P52292	KPNA2	Importin subunit alpha-1 OS=Homo sapiens OX=9606 GN=KPNA2 PE=1 SV=1 - [IMA1_HUMAN]	Y	N	N
Q6PKG0	LARP1	La-related protein 1 OS=Homo sapiens OX=9606 GN=LARP1 PE=1 SV=2 - [LARP1_HUMAN]	N	N	Y
Q92615	LARP4B	La-related protein 4B OS=Homo sapiens OX=9606 GN=LARP4B PE=1 SV=3 - [LAR4B_HUMAN]	Y	N	Y
P61326	MAGOH	Protein mago nashi homolog OS=Homo sapiens OX=9606 GN=MAGOH PE=1 SV=1 - [MGN_HUMAN]	Y	N	N
P43243	MATR3	Matrin-3 OS=Homo sapiens OX=9606 GN=MATR3 PE=1 SV=2 - [MATR3_HUMAN]	Y	N	N
C9JLW8	MCRIP1	Mapk-regulated corepressor-interacting protein 1 OS=Homo sapiens OX=9606 GN=MCRIP1 PE=1 SV=1 - [MCRI1_HUMAN]	Y	N	N
Q13434	MKRN4P	Putative E3 ubiquitin-protein ligase makorin-4 OS=Homo sapiens OX=9606 GN=MKRN4P PE=5 SV=1 - [MKRN4_HUMAN]	N	N	N
P52815	MRPL12	39S ribosomal protein L12, mitochondrial OS=Homo sapiens OX=9606 GN=MRPL12 PE=1 SV=2 - [RM12_HUMAN]	N	N	N
P82650	MRPS22	28S ribosomal protein S22, mitochondrial OS=Homo sapiens OX=9606 GN=MRPS22 PE=1 SV=1 - [RT22_HUMAN]	N	N	N
P82663	MRPS25	28S ribosomal protein S25, mitochondrial OS=Homo sapiens OX=9606 GN=MRPS25 PE=1 SV=1 - [RT25_HUMAN]	N	N	N
P82933	MRPS9	28S ribosomal protein S9, mitochondrial OS=Homo sapiens OX=9606 GN=MRPS9 PE=1 SV=2 - [RT09_HUMAN]	N	N	N

Accession	Gene Name	Description	Enriched in HIV-1 infection	Present in ivRBPome	Present in Hildebrandt AP
Q09161	NCBP1	Nuclear cap-binding protein subunit 1 OS=Homo sapiens OX=9606 GN=NCBP1 PE=1 SV=1 - [NCBP1_HUMAN]	Y	N	N
P52298	NCBP2	Nuclear cap-binding protein subunit 2 OS=Homo sapiens OX=9606 GN=NCBP2 PE=1 SV=1 - [NCBP2_HUMAN]	N	N	N
P19338	NCL	Nucleolin OS=Homo sapiens OX=9606 GN=NCL PE=1 SV=3 - [NUCL_HUMAN]	Y	Y	N
O43809	NUDT21	Cleavage and polyadenylation specificity factor subunit 5 OS=Homo sapiens OX=9606 GN=NUDT21 PE=1 SV=1 - [CPSF5_HUMAN]	N	N	N
P11940	PABPC1	Polyadenylate-binding protein 1 OS=Homo sapiens OX=9606 GN=PABPC1 PE=1 SV=2 - [PABP1_HUMAN]	Y	Y	Y
Q13310	PABPC4	Polyadenylate-binding protein 4 OS=Homo sapiens OX=9606 GN=PABPC4 PE=1 SV=1 - [PABP4_HUMAN]	Y	Y	Y
Q86U42	PABPN1	Polyadenylate-binding protein 2 OS=Homo sapiens OX=9606 GN=PABPN1 PE=1 SV=3 - [PABP2_HUMAN]	Y	N	N
Q99623	PHB2	Prohibitin-2 OS=Homo sapiens OX=9606 GN=PHB2 PE=1 SV=2 - [PHB2_HUMAN]	N	N	N
P25786	PSMA1	Proteasome subunit alpha type-1 OS=Homo sapiens OX=9606 GN=PSMA1 PE=1 SV=1 - [PSA1_HUMAN]	N	N	N
P25787	PSMA2	Proteasome subunit alpha type-2 OS=Homo sapiens OX=9606 GN=PSMA2 PE=1 SV=2 - [PSA2_HUMAN]	Y	N	N
P25788	PSMA3	Proteasome subunit alpha type-3 OS=Homo sapiens OX=9606 GN=PSMA3 PE=1 SV=2 - [PSA3_HUMAN]	Y	N	N
P25789	PSMA4	Proteasome subunit alpha type-4 OS=Homo sapiens OX=9606 GN=PSMA4 PE=1 SV=1 - [PSA4_HUMAN]	N	N	N
P28066	PSMA5	Proteasome subunit alpha type-5 OS=Homo sapiens OX=9606 GN=PSMA5 PE=1 SV=3 - [PSA5_HUMAN]	N	N	N
P60900	PSMA6	Proteasome subunit alpha type-6 OS=Homo sapiens OX=9606 GN=PSMA6 PE=1 SV=1 - [PSA6_HUMAN]	N	N	N
O14818	PSMA7	Proteasome subunit alpha type-7 OS=Homo sapiens OX=9606 GN=PSMA7 PE=1 SV=1 - [PSA7_HUMAN]	N	N	N
P20618	PSMB1	Proteasome subunit beta type-1 OS=Homo sapiens OX=9606 GN=PSMB1 PE=1 SV=2 - [PSB1_HUMAN]	Y	N	N
P28074	PSMB5	Proteasome subunit beta type-5 OS=Homo sapiens OX=9606 GN=PSMB5 PE=1 SV=3 - [PSB5_HUMAN]	N	N	N
P28072	PSMB6	Proteasome subunit beta type-6 OS=Homo sapiens OX=9606 GN=PSMB6 PE=1 SV=4 - [PSB6_HUMAN]	N	N	N
P62191	PSMC1	26S proteasome regulatory subunit 4 OS=Homo sapiens OX=9606 GN=PSMC1 PE=1 SV=1 - [PRS4_HUMAN]	N	N	N
P35998	PSMC2	26S proteasome regulatory subunit 7 OS=Homo sapiens OX=9606 GN=PSMC2 PE=1 SV=3 - [PRS7_HUMAN]	N	N	N
P17980	PSMC3	26S proteasome regulatory subunit 6A OS=Homo sapiens OX=9606 GN=PSMC3 PE=1 SV=3 - [PRS6A_HUMAN]	N	N	N
P62195	PSMC5	26S proteasome regulatory subunit 8 OS=Homo sapiens OX=9606 GN=PSMC5 PE=1 SV=1 - [PRS8_HUMAN]	N	N	N
Q99460	PSMD1	26S proteasome non-ATPase regulatory subunit 1 OS=Homo sapiens OX=9606 GN=PSMD1 PE=1 SV=2 - [PSMD1_HUMAN]	Y	N	N
O00231	PSMD11	26S proteasome non-ATPase regulatory subunit 11 OS=Homo sapiens OX=9606 GN=PSMD11 PE=1 SV=3 - [PSD11_HUMAN]	N	N	N
O00232	PSMD12	26S proteasome non-ATPase regulatory subunit 12 OS=Homo sapiens OX=9606 GN=PSMD12 PE=1 SV=3 - [PSD12_HUMAN]	Y	N	N
O00487	PSMD14	26S proteasome non-ATPase regulatory subunit 14 OS=Homo sapiens OX=9606 GN=PSMD14 PE=1 SV=1 - [PSDE_HUMAN]	N	N	N
Q13200	PSMD2	26S proteasome non-ATPase regulatory subunit 2 OS=Homo sapiens OX=9606 GN=PSMD2 PE=1 SV=3 - [PSMD2_HUMAN]	N	N	N
O43242	PSMD3	26S proteasome non-ATPase regulatory subunit 3 OS=Homo sapiens OX=9606 GN=PSMD3 PE=1 SV=2 - [PSMD3_HUMAN]	Y	N	N
Q15008	PSMD6	26S proteasome non-ATPase regulatory subunit 6 OS=Homo sapiens OX=9606 GN=PSMD6 PE=1 SV=1 - [PSMD6_HUMAN]	N	N	N
P26599	PTBP1	Polypyrimidine tract-binding protein 1 OS=Homo sapiens OX=9606 GN=PTBP1 PE=1 SV=1 - [PTBP1_HUMAN]	Y	N	N

Accession	Gene Name	Description	Enriched in HIV-1 infection	Present in ivRBPome	Present in Hildebrandt AP
Q9UKM9	RALY	RNA-binding protein Raly OS=Homo sapiens OX=9606 GN=RALY PE=1 SV=1 - [RALY_HUMAN]	N	N	N
Q9Y5S9	RBM8A	RNA-binding protein 8A OS=Homo sapiens OX=9606 GN=RBM8A PE=1 SV=1 - [RBM8A_HUMAN]	N	N	N
P38159	RBMX	RNA-binding motif protein, X chromosome OS=Homo sapiens OX=9606 GN=RBMX PE=1 SV=3 - [RBMX_HUMAN]	N	Y	N
Q14257	RCN2	Reticulocalbin-2 OS=Homo sapiens OX=9606 GN=RCN2 PE=1 SV=1 - [RCN2_HUMAN]	N	N	N
P27635	RPL10	60S ribosomal protein L10 OS=Homo sapiens OX=9606 GN=RPL10 PE=1 SV=4 - [RL10_HUMAN]	N	Y	Y
P62913	RPL11	60S ribosomal protein L11 OS=Homo sapiens OX=9606 GN=RPL11 PE=1 SV=2 - [RL11_HUMAN]	Y	N	N
P30050	RPL12	60S ribosomal protein L12 OS=Homo sapiens OX=9606 GN=RPL12 PE=1 SV=1 - [RL12_HUMAN]	N	N	N
P26373	RPL13	60S ribosomal protein L13 OS=Homo sapiens OX=9606 GN=RPL13 PE=1 SV=4 - [RL13_HUMAN]	Y	Y	N
P40429	RPL13A	60S ribosomal protein L13a OS=Homo sapiens OX=9606 GN=RPL13A PE=1 SV=2 - [RL13A_HUMAN]	Y	Y	N
P50914	RPL14	60S ribosomal protein L14 OS=Homo sapiens OX=9606 GN=RPL14 PE=1 SV=4 - [RL14_HUMAN]	Y	Y	N
P18621	RPL17	60S ribosomal protein L17 OS=Homo sapiens OX=9606 GN=RPL17 PE=1 SV=3 - [RL17_HUMAN]	N	N	Y
Q07020	RPL18	60S ribosomal protein L18 OS=Homo sapiens OX=9606 GN=RPL18 PE=1 SV=2 - [RL18_HUMAN]	Y	Y	N
Q02543	RPL18A	60S ribosomal protein L18a OS=Homo sapiens OX=9606 GN=RPL18A PE=1 SV=2 - [RL18A_HUMAN]	N	Y	N
P46778	RPL21	60S ribosomal protein L21 OS=Homo sapiens OX=9606 GN=RPL21 PE=1 SV=2 - [RL21_HUMAN]	Y	N	N
P35268	RPL22	60S ribosomal protein L22 OS=Homo sapiens OX=9606 GN=RPL22 PE=1 SV=2 - [RL22_HUMAN]	Y	N	N
P62829	RPL23	60S ribosomal protein L23 OS=Homo sapiens OX=9606 GN=RPL23 PE=1 SV=1 - [RL23_HUMAN]	N	N	N
P62750	RPL23A	60S ribosomal protein L23a OS=Homo sapiens OX=9606 GN=RPL23A PE=1 SV=1 - [RL23A_HUMAN]	N	Y	Y
P83731	RPL24	60S ribosomal protein L24 OS=Homo sapiens OX=9606 GN=RPL24 PE=1 SV=1 - [RL24_HUMAN]	Y	Y	N
P61254	RPL26	60S ribosomal protein L26 OS=Homo sapiens OX=9606 GN=RPL26 PE=1 SV=1 - [RL26_HUMAN]	N	N	Y
P61353	RPL27	60S ribosomal protein L27 OS=Homo sapiens OX=9606 GN=RPL27 PE=1 SV=2 - [RL27_HUMAN]	N	Y	N
P46776	RPL27A	60S ribosomal protein L27a OS=Homo sapiens OX=9606 GN=RPL27A PE=1 SV=2 - [RL27A_HUMAN]	Y	Y	N
P46779	RPL28	60S ribosomal protein L28 OS=Homo sapiens OX=9606 GN=RPL28 PE=1 SV=3 - [RL28_HUMAN]	N	Y	N
P39023	RPL3	60S ribosomal protein L3 OS=Homo sapiens OX=9606 GN=RPL3 PE=1 SV=2 - [RL3_HUMAN]	N	Y	Y
P62899	RPL31	60S ribosomal protein L31 OS=Homo sapiens OX=9606 GN=RPL31 PE=1 SV=1 - [RL31_HUMAN]	Y	Y	Y
P62910	RPL32	60S ribosomal protein L32 OS=Homo sapiens OX=9606 GN=RPL32 PE=1 SV=2 - [RL32_HUMAN]	N	N	N
P42766	RPL35	60S ribosomal protein L35 OS=Homo sapiens OX=9606 GN=RPL35 PE=1 SV=2 - [RL35_HUMAN]	Y	Y	Y
P18077	RPL35A	60S ribosomal protein L35a OS=Homo sapiens OX=9606 GN=RPL35A PE=1 SV=2 - [RL35A_HUMAN]	N	N	N
Q9Y3U8	RPL36	60S ribosomal protein L36 OS=Homo sapiens OX=9606 GN=RPL36 PE=1 SV=3 - [RL36_HUMAN]	Y	Y	N
P83881	RPL36A	60S ribosomal protein L36a OS=Homo sapiens OX=9606 GN=RPL36A PE=1 SV=2 - [RL36A_HUMAN]	Y	N	N

Accession	Gene Name	Description	Enriched in HIV-1 infection	Present in ivRBPome	Present in Hildebrandt AP
P62891	RPL39	60S ribosomal protein L39 OS=Homo sapiens OX=9606 GN=RPL39 PE=1 SV=2 - [RL39_HUMAN]	Y	N	N
P36578	RPL4	60S ribosomal protein L4 OS=Homo sapiens OX=9606 GN=RPL4 PE=1 SV=5 - [RL4_HUMAN]	Y	Y	N
Q02878	RPL6	60S ribosomal protein L6 OS=Homo sapiens OX=9606 GN=RPL6 PE=1 SV=3 - [RL6_HUMAN]	N	Y	N
P18124	RPL7	60S ribosomal protein L7 OS=Homo sapiens OX=9606 GN=RPL7 PE=1 SV=1 - [RL7_HUMAN]	Y	Y	N
P05388	RPLP0	60S acidic ribosomal protein P0 OS=Homo sapiens OX=9606 GN=RPLP0 PE=1 SV=1 - [RLA0_HUMAN]	N	N	N
P05386	RPLP1	60S acidic ribosomal protein P1 OS=Homo sapiens OX=9606 GN=RPLP1 PE=1 SV=1 - [RLA1_HUMAN]	N	N	N
P05387	RPLP2	60S acidic ribosomal protein P2 OS=Homo sapiens OX=9606 GN=RPLP2 PE=1 SV=1 - [RLA2_HUMAN]	N	N	N
P46783	RPS10	40S ribosomal protein S10 OS=Homo sapiens OX=9606 GN=RPS10 PE=1 SV=1 - [RS10_HUMAN]	Y	N	N
P62280	RPS11	40S ribosomal protein S11 OS=Homo sapiens OX=9606 GN=RPS11 PE=1 SV=3 - [RS11_HUMAN]	N	N	Y
P62277	RPS13	40S ribosomal protein S13 OS=Homo sapiens OX=9606 GN=RPS13 PE=1 SV=2 - [RS13_HUMAN]	Y	N	Y
P62263	RPS14	40S ribosomal protein S14 OS=Homo sapiens OX=9606 GN=RPS14 PE=1 SV=3 - [RS14_HUMAN]	Y	Y	N
P62244	RPS15A	40S ribosomal protein S15a OS=Homo sapiens OX=9606 GN=RPS15A PE=1 SV=2 - [RS15A_HUMAN]	Y	N	N
P62249	RPS16	40S ribosomal protein S16 OS=Homo sapiens OX=9606 GN=RPS16 PE=1 SV=2 - [RS16_HUMAN]	Y	N	N
P08708	RPS17	40S ribosomal protein S17 OS=Homo sapiens OX=9606 GN=RPS17 PE=1 SV=2 - [RS17_HUMAN]	Y	N	N
P62269	RPS18	40S ribosomal protein S18 OS=Homo sapiens OX=9606 GN=RPS18 PE=1 SV=3 - [RS18_HUMAN]	N	N	N
P39019	RPS19	40S ribosomal protein S19 OS=Homo sapiens OX=9606 GN=RPS19 PE=1 SV=2 - [RS19_HUMAN]	N	N	N
P15880	RPS2	40S ribosomal protein S2 OS=Homo sapiens OX=9606 GN=RPS2 PE=1 SV=2 - [RS2_HUMAN]	Y	Y	Y
P62266	RPS23	40S ribosomal protein S23 OS=Homo sapiens OX=9606 GN=RPS23 PE=1 SV=3 - [RS23_HUMAN]	Y	N	N
P62847	RPS24	40S ribosomal protein S24 OS=Homo sapiens OX=9606 GN=RPS24 PE=1 SV=1 - [RS24_HUMAN]	Y	Y	N
P62851	RPS25	40S ribosomal protein S25 OS=Homo sapiens OX=9606 GN=RPS25 PE=1 SV=1 - [RS25_HUMAN]	N	Y	N
P62854	RPS26	40S ribosomal protein S26 OS=Homo sapiens OX=9606 GN=RPS26 PE=1 SV=3 - [RS26_HUMAN]	N	N	Y
P42677	RPS27	40S ribosomal protein S27 OS=Homo sapiens OX=9606 GN=RPS27 PE=1 SV=3 - [RS27_HUMAN]	N	N	N
P62979	RPS27A	Ubiquitin-40S ribosomal protein S27a OS=Homo sapiens OX=9606 GN=RPS27A PE=1 SV=2 - [RS27A_HUMAN]	Y	N	N
Q71UM5	RPS27L	40S ribosomal protein S27-like OS=Homo sapiens OX=9606 GN=RPS27L PE=1 SV=3 - [RS27L_HUMAN]	N	N	N
P23396	RPS3	40S ribosomal protein S3 OS=Homo sapiens OX=9606 GN=RPS3 PE=1 SV=2 - [RS3_HUMAN]	Y	N	N
P61247	RPS3A	40S ribosomal protein S3a OS=Homo sapiens OX=9606 GN=RPS3A PE=1 SV=2 - [RS3A_HUMAN]	N	Y	N
P62701	RPS4X	40S ribosomal protein S4, X isoform OS=Homo sapiens OX=9606 GN=RPS4X PE=1 SV=2 - [RS4X_HUMAN]	N	Y	Y
P46782	RPS5	40S ribosomal protein S5 OS=Homo sapiens OX=9606 GN=RPS5 PE=1 SV=4 - [RS5_HUMAN]	N	N	N
P62753	RPS6	40S ribosomal protein S6 OS=Homo sapiens OX=9606 GN=RPS6 PE=1 SV=1 - [RS6_HUMAN]	N	Y	N

