Unveiling the protein-protein interactome of the RNA-binding protein Rev in HIV-1 infected T CD4+ lymphocytic cells

Catherine Truman

Trinity College

University of Oxford

A thesis presented for the degree of Doctor of Philosophy

Michaelmas 2021

Table of Contents

Abstract *5*

List of Figures *6*

List of Tables *8*

Abbreviations *9*

Acknowledgements *14*

1. Introduction *16*

1.1. The HIV-1 infection cycle **17**
1.1.1. The HIV-1 genome is alternatively spliced 17
1.1.2. HIV-1 virion structure and proteome 19
1.1.3. The HIV-1 infection cycle 20

1.2. The HIV-1 Rev protein **23**
1.2.1. Rev discovery and structure 23
1.2.2. Rev as an RNA export factor 24
1.2.3. Additional functions of Rev 28
1.2.4. Rev as a therapeutic target 33

1.3. RNA-binding proteins in HIV-1 infection **35**

1.4. Aims **37**

2. Materials *38*

2.1. Materials **38**
2.1.1. Chemicals and reagents 38
2.1.2. Consumables 39
2.1.3. Machines 39
2.1.4. Buffers and solutions 40
2.1.5. Commercial kits 40
2.1.6. Enzymes 41
2.1.7. Antibodies and dyes 41
2.1.8. Plasmids 42
2.1.9. Oligonucleotides and primers 42
2.1.10. Cell lines 43

3. Methods *44*

3.1. Cell biology **44**
3.1.1. Maintaining cells 44
3.1.2. Stable transfection of Flp-In T-Rex cells 44
3.1.3. Generating and titrating pseudotyped virus 45

3.2. Molecular biology techniques **45**
3.2.1. Cloning Rev-tagged HIV-1 replicons 45
3.2.2. Fusion PCR to generate ΔRev replicon 47
3.2.3. Plasmid preparation 48
3.2.4. X-tremeGENE 9 Transfection of HIV-1 plasmids 48
3.2.5. Western blotting and silver staining 49
3.2.6. Immunofluorescence 49
3.2.7. Single molecule fluorescence in situ hybridisation 50
3.2.8. Cloning RBP-GFP plasmids 51

3.3. Proteomic Analysis **52**
3.3.1. Tagged Rev immunoprecipitations 52

3.3.2. PURA-GFP and PURB-GFP immunoprecipitations 53
3.3.3. Testing bensonase RNA digestion 53
3.3.4. Mass spectrometry 54
3.4. Analysis 56
3.4.1. Mapping and comparing gene IDs 56
3.4.2. GO terms/STRING analysis 57
3.4.3. RNA-binding and RBD analysis 57
3.4.4. Drug interaction analyses 58
4. Establishing a system for researching the HIV-1 Rev protein under conditions approximating physiological infection 59
4.1. Introduction 60
4.2. Designing HIV-1 replicons expressing tagged Rev 61
4.3. Tagging of Rev does not affect HIV-1 gene expression 65
4.4. Assessing subcellular localisation of tagged Rev 67
4.5. Tagged Rev can effectively export HIV-1 gRNA 69
4.6. Discussion 72
4.6.1. Future Rev studies should utilise near-physiological systems 72
4.6.2. Rev-Halo is a powerful tool to study Rev localisation during infection 75
5. Revealing the landscape of Rev/host cell protein-protein interactions in HIV-1 infected T CD4+ lymphocytic cells 80
5.1. Introduction 81
5.2. Tagged Rev is efficiently purified using a Flag IP 83
5.3. The Rev interactome comprises 282 host cell proteins 92
5.4. Rev interacts with RBPs involved in splicing and translation 95
5.5. Testing the co-localisation of Rev-Halo and Rev interactors 101
5.6. Discussion 105
5.6.1. My Rev interactome addresses limitations of prior Rev IP studies 105
5.6.2. The shuttling of Rev and expression of vRNA is integral to Rev interactions 106
5.6.2. The Rev interactome: implications for Rev function 108
5.6.3. Future work 114
6. The Rev interactome overlaps heavily with the interactome of encapsidated HIV-1 RNA
6.1. Introduction 117
6.2. A substantial proportion of Rev interactors are encapsidated into HIV-1 virions 121
6.3. Rev delivers host cell proteins to viral complexes in the cytoplasm 125
6.4. Revealing the interactions of proteins encapsidated in HIV-1 virions 127
6.4.1. Introduction 127
6.4.2. Revealing the interactomes of encapsidated proteins during HIV-1 replication 128
6.5. Discussion 137
6.5.1. Rev acts as a molecular spiderweb to catch nuclear proteins for encapsidation 137
6.5.2. Rev promotes encapsidation of host cell proteins that are critical proviral factors 138
6.5.3. Staufen1 and Rev: assembling efficient sites of translation 139
7. General Discussion and future directions 142

7.1. Revisiting Rev functions in a more physiological system..... 142
7.2. Rev-mediated export may dictate protein packaging fate 144
7.3. Summary and future perspectives..... 147
Appendices..... 150
Bibliography..... 151

Abstract

Human immunodeficiency virus type I (HIV-1) is a retrovirus that causes acquired immunodeficiency syndrome (AIDS) and affects ~38 million people worldwide. The HIV-1 protein Rev escorts under-spliced viral RNAs to the cytoplasm, overriding nuclear retention imposed by the host cell. Thus, Rev permits full viral gene expression and particle production. However, Rev is challenging to study as it is highly disordered, expressed at low levels and encoded by an open reading frame that overlaps other viral genes. Therefore, previous Rev research was largely performed under non-physiological conditions, leaving the physiological interactions and functions of Rev uncertain. Here, I engineered HIV-1 constructs that encode tagged Rev protein expressed to physiological levels. I used these constructs to immunoprecipitate Rev interactors in CD4+ cells, revealing the physiological interactome of Rev. Through meta-analysis of this novel interactome and previously published datasets, I provide evidence for a new connection between Rev-mediated export and HIV-1 virion assembly. Furthermore, I investigate the interactions of two packaged proteins, PURA and PURB, during HIV-1 replication, to reveal possible therapeutic targets. My research uncovers the localisation, interactions and functions of Rev during HIV-1 replication under conditions better approximating physiological conditions than previously used. This work thus provides highly valuable datasets to instruct new avenues of research on the HIV-1 infection cycle.

List of Figures

Figure 1.1: The genome and proteome of HIV-1

Figure 1.2: The infection cycle of HIV-1

Figure 1.3: Rev structure and disorder

Figure 1.4: The RNA export cycle of Rev

Figure 4.1: Novel pNL4-3 replicons were engineered to express mCherry and tagged Rev proteins

Figure 4.2: Tagged Rev HIV-1 replicons lead to normal Gag/CA expression

Figure 4.3: Rev-Halo recapitulates the known nucleolar localisation of Rev and highlights spots at the nuclear membrane

Figure 4.4: Immunofluorescence and fluorescence *in situ* hybridisation of Rev-Halo and HIV-1 gRNA during HIV-1 replication

Figure 5.1: 3-way Venn diagram illustrating the overlap between Rev interactome studies

Figure 5.2: Immunoprecipitation of tagged Rev using different tags and their commercial antibodies

Figure 5.3: Rev-Flag is efficiently purified using an optimised Flag IP under conditions approximating physiological HIV-1 infection

Figure 5.4: The Rev interactome is enriched in known HIV-1 interactors

Figure 5.5: The Rev interactome is enriched in RBPs that are involved in vRNA metabolism

Figure 5.6: Rev interactors are mostly RBPs

Figure 5.7: Rev-Halo colocalises with ELMO1-GFP and is downregulated by ERGIC3-GFP

Figure 6.1: Uncovering host cell proteins encapsidated in HIV-1 virions

Figure 6.2: Comparison of the HIV-1 Rev interactome and ivRBPome

Figure 6.3: Proteins common to the Rev interactome and ivRBPome are RBPs that modify RNA and emerge from the nucleus

Figure 6.4: There is a high degree of conservation between proteins of the Rev, Staufen1 and ivRIC interactomes

Figure 6.5: PURA-GFP and PURB-GFP interactomes revealed during HIV-1 replication

Figure 6.6: PURA-GFP and PURB-GFP interactions may underpin proviral effects

Figure 7.1: Rev acts as a molecular spiderweb during HIV-1 infection.