Accession	Gene Name	Description	Enriched in HIV-1 infection	Present in ivRBPome	Present in Hildebrandt AP
P62081	RPS7	40S ribosomal protein S7 OS=Homo sapiens OX=9606 GN=RPS7 PE=1 SV=1 - [RS7_HUMAN]	Y	N	N
P62241	RPS8	40S ribosomal protein S8 OS=Homo sapiens OX=9606 GN=RPS8 PE=1 SV=2 - [RS8_HUMAN]	Y	Y	Y
P46781	RPS9	40S ribosomal protein S9 OS=Homo sapiens OX=9606 GN=RPS9 PE=1 SV=3 - [RS9_HUMAN]	Y	Y	Y
Q9Y224	RTRAF	RNA transcription, translation and transport factor protein OS=Homo sapiens OX=9606 GN=RTRAF PE=1 SV=1 - [RTRAF_HUMAN]	N	N	N
Q9Y467	SALL2	Sal-like protein 2 OS=Homo sapiens OX=9606 GN=SALL2 PE=1 SV=4 - [SALL2_HUMAN]	Y	N	N
P53007	SLC25A1	Tricarboxylate transport protein, mitochondrial OS=Homo sapiens OX=9606 GN=SLC25A1 PE=1 SV=2 - [TXTP_HUMAN]	Y	N	N
Q9UBX3	SLC25A10	Mitochondrial dicarboxylate carrier OS=Homo sapiens OX=9606 GN=SLC25A10 PE=1 SV=2 - [DIC_HUMAN]	Y	N	N
P08579	SNRPB2	U2 small nuclear ribonucleoprotein B" OS=Homo sapiens OX=9606 GN=SNRPB2 PE=1 SV=1 - [RU2B_HUMAN]	Y	N	N
P09234	SNRPC	U1 small nuclear ribonucleoprotein C OS=Homo sapiens OX=9606 GN=SNRPC PE=1 SV=1 - [RU1C_HUMAN]	N	N	N
P62318	SNRPD3	Small nuclear ribonucleoprotein Sm D3 OS=Homo sapiens OX=9606 GN=SNRPD3 PE=1 SV=1 - [SMD3_HUMAN]	N	N	N
A8MWD9	SNRPGP15	Putative small nuclear ribonucleoprotein G-like protein 15 OS=Homo sapiens OX=9606 GN=SNRPGP15 PE=5 SV=2 - [RUXGL_HUMAN]	N	N	N
P37108	SRP14	Signal recognition particle 14 kDa protein OS=Homo sapiens OX=9606 GN=SRP14 PE=1 SV=2 - [SRP14_HUMAN]	N	N	N
Q9UQ35	SRRM2	Serine/arginine repetitive matrix protein 2 OS=Homo sapiens OX=9606 GN=SRRM2 PE=1 SV=2 - [SRRM2_HUMAN]	N	N	N
Q07955	SRSF1	Serine/arginine-rich splicing factor 1 OS=Homo sapiens OX=9606 GN=SRSF1 PE=1 SV=2 - [SRSF1_HUMAN]	N	N	N
Q01130	SRSF2	Serine/arginine-rich splicing factor 2 OS=Homo sapiens OX=9606 GN=SRSF2 PE=1 SV=4 - [SRSF2_HUMAN]	Y	Y	N
P84103	SRSF3	Serine/arginine-rich splicing factor 3 OS=Homo sapiens OX=9606 GN=SRSF3 PE=1 SV=1 - [SRSF3_HUMAN]	N	N	N
Q13247	SRSF6	Serine/arginine-rich splicing factor 6 OS=Homo sapiens OX=9606 GN=SRSF6 PE=1 SV=2 - [SRSF6_HUMAN]	N	N	N
Q16629	SRSF7	Serine/arginine-rich splicing factor 7 OS=Homo sapiens OX=9606 GN=SRSF7 PE=1 SV=1 - [SRSF7_HUMAN]	Y	N	N
Q13242	SRSF9	Serine/arginine-rich splicing factor 9 OS=Homo sapiens OX=9606 GN=SRSF9 PE=1 SV=1 - [SRSF9_HUMAN]	Y	N	N
P05455	SSB	Lupus La protein OS=Homo sapiens OX=9606 GN=SSB PE=1 SV=2 - [LA_HUMAN]	N	N	N
Q9Y3F4	STRAP	Serine-threonine kinase receptor-associated protein OS=Homo sapiens OX=9606 GN=STRAP PE=1 SV=1 - [STRAP_HUMAN]	N	N	N
O60506	SYNCRIP	Heterogeneous nuclear ribonucleoprotein Q OS=Homo sapiens OX=9606 GN=SYNCRIP PE=1 SV=2 - [HNRPQ_HUMAN]	Y	Y	Y
Q3ZCQ8	TIMM50	Mitochondrial import inner membrane translocase subunit TIM50 OS=Homo sapiens OX=9606 GN=TIMM50 PE=1 SV=2 - [TIM50_HUMAN]	Y	N	N
Q13595	TRA2A	Transformer-2 protein homolog alpha OS=Homo sapiens OX=9606 GN=TRA2A PE=1 SV=1 - [TRA2A_HUMAN]	Y	N	N
P62995	TRA2B	Transformer-2 protein homolog beta OS=Homo sapiens OX=9606 GN=TRA2B PE=1 SV=1 - [TRA2B_HUMAN]	Y	N	N
Q13263	TRIM28	Transcription intermediary factor 1-beta OS=Homo sapiens OX=9606 GN=TRIM28 PE=1 SV=5 - [TIF1B_HUMAN]	Y	N	N
P26368	U2AF2	Splicing factor U2AF 65 kDa subunit OS=Homo sapiens OX=9606 GN=U2AF2 PE=1 SV=4 - [U2AF2_HUMAN]	N	N	Y
Q92900	UPF1	Regulator of nonsense transcripts 1 OS=Homo sapiens OX=9606 GN=UPF1 PE=1 SV=2 - [RENT1_HUMAN]	N	Y	Y
P67809	YBX1	Y-box-binding protein 1 OS=Homo sapiens OX=9606 GN=YBX1 PE=1 SV=3 - [YBOX1_HUMAN]	Y	Y	Y
P16989	YBX3	Y-box-binding protein 3 OS=Homo sapiens OX=9606 GN=YBX3 PE=1 SV=4 - [YBOX3_HUMAN]	N	Y	Y

10 Appendix 3

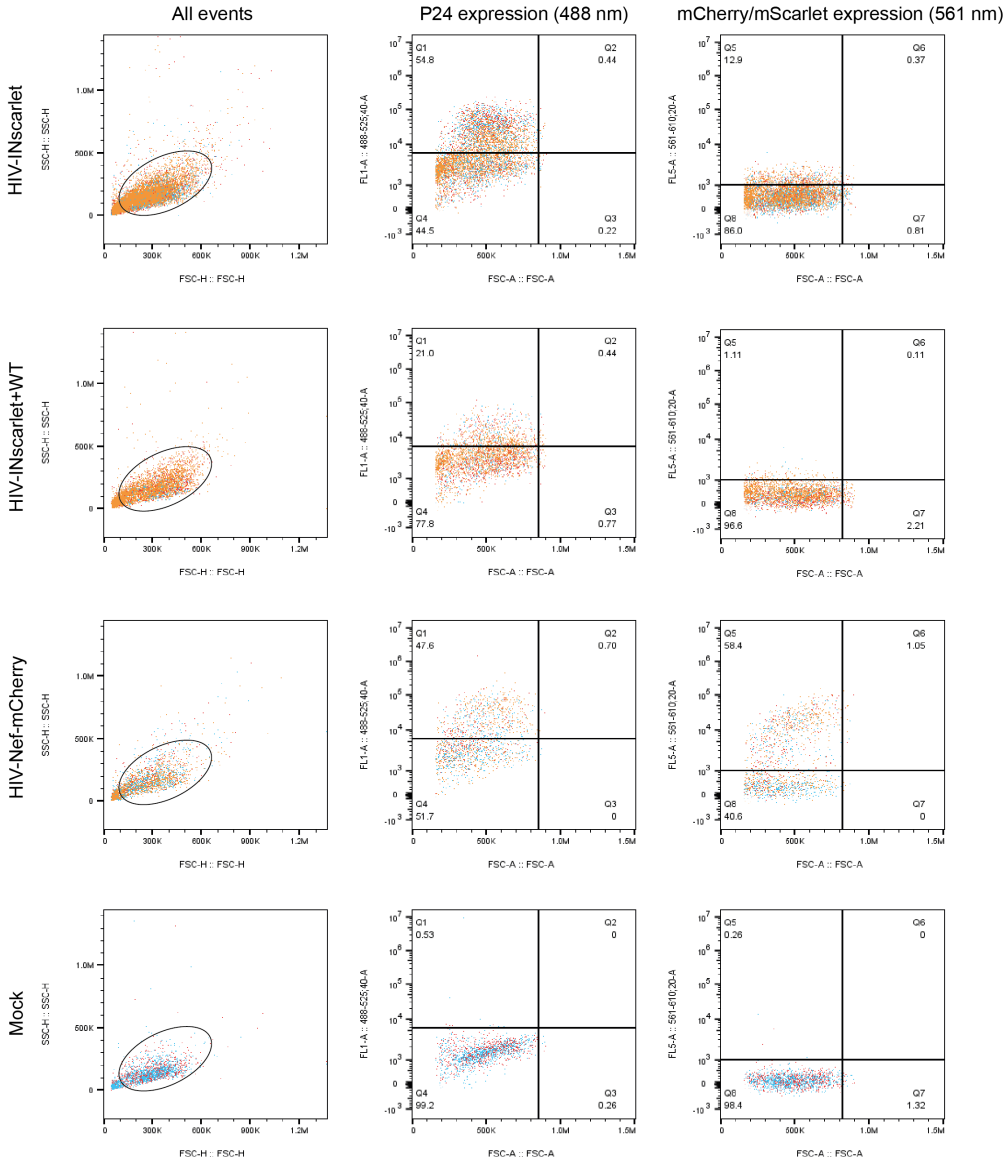

Figure 10.1: **Flow cytometry raw data for Figure 6.4B and C INscarlet P24 and mScarlet expression** SupT1 cells were infected with HIV-$1_{INscarlet}$, HIV-$1_{Nef\text{-}mCherry}$ and HIV-$1_{INscarlet/WT}$ (a WT complemented virus produced by co-transfecting cells with WT HIV-1 and HIV-$1_{INscarlet}$). At 48 hpi cells were fixed and intracellularly stained with anti-p24. The percentage of cells expressing HIV-1 p24 and mScarlet/mCherry was determined by flow cytometry. Each colour represents a different experimental repeat, n=3.

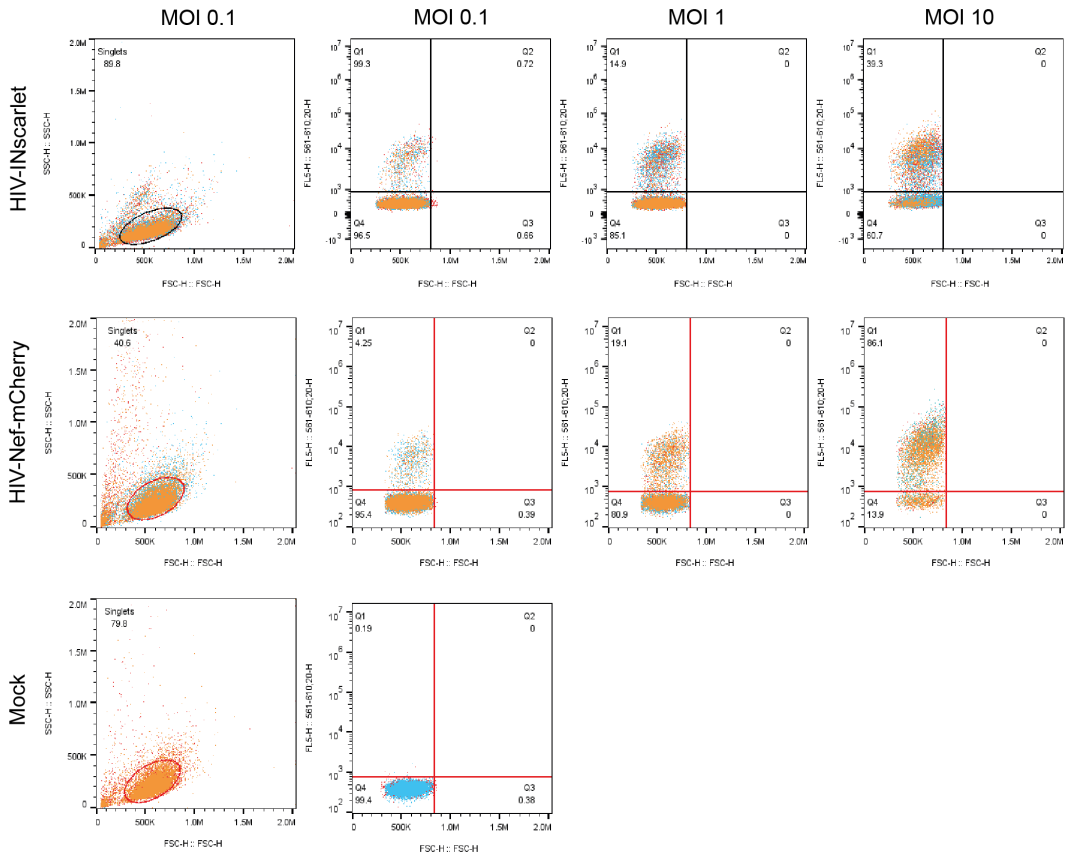

Figure 10.2: **Flow cytometry raw data for Figure 6.4D, second round of infection with HIV-1$_{\mathbf{INscarlet}}$** HIV-1$_{INscarlet}$ and HIV-1$_{Nef\text{-}mCherry}$ viral particles were collected from the supernatant of infected SupT1 cells, purified and used to infect fresh SupT1 cells at MOI of 0.1, 1 and 10. At 48 hpi the cells were fixed and the percentage of cells expressing mCherry/mScarlet was determined by flow cytometry. Gating strategy is shown for MOI 0.1 but was applied to all conditions. Each colour represents a different experimental repeat, n=3.

11 References

[1] World Health Organisation. Key facts and the latest estimates on the global hiv epidemic - 2020. https://www.who.int/teams/global-hiv-hepatitis-and-stis-programmes/hiv/strategic-information/hiv-data-and-statistics, 2020. Accessed: 2023-01-02.

[2] Eric J Arts and Daria J Hazuda. Hiv-1 antiretroviral drug therapy. *Cold Spring Harbor perspectives in medicine*, 2(4):a007161, 2012.

[3] Maria C Puertas, George Ploumidis, Michalis Ploumidis, Emilio Fumero, Bonaventura Clotet, Charles M Walworth, Christos J Petropoulos, and Javier Martinez-Picado. Pan-resistant hiv-1 emergence in the era of integrase strand-transfer inhibitors: a case report. *The Lancet Microbe*, 1(3):e130–e135, 2020.

[4] Milagros Moreno, Estrella Caballero, Raquel M Mateus, Filomena Samba, Eva Gil, and Vicenç Falcó. Hiv drug resistance in africa: an emerging problem that deserves urgent attention. *AIDS*, 31(11):1637–1639, 2017.

[5] Stefan HE Kaufmann, Anca Dorhoi, Richard S Hotchkiss, and Ralf Bartenschlager. Host-directed therapies for bacterial and viral infections. *Nature reviews Drug discovery*, 17(1):35–56, 2018.

[6] Jim Baggen, Els Vanstreels, Sander Jansen, and Dirk Daelemans. Cellular host factors for sars-cov-2 infection. *Nature Microbiology*, 6(10):1219–1232, 2021.

[7] Peter D Nagy and Judit Pogany. The dependence of viral rna replication on co-opted host factors. *Nature Reviews Microbiology*, 10(2):137–149, 2012.

[8] Katharina A Mayer, Johannes Stöckl, Gerhard J Zlabinger, and Guido A Gualdoni. Hijacking the supplies: metabolism as a novel facet of virus-host interaction. *Frontiers in immunology*, 10:1533, 2019.

[9] Kate Dicker, Aino I Järvelin, Manuel Garcia-Moreno, and Alfredo Castello. The importance of virion-incorporated cellular rna-binding proteins in viral particle

assembly and infectivity. In *Seminars in cell & developmental biology*, volume 111, pages 108–118. Elsevier, 2021.

[10] Korbinian Bösl, Aleksandr Ianevski, Thoa T Than, Petter I Andersen, Suvi Kuivanen, Mona Teppor, Eva Zusinaite, Uga Dumpis, Astra Vitkauskiene, Rebecca J Cox, et al. Common nodes of virus–host interaction revealed through an integrated network analysis. *Frontiers in immunology*, page 2186, 2019.

[11] David E Gordon, Joseph Hiatt, Mehdi Bouhaddou, Veronica V Rezelj, Svenja Ulferts, Hannes Braberg, Alexander S Jureka, Kirsten Obernier, Jeffrey Z Guo, Jyoti Batra, et al. Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science*, 370(6521):eabe9403, 2020.

[12] Neelanjana Ray. Maraviroc in the treatment of hiv infection. *Drug design, development and therapy*, pages 151–161, 2009.

[13] Mike Westby, Caroline Smith-Burchnell, Julie Mori, Marilyn Lewis, Michael Mosley, Mark Stockdale, Patrick Dorr, Giuseppe Ciaramella, and Manos Perros. Reduced maximal inhibition in phenotypic susceptibility assays indicates that viral strains resistant to the ccr5 antagonist maraviroc utilize inhibitor-bound receptor for entry. *Journal of virology*, 81(5):2359–2371, 2007.

[14] Devavrat Tripathi, Megha Sodani, Pramod Kumar Gupta, and Savita Kulkarni. Host directed therapies: Covid-19 and beyond. *Current Research in Pharmacology and Drug Discovery*, 2:100058, 2021.

[15] Manuel Garcia-Moreno, Marko Noerenberg, Shuai Ni, Aino I Järvelin, Esther González-Almela, Caroline E Lenz, Marcel Bach-Pages, Victoria Cox, Rosario Avolio, Thomas Davis, et al. System-wide profiling of rna-binding proteins uncovers key regulators of virus infection. *Molecular cell*, 74(1):196–211, 2019.

[16] Manuel Garcia-Moreno, Aino I Järvelin, and Alfredo Castello. Unconventional rna-binding proteins step into the virus–host battlefront. *Wiley Interdisciplinary Reviews: RNA*, 9(6):e1498, 2018.

[17] Tina Glisovic, Jennifer L Bachorik, Jeongsik Yong, and Gideon Dreyfuss. Rna-binding proteins and post-transcriptional gene regulation. *FEBS letters*, 582(14):1977–1986, 2008.

[18] Fátima Gebauer, Thomas Schwarzl, Juan Valcárcel, and Matthias W Hentze. Rna-binding proteins in human genetic disease. *Nature Reviews Genetics*, 22(3):185–198, 2021.

[19] Bruno Pereira, Marc Billaud, and Raquel Almeida. Rna-binding proteins in cancer: old players and new actors. *Trends in cancer*, 3(7):506–528, 2017.

[20] Ruben G de Bruin, Ton J Rabelink, Anton Jan van Zonneveld, and Eric P van der Veer. Emerging roles for rna-binding proteins as effectors and regulators of cardiovascular disease. *European heart journal*, 38(18):1380–1388, 2017.

[21] Masanori Yoshinaga and Osamu Takeuchi. Rna binding proteins in the control of autoimmune diseases. *Immunological medicine*, 42(2):53–64, 2019.

[22] Martin Turner, Alison Galloway, and Elena Vigorito. Noncoding rna and its associated proteins as regulatory elements of the immune system. *Nature immunology*, 15(6):484–491, 2014.

[23] Bradley M Lunde, Claire Moore, and Gabriele Varani. Rna-binding proteins: modular design for efficient function. *Nature reviews Molecular cell biology*, 8(6):479–490, 2007.

[24] William H Hudson and Eric A Ortlund. The structure, function and evolution of proteins that bind dna and rna. *Nature reviews Molecular cell biology*, 15(11):749–760, 2014.

[25] Guramrit Singh, Gabriel Pratt, Gene W Yeo, and Melissa J Moore. The clothes make the mrna: past and present trends in mrnp fashion. *Annual review of biochemistry*, 84:325, 2015.

[26] Thomas R Cech and Joan A Steitz. The noncoding rna revolution—trashing old rules to forge new ones. *Cell*, 157(1):77–94, 2014.

[27] Rastislav Horos, Magdalena Büscher, Rozemarijn Kleinendorst, Anne-Marie Alleaume, Abul K Tarafder, Thomas Schwarzl, Dmytro Dziuba, Christian Tischer, Elisabeth M Zielonka, Asli Adak, et al. The small non-coding vault rna1-1 acts as a riboregulator of autophagy. *Cell*, 176(5):1054–1067, 2019.

[28] Jacint G Sanchez, Konstantin MJ Sparrer, Cindy Chiang, Rebecca A Reis, Jessica J Chiang, Matthew A Zurenski, Yueping Wan, Michaela U Gack, and Owen Pornillos. Trim25 binds rna to modulate cellular anti-viral defense. *Journal of molecular biology*, 430(24):5280–5293, 2018.

[29] Thomas A Steitz. A structural understanding of the dynamic ribosome machine. *Nature reviews Molecular cell biology*, 9(3):242–253, 2008.

[30] Clemens Plaschka, Pei-Chun Lin, and Kiyoshi Nagai. Structure of a pre-catalytic spliceosome. *Nature*, 546(7660):617–621, 2017.

[31] Jakob Trendel, Thomas Schwarzl, Rastislav Horos, Ananth Prakash, Alex Bateman, Matthias W Hentze, and Jeroen Krijgsveld. The human rna-binding proteome and its dynamics during translational arrest. *Cell*, 176(1-2):391–403, 2019.

[32] Rayner ML Queiroz, Tom Smith, Eneko Villanueva, Maria Marti-Solano, Mie Monti, Mariavittoria Pizzinga, Dan-Mircea Mirea, Manasa Ramakrishna, Robert F Harvey, Veronica Dezi, et al. Comprehensive identification of rna–protein interactions in any organism using orthogonal organic phase separation (oops). *Nature biotechnology*, 37(2):169–178, 2019.

[33] Erika C Urdaneta, Carlos H Vieira-Vieira, Timon Hick, Hans-Herrmann Wessels, Davide Figini, Rebecca Moschall, Jan Medenbach, Uwe Ohler, Sander Granneman, Matthias Selbach, et al. Purification of cross-linked rna-protein complexes by phenol-toluol extraction. *Nature communications*, 10(1):1–17, 2019.

[34] Alexander G Baltz, Mathias Munschauer, Björn Schwanhäusser, Alexandra Vasile, Yasuhiro Murakawa, Markus Schueler, Noah Youngs, Duncan Penfold-

Brown, Kevin Drew, Miha Milek, et al. The mrna-bound proteome and its global occupancy profile on protein-coding transcripts. *Molecular cell*, 46(5):674–690, 2012.

[35] Alfredo Castello, Bernd Fischer, Katrin Eichelbaum, Rastislav Horos, Benedikt M Beckmann, Claudia Strein, Norman E Davey, David T Humphreys, Thomas Preiss, Lars M Steinmetz, et al. Insights into rna biology from an atlas of mammalian mrna-binding proteins. *Cell*, 149(6):1393–1406, 2012.

[36] Matthias W Hentze, Alfredo Castello, Thomas Schwarzl, and Thomas Preiss. A brave new world of rna-binding proteins. *Nature reviews Molecular cell biology*, 19(5):327–341, 2018.

[37] Alfredo Castello, Bernd Fischer, Christian K Frese, Rastislav Horos, Anne-Marie Alleaume, Sophia Foehr, Tomaz Curk, Jeroen Krijgsveld, and Matthias W Hentze. Comprehensive identification of rna-binding domains in human cells. *Molecular cell*, 63(4):696–710, 2016.

[38] Chongsheng He, Simone Sidoli, Robert Warneford-Thomson, Deirdre C Tatomer, Jeremy E Wilusz, Benjamin A Garcia, and Roberto Bonasio. High-resolution mapping of rna-binding regions in the nuclear proteome of embryonic stem cells. *Molecular cell*, 64(2):416–430, 2016.

[39] Aino I Järvelin, Marko Noerenberg, Ilan Davis, and Alfredo Castello. The new (dis) order in rna regulation. *Cell Communication and Signaling*, 14(1):1–22, 2016.

[40] Inbal Paz, Efrat Kligun, Barak Bengad, and Yael Mandel-Gutfreund. Bindup: a web server for non-homology-based prediction of dna and rna binding proteins. *Nucleic acids research*, 44(W1):W568–W574, 2016.

[41] Eckhard Jankowsky and Michael E Harris. Specificity and nonspecificity in rna–protein interactions. *Nature reviews Molecular cell biology*, 16(9):533–544, 2015.

[42] Donny D Licatalosi, Aldo Mele, John J Fak, Jernej Ule, Melis Kayikci, Sung Wook Chi, Tyson A Clark, Anthony C Schweitzer, John E Blume, Xuning Wang, et al. Hits-clip yields genome-wide insights into brain alternative rna processing. *Nature*, 456(7221):464–469, 2008.

[43] Julian König, Kathi Zarnack, Gregor Rot, Tomaž Curk, Melis Kayikci, Blaž Zupan, Daniel J Turner, Nicholas M Luscombe, and Jernej Ule. iclip reveals the function of hnrnp particles in splicing at individual nucleotide resolution. *Nature structural & molecular biology*, 17(7):909–915, 2010.

[44] Eric L Van Nostrand, Gabriel A Pratt, Alexander A Shishkin, Chelsea Gelboin-Burkhart, Mark Y Fang, Balaji Sundararaman, Steven M Blue, Thai B Nguyen, Christine Surka, Keri Elkins, et al. Robust transcriptome-wide discovery of rna-binding protein binding sites with enhanced clip (eclip). *Nature methods*, 13(6):508–514, 2016.

[45] Jan Rehwinkel. Rna sensing: the more rig-i the merrier? *EMBO reports*, 14(9):751–752, 2013.

[46] Joseph M Watts, Kristen K Dang, Robert J Gorelick, Christopher W Leonard, Julian W Bess Jr, Ronald Swanstrom, Christina L Burch, and Kevin M Weeks. Architecture and secondary structure of an entire hiv-1 rna genome. *Nature*, 460(7256):711–716, 2009.

[47] Jamal Tazi, Nadia Bakkour, Virginie Marchand, Lilia Ayadi, Amina Aboufirassi, and Christiane Branlant. Alternative splicing: regulation of hiv-1 multiplication as a target for therapeutic action. *The FEBS journal*, 277(4):867–876, 2010.

[48] Marina M O'Reilly, Mark T McNally, and Karen L Beemon. Two strong 5 splice sites and competing, suboptimal 3 splice sites involved in alternative splicing of human immunodeficiency virus type 1 rna. *Virology*, 213(2):373–385, 1995.

[49] Zhi-hai Si, Brad A Amendt, and C Martin Stoltzfus. Splicing efficiency of human immunodeficiency virus type 1 tat rna is determined by both a suboptimal 3

splice site and a 10 nucleotide exon splicing silencer element located within tat exon 2. *Nucleic acids research*, 25(4):861–867, 1997.

[50] Patricia S Bilodeau, Jeffrey K Domsic, Akila Mayeda, Adrian R Krainer, and C Martin Stoltzfus. Rna splicing at human immunodeficiency virus type 1 3 splice site a2 is regulated by binding of hnrnp a/b proteins to an exonic splicing silencer element. *Journal of virology*, 75(18):8487–8497, 2001.

[51] Alan M Zahler, Christian K Damgaard, Jorgen Kjems, and Massimo Caputi. Sc35 and heterogeneous nuclear ribonucleoprotein a/b proteins bind to a juxtaposed exonic splicing enhancer/exonic splicing silencer element to regulate hiv-1 tat exon 2 splicing. *Journal of Biological Chemistry*, 279(11):10077–10084, 2004.

[52] Kathryn Woolaway, Kengo Asai, Andrew Emili, and Alan Cochrane. hnrnp e1 and e2 have distinct roles in modulating hiv-1 gene expression. *Retrovirology*, 4(1):1–18, 2007.

[53] Bing Chen. Molecular mechanism of hiv-1 entry. *Trends in microbiology*, 27(10):878–891, 2019.

[54] Simone Mattei, Bärbel Glass, Wim JH Hagen, Hans-Georg Kräusslich, and John AG Briggs. The structure and flexibility of conical hiv-1 capsids determined within intact virions. *Science*, 354(6318):1434–1437, 2016.

[55] David A Jacques, William A McEwan, Laura Hilditch, Amanda J Price, Greg J Towers, and Leo C James. Hiv-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated dna synthesis. *Nature*, 536(7616):349–353, 2016.

[56] Elena Chertova, Oleg Chertov, Lori V Coren, James D Roser, Charles M Trubey, Julian W Bess, Raymond C Sowder, Eugene Barsov, Brian L Hood, Robert J Fisher, et al. Proteomic and biochemical analysis of purified human immunodeficiency virus type 1 produced from infected monocyte-derived macrophages. *Journal of virology*, 80(18):9039–9052, 2006.

[57] Steven Santos, Yuri Obukhov, Sergei Nekhai, Michael Bukrinsky, and Sergey Iordanskiy. Virus-producing cells determine the host protein profiles of hiv-1 virion cores. *Retrovirology*, 9(1):65, 2012.

[58] Christelle Brégnard, Alessia Zamborlini, Marjorie Leduc, Philippe Chafey, Luc Camoin, Ali Saïb, Serge Benichou, Olivier Danos, and Stéphane Basmaciogullari. Comparative proteomic analysis of hiv-1 particles reveals a role for ezrin and ehd4 in the nef-dependent increase of virus infectivity. *Journal of virology*, 87(7):3729–3740, 2013.

[59] Jérôme Denard, Stéphanie Rundwasser, Nicolas Laroudie, Florence Gonnet, Luigi Naldini, Marina Radrizzani, Anne Galy, Otto-Wilhelm Merten, Olivier Danos, and Fedor Svinartchouk. Quantitative proteomic analysis of lentiviral vectors using 2-de. *Proteomics*, 9(14):3666–3676, 2009.

[60] Andrew CS Saphire, Philippe A Gallay, and Steven J Bark. Proteomic analysis of human immunodeficiency virus using liquid chromatography/tandem mass spectrometry effectively distinguishes specific incorporated host proteins. *Journal of proteome research*, 5(3):530–538, 2006.

[61] David E Ott, Lori V Coren, Donald G Johnson, Bradley P Kane, Raymond C Sowder II, Young D Kim, Robert J Fisher, Xiao Zhen Zhou, Kun Ping Lu, and Louis E Henderson. Actin-binding cellular proteins inside human immunodeficiency virus type 1. *Virology*, 266(1):42–51, 2000.

[62] Nathalie Arhel, Auguste Genovesio, Kyeong Kim, Sarah Miko, Emmanuelle Perret, Jean-Christophe Olivo-Marin, Spencer Shorte, and Pierre Charneau. Quantitative four-dimensional tracking of cytoplasmic and nuclear hiv-1 complexes. *Nature methods*, 3(10):817–824, 2006.

[63] Adarsh Dharan, Silvana Opp, Omar Abdel-Rahim, Sevnur Komurlu Keceli, Sabrina Imam, Felipe Diaz-Griffero, and Edward M Campbell. Bicaudal d2 facilitates the cytoplasmic trafficking and nuclear import of hiv-1 genomes during

infection. *Proceedings of the National Academy of Sciences*, 114(50):E10707–E10716, 2017.

[64] Juliette Fernandez, Débora M Portilho, Anne Danckaert, Sandie Munier, Andreas Becker, Pascal Roux, Anaba Zambo, Spencer Shorte, Yves Jacob, Pierre-Olivier Vidalain, et al. Microtubule-associated proteins 1 (map1) promote human immunodeficiency virus type i (hiv-1) intracytoplasmic routing to the nucleus. *Journal of Biological Chemistry*, 290(8):4631–4646, 2015.

[65] Viacheslav Malikov, Eveline Santos Da Silva, Vladimir Jovasevic, Geoffrey Bennett, Daniel A de Souza Aranha Vieira, Bianca Schulte, Felipe Diaz-Griffero, Derek Walsh, and Mojgan H Naghavi. Hiv-1 capsids bind and exploit the kinesin-1 adaptor fez1 for inward movement to the nucleus. *Nature communications*, 6(1):1–13, 2015.

[66] Masha Sorin, Jennifer Cano, Supratik Das, Sheeba Mathew, Xuhong Wu, Kelvin P Davies, Xuanling Shi, S-W Grace Cheng, David Ott, and Ganjam V Kalpana. Recruitment of a sap18-hdac1 complex into hiv-1 virions and its requirement for viral replication. *PLoS pathogens*, 5(6):e1000463, 2009.

[67] Fadila Larguet, Clément Caté, Benoit Barbeau, Eric Rassart, and Elsy Edouard. Histone deacetylase 1 interacts with hiv-1 integrase and modulates viral replication. *Virology journal*, 16(1):1–17, 2019.

[68] Wei-Shau Hu and Stephen H Hughes. Hiv-1 reverse transcription. *Cold Spring Harbor perspectives in medicine*, 2(10):a006882, 2012.

[69] Youichi Suzuki and Robert Craigie. The road to chromatin—nuclear entry of retroviruses. *Nature Reviews Microbiology*, 5(3):187–196, 2007.

[70] Guillermo Blanco-Rodriguez, Anastasia Gazi, Blandine Monel, Stella Frabetti, Viviana Scoca, Florian Mueller, Olivier Schwartz, Jacomine Krijnse-Locker, Pierre Charneau, and Francesca Di Nunzio. Remodeling of the core leads hiv-1 preintegration complex into the nucleus of human lymphocytes. *Journal of virology*, 94(11):e00135–20, 2020.

[71] Michael D Miller, Chris M Farnet, and Frederic D Bushman. Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. *Journal of virology*, 71(7):5382–5390, 1997.

[72] Ophélie Cosnefroy, Philip J Murray, and Kate N Bishop. Hiv-1 capsid uncoating initiates after the first strand transfer of reverse transcription. *Retrovirology*, 13(1):58, 2016.

[73] Sanela Rankovic, Janani Varadarajan, Ruben Ramalho, Christopher Aiken, and Itay Rousso. Reverse transcription mechanically initiates hiv-1 capsid disassembly. *Journal of virology*, 91(12):e00289–17, 2017.

[74] Ashwanth C Francis and Gregory B Melikyan. Single hiv-1 imaging reveals progression of infection through ca-dependent steps of docking at the nuclear pore, uncoating, and nuclear transport. *Cell host & microbe*, 23(4):536–548, 2018.