List of Tables

Table 1.1. A summary of the physiological functions that Rev has been posited to influence, except RNA export

Table 2.1. Chemicals and reagents used in this study

Table 2.2. Consumables used in this study

Table 2.3. Machines used in this study

Table 2.4. Buffers used in this study

Table 2.5. Commercial kits used in this study

Table 2.6. Enzymes used in this study

Table 2.7. Antibodies and dyes used in this study

Table 2.8. Plasmids used in this study

Table 2.9. Oligonucleotides and primers used in this study

Table 2.10. Parental cell lines used in this study

Table 2.11. Genetically modified cell lines used in this study

Table 3.1: Reaction mix for restricting HIV-1 vector

Table 3.2: Reaction mix for ligation

Tables 3.4, 3.5: PCR thermocycles and reagent mixture used to generate fusion PCR products

Tables 3.5, 3.6: Reaction mix for amplifying RBP ORFs by PCR

Table 3.7: RBDs classified as ‘classified’ in domain analysis

Table 5.1. Experimental specifications of the 3 publicly available Rev IP studies

Table 5.2. HNRNPs and SR proteins identified in the Rev-Flag interactome

Table 6.1. The 38 proteins shared between Gag, ivRIC, Staufen1 and Rev interactomes

Abbreviations

Å - Angstrom

aa – Amino acid

AEBSF – 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride

AIDS – Acquired immunodeficiency syndrome

AMD – ARE-mediated decay

AP – Affinity purification

ARE – AU-rich element

ARM – Arginine-rich motif

ART – Anti-retroviral therapy

bp – Base pairs

CA – HIV-1 Capsid

CBC – Cap-binding complex

cDNA – Complementary DNA

cm – Centimeters

CTE – Constitutive transport element

CTD – C-terminal domain

DAPI - 4',6-Diamidino-2-Phenylindole, Dihydrochloride

DMEM – Dulbecco's modified eagle medium

DNA – Deoxyribonucleic acid

ds – Double-stranded

DTT – Dithiothreitol

FASP – Filter-aided sample preparation protocol

FBS – Foetal bovine serum

GO – Gene ontology

gRNA – Genomic RNA

h – Hours

HERV-K - Human endogenous retrovirus family

HGNC – Hugo gene nomenclature committee

HIV-1 – Human immunodeficiency virus type I

HNRNP - Heterogeneous nuclear ribonucleoproteins

hpi – Hours post-infection

hpt – Hours post-transfection/transduction

IDR – Intrinsically disordered region

IF – Immunofluorescence

INS – Instability

INT – HIV-1 Integrase

IP – Immunoprecipitation

IRES – Internal Ribosome entry site

(iv)RIC – (In-virion) RNA interactome capture

kb – Kilobase

Kd – Dissociation constant

kDa – Kilodalton

KH – K-homology

LTR – Long-terminal repeat

M – Molar

MA – HIV-1 Matrix

min – Minutes

ml - Millilitre

mM – Millimolar

mRNA – messenger RNA

MS – Mass spectrometry

m/z – Mass-to-charge ratio

NC – HIV-1 Nucleocapsid

NCBI – National center for biotechnology information

NEB – New England biolabs

NES – Nuclear export signal

ng – Nanograms

NLS – Nuclear localisation signal

NTD – N-terminal domain

nm - Nanometers

nM – Nanomoles

NMD – Nonsense-mediated decay

NPC – Nuclear pore complex

ORF – Open reading frame

PBS – Phosphate-buffered saline

PBST – Phosphate-buffered saline with 0.1% Tween20

PBSTX – Phosphate-buffered saline with 0.1% TritonX-100

PCR – Polymerase chain reaction

PFA – Paraformaldehyde

pg - Picograms

PIC – Pre-integration complex

pmol – Picomoles

PPI – Protein-protein interaction

PR – HIV-1 Protease

PS – Penicillin/Streptomycin

RBD – RNA-binding domain

RBP – RNA-binding protein

RBPome – RNA-binding proteome

RNA – Ribonucleic acid

RNP – Ribonucleoprotein

rpm – Revolutions per minute

RPMI – Roswell park memorial institute

RRE – Rev-response element

RRM – RNA recognition motif

rRNA – Ribosomal RNA

RT – HIV-1 Reverse Transcriptase

RTC – Reverse transcription complex

SDS – Sodium dodecyl sulfate

SG – Stress granule

smFISH – Single molecular fluorescence *in situ* hybridisation

sn – Small nuclear

SR – Serine-arginine

SRP – Signal recognition particle

ss – Single-stranded

SSC – Saline-sodium citrate solution

SU - Surface

TAE – Tric base, acetic acid and EDTA

TM - Transmembrane

$\mathbf{T_m}$ – Annealing temperature

tRNA – Transfer RNA

µg – Microgram

µl – Microlitre

UV – Ultraviolet

V - Volts

VSVG – Vestitular stomatitis virus glycoprotein

vRNA – Viral RNA

xg – Times gravity

ZF – Zinc finger

°C – Degree centigrade

Acknowledgements

I would like to thank Alfredo Castello for his unending guidance and constant optimism. For giving me opportunities, such as presenting at conferences and writing reviews, and for his help in preparing this manuscript. Thank you for the opportunity to grow as an intellectual and for always reminding me, even in 5-minute conversations, why we should remain passionate about science. To any future DPhil studies considering candidacy in the Castello lab, do it. You don't often get to work for a genius!

I would like to thank Ilan Davis for his advice, support and kindness. Thank you to all members of the Castello and Davis labs for fielding my many questions and helping me always. Special thanks to Aino Jarvelin, who led me to coding, and to Honglin Chen and Shabaz Mohammed for all their help with mass spectrometry experiments. Thanks to Marko Noerenberg, my favourite scientist who I hope recruits me when he is a world-famous PI. Thanks to Manuel Garcia-Moreno for all his amazing work on helping me unveil the mysteries of Rev and for building a manuscript with me.

To my adventuring companions Elly, Will, Simon, Charlie and Scott: our stories are my favourite part of Oxford. I cannot thank you enough for giving me characters to escape into. Will happen, happening, happened. Special thanks to Wills, who filled these last 2 years with belly-laughs and late-night conversations that kept me sane.

Thanks to my very best friend in the whole wide world and future folk star, Chuck. Our flat in Oxford is my Good Place and I cannot imagine a more supportive, beautiful, and

kind human as you. Thank you for every single thing you do. I am so proud of you. I will stop hiding behind doors to scare you.

Thanks to my dad Chris, who doubles as my best friend. Weekends away at yours were a magical portal from the stress of the lab to a room filled with laughter, video games and Rush playing on some very high-quality speakers. No one has done more for me than you. You're the best Dad in the world, and I dedicate my doctorate to you. Also to Luke.

Thanks to my mum Toni, who has always been proud of me. Thank you for all your support, and for caring even when I made you do it from afar. When I think about my passion for science, I think about the small anatomy manakin you used to teach me about physiology. I am so lucky you were my first teacher, I wouldn't be here otherwise.

1. Introduction

Human immunodeficiency virus type I (HIV-1) is a retrovirus that infects CD4+ T lymphocytes and macrophages, leading to a gradual loss of CD4+ cells and subsequent immunodepression termed acquired immunodeficiency syndrome (AIDS). As of 2019, ~38 million people globally were living with HIV-1 (1). HIV-1 infection is currently treated using life-long anti-retroviral therapy (ART) (2), which suppresses but does not fully eliminate the virus. Widely used first-line ART include a cocktail of compounds that target the viral enzymes, particularly protease (PR), reverse transcriptase (RT) and integrase (INT) (2). The inability to regularly monitor viral loads and the limited availability of ART in developing nations has resulted in HIV-1 positive individuals adhering to sub-optimal ART combinations (3). Ultimately, this has caused resistant viral serotypes to emerge (4). In the search for new therapies, scientists are expanding their interest towards other viral and cellular proteins. The topic of this study is Rev, an essential yet poorly understood HIV-1 protein. Discerning the interactions and functions of Rev during physiological infection will advance understanding of the HIV-1 infection cycle and, critically, may reveal new therapeutic targets. Rev is also a functional homolog of other viral factors such as human endogenous retrovirus family (HERV-K) Rec protein and human T-cell leukaemia virus Rex protein (5). Therefore, knowledge of Rev functions and interactions in the viral cycle will likely inform the functions of these other viral factors.