[75] Vojtech Zila, Erica Margiotta, Beata Turonova, Thorsten G Müller, Christian E Zimmerli, Simone Mattei, Matteo Allegretti, Kathleen Börner, Jona Rada, Barbara Müller, et al. Cone-shaped hiv-1 capsids are transported through intact nuclear pores. *bioRxiv*, 2020.

[76] Ryan C Burdick, Chenglei Li, MohamedHusen Munshi, Jonathan MO Rawson, Kunio Nagashima, Wei-Shau Hu, and Vinay K Pathak. Hiv-1 uncoats in the nucleus near sites of integration. *Proceedings of the National Academy of Sciences*, 117(10):5486–5493, 2020.

[77] Adarsh Dharan, Niklas Bachmann, Sarah Talley, Virginia Zwikelmaier, and Edward M Campbell. Nuclear pore blockade reveals that hiv-1 completes reverse transcription and uncoating in the nucleus. *Nature Microbiology*, pages 1–8, 2020.

[78] Ryan C Burdick, Krista A Delviks-Frankenberry, Jianbo Chen, Sanath K Janaka, Jaya Sastri, Wei-Shau Hu, and Vinay K Pathak. Dynamics and regulation of

nuclear import and nuclear movements of hiv-1 complexes. *PLoS pathogens*, 13(8):e1006570, 2017.

[79] David Alejandro Bejarano, Ke Peng, Vibor Laketa, Kathleen Börner, K Laurence Jost, Bojana Lucic, Bärbel Glass, Marina Lusic, Barbara Müller, and Hans-Georg Kräusslich. Hiv-1 nuclear import in macrophages is regulated by cpsf6-capsid interactions at the nuclear pore complex. *Elife*, 8:e41800, 2019.

[80] Elena Rensen, Florian Mueller, Viviana Scoca, Jyotsana Parmar, Philippe Souque, Christophe Zimmer, and Francesca Di Nunzio. Clustering and reverse transcription of hiv-1 genomes in nuclear niches of macrophages. *bioRxiv*, 2020.

[81] Kenneth A Matreyek and Alan Engelman. Viral and cellular requirements for the nuclear entry of retroviral preintegration nucleoprotein complexes. *Viruses*, 5(10):2483–2511, 2013.

[82] Gary P Wang, Angela Ciuffi, Jeremy Leipzig, Charles C Berry, and Frederic D Bushman. Hiv integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome research*, 17(8):1186–1194, 2007.

[83] Erik Serrao, Lavanya Krishnan, Ming-Chieh Shun, Xiang Li, Peter Cherepanov, Alan Engelman, and Goedele N Maertens. Integrase residues that determine nucleotide preferences at sites of hiv-1 integration: implications for the mechanism of target dna binding. *Nucleic acids research*, 42(8):5164–5176, 2014.

[84] Astrid RW Schröder, Paul Shinn, Huaming Chen, Charles Berry, Joseph R Ecker, and Frederic Bushman. Hiv-1 integration in the human genome favors active genes and local hotspots. *Cell*, 110(4):521–529, 2002.

[85] Jun Katahira. Nuclear export of messenger rna. *Genes*, 6(2):163–184, 2015.

[86] Matthew D Daugherty, David S Booth, Bhargavi Jayaraman, Yifan Cheng, and Alan D Frankel. Hiv rev response element (rre) directs assembly of the rev

homooligomer into discrete asymmetric complexes. *Proceedings of the National Academy of Sciences*, 107(28):12481–12486, 2010.

[87] Jianbo Chen, Chijioke Umunnakwe, David Q Sun, Olga A Nikolaitchik, Vinay K Pathak, Ben Berkhout, Atze T Das, and Wei-Shau Hu. Impact of nuclear export pathway on cytoplasmic hiv-1 rna transport mechanism and distribution. *Mbio*, 11(6):e01578–20, 2020.

[88] Jianbo Chen, Yang Liu, Bin Wu, Olga A Nikolaitchik, Preeti R Mohan, Jiji Chen, Vinay K Pathak, and Wei-Shau Hu. Visualizing the translation and packaging of hiv-1 full-length rna. *Proceedings of the National Academy of Sciences*, 117(11):6145–6155, 2020.

[89] Olga A Nikolaitchik, Kari A Dilley, William Fu, Robert J Gorelick, S-H Sheldon Tai, Ferri Soheilian, Roger G Ptak, Kunio Nagashima, Vinay K Pathak, and Wei-Shau Hu. Dimeric rna recognition regulates hiv-1 genome packaging. *PLoS pathogens*, 9(3):e1003249, 2013.

[90] Emma Poole, Padraig Strappe, Hoi-Ping Mok, Ray Hicks, and Andrew M L Lever. Hiv-1 gag–rna interaction occurs at a perinuclear/centrosomal site; analysis by confocal microscopy and fret. *Traffic*, 6(9):741–755, 2005.

[91] Jamil S Saad, Jaime Miller, Janet Tai, Andrew Kim, Ruba H Ghanam, and Michael F Summers. Structural basis for targeting hiv-1 gag proteins to the plasma membrane for virus assembly. *Proceedings of the National Academy of Sciences*, 103(30):11364–11369, 2006.

[92] Eric O Freed. Hiv-1 assembly, release and maturation. *Nature Reviews Microbiology*, 13(8):484–496, 2015.

[93] Christopher Sumner and Akira Ono. Relationship between hiv-1 gag multimerization and membrane binding. *Viruses*, 14(3):622, 2022.

[94] Mary Ann Checkley, Benjamin G Luttge, and Eric O Freed. Hiv-1 envelope glycoprotein biosynthesis, trafficking, and incorporation. *Journal of molecular biology*, 410(4):582–608, 2011.

[95] Ali Tavassoli, Quan Lu, Jongsik Gam, Hui Pan, Stephen J Benkovic, and Stanley N Cohen. Inhibition of hiv budding by a genetically selected cyclic peptide targeting the gag- tsg101 interaction. *ACS chemical biology*, 3(12):757–764, 2008.

[96] Pei-I Ku, Mourad Bendjennat, Jeff Ballew, Michael B Landesman, and Saveez Saffarian. Alix is recruited temporarily into hiv-1 budding sites at the end of gag assembly. *PLoS One*, 9(5):e96950, 2014.

[97] Eiji Morita, Virginie Sandrin, John McCullough, Angela Katsuyama, Ira Baci Hamilton, and Wesley I Sundquist. Escrt-iii protein requirements for hiv-1 budding. *Cell host & microbe*, 9(3):235–242, 2011.

[98] Benedikt M Beckmann, Alfredo Castello, and Jan Medenbach. The expanding universe of ribonucleoproteins: of novel rna-binding proteins and unconventional interactions. *Pflügers Archiv-European Journal of Physiology*, 468(6):1029–1040, 2016.

[99] Anna Kristina P Serquiña, Suman R Das, Elena Popova, Ogooluwa A Ojelabi, Christian K Roy, and Heinrich G Göttlinger. Upf1 is crucial for the infectivity of human immunodeficiency virus type 1 progeny virions. *Journal of virology*, 87(16):8853–8861, 2013.

[100] Caroline Weydert, Bart van Heertum, Lieve Dirix, Stephanie De Houwer, Flore De Wit, Jan Mast, Steven J Husson, Katrien Busschots, Renate Koenig, Rik Gijsbers, et al. Y-box-binding protein 1 supports the early and late steps of hiv replication. *PloS one*, 13(7), 2018.

[101] Samantha Brady, Gatikrushna Singh, Cheryl Bolinger, Zhenwei Song, Ioana Boeras, Kexin Weng, Bria Trent, William Clay Brown, Kamalendra Singh,

Kathleen Boris-Lawrie, et al. Virion-associated, host-derived dhx9/rna helicase a enhances the processivity of hiv-1 reverse transcriptase on genomic rna. *Journal of Biological Chemistry*, 294(30):11473–11485, 2019.

[102] Daniel J Rawle, Dongsheng Li, Joakim E Swedberg, Lu Wang, Dinesh C Soares, and David Harrich. Hiv-1 uncoating and reverse transcription require eef1a binding to surface-exposed acidic residues of the reverse transcriptase thumb domain. *MBio*, 9(2):e00316–18, 2018.

[103] Dongsheng Li, Ting Wei, Daniel J Rawle, Fangyun Qin, Rui Wang, Dinesh C Soares, Hongping Jin, Haran Sivakumaran, Min-Hsuan Lin, Kirsten Spann, et al. Specific interaction between eef1a and hiv rt is critical for hiv-1 reverse transcription and a potential anti-hiv target. *PLoS pathogens*, 11(12), 2015.

[104] Alberto De Iaco and Jeremy Luban. Cyclophilin a promotes hiv-1 reverse transcription but its effect on transduction correlates best with its effect on nuclear entry of viral cdna. *Retrovirology*, 11(1):11, 2014.

[105] Danni Jin and Karin Musier-Forsyth. Role of host trnas and aminoacyl-trna synthetases in retroviral replication. *Journal of Biological Chemistry*, 294(14):5352–5364, 2019.

[106] Ryan Burdick, Jessica L Smith, Chawaree Chaipan, Yeshitila Friew, Jianbo Chen, Narasimhan J Venkatachari, Krista A Delviks-Frankenberry, Wei-Shau Hu, and Vinay K Pathak. P body-associated protein mov10 inhibits hiv-1 replication at multiple stages. *Journal of virology*, 84(19):10241–10253, 2010.

[107] Kumkum Rahman Mouree, Naoki Kishimoto, Nozomi Iga, Chie Kirihara, Kengo Yamamoto, Nobutoki Takamune, and Shogo Misumi. Virion-packaged pyruvate kinase muscle type 2 affects reverse transcription efficiency of human immunodeficiency virus type 1 by blocking virion recruitment of trnalys3. *Biological and Pharmaceutical Bulletin*, 41(4):612–618, 2018.

[108] Naoki Kishimoto, Nozomi Iga, Kengo Yamamoto, Nobutoki Takamune, and Shogo Misumi. Virion-incorporated alpha-enolase suppresses the early

stage of hiv-1 reverse transcription. *Biochemical and biophysical research communications*, 484(2):278–284, 2017.

[109] Wael Kamel, Marko Noerenberg, Berati Cerikan, Honglin Chen, Aino I Järvelin, Mohamed Kammoun, Jeffrey Y Lee, Ni Shuai, Manuel Garcia-Moreno, Anna Andrejeva, et al. Global analysis of protein-rna interactions in sars-cov-2-infected cells reveals key regulators of infection. *Molecular cell*, 81(13):2851–2867, 2021.

[110] Honglin Zhou, Min Xu, Qian Huang, Adam T Gates, Xiaohua D Zhang, John C Castle, Erica Stec, Marc Ferrer, Berta Strulovici, Daria J Hazuda, et al. Genome-scale rnai screen for host factors required for hiv replication. *Cell host & microbe*, 4(5):495–504, 2008.

[111] Man Lung Yeung, Laurent Houzet, Venkat SRK Yedavalli, and Kuan-Teh Jeang. A genome-wide short hairpin rna screening of jurkat t-cells for human proteins contributing to productive hiv-1 replication. *Journal of Biological Chemistry*, 284(29):19463–19473, 2009.

[112] Renate König, Yingyao Zhou, Daniel Elleder, Tracy L Diamond, Ghislain MC Bonamy, Jeffrey T Irelan, Chih-yuan Chiang, Buu P Tu, Paul D De Jesus, Caroline E Lilley, et al. Global analysis of host-pathogen interactions that regulate early-stage hiv-1 replication. *Cell*, 135(1):49–60, 2008.

[113] Abraham L Brass, Derek M Dykxhoorn, Yair Benita, Nan Yan, Alan Engelman, Ramnik J Xavier, Judy Lieberman, and Stephen J Elledge. Identification of host proteins required for hiv infection through a functional genomic screen. *science*, 319(5865):921–926, 2008.

[114] Anna Kula, Jessica Guerra, Anna Knezevich, Danijela Kleva, Michael P Myers, and Alessandro Marcello. Characterization of the hiv-1 rna associated proteome identifies matrin 3 as a nuclear cofactor of rev function. *Retrovirology*, 8(1):1–15, 2011.

[115] Virginie Marchand, Maryline Santerre, Christelle Aigueperse, Laetitia Fouillen, Jean-Michel Saliou, Alain Van Dorsselaer, Sarah Sanglier-Cianferani, Christiane Branlant, and Yuri Motorin. Identification of protein partners of the human immunodeficiency virus 1 tat/rev exon 3 leads to the discovery of a new hiv-1 splicing regulator, protein hnrnp k. *RNA biology*, 8(2):325–342, 2011.

[116] Rachel A Knoener, Jordan T Becker, Mark Scalf, Nathan M Sherer, and Lloyd M Smith. Elucidating the in vivo interactome of hiv-1 rna by hybridization capture and mass spectrometry. *Scientific reports*, 7(1):1–16, 2017.

[117] Rachel Knoener, Edward Evans III, Jordan T Becker, Mark Scalf, Bayleigh Benner, Nathan M Sherer, and Lloyd M Smith. Identification of host proteins differentially associated with hiv-1 rna splice variants. *Elife*, 10, 2021.

[118] Reuben S Harris and Jaquelin P Dudley. Apobecs and virus restriction. *Virology*, 479:131–145, 2015.

[119] Madison B Adolph, Anjuman Ara, and Linda Chelico. Apobec3 host restriction factors of hiv-1 can change the template switching frequency of reverse transcriptase. *Journal of molecular biology*, 431(7):1339–1352, 2019.

[120] Sebla B Kutluay and Paul D Bieniasz. Analysis of hiv-1 gag-rna interactions in cells and virions by clip-seq. *HIV Protocols*, pages 119–131, 2016.

[121] Pooja Yadav, Souvik Sur, Dipen Desai, Smita Kulkarni, Vartika Sharma, and Vibha Tandon. Interaction of hiv-1 integrase with polypyrimidine tract binding protein and associated splicing factor (psf) and its impact on hiv-1 replication. *Retrovirology*, 16:1–18, 2019.

[122] Lisa Muniz, Sylvain Egloff, Bettina Ughy, Beáta E Jády, and Tamás Kiss. Controlling cellular p-tefb activity by the hiv-1 transcriptional transactivator tat. *PLoS pathogens*, 6(10):e1001152, 2010.

[123] Brian J Krueger, Katayoun Varzavand, Jeffrey J Cooper, and David H Price. The mechanism of release of p-tefb and hexim1 from the 7sk snrnp by viral

and cellular activators includes a conformational change in 7sk. *PloS one*, 5(8):e12335, 2010.

[124] Venkat SRK Yedavalli, Christine Neuveut, Ya-hui Chi, Lawrence Kleiman, and Kuan-Teh Jeang. Requirement of ddx3 dead box rna helicase for hiv-1 rev-rre export function. *Cell*, 119(3):381–392, 2004.

[125] Sylvain de Breyne, Nathalie Chamond, Didier Decimo, Mary-Anne Trabaud, Patrice Andre, Bruno Sargueil, and Théophile Ohlmann. In vitro studies reveal that different modes of initiation on hiv-1 mrna have different levels of requirement for eukaryotic initiation factor 4f. *The FEBS journal*, 279(17):3098–3111, 2012.

[126] Nicolas Locker, Nathalie Chamond, and Bruno Sargueil. A conserved structure within the hiv gag open reading frame that controls translation initiation directly recruits the 40s subunit and eif3. *Nucleic acids research*, 39(6):2367–2377, 2011.

[127] Guerline Clerzius, Jean-François Gélinas, and Anne Gatignol. Multiple levels of pkr inhibition during hiv-1 replication. *Reviews in medical virology*, 21(1):42–53, 2011.

[128] Viraj R Sanghvi and Laura F Steel. The cellular tar rna binding protein, trbp, promotes hiv-1 replication primarily by inhibiting the activation of double-stranded rna-dependent kinase pkr. *Journal of virology*, 85(23):12614–12621, 2011.

[129] Laurent Chatel-Chaix, Jean-Francois Clément, Catherine Martel, Véronique Bériault, Anne Gatignol, Luc DesGroseillers, and Andrew J Mouland. Identification of staufen in the human immunodeficiency virus type 1 gag ribonucleoprotein complex and a role in generating infectious viral particles. *Molecular and cellular biology*, 24(7):2637–2648, 2004.

[130] Levon G Abrahamyan, Laurent Chatel-Chaix, Lara Ajamian, Miroslav P Milev, Anne Monette, Jean-François Clément, Rujun Song, Martin Lehmann, Luc

DesGroseillers, Michael Laughrea, et al. Novel staufen1 ribonucleoproteins prevent formation of stress granules but favour encapsidation of hiv-1 genomic rna. *Journal of cell science*, 123(3):369–383, 2010.

[131] Miroslav P Milev, Mukunthan Ravichandran, Morgan F Khan, David C Schriemer, and Andrew J Mouland. Characterization of staufen1 ribonucleoproteins by mass spectrometry and biochemical analyses reveal the presence of diverse host proteins associated with human immunodeficiency virus type 1. *Frontiers in microbiology*, 3:367, 2012.

[132] Cagan Gurer, Andrea Cimarelli, and Jeremy Luban. Specific incorporation of heat shock protein 70 family members into primate lentiviral virions. *Journal of virology*, 76(9):4666–4670, 2002.

[133] Sergei Popov, Elena Popova, Michio Inoue, and Heinrich G Gottlinger. Human immunodeficiency virus type 1 gag engages the bro1 domain of alix/aip1 through the nucleocapsid. *Journal of virology*, 82(3):1389–1398, 2008.

[134] Paola Sette, Vincent Dussupt, and Fadila Bouamr. Identification of the hiv-1 nc binding interface in alix bro1 reveals a role for rna. *Journal of virology*, 86(21):11608–11615, 2012.

[135] Manuel Garcia-Moreno, Robin Truman, Honglin Chen, Louisa Iselin, Caroline E Lenz, Jeff Lee, Kate Dicker, Marko Noerenberg, Thibault JM Sohier, Natasha Palmalux, et al. Incorporation of genome-bound cellular proteins into hiv-1 particles regulates viral infection. *bioRxiv*, pages 2023–06, 2023.

[136] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.

[137] Mihaly Varadi, Stephen Anyango, Mandar Deshpande, Sreenath Nair, Cindy Natassia, Galabina Yordanova, David Yuan, Oana Stroe, Gemma Wood, Agata

Laydon, et al. Alphafold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic acids research*, 50(D1):D439–D444, 2022.

[138] Paul Shannon, Andrew Markiel, Owen Ozier, Nitin S Baliga, Jonathan T Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*, 13(11):2498–2504, 2003.

[139] Huaiyu Mi, Anushya Muruganujan, Dustin Ebert, Xiaosong Huang, and Paul D Thomas. Panther version 14: more genomes, a new panther go-slim and improvements in enrichment analysis tools. *Nucleic acids research*, 47(D1):D419–D426, 2019.

[140] Chris Allan, Jean-Marie Burel, Josh Moore, Colin Blackburn, Melissa Linkert, Scott Loynton, Donald MacDonald, William J Moore, Carlos Neves, Andrew Patterson, et al. Omero: flexible, model-driven data management for experimental biology. *Nature methods*, 9(3):245–253, 2012.

[141] Schrödinger, LLC. The PyMOL molecular graphics system, version 1.8. November 2015.

[142] Damian Szklarczyk, Andrea Franceschini, Stefan Wyder, Kristoffer Forslund, Davide Heller, Jaime Huerta-Cepas, Milan Simonovic, Alexander Roth, Alberto Santos, Kalliopi P Tsafou, et al. String v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic acids research*, 43(D1):D447–D452, 2015.

[143] Robert H Kutner, Xian-Yang Zhang, and Jakob Reiser. Production, concentration and titration of pseudotyped hiv-1-based lentiviral vectors. *Nature protocols*, 4(4):495, 2009.

[144] John L Spouge. Statistical analysis of sparse infection data and its implications for retroviral treatment trials in primates. *Proceedings of the National Academy of Sciences*, 89(16):7581–7585, 1992.

[145] Imre Gaspar, Frank Wippich, and Anne Ephrussi. Enzymatic production of single-molecule fish and rna capture probes. *Rna*, 23(10):1582–1591, 2017.

[146] Arthur Imbert, Wei Ouyang, Adham Safieddine, Emeline Coleno, Christophe Zimmer, Edouard Bertrand, Thomas Walter, and Florian Mueller. Fish-quant v2: a scalable and modular tool for smfish image analysis. *Rna*, 28(6):786–795, 2022.

[147] Gregory Berger, Stéphanie Durand, Guillaume Fargier, Xuan-Nhi Nguyen, Stéphanie Cordeil, Serge Bouaziz, Delphine Muriaux, Jean-Luc Darlix, and Andrea Cimarelli. Apobec3a is a specific inhibitor of the early phases of hiv-1 infection in myeloid cells. *PLoS pathogens*, 7(9):e1002221, 2011.

[148] Alfredo Castello, Rastislav Horos, Claudia Strein, Bernd Fischer, Katrin Eichelbaum, Lars M Steinmetz, Jeroen Krijgsveld, and Matthias W Hentze. System-wide identification of rna-binding proteins by interactome capture. *Nature protocols*, 8(3):491–500, 2013.

[149] Illya G Pashev, Stephan I Dimitrov, and Dimitar Angelov. Crosslinking proteins to nucleic acids by ultraviolet laser irradiation. *Trends in biochemical sciences*, 16:323–326, 1991.

[150] Todd A Gray, Lidia Hernandez, Alisoun H Carey, Melanie A Schaldach, Megan J Smithwick, Kristina Rus, Jennifer A Marshall Graves, Colin L Stewart, and Robert D Nicholls. The ancient source of a distinct gene family encoding proteins featuring ring and c3h zinc-finger motifs with abundant expression in developing brain and nervous system. *Genomics*, 66(1):76–86, 2000.

[151] Hyunji Han, Sehyun Chae, Daehee Hwang, and Jaewhan Song. Attenuating mkrn1 e3 ligase-mediated ampkα suppression increases tolerance against metabolic stresses in mice. *Cell stress*, 2(11):325, 2018.

[152] Eun-Woo Lee, Min-Sik Lee, Suzanne Camus, Jaewang Ghim, Mi-Ran Yang, Wonkyung Oh, Nam-Chul Ha, David P Lane, and Jaewhan Song. Differential

regulation of p53 and p21 by mkrn1 e3 ligase controls cell cycle arrest and apoptosis. *The EMBO journal*, 28(14):2100–2113, 2009.

[153] Aram Ko, Ji-Young Shin, Jinho Seo, Kang-Duck Lee, Eun-Woo Lee, Min-Sik Lee, Han-Woong Lee, Il-Ju Choi, Jin Sook Jeong, Kyung-Hee Chun, et al. Acceleration of gastric tumorigenesis through mkrn1-mediated posttranslational regulation of p14arf. *Journal of the National Cancer Institute*, 104(21):1660–1672, 2012.

[154] Min-Sik Lee, Man-Hyung Jeong, Hyun-Woo Lee, Hyun-Ji Han, Aram Ko, Stephen M Hewitt, Jae-Hoon Kim, Kyung-Hee Chun, Joon-Yong Chung, Cheolju Lee, et al. Pi3k/akt activation induces pten ubiquitination and destabilization accelerating tumourigenesis. *Nature communications*, 6(1):1–14, 2015.

[155] Jun Hyun Kim, Sun-Mi Park, Mi Ran Kang, Sue-Young Oh, Tae H Lee, Mark T Muller, and In Kwon Chung. Ubiquitin ligase mkrn1 modulates telomere length homeostasis through a proteolysis of htert. *Genes & development*, 19(7):776–781, 2005.

[156] Jose Salvatico, Joo Hee Kim, In Kwon Chung, and Mark T Muller. Differentiation linked regulation of telomerase activity by makorin-1. *Molecular and cellular biochemistry*, 342(1):241–250, 2010.

[157] Josephat Omwancha, Xiao-Feng Zhou, Shao-Yong Chen, Timour Baslan, Christopher J Fisher, Zhe Zheng, Changmeng Cai, and Lirim Shemshedini. Makorin ring finger protein 1 (mkrn1) has negative and positive effects on rna polymerase ii-dependent transcription. *Endocrine*, 29(2):363–373, 2006.

[158] Paul A Cassar, Richard L Carpenedo, Payman Samavarchi-Tehrani, Jonathan B Olsen, Chang Jun Park, Wing Y Chang, Zhaoyi Chen, Chandarong Choey, Sean Delaney, Huishan Guo, et al. Integrative genomics positions mkrn1 as a novel ribonucleoprotein within the embryonic stem cell gene regulatory network. *EMBO reports*, 16(10):1334–1357, 2015.

[159] Hatmone Miroci, Claudia Schob, Stefan Kindler, Janin Ölschläger-Schütt, Susanne Fehr, Tassilo Jungenitz, Stephan W Schwarzacher, Claudia Bagni, and Evita Mohr. Makorin ring zinc finger protein 1 (mkrn1), a novel poly (a)-binding protein-interacting protein, stimulates translation in nerve cells. *Journal of Biological Chemistry*, 287(2):1322–1334, 2012.

[160] Andrea Hildebrandt, Mirko Brüggemann, Cornelia Rücklé, Susan Boerner, Jan B Heidelberger, Anke Busch, Heike Hänel, Andrea Voigt, Martin M Möckel, Stefanie Ebersberger, et al. The rna-binding ubiquitin ligase mkrn1 functions in ribosome-associated quality control of poly (a) translation. *Genome biology*, 20(1):1–20, 2019.

[161] Tongtong Wang, Qian Du, Xingchen Wu, Yingying Niu, Lijuan Guan, Zhenyu Wang, Xiaomin Zhao, Shan-Lu Liu, Dewen Tong, and Yong Huang. Porcine mkrn1 modulates the replication and pathogenesis of porcine circovirus type 2 by inducing capsid protein ubiquitination and degradation. *Journal of virology*, 92(11):e00100–18, 2018.

[162] Aram Ko, Eun-Woo Lee, Jung-Yong Yeh, Mi-Ran Yang, Wonkyung Oh, Jin-San Moon, and Jaewhan Song. Mkrn1 induces degradation of west nile virus capsid protein by functioning as an e3 ligase. *Journal of virology*, 84(1):426–436, 2010.

[163] Meenu Subrahmanian, Jeya Marimuthu, Thiagarajan Sairam, and Ramalingam Sankaran. In vitro ubiquitination of mycobacterium tuberculosis by e3 ubiquitin ligase, mkrn1. *Biotechnology Letters*, 42(8):1527–1534, 2020.

[164] Raviteja Inturi, Kwangchol Mun, Katrin Singethan, Sabrina Schreiner, and Tanel Punga. Human adenovirus infection causes cellular e3 ubiquitin ligase mkrn1 degradation involving the viral core protein pvii. *Journal of virology*, 92(3):e01154–17, 2018.

[165] Huifang Lv, Wang Dong, Zhi Cao, Jihui Lin, Yueling Ouyang, Kangkang Guo, Cunfa Li, and Yanming Zhang. Classical swine fever virus non-structural protein 4b binds tank-binding kinase 1. *Journal of biosciences*, 43(5):947–957, 2018.

[166] Adam J Fletcher, Marina Vaysburd, Sarah Maslen, Jingwei Zeng, J Mark Skehel, Greg J Towers, and Leo C James. Trivalent ring assembly on retroviral capsids activates trim5 ubiquitination and innate immune signaling. *Cell host & microbe*, 24(6):761–775, 2018.

[167] Sergei Popov, Elena Popova, Michio Inoue, Yuanfei Wu, and Heinrich Göttlinger. Hiv-1 gag recruits pacsin2 to promote virus spreading. *Proceedings of the National Academy of Sciences*, 115(27):7093–7098, 2018.

[168] Sakshi Arora, Sachin Verma, and Akhil C Banerjea. Hiv-1 vpr redirects host ubiquitination pathway. *Journal of virology*, 88(16):9141–9152, 2014.

[169] Vanessa Brès, Rosemary E Kiernan, Laetitia K Linares, Christine Chable-Bessia, Olga Plechakova, Céline Tréand, Stephane Emiliani, Jean-Marie Peloponese, Kuan-Teh Jeang, Olivier Coux, et al. A non-proteolytic role for ubiquitin in tat-mediated transactivation of the hiv-1 promoter. *Nature cell biology*, 5(8):754–761, 2003.

[170] Dohun Pyeon, Vivian K Rojas, Lenore Price, Seongcheol Kim, Meharvan Singh, and In-Woo Park. Hiv-1 impairment via ube3a and hiv-1 nef interactions utilizing the ubiquitin proteasome system. *Viruses*, 11(12):1098, 2019.

[171] David E Ott, Lori V Coren, Terry D Copeland, Bradley P Kane, Donald G Johnson, Raymond C Sowder, Yoshiyuki Yoshinaka, Stephen Oroszlan, Larry O Arthur, and Louis E Henderson. Ubiquitin is covalently attached to the p6gagproteins of human immunodeficiency virus type 1 and simian immunodeficiency virus and to the p12gag protein of moloney murine leukemia virus. *Journal of Virology*, 72(4):2962–2968, 1998.

[172] Rubén A Bartolomé, Irene García-Palmero, Sofía Torres, María López-Lucendo, Irina V Balyasnikova, and J Ignacio Casal. Il13 receptor α2 signaling requires a scaffold protein, fam120a, to activate the fak and pi3k pathways in colon cancer metastasis. *Cancer research*, 75(12):2434–2444, 2015.

[173] Masamitsu Tanaka, Kazuki Sasaki, Reiko Kamata, Yukari Hoshino, Kazuyoshi Yanagihara, and Ryuichi Sakai. A novel rna-binding protein, ossa/c9orf10, regulates activity of src kinases to protect cells from oxidative stress-induced apoptosis. *Molecular and cellular biology*, 29(2):402–413, 2009.

[174] Timothy J Kelly, Hiroshi I Suzuki, Jesse R Zamudio, Megumu Suzuki, and Phillip A Sharp. Sequestration of microrna-mediated target repression by the ago2-associated rna-binding protein fam120a. *RNA*, 25(10):1291–1297, 2019.

[175] Jialing Huang, Fengxiang Wang, Elias Argyris, Keyang Chen, Zhihui Liang, Heng Tian, Wenlin Huang, Kathleen Squires, Gwen Verlinghieri, and Hui Zhang. Cellular microrna contribute to hiv-1 latency in resting primary cd4+ t lymphocytes. *Nature medicine*, 13(10):1241–1247, 2007.

[176] Robin Nathans, Chia-ying Chu, Anna Kristina Serquina, Chih-Chung Lu, Hong Cao, and Tariq M Rana. Cellular microrna and p bodies modulate host-hiv-1 interactions. *Molecular cell*, 34(6):696–709, 2009.

[177] Robinson Triboulet, Bernard Mari, Yea-Lih Lin, Christine Chable-Bessia, Yamina Bennasser, Kevin Lebrigand, Bruno Cardinaud, Thomas Maurin, Pascal Barbry, Vincent Baillat, et al. Suppression of microrna-silencing pathway by hiv-1 during virus replication. *Science*, 315(5818):1579–1582, 2007.

[178] Karen Chiang, Tzu-Ling Sung, and Andrew P Rice. Regulation of cyclin t1 and hiv-1 replication by microrna in resting cd4+ t lymphocytes. *Journal of virology*, 86(6):3244–3252, 2012.

[179] Tzu-Ling Sung and Andrew P Rice. mir-198 inhibits hiv-1 gene expression and replication in monocytes and its mechanism of action appears to involve repression of cyclin t1. *PLoS pathogens*, 5(1):e1000263, 2009.

[180] Kate Hsu, Jegatheesan Seharaseyon, Peihong Dong, Stephan Bour, and Eduardo Marbán. Mutual functional destruction of hiv-1 vpu and host task-1 channel. *Molecular cell*, 14(2):259–267, 2004.

[181] Huabiao Chen, Chun Li, Jinghe Huang, Thai Cung, Katherine Seiss, Jill Beamon, Mary F Carrington, Lindsay C Porter, Patrick S Burke, Yue Yang, et al. Cd4+ t cells from elite controllers resist hiv-1 infection by selective upregulation of p21. *The Journal of clinical investigation*, 121(4):1549–1560, 2011.

[182] Martin Lehmann, Miroslav P Milev, Levon Abrahamyan, Xiao-Jian Yao, Nelly Pante, and Andrew J Mouland. Intracellular transport of human immunodeficiency virus type 1 genomic rna and viral production are dependent on dynein motor function and late endosome positioning. *Journal of Biological Chemistry*, 284(21):14572–14585, 2009.

[183] WH Reeves, S Narain, and M Satoh. Henry kunkel, stephanie smith, clinical immunology, and split genes. *Lupus*, 12(3):213–217, 2003.

[184] Natalia V Lekontseva, Elena A Stolboushkina, and Alexey D Nikulin. Diversity of lsm family proteins: similarities and differences. *Biochemistry (Moscow)*, 86(1):S38–S49, 2021.

[185] Wei-hong Yang, Jiang Hong Yu, Tod Gulick, Kenneth D Bloch, and Donald B Bloch. Rna-associated protein 55 (rap55) localizes to mrna processing bodies and stress granules. *Rna*, 12(4):547–554, 2006.

[186] Xiao-Man Liu, Liang Ma, and Randy Schekman. Selective sorting of microrna into exosomes by phase-separated ybx1 condensates. *Elife*, 10:e71982, 2021.

[187] Danae Campos-Melo, Zachary CE Hawley, Cristian A Droppelmann, and Michael J Strong. The integral role of rna in stress granule formation and function. *Frontiers in Cell and Developmental Biology*, 9:621779, 2021.

[188] Bobo Wing-Yee Mok, Wenjun Song, Pui Wang, Hung Tai, Yixin Chen, Min Zheng, Xi Wen, Siu-Ying Lau, Wai Lan Wu, Ken Matsumoto, et al. The ns1 protein of influenza a virus interacts with cellular processing bodies and stress granules through rna-associated protein 55 (rap55) during virus infection. *Journal of virology*, 86(23):12695–12707, 2012.

[189] Ying Li, Rui Chen, Qian Zhou, Zhisheng Xu, Chao Li, Shuai Wang, Aiping Mao, Xiaodong Zhang, Weiwu He, and Hong-Bing Shu. Lsm14a is a processing body-associated sensor of viral nucleic acids that initiates cellular antiviral response in the early phase of viral infection. *Proceedings of the National Academy of Sciences*, 109(29):11770–11775, 2012.

[190] Tian-Tian Liu, Qing Yang, Mi Li, Bo Zhong, Yong Ran, Li-Li Liu, Yan Yang, Yan-Yi Wang, and Hong-Bing Shu. Lsm14a plays a critical role in antiviral immune responses by regulating mita level in a cell-specific manner. *The Journal of Immunology*, 196(12):5101–5111, 2016.

[191] Matthew A Pearson, David Reczek, Anthony Bretscher, and P Andrew Karplus. Structure of the erm protein moesin reveals the ferm domain fold masked by an extended actin binding tail domain. *Cell*, 101(3):259–270, 2000.

[192] Ildikó Kristó, Csaba Bajusz, Barbara N Borsos, Tibor Pankotai, Joseph Dopie, Ferenc Jankovics, Maria K Vartiainen, Miklós Erdélyi, and Péter Vilmos. The actin binding cytoskeletal protein moesin is involved in nuclear mrna export. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1864(10):1589–1604, 2017.

[193] Csaba Bajusz, Ildikó Kristó, Csilla Abonyi, Tomáš Venit, Viktor Vedelek, Tamás Lukácsovich, Attila Farkas, Péter Borkúti, Zoltán Kovács, Izabella Bajusz, et al. The nuclear activity of the actin-binding moesin protein is necessary for gene expression in drosophila. *The FEBS Journal*, 288(16):4812–4832, 2021.

[194] Shirsendu Ghosh, Vincenzo Di Bartolo, Liron Tubul, Eyal Shimoni, Elena Kartvelishvily, Tali Dadosh, Sara W Feigelson, Ronen Alon, Andres Alcover, and Gilad Haran. Erm-dependent assembly of t cell receptor signaling and co-stimulatory molecules on microvilli prior to activation. *Cell reports*, 30(10):3434–3447, 2020.