1.1. The HIV-1 infection cycle

1.1.1. The HIV-1 genome is alternatively spliced

The HIV-1 genome is a single-stranded (ss) RNA molecule of approximately 9.2 kilobases (kb) and is carried inside a HIV-1 virion as a non-covalent dimer (6). Upon entry into a new cell, the HIV-1 genome is reverse transcribed into DNA, irreversibly integrated into the host cell genome and then transcribed by host RNA polymerase II via a promoter in the long terminal repeat (LTR) region. The full-length genome encodes 9 open reading frames (ORFs) that generate at least 15 mature proteins through polyprotein processing by HIV-1 protease (PR) and alternative splicing (**Figure 1.1, A**). Specifically, there are 3 levels of splicing based on clustering of donor and acceptor splice sites in 2 major intronic regions in the HIV-1 genome (**Figure 1.1, B**). The polycistronic ~9 kb unspliced genomic (g)RNA encodes the Gag protein and Gag-Pol polyprotein, which are the structural and enzymatic proteins required for virus formation and propagation respectively (7). Gag-Pol is produced by a ribosomal frameshift event, which regulates the stoichiometric ratio between Gag-Pol and Gag proteins (6). These polyproteins are proteolytically processed to produce the viral enzymes PR, reverse transcriptase (RT) and integrase (INT), as well as structural components matrix (MA), capsid (CA), nucleocapsid (NC) and adaptor protein p6. The ~4 kb singly-spliced species, which are generated by the removal of the major 5' intron, encode the envelope transmembrane (TM) and surface proteins (SU) as well as the auxiliary factors Vpr, Vpu and Vif. Finally, the ~2 kb fully spliced RNAs encode for the regulatory proteins Tat, Rev and Nef and have both 5' and 3' introns removed (17). The full complement of HIV-1 splice products is not fully known yet. Clinical isolates of HIV-1 produce over 100 processed viral (v)RNAs by alternative splicing (8), and novel proteins formed through cryptic splice sites have been regularly identified since the

establishment of the HIV-1 genome (9-11). This splicing process is highly specific. More frequently recognised splice sites outcompete suboptimal sites in the HIV-1 genome to generate a specific stoichiometry of RNAs (12-14). The efficiency of these sites is also highly moderated by the binding of host cell proteins to *cis*-acting signals in the genome (15-17). Changes in the expression of these splicing factors correlate with viral protein levels, suggesting that HIV-1 expression is regulated through recruitment of RNA-binding proteins (RBPs) to vRNA (18). Changes in splice site efficiency and host cell splicing factor expression drastically impair HIV-1 infection (19), demonstrating the need for a defined stoichiometry of HIV-1 RNAs.

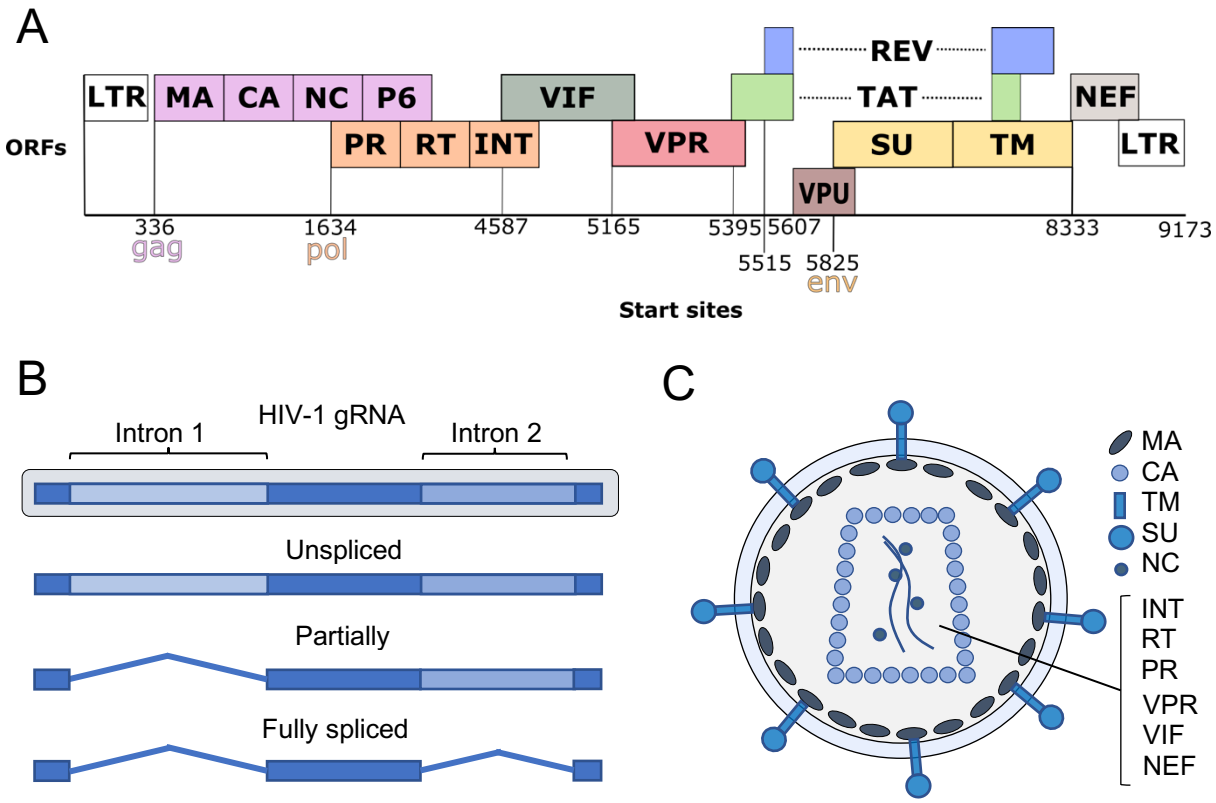


Figure 1.1: The genome and proteome of HIV-1. (**A**) Schematic of the HIV-1 NL4-3 genome. Start sites of ORFs are numbered according to the NL4-3 HIV-1 sequence. (**B**) Schematic of the major introns (dark blue bars) and exons (light blue bars) in NL4-3 HIV-1 gRNA and the 3 levels of splicing used to produce all mRNA transcripts. Spliced regions are represented as bent lines. (**C**) Schematic of a mature HIV-1 virion. Proteins are labelled according to the legend.

1.1.2. HIV-1 virion structure and proteome

HIV-1 virions are spherical particles approximately 100 nanometres (nm) in diameter (**Figure 1.1, C**) (20). The outer lipid membrane, which protects the inner CA core, is derived from the plasma membrane of the cell from which the immature virus budded (henceforth called the 'producer' cell). Viral envelope proteins are embedded in the outer membrane of the virion and form 'spikes' on its surface. Envelope TM trimers are anchored to epitope-carrying envelope SU trimers, which bind to CD4 receptors and chemokine co-receptors CCR5 or CXCR4 on the surface of immune cells. The Gag-derived MA protein sits beneath the lipid membrane, tethered to envelope proteins, and conceals the inner CA-coated core. This core is comprised of approximately 1500-2000 Gag-derived CA proteins assembled into pentamers and hexamers (21). Critically, the CA protects 2 molecules of non-covalently bound HIV-1 gRNA, which are in turn bound by viral enzymes RT and INT. Encapsidation of these enzymes into virions in the producer cell is crucial, as early infection steps such as reverse transcription must occur in the protected virion core to avoid host cell immune sensors (22).

HIV-1 particles exist in two morphological states. Immature virions are not infective and contain a spherical core comprised of the Gag polyprotein. Cleavage of Gag/Gag-Pol by PR triggers a morphological rearrangement to produce the fully formed complement of Gag/Gag-Pol derived proteins, which assemble to produce a cone-shaped inner CA core. This assembly marks the maturation of the fully infective virus. Many host cell chaperone proteins, such as heat shock proteins, have been found in viral cores (23-25). Since these proteins facilitate protein folding and assembly, they may be recruited to aid the maturation of viral particles and assembly

of cleaved Gag subunits.

1.1.3. The HIV-1 infection cycle

HIV-1 infection begins with the attachment of viral envelope SU glycoproteins to CD4 receptors on immune cells (**Figure 1.2**). Concomitant with CD4 binding, SU proteins also bind to a chemokine co-receptor, CCX5 or CCRC4. This permits a conformational change in the attached TM envelope protein, which subsequently embeds into the plasma membrane of the host cell, triggering membrane fusion (26). The CA core of the virion, enveloping 2 copies of the RNA genome and viral proteins, enters the cytoplasm of the host cell. The core then travels towards the nucleus, viral RT transcribes the positive-sense ssRNA genome to double-stranded (ds) proviral DNA (27). HIV-1 gRNA is bound by a cellular $tRNA^{Lys,3}$ molecule to prime the first step of complementary (c)DNA synthesis. RT simultaneously degrades the RNA genome due to the presence of RNase H activity (27). At this point, the vRNA/ribonucleoprotein (RNP) complex is termed the 'reverse transcription complex' (RTC). The RTC is comprised of the viral genome and many other viral proteins, such as NC, which facilitates reverse transcription through chaperone activity (28). In later stages of reverse transcription, this transitions into the pre-integration complex (PIC) which is distinct in protein composition and is marked by changes in the CA core structure (29).

While reverse transcription is believed to be initialised in the cytoplasm, the location and timing of CA disassembly, which liberates reverse-transcribed vDNA, is under debate. CA uncoating was long believed to happen in the cytoplasm, due to the CA core size (~60 nm (30)) exceeding the maximum entry size of nuclear pore complexes (NPCs) (31). Two popularised models envisioned that the CA core either (a) disassembled immediately after entering the host cell cytoplasm or (b) partially

uncoated in the cytoplasm, but remained associated with the PIC up to the nuclear pore (22). However, recent evidence suggests that core uncoating may primarily occur in the nucleus of infected cells, to release reverse transcription products near integration sites (32). This model is supported by a biological necessity to shield viral nucleic acid intermediates from innate receptors in the cytoplasm (33). Furthermore, CA molecules have been shown to interact with members of the NPC (34, 35), suggesting a role in nuclear entry. CA structures, including PICs associated with CA proteins, have recently been identified in the nucleus (29, 32, 36, 37). Moreover, recent data identified morphological changes in CA proteins which allow the CA-associated PIC to fit through NPCs (29). This model necessitates that all viral and host cell factors necessary for early HIV-1 gRNA metabolism must be encapsidated inside the virion during assembly in the producer cell. If true, a thorough characterisation of the HIV-1 virion proteome could uncover which host cell proteins are required for efficient replication, particularly in early phase infection.

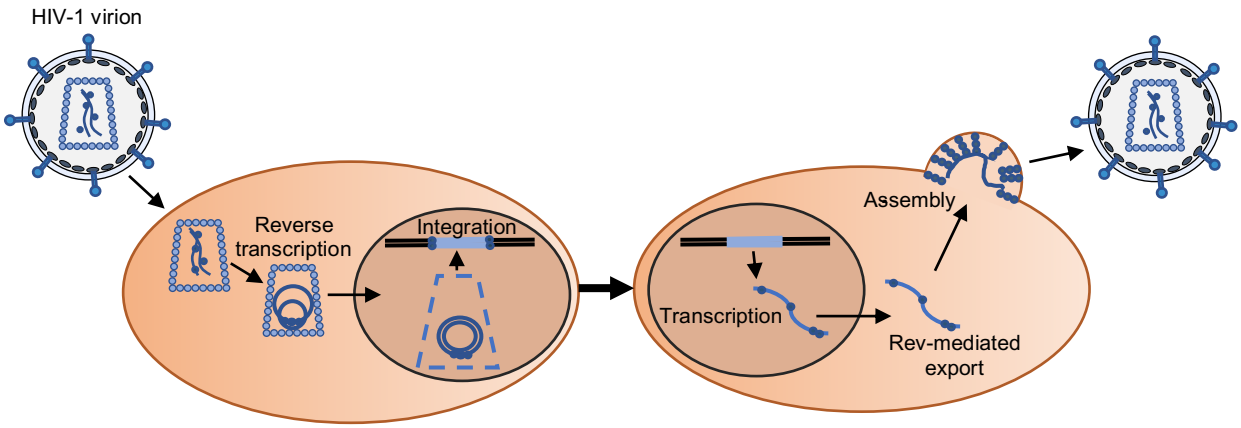


Figure 2.2: The infection cycle of HIV-1. Schematic of the HIV-1 infection cycle, which begins with HIV-1 virion (top left) binding to receptors on CD4+ immune cells. Following membrane fusion, the viral capsid core is deposited into the cytoplasm and the HIV-1 RNA genome is reverse transcribed by viral Reverse Transcriptase. The viral DNA is then integrated into the host cell genome in the nucleus and transcribed to generated mRNAs. Viral mRNAs are exported to the cytoplasm by Rev, translated and finally assemble at the plasma membrane. The host cell endosomal sorting complex required for transport then performs membrane scission, allowing a new particle to bud off from the infected cell.

Following core disassembly, the viral ssDNA genome is integrated into the host genome by viral INT, and after a period of indeterminate latency, host RNA polymerases transcribe vRNAs to produce differentially spliced vRNAs. These RNAs are exported to the cytoplasm by cellular pathways and by the viral Rev protein (**section 1.2**) where they are translated. HIV-1 gRNA acts as both messenger (m)RNA for translation and a genome to be packaged in progeny virions. How these nucleic acids are distinguished is unclear. Nevertheless, nascently translated Gag polyprotein and the HIV-1 gRNA interact with each other to regulate late-phase assembly. Firstly, newly synthesised Gag polyprotein undergoes oligomerization in the cytoplasm (38) and binds to target gRNA dimers at a packaging signal known as *psi* at the 5' end of the RNA (39) using it's RNA-binding NC domain. This HIV-1 RNP initiates at juxtanuclear and cytoplasmic sites (40). It is then transported to HIV-1 assembly sites at the plasma membrane for packaging into viral particles (41). How the Gag/vRNA complex reaches the plasma membrane is unknown, but live imaging studies suggest that exposure of a myristate group on the MA domain (which occurs only after binding to vRNA (42)) controls this targeting (43) and is responsible for anchoring Gag to the plasma membrane during viral assembly (44). Gag multimerisation continues at the plasma membrane. The CA domain stabilises these interactions (45) and the p6 domain recruits subunits of the endosomal sorting complex required for transport to carry out membrane fission. This allows a new particle of HIV-1 to bud off from the infected cell (46, 47). Inside the virion, HIV-1 PR cleaves the Gag polyprotein to form individual viral proteins, which spontaneously assemble into a mature CA core, forming an infective particle.