[195] Marta Barrero-Villar, José Román Cabrero, Mónica Gordón-Alonso, Jonathan Barroso-González, Susana Álvarez-Losada, M Ángeles Muñoz-Fernández,

Francisco Sánchez-Madrid, and Agustín Valenzuela-Fernández. Moesin is required for hiv-1-induced cd4-cxcr4 interaction, f-actin redistribution, membrane fusion and viral infection in lymphocytes. *Journal of cell science*, 122(1):103–113, 2009.

[196] Emily JH Chen, Meredith H Shaffer, Edward K Williamson, Yanping Huang, and Janis K Burkhardt. Ezrin and moesin are required for efficient t cell adhesion and homing to lymphoid organs. *PloS one*, 8(2):e52368, 2013.

[197] Juan M Serrador, Marta Nieto, José L Alonso-Lebrero, Miguel A del Pozo, Javier Calvo, Heinz Furthmayr, Reinhard Schwartz-Albiez, Francisco Lozano, Roberto Gonzalez-Amaro, Paloma Sanchez-Mateos, et al. Cd43 interacts with moesin and ezrin and regulates its redistribution to the uropods of t lymphocytes at the cell-cell contacts. *Blood, The Journal of the American Society of Hematology*, 91(12):4632–4644, 1998.

[198] M Stevenson, TL Stanwick, MP Dempsey, and CA Lamonica. Hiv-1 replication is controlled at the level of t cell activation and proviral integration. *The EMBO journal*, 9(5):1551–1560, 1990.

[199] Mojgan H Naghavi, Susana Valente, Theodora Hatziioannou, Kenia De Los Santos, Ying Wen, Christina Mott, Gregg G Gundersen, and Stephen P Goff. Moesin regulates stable microtubule formation and limits retroviral infection in cultured cells. *The EMBO journal*, 26(1):41–52, 2007.

[200] Yoshinao Kubo, Hiroaki Yoshii, Haruka Kamiyama, Chika Tominaga, Yuetsu Tanaka, Hironori Sato, and Naoki Yamamoto. Ezrin, radixin, and moesin (erm) proteins function as pleiotropic regulators of human immunodeficiency virus type 1 infection. *Virology*, 375(1):130–140, 2008.

[201] Steven J Anderson, Steven D Levin, and Roger M Perlmutter. Involvement of the protein tyrosine kinase p56lck in t cell signaling and thymocyte development. In *Advances in immunology*, volume 56, pages 151–178. Elsevier, 1994.

[202] Noah Isakov and Brigitte Biesinger. Lck protein tyrosine kinase is a key regulator of t-cell activation and a target for signal intervention by herpesvirus saimiri and other viral gene products. *European Journal of Biochemistry*, 267(12):3413–3421, 2000.

[203] Shengduo Liu, Shasha Chen, Xinran Li, Shiying Wu, Qian Zhang, Qiuheng Jin, Lin Hu, Ruyuan Zhou, Zhengyang Yu, Fansen Meng, et al. Lck/hck/fgr-mediated tyrosine phosphorylation negatively regulates tbk1 to restrain innate antiviral responses. *Cell host & microbe*, 21(6):754–768, 2017.

[204] Qianru Wei, Joanna Brzostek, Shvetha Sankaran, Javier Casas, Lois Shi-Qi Hew, Jiawei Yap, Xiang Zhao, Lukasz Wojciech, and Nicholas RJ Gascoigne. Lck bound to coreceptor is less active than free lck. *Proceedings of the National Academy of Sciences*, 117(27):15809–15817, 2020.

[205] Peter W Kim, Zhen-Yu J Sun, Stephen C Blacklow, Gerhard Wagner, and Michael J Eck. A zinc clasp structure tethers lck to t cell coreceptors cd4 and cd8. *Science*, 301(5640):1725–1728, 2003.

[206] Melany J Wagner and James R Smiley. Herpes simplex virus requires vp11/12 to induce phosphorylation of the activation loop tyrosine (y394) of the src family kinase lck in t lymphocytes. *Journal of virology*, 83(23):12452–12461, 2009.

[207] Laurence Briant, Veronique Robert-Hebmann, Claire Acquaviva, Annegret Pelchen-Matthews, Mark Marsh, and Christian Devaux. The protein tyrosine kinase p56 lck is required for triggering nf-κb activation upon interaction of human immunodeficiency virus type 1 envelope glycoprotein gp120 with cell surface cd4. *Journal of virology*, 72(7):6207–6214, 1998.

[208] Konstantina Nika, Cristiana Soldani, Mogjiborahman Salek, Wolfgang Paster, Adrian Gray, Ruth Etzensperger, Lars Fugger, Paolo Polzella, Vincenzo Cerundolo, Omer Dushek, et al. Constitutively active lck kinase in t cells drives antigen receptor signal transduction. *Immunity*, 32(6):766–777, 2010.

[209] Louise McNeill, Robert J Salmond, Joanne C Cooper, Céline K Carret, Robin L Cassady-Cain, Marta Roche-Molina, Panna Tandon, Nick Holmes, and Denis R Alexander. The differential regulation of lck kinase phosphorylation sites by cd45 is critical for t cell receptor signaling responses. *Immunity*, 27(3):425–437, 2007.

[210] Claudia Haller, Susanne Rauch, and Oliver T Fackler. Hiv-1 nef employs two distinct mechanisms to modulate lck subcellular localization and tcr induced actin remodeling. *PloS one*, 2(11):e1212, 2007.

[211] Xiaoyu Pan, Jochen M Rudolph, Libin Abraham, Anja Habermann, Claudia Haller, Jacomine Krijnse-Locker, and Oliver T Fackler. Hiv-1 nef compensates for disorganization of the immunological synapse by inducing trans-golgi network–associated lck signaling. *Blood, The Journal of the American Society of Hematology*, 119(3):786–797, 2012.

[212] Iratxe del Río-Iñiguez, Elena Vázquez-Chávez, Céline Cuche, Vincenzo Di Bartolo, Jérôme Bouchet, and Andrés Alcover. Hiv-1 nef hijacks lck and rac1 endosomal traffic to dually modulate signaling-mediated and actin cytoskeleton–mediated t cell functions. *The Journal of Immunology*, 201(9):2624–2640, 2018.

[213] Amy B Strasner, Malini Natarajan, Tom Doman, Douglas Key, Avery August, and Andrew J Henderson. The src kinase lck facilitates assembly of hiv-1 at the plasma membrane. *The Journal of Immunology*, 181(5):3706–3713, 2008.

[214] Hiroyuki Nishiyama, Katsuhiko Itoh, Yoshiyuki Kaneko, Masamichi Kishishita, Osamu Yoshida, and Jun Fujita. A glycine-rich rna-binding protein mediating cold-inducible suppression of mammalian cell growth. *The Journal of cell biology*, 137(4):899–908, 1997.

[215] Christian Brochu, Miguel A Cabrita, Brian D Melanson, Jeffrey D Hamill, Rosanna Lau, MA Christine Pratt, and Bruce C McKay. Nf-κb-dependent role for cold-inducible rna binding protein in regulating interleukin 1β. *PloS one*, 8(2):e57426, 2013.

[216] Frederic De Leeuw, Tong Zhang, Corinne Wauquier, Georges Huez, Véronique Kruys, and Cyril Gueydan. The cold-inducible rna-binding protein migrates from the nucleus to cytoplasmic stress granules by a methylation-dependent mechanism and acts as a translational repressor. *Experimental cell research*, 313(20):4130–4144, 2007.

[217] Juan Pablo Idrovo, Asha Jacob, Weng Lang Yang, Zhimin Wang, Hao Ting Yen, Jeffrey Nicastro, Gene F Coppa, and Ping Wang. A deficiency in cold-inducible rna-binding protein accelerates the inflammation phase and improves wound healing. *International journal of molecular medicine*, 37(2):423–428, 2016.

[218] Daniel A Lujan, Joey L Ochoa, and Rebecca S Hartley. Cold-inducible rna binding protein in cancer and inflammation. *Wiley Interdisciplinary Reviews: RNA*, 9(2):e1462, 2018.

[219] Xiao Xiao, Wentao Zhang, Deping Hua, Lilin Zhang, Wei Meng, Jinhai Huang, and Lei Zhang. Cold-inducible rna-binding protein (cirbp) promotes porcine reproductive and respiratory syndrome virus (prrsv)-induced inflammatory response. *International Immunopharmacology*, 86:106728, 2020.

[220] Nandan S Gokhale, Alexa BR McIntyre, Melissa D Mattocks, Christopher L Holley, Helen M Lazear, Christopher E Mason, and Stacy M Horner. Altered m6a modification of specific cellular transcripts affects flaviviridae infection. *Molecular Cell*, 77(3):542–555, 2020.

[221] Daoliang Lan, Cheng Tang, Mingyi Li, and Hua Yue. Screening and identification of differentially expressed genes from chickens infected with newcastle disease virus by suppression subtractive hybridization. *Avian pathology*, 39(3):151–159, 2010.

[222] Katherine Kedzierska and Suzanne M Crowe. Cytokines and hiv-1: interactions and clinical implications. *antiviral Chemistry and Chemotherapy*, 12(3):133–150, 2001.

[223] Elizabeth Jurrus, Dave Engel, Keith Star, Kyle Monson, Juan Brandi, Lisa E Felberg, David H Brookes, Leighton Wilson, Jiahui Chen, Karina Liles, et al. Improvements to the apbs biomolecular solvation software suite. *Protein Science*, 27(1):112–128, 2018.

[224] Kimitoshi Kohno, Hiroto Izumi, Takeshi Uchiumi, Megumi Ashizuka, and Michihiko Kuwano. The pleiotropic functions of the y-box-binding protein, yb-1. *Bioessays*, 25(7):691–698, 2003.

[225] Ken Matsumoto and Alan P Wolffe. Gene regulation by y-box proteins: coupling control of transcription and translation. *Trends in cell biology*, 8(8):318–323, 1998.

[226] Mei Nie, Maria S Balda, and Karl Matter. Stress-and rho-activated zo-1–associated nucleic acid binding protein binding to p21 mrna mediates stabilization, translation, and cell survival. *Proceedings of the National Academy of Sciences*, 109(27):10897–10902, 2012.

[227] Sébastien Dupasquier, Anne-Sophie Delmarcelle, Etienne Marbaix, Jean-Pierre Cosyns, Pierre J Courtoy, and Christophe E Pierreux. Validation of housekeeping gene and impact on normalized gene expression in clear cell renal cell carcinoma: critical reassessment of ybx3/zonab/csda expression. *BMC molecular biology*, 15(1):1–13, 2014.

[228] Xiaoqin Fan, Xina Xie, Ming Yang, Yujie Wang, Hanwei Wu, Tingting Deng, Xin Weng, Weiping Wen, and Guohui Nie. Ybx3 mediates the metastasis of nasopharyngeal carcinoma via pi3k/akt signaling. *Frontiers in Oncology*, 11:617621, 2021.

[229] Hai-Long Qin, Xue-Jun Wang, Bi-Xian Yang, Bin Du, and Xue-Lin Yun. Notoginsenoside r1 attenuates breast cancer progression by targeting ccnd2 and ybx3. *Chinese Medical Journal*, 134(05):546–554, 2021.

[230] Yiming Sun, Zhixi Li, Wensheng Wang, Xiuyang Zhang, Wenjing Li, Guangsheng Du, Jiuheng Yin, Weidong Xiao, and Hua Yang. Identification and

verification of ybx3 and its regulatory gene heih as an oncogenic system: A multidimensional analysis in colon cancer. *Frontiers in Immunology*, 13, 2022.

[231] Shinichi Kudo, Marie-Genevieve Mattei, and Minoru Fukuda. Characterization of the gene for dbpa, a family member of the nucleic-acid-binding proteins containing a cold-shock domain. *European journal of biochemistry*, 231(1):72–82, 1995.

[232] Leeanne S Coles, M Antonetta Bartley, Andrew Bert, Julie Hunter, Steven Polyak, Peter Diamond, Mathew A Vadas, and Gregory J Goodall. A multi-protein complex containing cold shock domain (y-box) and polypyrimidine tract binding proteins forms on the vascular endothelial growth factor mrna: Potential role in mrna stabilization. *European journal of biochemistry*, 271(3):648–660, 2004.

[233] Holly G Davies, Flaviano Giorgini, Mark A Fajardo, and Robert E Braun. A sequence-specific rna binding complex expressed in murine germ cells contains msy2 and msy4. *Developmental biology*, 221(1):87–100, 2000.

[234] Amy Cooke, Thomas Schwarzl, Ina Huppertz, Gertjan Kramer, Panagiotis Mantas, Anne-Marie Alleaume, Wolfgang Huber, Jeroen Krijgsveld, and Matthias W Hentze. The rna-binding protein ybx3 controls amino acid levels by regulating slc mrna abundance. *Cell reports*, 27(11):3097–3106, 2019.

[235] Antonio J Pérez-Pulido, Gualberto Asencio-Cortés, Ana M Brokate-Llanos, Gloria Brea-Calvo, Rosario Rodríguez-Griñolo, Andrés Garzón, and Manuel J Muñoz. Serial co-expression analysis of host factors from sars-cov viruses highly converges with former high-throughput screenings and proposes key regulators. *Briefings in bioinformatics*, 22(2):1038–1052, 2021.

[236] Zhenqiao Qin, Xiao Qu, Lei Lei, Lulai Xu, and Zishu Pan. Y-box-binding protein 3 (ybx3) restricts influenza a virus by interacting with viral ribonucleoprotein complex and imparing its function. *Journal of General Virology*, 101(4):385–398, 2020.

[237] Philippe E Mangeot, Valérie Risson, Floriane Fusil, Aline Marnef, Emilie Laurent, Juliana Blin, Virginie Mournetas, Emmanuelle Massouridès, Thibault JM Sohier, Antoine Corbin, et al. Genome editing in primary cells and in vivo using viral-derived nanoblades loaded with cas9-sgrna ribonucleoproteins. *Nature communications*, 10(1):1–15, 2019.

[238] Jovan Nikolic, Laura Belot, Hélène Raux, Pierre Legrand, Yves Gaudin, and Aurélie A Albertini. Structural basis for the recognition of ldl-receptor family members by vsv glycoprotein. *Nature communications*, 9(1):1–12, 2018.

[239] John AG Briggs, Martha N Simon, Ingolf Gross, Hans-Georg Kräusslich, Stephen D Fuller, Volker M Vogt, and Marc C Johnson. The stoichiometry of gag protein in hiv-1. *Nature structural & molecular biology*, 11(7):672–675, 2004.

[240] Shingo Iwami, Junko S Takeuchi, Shinji Nakaoka, Fabrizio Mammano, François Clavel, Hisashi Inaba, Tomoko Kobayashi, Naoko Misawa, Kazuyuki Aihara, Yoshio Koyanagi, et al. Cell-to-cell infection by hiv contributes over half of virus infection. *Elife*, 4:e08150, 2015.

[241] Christopher Aiken. Pseudotyping human immunodeficiency virus type 1 (hiv-1) by the glycoprotein of vesicular stomatitis virus targets hiv-1 entry to an endocytic pathway and suppresses both the requirement for nef and the sensitivity to cyclosporin a. *Journal of virology*, 71(8):5871–5877, 1997.

[242] Bangdong L Wei, Paul W Denton, Eduardo O'Neill, Tianci Luo, John L Foster, and J Victor Garcia. Inhibition of lysosome and proteasome function enhances human immunodeficiency virus type 1 infection. *Journal of virology*, 79(9):5705–5712, 2005.

[243] Akash Patnaik, Vincent Chau, and John W Wills. Ubiquitin is part of the retrovirus budding machinery. *Proceedings of the National Academy of Sciences*, 97(24):13069–13074, 2000.

[244] Paola Sette, Kunio Nagashima, Robert C Piper, and Fadila Bouamr. Ubiquitin conjugation to gag is essential for escrt-mediated hiv-1 budding. *Retrovirology*, 10(1):79, 2013.

[245] Hyo-Young Chung, Eiji Morita, Uta von Schwedler, Barbara Muller, Hans-Georg Krausslich, and Wesley I Sundquist. Nedd4l overexpression rescues the release and infectivity of human immunodeficiency virus type 1 constructs lacking ptap and ypxl late domains. *Journal of virology*, 82(10):4884–4897, 2008.

[246] Yoshiko Usami, Sergei Popov, Elena Popova, and Heinrich G Gottlinger. Efficient and specific rescue of human immunodeficiency virus type 1 budding defects by a nedd4-like ubiquitin ligase. *Journal of virology*, 82(10):4898–4907, 2008.

[247] Paola Sette, Joshua A Jadwin, Vincent Dussupt, Nana F Bello, and Fadila Bouamr. The escrt-associated protein alix recruits the ubiquitin ligase nedd4-1 to facilitate hiv-1 release through the lypxnl l domain motif. *Journal of virology*, 84(16):8181–8192, 2010.

[248] Kalle Saksela, GENHONG Cheng, and David Baltimore. Proline-rich (pxxp) motifs in hiv-1 nef bind to sh3 domains of a subset of src kinases and are required for the enhanced growth of nef+ viruses but not for down-regulation of cd4. *The EMBO journal*, 14(3):484–491, 1995.

[249] Anastassia D Blagoveshchenskaya, Laurel Thomas, Sylvain F Feliciangeli, Chien-Hui Hung, and Gary Thomas. Hiv-1 nef downregulates mhc-i by a pacs-1-and pi3k-regulated arf6 endocytic pathway. *Cell*, 111(6):853–866, 2002.

[250] Tobias Brandmann, Hana Fakim, Zoya Padamsi, Ji-Young Youn, Anne-Claude Gingras, Marc R Fabian, and Martin Jinek. Molecular architecture of lsm 14 interactions involved in the assembly of mrna silencing complexes. *The EMBO Journal*, 37(7):e97869, 2018.

[251] Shringar Rao, Alessandro Cinti, Abdelkrim Temzi, Raquel Amorim, Ji Chang You, and Andrew J Mouland. Hiv-1 nc-induced stress granule assembly and

translation arrest are inhibited by the dsrna binding protein staufen1. *Rna*, 24(2):219–236, 2018.