1.2. The HIV-1 Rev protein

1.2.1. Rev discovery and structure

Rev was discovered through mutations of the overlapping HIV-1 Tat gene, which specifically increased the production of fully spliced vRNA and eliminated production of under-spliced transcripts (48). These effects could not be reversed by *in trans* complementation with Tat, suggesting that they were Tat-independent (49). They were attributed to a novel HIV-1 gene, initially termed trs/art, later named Regulator of Expression of Virion proteins (Rev). Rev comprises 116 amino acids and is highly disordered (**Figure 1.3, A**)**.** It harbours a folded N-terminal domain (NTD) and unstructured C-terminal domain (CTD). The 65-residue NTD adopts an anti-parallel helix-turn-helix organisation and encodes an arginine-rich RNA-binding motif (ARM) encompassing a nuclear/nucleolar localisation signal (NLS) and two flanking oligomerisation motifs, which allow Rev to multimerise (**Figure 1.3**) (50). Crystal structures have shown that Rev self-associates through 3 types of homotypic interactions termed A-A, B-B and C-C interfaces (51-53). A-A and B-B interactions occur when same-face alpha helices associate in a V motif (**Figure 1.3, B**) and exhibit some degree of structural malleability. C-C pairings occur at the loop ends of alpha helices by a proline interaction and exhibit restricted flexibility (51-53) The disordered nature of the Rev CTD has made it challenging to characterise biochemically and structurally, although it has been proposed to fold into beta sheets in certain contexts such as filament formation (52, 54, 55). Notably, the CTD encodes a leucine-rich nuclear export signal (NES), sometimes referred to as the ‘activation domain.’ Rev employs its NES and NLS to traverse the nuclear pore and move between the nucleus

and cytoplasm. This shuttling ability is critical for allowing Rev to export under-spliced vRNAs and return to the nucleus (56).

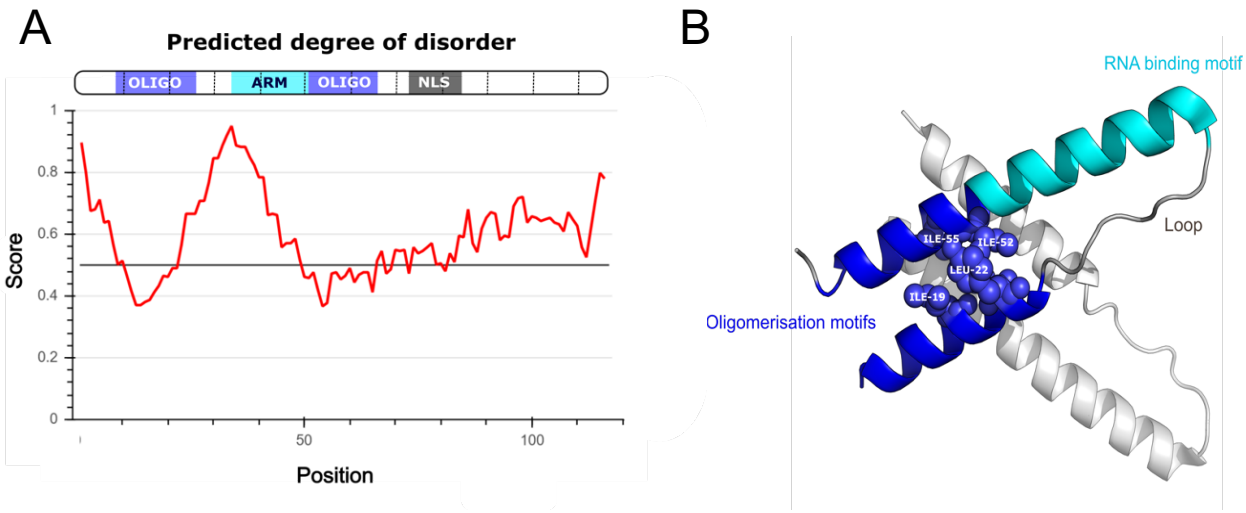


Figure 1.3: Rev structure and disorder. (**A**) The Rev protein is highly disordered, as predicted using IUPred2 (57) where a score >0.5 indicates disorder. (**B**) The Rev NTD is composed of a helix-turn-helix motif and dimerises in a V-like shape. An ARM is flanked by oligomerisation motifs, which allow Rev to multimerise. NTD helices are stabilised by core hydrophobic interactions; some contributing residues are labelled (PDB: 6BSY).

1.2.2. Rev as an RNA export factor

HIV-1 RNA is considered fully spliced when the two large intronic sequences present in the genome are removed. These fully spliced transcripts leave the nucleus using the cellular TAP/NXF1-mediated pathway (58). However, under-spliced transcripts are retained in the nucleus, in an analogous manner to unspliced cellular mRNA (59, 60). To bypass nuclear retention, retroviral under-spliced RNAs harbour regulatory elements that recruit the cellular export machinery either directly or through an adaptor viral protein. For example, Mason-Pfizer monkey virus RNA harbours a constitutive transport element (CTE) that allows direct recruitment of TAP/NXF1 export machinery (61). By contrast, under-spliced HIV-1 RNAs contain an RNA structure termed the Rev response element (RRE) which recruits the viral export factor Rev (62). Rev is expressed from a fully spliced RNA and binds to the RRE to elicit the export of the

under-spliced vRNAs to the cytoplasm through the recruitment of host factors (**Figure 1.4, A**). Once Rev protein is produced, it is imported into the nucleus by direct interaction with cellular importins beta, transportin, importin 5 and importin 7 through its NLS (63, 64). Ribosome assembly factor B23 has been shown to aid in this import (65, 66). In the nucleus, Rev interacts with the 351-nucleotide (nt) RRE in the second intron of under-spliced (~9 and ~4-kb) vRNAs (67-75) with high affinity (reported Kd = 0.3-5 nanomolar (nM)). The arginine residues of the ARM motif aligned in an alpha helix insert into a major groove in the RNA (**Figure 1.4, B**) (71, 73, 76-79).

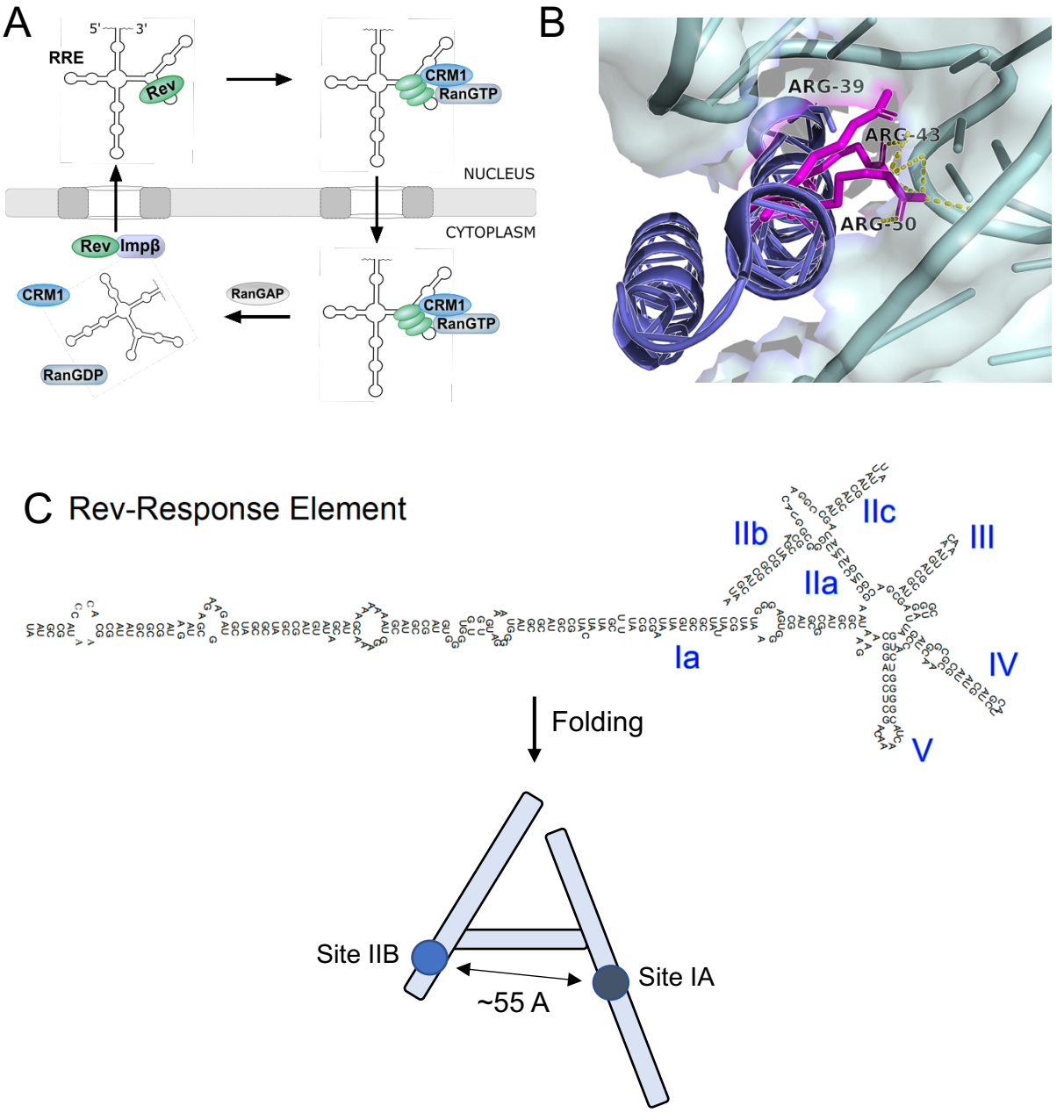


Figure 1.4: The RNA export cycle of Rev. (**A**) Rev binds the Rev response element on under-spliced RNA at stem IIB, multimerises and recruits CRM1 and RanGTP to cross the nuclear pore complex. In the cytoplasm, this complex is dissociated by RanGAP promoting RanGTP hydrolysis, freeing viral RNA. Rev re-enters into the nucleus by binding importin-ß. (**B**) A Rev dimer was crystallised in complex with the Rev response element; Rev binds stem IIB of the RRE using its arginine-rich motif, which contacts the negatively charged RNA backbone (PDB: 4PMI (53)). (**C**) The structure of the 351-nt Rev response element; regions are named using Roman numerals. The 'A-like' folded conformation of the RRE is shown, as determined by using X-ray scattering experiments (80).

How Rev assembles on the RRE remains unclear. Some groups postulate that Rev binds the RRE as a monomer and oligomerises thereafter in a co-operative manner along lower affinity sites (70, 73, 77, 78, 81-85). Other groups suggest that Rev exists as a multimer in solution and binds the RRE as such, where gel shift and fractionation experiments show Rev-RRE interacting at distinct stoichiometries (63, 71, 72, 76, 86, 87). Technological advances have led to a series of recent Rev and RRE structures, which provide clues to Rev-RRE stoichiometry. 3D small-angle X-ray scattering reconstructions show that the RRE adopts an 'A-like' topology (80), where the primary binding site for Rev on stem loop IIB maps to one of the 'A' legs, and site IA, a posited secondary Rev binding site (83, 88), maps to the other leg (**Figure 1.4, C**). These sites are separated by ~55 angstrom (Å), which matches the span of a Rev dimer (80). It has been posited that a Rev dimer first binds the RRE at these sites; when either leg of the 'A' shape is removed, the RRE is non-functional, supporting this model (89). Similarly, when the 'crossbar' of the A is extended, spanning greater than 55 Å, Rev/RRE binding is strongly impaired (89). A different model envisions 2 Rev dimers associating through a C-C interface, each dimer binding the RRE separately at sites stem IIB and stem IA (52). This model is supported by a new tetramer crystal structure, which indicates that B-B Rev dimers bind through C-C interfaces to form tetramers, which bind stem IIB and IA sites (90). Nucleotides G47, G48 and A73 in stem loop IIB

of the RRE form non-canonical base pairs that present a wider groove to facilitate Rev binding. Crystal structures have identified that these ‘wobble’ pairings likely help to orient the interaction of a second Rev monomer, allowing it to reach additional junction sites along stem II (53). Accordingly, mutation of these nucleotides reduces the Rev-stem loop IIB interaction, increasing the dissociation constant by over 2 μM in the case of A73 (78). Regardless, Rev oligomerisation appears to be a pre-requisite for successful RNA export since (i) Rev mutants unable to oligomerise are defective for export (67, 87, 91, 92) and (ii) stem loop IIB alone is insufficient to trigger RNA export, suggesting a requirement for oligomerisation along low-affinity binding sites (83, 93).

After Rev binds the RRE, host cell exportin chromosome region maintenance gene 1 (CRM1/XPOI) and RanGTP are recruited to the Rev NES (94, 95). Rev interacts with CRM1 non-canonically; a recent electron microscopy structure indicates that CRM1 interacts with Rev as a dimer (96). Crystal structures show that the Rev NES binds CRM1 in a linear, unorthodox fashion, where the NES is spaced out using proline residues to reach all 5 hydrophobic binding pockets of CRM1 (97). It also appears that RRE-bound Rev dimers spatially orient their CTDs to optimally recruit CRM1 (90). In this RNA-bound state, Rev’s NLS is occluded while the NES remains accessible, facilitating CRM1 recruitment and export (63). While no other proteins were known to contribute to Rev/CRM1-mediated export for many years, proteins have been identified recently which may be part of this complex, including phosphoproteins (98), RNA helicases (99, 100), nucleoporins (101), and additional factors (102-105). The Rev/vRNP complex then traverses the hydrophobic channel of the nuclear pore and, once in the cytoplasm, disassembles upon RanGTP hydrolysis. Liberated vRNAs can then be translated and Rev can be recycled back to the nucleus by importin recruitment, which occurs exclusively following RNA dissociation (63). In this way, Rev

tightly couples the transcription of vRNAs with their nuclear export, enabling expression of all viral genes in the right quantities and at the right time.

1.2.3. Additional functions of Rev

Blocking host cell export pathways

In addition to exporting vRNAs from the nucleus, Rev may drive under-spliced HIV-1 RNAs towards the CRM1 export pathway. Rev reduces the levels of cellular and vRNAs exported by CRM1-independent pathways by blocking TAP/NXF1, the main pathway which exports the bulk of cellular mRNAs. Taniguchi *et al* showed that Rev inhibits the TAP/NXF1-specific export of RNAs containing the RRE (106). Overexpression of TAP/NXF1 components decreased levels of under-spliced vRNA, an effect that was reversed by Rev overexpression (106). This suggests that Rev downregulates the TAP/NXF1 pathway, perhaps to bypass cellular checkpoints that induce the degradation of under-spliced vRNA. Since fully spliced viral mRNAs such as *rev* use TAP/NXF1 for nuclear export, accumulation of Rev protein may lead to the suppression of the TAP/NXF1 export pathway to increase the ratio of under-spliced to spliced transcripts in the cytoplasm. Rev was proposed to interfere with the association of TAP/NXF1 components, such as ALY/REF, with RNA (106). Although the exact mechanism remains unknown, Taniguchi *et al* propose a model in which RNA-bound Rev interacts with the nuclear cap-binding complex (CBC) to inhibit its interaction with ALY/REF and subsequently suppress downstream TAP/NXF1 export (106). Rev can indeed bind subunits of the nuclear CBC (104), which lends credence to this model.

Stabilisation of RNA

Rev may have a role in stabilising vRNAs. Rev binding to the RRE appears to overcome the inhibitory effects of instability ('INS') sequences in HIV-1 RNAs. These INS regions have been identified in the *env, gag* and *pol* coding regions and promote nuclear retention, instability and reduced polysome loading of vRNA (107-112), independently of splicing. INS regions are AU-rich, which confers a different codon usage to that of human mRNAs (113-115). The presence of non-optimal codons promotes mRNA instability, likely due to ribosome stalling and inefficient translation at these sites. The mutagenesis of an AU-rich INS site in *gagpol* increased protein expression by ameliorating steady-state mRNA levels (116, 117), illustrating the effects of these regions. High A and low C content produces a codon bias on HIV-1 RNA that may also decrease protein expression because of a lack of cognate tRNAs in steady-state cellular conditions, inducing long 'waiting times' for the ribosome to engage with the correct transfer (t)RNA (118, 119). Indeed, circumventing INS sites with codon-optimization increases HIV-1 envelope protein expression, independently of effects on RNA export or stability (115). The Rev-RRE interaction seems necessary to counteract INS-mediated effects (107-111), since mutations nullifying INS regions can switch HIV-1 protein expression from Rev-dependent to Rev-independent (116, 117, 120). How Rev helps to overcome INS signatures is unknown. However, it has been demonstrated that enhanced expression of codon-optimized *gag* and *pol* genes results predominantly from an increase in cytoplasmic mRNA (117, 120). When cells were pre-treated with leptomycin B, an inhibitor of Rev-RRE export partner CRM1, expression of Rev-dependent HIV-1 proteins was significantly reduced. Expression of codon-optimized counterparts, however, was not affected (75). This suggests that codon-optimization allows HIV-1 RNAs to use other export pathways, triggering this increase in cytoplasmic RNA. Thus, INS regions may cause nuclear retention of HIV-

1 RNAs, which Rev-RRE binding and export overcomes. It has also been proposed that certain host proteins can bind AU-rich INS sequences, and that these interactions regulate HIV-1 RNA stability (121-125).