[252] Alissa Bukrinskaya, Beda Brichacek, Angela Mann, and Mario Stevenson. Establishment of a functional human immunodeficiency virus type 1 (hiv-1) reverse transcription complex involves the cytoskeleton. *The Journal of experimental medicine*, 188(11):2113–2125, 1998.

[253] Jianbo Chen, David Grunwald, Luca Sardo, Andrea Galli, Sergey Plisov, Olga A Nikolaitchik, De Chen, Stephen Lockett, Daniel R Larson, Vinay K Pathak, et al. Cytoplasmic hiv-1 rna is mainly transported by diffusion in the presence or absence of gag protein. *Proceedings of the National Academy of Sciences*, 111(48):E5205–E5213, 2014.

[254] Michelle S Itano, Marina Bleck, Daniel S Johnson, and Sanford M Simon. Readily accessible multiplane microscopy: 3d tracking the hiv-1 genome in living cells. *Traffic*, 17(2):179–186, 2016.

[255] Hiroyuki Sasaki, Hiroshi Ozaki, Hideaki Karaki, and Yoshiaki Nonomura. Actin filaments play an essential role for transport of nascent hiv-1 proteins in host cells. *Biochemical and biophysical research communications*, 316(2):588–593, 2004.

[256] F Javier Ibarrondo, Ruth Choi, Yong-Zhi Geng, Jude Canon, Osvaldo Rey, Gayle Cocita Baldwin, and Paul Krogstad. Hiv type 1 gag and nucleocapsid proteins: cytoskeletal localization and effects on cell motility. *AIDS research and human retroviruses*, 17(16):1489–1500, 2001.

[257] Haruka Kamiyama, Mai Izumida, Yuria Umemura, Hideki Hayashi, Toshifumi Matsuyama, and Yoshinao Kubo. Role of ezrin phosphorylation in hiv-1 replication. *Frontiers in Microbiology*, 9:1912, 2018.

[258] Pascal RA Meylan, John C Guatelli, James R Munis, Douglas D Richman, and Richard S Kornbluth. Mechanisms for the inhibition of hiv replication by

interferons-α,-β, and-γ, in primary human macrophages. *Virology*, 193(1):138–148, 1993.

[259] Lilen C Sarol, Kenichi Imai, Kaori Asamitsu, Toshifumi Tetsuka, Nina G Barzaga, and Takashi Okamoto. Inhibitory effects of ifn-γ on hiv-1 replication in latently infected cells. *Biochemical and biophysical research communications*, 291(4):890–896, 2002.

[260] Najla Nasr, Susan Maddocks, Stuart G Turville, Andrew N Harman, Natalie Woolger, Karla J Helbig, John Wilkinson, Chris R Bye, Thomas K Wright, Dharshini Rambukwelle, et al. Hiv-1 infection of human macrophages directly induces viperin which inhibits viral production. *Blood, The Journal of the American Society of Hematology*, 120(4):778–788, 2012.

[261] DN Lyabin, IA Eliseeva, EA Smolin, AN Doronin, KS Budkina, IV Kulakovskiy, and LP Ovchinnikov. Yb-3 substitutes yb-1 in global mrna binding. *RNA biology*, 17(4):487–499, 2020.

[262] Tom Li Stephen, Bridget S Wilson, and Terri M Laufer. Subcellular distribution of lck during cd4 t-cell maturation in the thymic medulla regulates the t-cell activation threshold. *Proceedings of the National Academy of Sciences*, 109(19):7415–7420, 2012.

[263] Olivia Muriel, Alejandra Tomas, Cameron C Scott, and Jean Gruenberg. Moesin and cortactin control actin-dependent multivesicular endosome biogenesis. *Molecular biology of the cell*, 27(21):3305–3316, 2016.

[264] Ming Li, Feng Gao, John R Mascola, Leonidas Stamatatos, Victoria R Polonis, Marguerite Koutsoukos, Gerald Voss, Paul Goepfert, Peter Gilbert, Kelli M Greene, et al. Human immunodeficiency virus type 1 env clones from acute and early subtype b infections for standardized assessments of vaccine-elicited neutralizing antibodies. *Journal of virology*, 79(16):10108–10125, 2005.

[265] Mark K Louder, Anna Sambor, Elena Chertova, Tai Hunte, Sarah Barrett, Fallon Ojong, Eric Sanders-Buell, Susan Zolla-Pazner, Francine E McCutchan,

James D Roser, et al. Hiv-1 envelope pseudotyped viral vectors and infectious molecular clones expressing the same envelope glycoprotein have a similar neutralization phenotype, but culture in peripheral blood mononuclear cells is associated with decreased neutralization sensitivity. *Virology*, 339(2):226–238, 2005.

[266] Miroslawa Bilska, Haili Tang, and David C Montefiori. Potential risk of replication-competent virus in hiv-1 env-pseudotyped virus preparations. *AIDS Research and Human Retroviruses*, 33(4):368–372, 2017.

[267] Akira Ono. Hiv-1 assembly at the plasma membrane: Gag trafficking and localization. *Future Virology*, 4(3):241–245, 2009.

[268] Wesley I Sundquist and Hans-Georg Kräusslich. Hiv-1 assembly, budding, and maturation. *Cold Spring Harbor perspectives in medicine*, 2(7):a006924, 2012.

[269] Sebla B Kutluay and Paul D Bieniasz. Analysis of the initiating events in hiv-1 particle assembly and genome packaging. *PLoS pathogens*, 6(11):e1001200, 2010.

[270] Siddhartha AK Datta, Frank Heinrich, Sindhu Raghunandan, Susan Krueger, Joseph E Curtis, Alan Rein, and Hirsh Nanda. Hiv-1 gag extension: conformational changes require simultaneous interaction with membrane and nucleic acid. *Journal of molecular biology*, 406(2):205–214, 2011.

[271] JAG Briggs, JD Riches, Barbel Glass, V Bartonova, Giulia Zanetti, and H-G Kräusslich. Structure and assembly of immature hiv. *Proceedings of the National Academy of Sciences*, 106(27):11090–11095, 2009.

[272] James H Hurley and Phyllis I Hanson. Membrane budding and scission by the escrt machinery: it's all in the neck. *Nature reviews Molecular cell biology*, 11(8):556–566, 2010.

[273] James H Hurley. The escrt complexes. *Critical reviews in biochemistry and molecular biology*, 45(6):463–487, 2010.

[274] Juan Martin-Serrano, Anton Yaravoy, David Perez-Caballero, and Paul D Bieniasz. Divergent retroviral late-budding domains recruit vacuolar protein sorting factors by using alternative adaptor proteins. *Proceedings of the National Academy of Sciences*, 100(21):12414–12419, 2003.

[275] Charles Langelier, Uta K von Schwedler, Robert D Fisher, Ivana De Domenico, Paul L White, Christopher P Hill, Jerry Kaplan, Diane Ward, and Wesley I Sundquist. Human escrt-ii complex and its role in human immunodeficiency virus type 1 release. *Journal of virology*, 80(19):9465–9480, 2006.

[276] Sara Marie Øie Solbak, Tove Ragna Reksten, Friedrich Hahn, Victor Wray, Petra Henklein, Peter Henklein, Øyvind Halskau, Ulrich Schubert, and Torgils Fossen. Hiv-1 p6—a structured to flexible multifunctional membrane-interacting protein. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1828(2):816–823, 2013.

[277] Julia JM Eekels, Dirk Geerts, Rienk E Jeeninga, and Ben Berkhout. Long-term inhibition of hiv-1 replication with rna interference against cellular co-factors. *Antiviral research*, 89(1):43–53, 2011.

[278] Young Jun Im, Lillian Kuo, Xuefeng Ren, Patricia V Burgos, Xue Zhi Zhao, Fa Liu, Terrence R Burke, Juan S Bonifacino, Eric O Freed, and James H Hurley. Crystallographic and functional analysis of the escrt-i/hiv-1 gag ptap interaction. *Structure*, 18(11):1536–1547, 2010.

[279] Bettina Strack, Arianna Calistri, Stewart Craig, Elena Popova, and Heinrich G Göttlinger. Aip1/alix is a binding partner for hiv-1 p6 and eiav p9 functioning in virus budding. *Cell*, 114(6):689–699, 2003.

[280] Suman Peel, Pauline Macheboeuf, Nicolas Martinelli, and Winfried Weissenhorn. Divergent pathways lead to escrt-iii-catalyzed membrane fission. *Trends in biochemical sciences*, 36(4):199–210, 2011.

[281] Juan Martin-Serrano. The role of ubiquitin in retroviral egress. *Traffic*, 8(10):1297–1303, 2007.

[282] Viola Baumgärtel, Sergey Ivanchenko, Aurélie Dupont, Mikhail Sergeev, Paul W Wiseman, Hans-Georg Kräusslich, Christoph Bräuchle, Barbara Müller, and Don C Lamb. Live-cell visualization of dynamics of hiv budding site interactions with an escrt component. *Nature cell biology*, 13(4):469–474, 2011.

[283] Nolwenn Jouvenet, Maria Zhadina, Paul D Bieniasz, and Sanford M Simon. Dynamics of escrt protein recruitment during retroviral assembly. *Nature cell biology*, 13(4):394–401, 2011.

[284] Gur Fabrikant, Suman Lata, James D Riches, John AG Briggs, Winfried Weissenhorn, and Michael M Kozlov. Computational model of membrane fission catalyzed by escrt-iii. *PLoS computational biology*, 5(11):e1000575, 2009.

[285] Lars-Anders Carlson, Qing-Tao Shen, Mark Remec Pavlin, and James H Hurley. Escrt filaments as spiral springs. *Developmental cell*, 35(4):397–398, 2015.

[286] Jaime Agudo-Canalejo and Reinhard Lipowsky. Domes and cones: Adhesion-induced fission of membranes by escrt proteins. *PLoS computational biology*, 14(8):e1006422, 2018.

[287] Sara Ghazi-Tabatabai, Suraj Saksena, Judith M Short, Ajaybabu V Pobbati, Dmitry B Veprintsev, R Anthony Crowther, Scott D Emr, Edward H Egelman, and Roger L Williams. Structure and disassembly of filaments formed by the escrt-iii subunit vps24. *Structure*, 16(9):1345–1356, 2008.

[288] Ulrich Schubert, David E Ott, Elena N Chertova, Reinhold Welker, Uwe Tessmer, Michael F Princiotta, Jack R Bennink, Hans-Georg Kräusslich, and Jonathan W Yewdell. Proteasome inhibition interferes with gag polyprotein processing, release, and maturation of hiv-1 and hiv-2. *Proceedings of the National Academy of Sciences*, 97(24):13057–13062, 2000.

[289] Bettina Strack, Arianna Calistri, Molly A Accola, Giorgio Palù, and Heinrich G Göttlinger. A role for ubiquitin ligase recruitment in retrovirus release. *Proceedings of the National Academy of Sciences*, 97(24):13063–13068, 2000.

[290] Stefanie Jager, Eva Gottwein, and Hans-Georg Krausslich. Ubiquitination of human immunodeficiency virus type 1 gag is highly dependent on gag membrane association. *Journal of virology*, 81(17):9193–9201, 2007.

[291] Eva Gottwein, Stefanie Jager, Anja Habermann, and Hans-Georg Krausslich. Cumulative mutations of ubiquitin acceptor sites in human immunodeficiency virus type 1 gag cause a late budding defect. *Journal of virology*, 80(13):6267–6275, 2006.

[292] Eric R Weiss, Elena Popova, Hikaru Yamanaka, Hyung Cheol Kim, Jon M Huibregtse, and Heinrich Göttlinger. Rescue of hiv-1 release by targeting widely divergent nedd4-type ubiquitin ligases and isolated catalytic hect domains to gag. *PLoS pathogens*, 6(9):e1001107, 2010.

[293] Maria Zhadina and Paul D Bieniasz. Functional interchangeability of late domains, late domain cofactors and ubiquitin in viral budding. *PLoS pathogens*, 6(10):e1001153, 2010.

[294] Pamela J Plant, Herman Yeger, Olivier Staub, Perry Howard, and Daniela Rotin. The c2 domain of the ubiquitin protein ligase nedd4 mediates ca2+-dependent plasma membrane localization. *Journal of Biological Chemistry*, 272(51):32329–32336, 1997.

[295] Vincent Blot, Fabien Perugi, Bernard Gay, Marie-Christine Prévost, Laurence Briant, Frédéric Tangy, Hugues Abriel, Olivier Staub, Marie-Christine Dokhélar, and Claudine Pique. Nedd4. 1-mediated ubiquitination and subsequent recruitment of tsg101 ensure htlv-1 gag trafficking towards the multivesicular body pathway prior to virus budding. *Journal of cell science*, 117(11):2357–2367, 2004.

[296] Jiro Yasuda, Eric Hunter, Mitsuyoshi Nakao, and Hisatoshi Shida. Functional involvement of a novel nedd4-like ubiquitin ligase on retrovirus budding. *EMBO reports*, 3(7):636–640, 2002.

[297] Marcy L Vana, Yi Tang, Aiping Chen, Gisselle Medina, Carol Carter, and Jonathan Leis. Role of nedd4 and ubiquitination of rous sarcoma virus gag in budding of virus-like particles from cells. *Journal of virology*, 78(24):13943–13953, 2004.

[298] Wesley I Sundquist, Heidi L Schubert, Brian N Kelly, Gina C Hill, James M Holton, and Christopher P Hill. Ubiquitin recognition by the human tsg101 protein. *Molecular cell*, 13(6):783–789, 2004.

[299] Anjali Joshi, Utpal Munshi, Sherimay D Ablan, Kunio Nagashima, and Eric O Freed. Functional replacement of a retroviral late domain by ubiquitin fusion. *Traffic*, 9(11):1972–1983, 2008.

[300] Robert D Fisher, Hyo-Young Chung, Qianting Zhai, Howard Robinson, Wesley I Sundquist, and Christopher P Hill. Structural and biochemical studies of alix/aip1 and its role in retrovirus budding. *Cell*, 128(5):841–852, 2007.

[301] Dara P Dowlatshahi, Virginie Sandrin, Sandro Vivona, Thomas A Shaler, Stephen E Kaiser, Francesco Melandri, Wesley I Sundquist, and Ron R Kopito. Alix is a lys63-specific polyubiquitin binding protein that functions in retrovirus budding. *Developmental cell*, 23(6):1247–1254, 2012.

[302] Tal Keren-Kaplan, Ilan Attali, Michael Estrin, Lillian S Kuo, Efrat Farkash, Moran Jerabek-Willemsen, Noa Blutraich, Shay Artzi, Aviyah Peri, Eric O Freed, et al. Structure-based in silico identification of ubiquitin-binding domains provides insights into the alix-v: ubiquitin complex and retrovirus budding. *The EMBO journal*, 32(4):538–551, 2013.

[303] Atsushi Okumura, Gengshi Lu, Ian Pitha-Rowe, and Paula M Pitha. Innate antiviral response targets hiv-1 release by the induction of ubiquitin-like protein isg15. *Proceedings of the National Academy of Sciences*, 103(5):1440–1445, 2006.

[304] Owen Pornillos, Steven L Alam, Darrell R Davis, and Wesley I Sundquist. Structure of the tsg101 uev domain in complex with the ptap motif of the hiv-1 p6 protein. *Nature structural biology*, 9(11):812–817, 2002.

[305] Stephen O'Gorman, DANIEL T Fox, and Geoffrey M Wahl. Recombinase-mediated gene activation and site-specific integration in mammalian cells. *Science*, 251(4999):1351–1355, 1991.

[306] Robert B DuBridge, P Tang, H Chao Hsia, PAHIK-MOOI Leong, JH Miller, and MP Calos. Analysis of mutation in human cells by using an epstein-barr virus shuttle system. *Molecular and cellular biology*, 7(1):379–387, 1987.

[307] Dahae Hailey Bae, Michael Marino, Brian Iaffaldano, Sydney Fenstermaker, Sandra Afione, Takele Argaw, Jacob McCright, Anna Kwilas, John A Chiorini, Andrew E Timmons, et al. Design and testing of vector-producing hek293t cells bearing a genomic deletion of the sv40 t antigen coding region. *Molecular Therapy-Methods & Clinical Development*, 18:631–638, 2020.

[308] Patrick Salmon and Didier Trono. Production and titration of lentiviral vectors. *Current protocols in human genetics*, 54(1):12–10, 2007.

[309] Tamsin J Samuels, Yoav Arava, Aino I Järvelin, Francesca Robertson, Jeffrey Y Lee, Lu Yang, Ching-Po Yang, Tzumin Lee, David Ish-Horowicz, and Ilan Davis. Neuronal upregulation of prospero protein is driven by alternative mrna polyadenylation and syncrip-mediated mrna stabilisation. *Biology open*, 9(5):bio049684, 2020.

[310] Jeffrey Y Lee, Peter AC Wing, Dalia S Gala, Marko Noerenberg, Aino I Järvelin, Joshua Titlow, Xiaodong Zhuang, Natasha Palmalux, Louisa Iselin, Mary Kay Thompson, et al. Absolute quantitation of individual sars-cov-2 rna molecules provides a new paradigm for infection dynamics and variant differences. *Elife*, 11:e74153, 2022.

[311] Ya-Ching Fang, Ssu-Ju Fu, Po-Hao Hsu, Pei-Tzu Chang, Jing-Jia Huang, Yi-Chih Chiu, Yi-Fan Liao, Guey-Mei Jow, Chih-Yung Tang, and Chung-Jiuan Jeng.

Identification of mkrn1 as a second e3 ligase for eag1 potassium channels reveals regulation via differential degradation. *Journal of Biological Chemistry*, 296, 2021.

[312] Min-Sik Lee, Hyun-Ji Han, Su Yeon Han, Il Young Kim, Sehyun Chae, Choong-Sil Lee, Sung Eun Kim, Seul Gi Yoon, Jun-Won Park, Jung-Hoon Kim, et al. Loss of the e3 ubiquitin ligase mkrn1 represses diet-induced metabolic syndrome through ampk activation. *Nature communications*, 9(1):3404, 2018.