Splicing inhibition

Rev may inhibit splicing in an RRE-dependent manner. The Rev ARM peptide inhibited splicing of RRE-containing RNA almost 15-fold more than control RNA lacking the RRE (126). Incubation of a ß-globin-RRE pre-mRNA with the ARM peptide across different time points led to splicing inhibition only when Rev was added early, suggesting that the Rev/RRE interaction interferes with initial spliceosome assembly on the RNA (126). Fractionation and sucrose centrifugation of these splicing extracts revealed that addition of Rev to RRE-containing pre-mRNA removed the 60S peaks corresponding to fully assembled spliceosomes (126), and caused an accumulation of a 45-50S splicing-deficient complex. The same group discovered that the Rev ARM blocked binding of U4/U5/U6 tri-snRNP in an RRE-dependent manner (127). Rev may interact with host cell splicing factors to mediate these effects. A recent study identified proteins responsible for HIV-1 RNA nuclear retention through genome-wide CRISPR/Cas9 knock-out experiments. The majority of identified proteins were factors involved in pre-mRNA splicing and associated with the spliceosome (128). Rev may thus suppress early spliceosome formation by interacting with these factors, promoting intron retention. For example, SF2/ASF is an essential splicing factor that binds RNAs with a 5' splice site to aid U1 snRNP docking (129). SF2/ASF binds the Rev-bound RRE *in vitro* and its overexpression can inhibit Rev function and HIV-1 gene expression in a dose-dependent manner (129). It also regulates HIV-1 gene expression (17, 130, 131). Rev could sequester this splicing factor, alongside other

spliceosome-associated components, to stop full spliceosome formation. Importantly, SF2/ASF overexpression alters the alternative splicing pattern of HIV-1 RNAs (132). Rev may therefore bind this factor during HIV-1 infection to ensure optimal stoichiometry of HIV-1 RNAs. Infection increases the proportion of introns globally within cellular RNAs in primary T cells (133), aligning well with this inhibitory activity of Rev on SF2/ASF. Collectively, these posited roles paint a picture of Rev working against host cell regulation to repress splicing and promote the expression of under-spliced HIV-1 RNA.

Translation

Rev has been posited to influence the translation efficiency of vRNA. Early research highlighted disproportionate increases in viral envelope protein expression relative to total mRNA levels in Rev and Tat-transfected cell lines (134). Additional studies confirmed large discrepancies between cytoplasmic *gag* mRNA and protein levels in the absence of Rev (135-138), suggesting that Rev may affect translation. This may be through increasing the association of vRNAs with polysomes. One study found that 90% of cytoplasmic, singly-spliced *env* RNA associated with monosomes when infected with HIV-1 lacking Rev (HIV-$1_{Rev(-)}$), whereas over 75% associated with polysomes when infected with wildtype Rev HIV-1 (136). Conversely, polysome association with fully spliced vRNAs such as *tat* was unaffected by the presence or absence or Rev, indicating efficient assembly of ribosomes on these vRNAs in a Rev-independent manner. Interestingly, Rev-dependent under-spliced vRNA/polysomes association was found to be dependent on the presence of the RRE, suggesting a requirement for the Rev/RRE interaction (135). However, tethering the leucine-rich domain of Rev, which interacts with CRM1, to HIV-1 RNA allowed efficient Gag

production in the absence of Rev (139). These results indicate that CRM1-mediated export alone is sufficient for triggering HIV-1 RNA translation downstream through an export/translation coupling mechanism. Intriguingly, a conserved Rev-binding site was discovered in the 5' UTR of HIV-1 RNA (140). Rev enhanced the translation of RNA reporters containing this site *in vitro* (141). However, these findings were not recapitulated by the same group in COS-1 cells (142). Whether Rev controls HIV-1 RNA translation and if so, the mechanism by which it does so remain ultimately unclear.

Encapsidation

The Rev/RRE interaction has been proposed to promote vRNA packaging (143-146). The ratio of HIV-1 gRNA in virions versus cytoplasmic gRNA was measured using a HIV-1 chimeric construct containing the RRE in the presence or absence of Rev (147). Absence of Rev caused a 10-fold decrease in genome packaging compared to Rev inclusion (147). However, the RRE-RNA construct used did not contain canonical INS sites, and therefore was probably not dependent on Rev for nuclear export, unlike wildtype HIV-1 gRNA. Thus, the same group repeated this work with an almost full-length HIV-$1_{Rev(-)}$ expression plasmid (148). They found that inclusion of Rev induced a 4,500-fold increase in packaged HIV-1 gRNA, and only increased cytoplasmic gRNA levels 5-fold (148). The same enhancement in RNA packaging was observed when a chimeric HIV-$1_{Rev(-)}$ genomic RNA including MS2 stem-loops was co-transfected with MS2-TAP (148), suggesting that is not necessarily Rev, but the export process itself which boosts gRNA packaging. Regardless, the Rev-RRE interaction was more efficient at promoting packaging than MS2-TAP (146, 147). The mechanistic details of how Rev confers increased genome encapsidation remain unknown. It is possible that certain export pathways lead to the formation (or avoidance) or specific RNP

complexes which may promote or hinder downstream packaging. Indeed, helicase DDX24 has been shown to directly interact with Rev and increase RNA packaging only in the context of Rev/RRE export (149).

Table 1.1. A summary of the physiological functions that Rev has been posited to influence, except RNA export.

Posited Rev function	References
Rev may block the TAP/NXF1 pathway	(106)
Rev may stabilise vRNAs	(109, 110)
Rev may inhibit splicing	(126, 127)
Rev may promote translation of under-spliced vRNAs	(134-138)
Rev may promote encapsidation of the viral genome	(143-148)

1.2.4. Rev as a therapeutic target

Rev is critical for HIV-1 gene expression and therefore represents a potential antiviral target. Despite this, there are currently no Rev-based therapeutics in clinical use. It has long been known that dominant negative mutants of Rev can abrogate wildtype Rev function in lymphocytic cells (150, 151). For example, 'Rev M10' contains 2 point mutations in the Rev NES which completely abrogates its trans-activation function, while competing with wildtype Rev for binding to the RRE (91). M10 has been investigated in clinical trials (152-154). However, it is challenging to deliver into cells, and resistant strains of HIV-1 with altered RREs arose after constitutive M10 expression (155).

Other therapeutic approaches target the Rev/RRE interaction itself using small molecules. An early iteration of this approach used aminoglycoside antibiotics which specifically bound to the RRE at the Rev binding site (156), blocking Rev's interaction and inhibiting HIV-1 gene expression (157). Since then, a series of small molecules

able to inhibit Rev/RRE association have been reported, including aminoglycosides, antisense nucleic acids (158, 159), synthetic diphenylfuran cations (160, 161), RNA aptamers (162, 163), metallopeptides (164, 165) and several pre-existing drug compounds (166). Several of these agents bind to the RRE synonymously to Rev, inserting basic regions into the same wobble-base groove in the RRE. Peptide ligands have been developed which similarly adopt the same alpha helicity as the Rev ARM; in some cases, these ligands are able to bind to the RRE with higher affinity than Rev itself (about 7-fold) and can successfully block HIV-1 replication (167-171). Other small molecules, including 8-azaguanine, suppress viral gene expression by redirecting localisation of Rev to the cytoplasm, impairing its function (172). While these agents are antiviral, off-target effects often render them toxic for human cells. Moreover, those that rely on structure specificity inadvertently apply a selection pressure for resistant serotypes (173). It also remains challenging to deliver these treatments to target cells.