[313] Cristina Garcia-Barcena, Nerea Osinalde, Juanma Ramirez, and Ugo Mayor. How to inactivate human ubiquitin e3 ligases by mutation. *Frontiers in cell and developmental biology*, 8:39, 2020.

[314] Nolwenn Jouvenet, Sanford M Simon, and Paul D Bieniasz. Imaging the interaction of hiv-1 genomes and gag during assembly of individual viral particles. *Proceedings of the National Academy of Sciences*, 106(45):19114–19119, 2009.

[315] Heinrich G Göttlinger, Tatyana Dorfman, Joseph G Sodroski, and William A Haseltine. Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. *Proceedings of the National Academy of Sciences*, 88(8):3195–3199, 1991.

[316] Sandrine Jacquenet, Didier Decimo, Delphine Muriaux, and Jean-Luc Darlix. Dual effect of the sr proteins asf/sf2, sc35 and 9g8 on hiv-1 rna splicing and virion production. *Retrovirology*, 2:1–13, 2005.

[317] Delphine Ropers, Lilia Ayadi, Renata Gattoni, Sandrine Jacquenet, Laurence Damier, Christiane Branlant, and James Stévenin. Differential effects of the sr proteins 9g8, sc35, asf/sf2, and srp40 on the utilization of the a1 to a5 splicing sites of hiv-1 rna. *Journal of Biological Chemistry*, 279(29):29963–29973, 2004.

[318] Jodi Pongoski, Kengo Asai, and Alan Cochrane. Positive and negative modulation of human immunodeficiency virus type 1 rev function by cis and trans regulators of viral rna splicing. *Journal of virology*, 76(10):5108–5120, 2002.

[319] Matteo Scalabrin, Ilaria Frasson, Emanuela Ruggiero, Rosalba Perrone, Elena Tosoni, Sara Lago, Martina Tassinari, Giorgio Palù, and Sara N Richter. The cellular protein hnrnp a2/b1 enhances hiv-1 transcription by unfolding ltr promoter g-quadruplexes. *Scientific reports*, 7(1):1–13, 2017.

[320] Nicole Lund, Miroslav P Milev, Raymond Wong, Tharmila Sanmuganantham, Kathryn Woolaway, Benoit Chabot, Sherif Abou Elela, Andrew J Mouland, and Alan Cochrane. Differential effects of hnrnp d/auf1 isoforms on hiv-1 gene expression. *Nucleic acids research*, 40(8):3663–3675, 2012.

[321] Xin Mu, Wei Li, Xinlu Wang, and Guangxia Gao. Yb-1 stabilizes hiv-1 genomic rna and enhances viral production. *Protein & cell*, 4:591–597, 2013.

[322] Michelle Vincendeau, Daniel Nagel, Jara K Brenke, Ruth Brack-Werner, and Kamyar Hadian. Heterogenous nuclear ribonucleoprotein q increases protein expression from hiv-1 rev-dependent transcripts. *Virology journal*, 10(1):1–7, 2013.

[323] Ioana Boeras, Zhenwei Song, Andrew Moran, Jarryd Franklin, William Clay Brown, Marc Johnson, Kathleen Boris-Lawrie, and Xiao Heng. Dhx9/rha binding to the pbs-segment of the genomic rna during hiv-1 assembly bolsters virion infectivity. *Journal of molecular biology*, 428(11):2418–2429, 2016.

[324] Hala El Mekdad, Emmanuel Boutant, Hassan Karnib, Marina E Biedma, Kamal Kant Sharma, Iuliia Malytska, Géraldine Laumond, Marion Roy, Eléonore Réal, Jean-Christophe Paillart, et al. Characterization of the interaction between the hiv-1 gag structural polyprotein and the cellular ribosomal protein l7 and its implication in viral nucleic acid remodeling. *Retrovirology*, 13(1):1–14, 2016.

[325] Anne Monette, Fernando Valiente-Echeverría, Matias Rivero, Éric A Cohen, Marcelo Lopez-Lastra, and Andrew J Mouland. Dual mechanisms of translation initiation of the full-length hiv-1 mrna contribute to gag synthesis. *PloS one*, 8(7):e68108, 2013.

[326] Alfredo Castello, David Franco, Pablo Moral-Lopez, Juan J Berlanga, Enrique Alvarez, Eckard Wimmer, and Luis Carrasco. Hiv-1 protease inhibits cap- and poly (a)-dependent translation upon eif4gi and pabp cleavage. *PloS one*, 4(11):e7997, 2009.

[327] Delphine Muriaux, Jane Mirro, Demetria Harvin, and Alan Rein. Rna is a structural element in retrovirus particles. *Proceedings of the National Academy of Sciences*, 98(9):5246–5251, 2001.

[328] David McDonald, Marie A Vodicka, Ginger Lucero, Tatyana M Svitkina, Gary G Borisy, Michael Emerman, and Thomas J Hope. Visualization of the intracellular behavior of hiv in living cells. *The Journal of cell biology*, 159(3):441–452, 2002.

[329] Benjamin M Dale, Gregory P McNerney, Wolfgang Hübner, Thomas R Huser, and Benjamin K Chen. Tracking and quantitation of fluorescent hiv during cell-to-cell transmission. *Methods*, 53(1):20–26, 2011.

[330] Edward M Campbell, Omar Perez, Jenny L Anderson, and Thomas J Hope. Visualization of a proteasome-independent intermediate during restriction of hiv-1 by rhesus trim5α. *The Journal of cell biology*, 180(3):549–561, 2008.

[331] Suzannah J Rihn, Sam J Wilson, Nick J Loman, Mudathir Alim, Saskia E Bakker, David Bhella, Robert J Gifford, Frazer J Rixon, and Paul D Bieniasz. Extreme genetic fragility of the hiv-1 capsid. *PLoS pathogens*, 9(6):e1003461, 2013.

[332] Irena Zurnic Bönisch, Lieve Dirix, Veerle Lemmens, Doortje Borrenberghs, Flore De Wit, Frank Vernaillen, Susana Rocha, Frauke Christ, Jelle Hendrix, Johan Hofkens, et al. Capsid-labelled hiv to investigate the role of capsid during nuclear import and integration. *Journal of virology*, 94(7):e01024–19, 2020.

[333] Sandra Schifferdecker, Vojtech Zila, Thorsten G Müller, Volkan Sakin, Maria Anders-Össwein, Vibor Laketa, Hans-Georg Kräusslich, and Barbara Müller. Direct capsid labeling of infectious hiv-1 by genetic code expansion allows detection of largely complete nuclear capsids and suggests nuclear entry of hiv-1 complexes via common routes. *MBio*, 13(5):e01959–22, 2022.

[334] Ashwanth C Francis, Mariana Marin, Jiong Shi, Christopher Aiken, and Gregory B Melikyan. Time-resolved imaging of single hiv-1 uncoating in vitro and in living cells. *PLoS pathogens*, 12(6):e1005709, 2016.

[335] Wang Peng, Jiong Shi, Chantal L Márquez, Derrick Lau, James Walsh, KM Rifat Faysal, Chang H Byeon, In-Ja L Byeon, Christopher Aiken, and Till Böcking. Functional analysis of the secondary hiv-1 capsid binding site in the host protein cyclophilin a. *Retrovirology*, 16:1–13, 2019.

[336] Tanay M Desai, Mariana Marin, Chetan Sood, Jiong Shi, Fatima Nawaz, Christopher Aiken, and Gregory B Melikyan. Fluorescent protein-tagged vpr dissociates from hiv-1 core after viral fusion and rapidly enters the cell nucleus. *Retrovirology*, 12:1–20, 2015.

[337] Candida F Pereira, Paula C Ellenberg, Kate L Jones, Tara L Fernandez, Redmond P Smyth, David J Hawkes, Marcel Hijnen, Valerie Vivet-Boudou, Roland Marquet, Iain Johnson, et al. Labeling of multiple hiv-1 proteins with the biarsenical-tetracysteine system. *PloS one*, 6(2):e17016, 2011.

[338] Julian C-H Chen, Jolanta Krucinski, Larry JW Miercke, Janet S Finer-Moore, Ann H Tang, Andrew D Leavitt, and Robert M Stroud. Crystal structure of the hiv-1 integrase catalytic core and c-terminal domains: a model for viral dna binding. *Proceedings of the National Academy of Sciences*, 97(15):8233–8238, 2000.

[339] Jian-Yong Wang, Hong Ling, Wei Yang, and Robert Craigie. Structure of a two-domain fragment of hiv-1 integrase: implications for domain organization in the intact protein. *The EMBO journal*, 20(24):7333–7343, 2001.

[340] Peter Cherepanov, Goedele Maertens, Paul Proost, Bart Devreese, Jozef Van Beeumen, Yves Engelborghs, Erik De Clercq, and Zeger Debyser. Hiv-1 integrase forms stable tetramers and associates with ledgf/p75 protein in human cells. *Journal of Biological Chemistry*, 278(1):372–381, 2003.

[341] Angela Ciuffi, Manuel Llano, Eric Poeschla, Christian Hoffmann, Jeremy Leipzig, Paul Shinn, Joseph R Ecker, and Frederic Bushman. A role for ledgf/p75 in targeting hiv dna integration. *Nature medicine*, 11(12):1287–1289, 2005.

[342] João I Mamede, Joseph Griffin, Stéphanie Gambut, and Thomas J Hope. A new generation of functional tagged proteins for hiv fluorescence imaging. *Viruses*, 13(3):386, 2021.

[343] Sofia Benkhelifa-Ziyyat, Stéphanie Bucher, Maria-Antonietta Zanta-Boussif, Julie Pasquet, and Olivier Danos. Changes in the accessibility of the hiv-1 integrase c-terminus in the presence of cellular proteins. *Retrovirology*, 7(1):1–10, 2010.

[344] Alberto Albanese, Daniele Arosio, Mariaelena Terreni, and Anna Cereseto. Hiv-1 pre-integration complexes selectively target decondensed chromatin in the nuclear periphery. *PloS one*, 3(6):e2413, 2008.

[345] Ashwanth Christopher Francis, Cristina Di Primio, Valentina Quercioli, Paola Valentini, Annegret Boll, Gabriele Girelli, Francesca Demichelis, Daniele Arosio, and Anna Cereseto. Second generation imaging of nuclear/cytoplasmic hiv-1 complexes. *AIDS research and human retroviruses*, 30(7):717–726, 2014.

[346] Stephen Hare, Francesca Di Nunzio, Alfred Labeja, Jimin Wang, Alan Engelman, and Peter Cherepanov. Structural basis for functional tetramerization of lentiviral integrase. *PLoS pathogens*, 5(7):e1000515, 2009.

[347] Amy E Hulme, Z Kelley, Deirdre Foley, and Thomas J Hope. Complementary assays reveal a low level of ca associated with viral complexes in the nuclei of hiv-1-infected cells. *Journal of virology*, 89(10):5350–5361, 2015.

[348] Jennifer L Elliott and Sebla B Kutluay. Going beyond integration: The emerging role of hiv-1 integrase in virion morphogenesis. *Viruses*, 12(9):1005, 2020.

[349] Alan Engelman, George Englund, Jan M Orenstein, Malcolm A Martin, and Robert Craigie. Multiple effects of mutations in human immunodeficiency virus type 1 integrase on viral replication. *Journal of virology*, 69(5):2729–2736, 1995.

[350] Jacques J Kessl, Sebla B Kutluay, Dana Townsend, Stephanie Rebensburg, Alison Slaughter, Ross C Larue, Nikoloz Shkriabai, Nordine Bakouche, James R Fuchs, Paul D Bieniasz, et al. Hiv-1 integrase binds the viral rna genome and is essential during virion morphogenesis. *Cell*, 166(5):1257–1268, 2016.

[351] Jennifer L Elliott, Jenna E Eschbach, Pratibha C Koneru, Wen Li, Maritza Puray-Chavez, Dana Townsend, Dana Q Lawson, Alan N Engelman, Mamuka Kvaratskhelia, and Sebla B Kutluay. Integrase-rna interactions underscore the critical role of integrase in hiv-1 virion morphogenesis. *Elife*, 9:e54311, 2020.

[352] Daphne S Bindels, Lindsay Haarbosch, Laura Van Weeren, Marten Postma, Katrin E Wiese, Marieke Mastop, Sylvain Aumonier, Guillaume Gotthard, Antoine Royant, Mark A Hink, et al. mscarlet: a bright monomeric red fluorescent protein for cellular imaging. *Nature methods*, 14(1):53–56, 2017.

[353] Claudia Muratori, Paola d'Aloja, Fabiana Superti, Antonella Tinari, Nathalie Sol-Foulon, Sandra Sparacio, Valerie Bosch, Olivier Schwartz, and Maurizio Federico. Generation and characterization of a stable cell population releasing fluorescent hiv-1-based virus like particles in an inducible way. *BMC biotechnology*, 6(1):1–17, 2006.

[354] Pejman Mohammadi, Sébastien Desfarges, István Bartha, Beda Joos, Nadine Zangger, Miguel Munoz, Huldrych F Günthard, Niko Beerenwinkel, Amalio Telenti, and Angela Ciuffi. 24 hours in the life of hiv-1 in a t cell line. *PLoS pathogens*, 9(1):e1003161, 2013.

[355] Gregory B Melikyan, Richard JO Barnard, Levon G Abrahamyan, Walther Mothes, and John AT Young. Imaging individual retroviral fusion events: from hemifusion to pore formation and growth. *Proceedings of the National Academy of Sciences*, 102(24):8728–8733, 2005.

[356] João I Mamede, Gianguido C Cianci, Meegan R Anderson, and Thomas J Hope. Early cytoplasmic uncoating is associated with infectivity of hiv-1. *Proceedings of the National Academy of Sciences*, 114(34):E7169–E7178, 2017.

[357] Soumajit Mukherjee, Emmanuel Boutant, Eleonore Réal, Yves Mély, and Halina Anton. Imaging viral infection by fluorescence microscopy: Focus on hiv-1 early stage. *Viruses*, 13(2):213, 2021.

[358] Chuang Liu, Juan R Perilla, Jiying Ning, Manman Lu, Guangjin Hou, Ruben Ramalho, Benjamin A Himes, Gongpu Zhao, Gregory J Bedwell, In-Ja Byeon, et al. Cyclophilin a stabilizes the hiv-1 capsid through a novel non-canonical binding site. *Nature communications*, 7(1):10714, 2016.

[359] David S Booth, Yifan Cheng, and Alan D Frankel. The export receptor crm1 forms a dimer to promote nuclear export of hiv rna. *Elife*, 3:e04121, 2014.

[360] Darja Pollpeter, Maddy Parsons, Andrew E Sobala, Sashika Coxhead, Rupert D Lang, Annie M Bruns, Stelios Papaioannou, James M McDonnell, Luis Apolonia, Jamil A Chowdhury, et al. Deep sequencing of hiv-1 reverse transcripts reveals the multifaceted antiviral functions of apobec3g. *Nature microbiology*, 3(2):220–233, 2018.

[361] JM Moreno, JM Sanchez-Montero, JV Sinisterra, and LB Nielsen. Contribution to the study of the enzymatic activity of benzonase. *Journal of molecular catalysis*, 69(3):419–427, 1991.

[362] Paul D Bieniasz and Sebla B Kutluay. Clip-related methodologies and their application to retrovirology. *Retrovirology*, 15(1):1–14, 2018.

[363] Sebla B Kutluay, Trinity Zang, Daniel Blanco-Melo, Chelsea Powell, David Jannain, Manel Errando, and Paul D Bieniasz. Global changes in the rna binding specificity of hiv-1 gag regulate virion genesis. *Cell*, 159(5):1096–1109, 2014.

[364] Luis Apolonia, Reiner Schulz, Tomaž Curk, Paula Rocha, Chad M Swanson, Torsten Schaller, Jernej Ule, and Michael H Malim. Promiscuous rna binding

ensures effective encapsidation of apobec3 proteins by hiv-1. *PLoS pathogens*, 11(1):e1004609, 2015.

[365] Ashley York, Sebla B Kutluay, Manel Errando, and Paul D Bieniasz. The rna binding specificity of human apobec3 proteins resembles that of hiv-1 nucleocapsid. *PLoS pathogens*, 12(8):e1005833, 2016.

[366] Yafeng Dou, Yan Xie, Lingyun Zhang, Sheng Liu, Dandan Xu, Yongshuai Li, and Xiao-Lian Zhang. Host mkrn1-mediated mycobacterial ppe protein ubiquitination suppresses innate immune response. *Frontiers in Immunology*, page 2100, 2022.

[367] Andrew D Jacobson, Nan-Yan Zhang, Ping Xu, Ke-Jun Han, Seth Noone, Junmin Peng, and Chang-Wei Liu. The lysine 48 and lysine 63 ubiquitin conjugates are processed differently by the 26 s proteasome. *Journal of Biological Chemistry*, 284(51):35485–35494, 2009.

[368] James A Nathan, Hyoung Tae Kim, Lily Ting, Steven P Gygi, and Alfred L Goldberg. Why do cellular proteins linked to k63-polyubiquitin chains not associate with proteasomes? *The EMBO journal*, 32(4):552–565, 2013.

[369] Amit Fulzele and Eric J Bennett. Ubiquitin digly proteomics as an approach to identify and quantify the ubiquitin-modified proteome. *The Ubiquitin Proteasome System: Methods and Protocols*, pages 363–384, 2018.

[370] Roland Hjerpe, Fabienne Aillet, Fernando Lopitz-Otsoa, Valerie Lang, Patrick England, and Manuel S Rodriguez. Efficient protection and isolation of ubiquitylated proteins using tandem ubiquitin-binding entities. *EMBO reports*, 10(11):1250–1258, 2009.

[371] Martin Steger, Özge Karayel, and Vadim Demichev. Ubiquitinomics: History, methods, and applications in basic research and drug discovery. *Proteomics*, 22(15-16):2200074, 2022.

[372] Raviteja Inturi, Kwangchol Mun, Katrin Singethan, Sabrina Schreiner, and Tanel Punga. Human adenovirus infection causes cellular e3 ubiquitin ligase mkrn1 degradation involving the viral core protein pvii. *Journal of virology*, 92(3):e01154–17, 2018.