As CRM1 is required to escort RNA-bound Rev across the nuclear envelope, blocking the CRM1/Rev interaction is a potential antiviral opportunity. It is well established that CRM1-inhibiting drugs, such an anti-fungal agent leptomycin B, can restrict HIV-1 replication (174, 175). However, CRM1 typically exports host proteins and is instrumental for exporting ribosomal subunits (176). Therefore, CRM1 inhibition affects downstream targets and the cellular environment and, indeed, leptomycin B is toxic to human cells. However, small molecular inhibitor KPT-185 was able to both restrict HIV-1 replication and AIDS-induced primary effusion lymphoma in primary cells by blocking Rev/CRM1 interactions (177) and elicited cytotoxic effects only at concentrations 850-fold higher than the active concentration (177).

Targeting other Rev-host cofactor interactions is similarly difficult but not impossible: Campos *et al* developed ABX464 (178), a drug that restricts HIV-1

replication in mice. ABX464 binds to and stabilises the CBC complex, enhancing RNA export by the TAP/NXF1 export pathway of fully spliced RNAs, antagonising Rev posited inhibition of this pathway (106). Importantly, while it changes the levels of spliced/unspliced viral RNA, ABX464 does not affect cellular RNA (178). It has since completed three phase II clinical trials and has successfully restricted HIV-1 replication *in vivo.* This highlights the importance of reproducibly defining the Rev interactome: it may unlock dozens of potential therapeutic targets.

1.3. RNA-binding proteins in HIV-1 infection

RBPs are proteins that bind to sequence or structural motifs of RNA with RNA-Binding Domains (RBDs) to form RNP complexes. RNP complexes are dynamic and modulate crucial cellular processes beyond RNA metabolism and gene expression, including immunity (179-182). Although RBPs were once thought to be limited to proteins with a conventional subset of well-defined RBDs, such as RNA recognition motifs, (RRMs), dsRBDs, K-Homology (KH) domains, and Zinc Fingers (ZF), novel experimental approaches to elucidate the RNA-binding proteome (RBPome) have massively broadened the scope of known RNA interactors in recent years (183, 184). RNA interactome capture (RIC) was developed by Castello *et al* (185, 186) to this end. RIC exploits the polyadenylation of cellular RNAs by crosslinking RNA to proteins using ultraviolent (UV) light and capturing bound RNAs with oligo(dT) beads, which can be identified by quantitative proteomics. RIC was employed to reveal hundreds of novel RBPs from human to yeast (185). Interestingly, about half of known cellular RBPs comprise proteins with no conventional RBD or characterised relationships to RNA metabolism. Moreover, many of these novel RBPs harbour an intrinsically disordered region (IDR), which provide a malleable interface to interact with RNA. A high

proportion of human RBPs were found to contain IDRs (187, 188). A subset of RIC-identified RBPs included enzymes that can interact with RNA via basic cavities in their surface (183). These RNA-binding enzymes comprise kinases, E3-ubiquitin ligases and members of metabolic pathways (189). These findings have had wide-reaching implications in the field of RNA research, suggesting that RNA can affect metabolic pathways and that RBP regulation operates in far more cellular pathways than considered prior.

Many RBPs play important roles in the viral infection cycle (190). Viruses often harbour relatively small genomes with limited coding capacity. Therefore, they often form interactions with host cells RBPs which can bind to the vRNA genome and aid in important steps of the replication cycle. Indeed, recent work by Garcia-Moreno *et al* discovered a pervasive remodelling of the RBPome by Sindbis virus and SARS-CoV-2 infection, with almost 200 RBPs exhibiting enhanced or inhibited RNA-binding activity upon infection (191, 192). Importantly, many of these RBPs relocalised to viral replication factories in the cytoplasm to elicit pro or antiviral functions. The regulation of RBPs by vRNAs, a concept termed 'riboregulation' (184), is another example of how host cell RBPs are hijacked by viruses during infection cycles. For example, the subgenomic RNA of Dengue virus is able to sequester the host cell protein TRIM25, attenuating the antiviral activity of this protein (193). Similarly, the Tat-responsive region (TAR) in HIV-1 gRNA can bind to protein kinase R and prevent its activation by autophosphorylation (194). While the full scope of HIV-1 RNA/host cell RBP interactions is unknown, recent studies by Knoener *et al* identified over 200 host cell proteins bound to different splice variants of HIV-1 RNA (195, 196). siRNA-mediated knockdown of over 80 of these proteins impacted HIV-1 gene expression,

demonstrating the important pro and antiviral effects of vRNA/host cell RBP interactions (195, 196).

1.4. Aims

The critical roles of Rev in HIV-1 gene expression are posited to extend beyond RNA export and are evidenced to include the regulation of vRNA splicing, stability and translation. How Rev influences these processes, however, remains poorly characterised. Research into Rev has been impaired by technical difficulties of the Rev protein, which I discuss in depth in chapter 4. Consequently, researchers have been forced to utilise systems that do not fully recapitulate physiological HIV-1 infection. This has led to contrasting data, leaving the physiological functions and interactions of Rev largely unclear. My project addresses these issues by:

- Establishing a novel approach to study Rev in conditions approximating physiological HIV-1 infection and using this approach to reveal the Rev interactome.
- Performing meta-analysis of the Rev interactome and published HIV-1 datasets to characterise the roles of Rev during HIV-1 replication and identify host cell proteins that may be important for HIV-1 infection.
- Characterising the localisations and interactions of Rev interactors to understand their roles in HIV-1 infection.

2. Materials

2.1. Materials

2.1.1. Chemicals and reagents

Table 2.1. Chemicals and reagents used in this study

REAGENT	SUPPLIER	IDENTIFIER
100 bp ladder	NEB	B7025
2 log DNA ladder	NEB	N3200S
6x purple loading dye	NEB	B7024S
AEBSF hydrochloride	BioChemica	A14210100
Agarose	MP Biomedicals	AGAP0100
Anti-Flag M2 magnetic beads	Sigma Aldrich	F3165
Bensonase	EMD Millipore	70746
Blasticidin	TOKU-E	B001
Calcium chloride	Sigma Aldrich	499609
Chloroquine	ACROS Organics	50-63-5
DAPI	Thermo Fisher	D1306
DMEM	Gibco	11995-065
Doxycycline	Sigma Aldrich	D9891
DTT	Sigma Aldrich	D1532
E. coli Poly(A) Polymerase	NEB	M0276S
Ethanol absolute	VWR chemicals	20821.33
FBS	Sigma Aldrich	F9665
Formaldehyde solution 16%	Thermo Fisher	28908
GFP trap beads	Chromotek	gta-10
Halo-trap agarose	Chromotek	ota-10
Hydrochloric acid, 37%	Thermo Fisher	7647-01-0
Hygromycin B	TOKU-E	H011
Midori green	Bulldog bio	MG04
Milk powder	Sigma Aldrich	70166
NuPAGE 4x LDS sample buffer	Invitrogen	NP0008
Oligo(dT)25 beads	NEB	S1419S
Opti-MEM reduced serum medium	Thermo Fisher	31985062
PBS 1x, pH 7.4	Gibco	10010023
Phenol:Chloroform:Isoamyl alcohol pH 6.6	Thermo Fisher	AM9730

Pierce control agarose resin	Thermo Fisher	26150
Polybrene	Sigma Aldrich	TR1003G
Proteinase K	Roche	3115887001
PS	Sigma Aldrich	P4458
RPMI 1640	Lonza Bioscience	BE12-702F
Ribolock RNase inhibitor	Thermo Fisher	EO0381
Triton X-100	MP Biomedicals	4807423
Tween-20	Sigma Aldrich	P1379
Trypan Blue Stain (0.4%)	Invitrogen	T10282
Trypsin EDTA	Sigma Aldrich	T3924
Turbo DNase	Thermo Fisher	AM2239
Spot peptide	Chromotek	ep-1
Spot-trap agarose	Chromotek	eta-10
Stable competent E. coli	NEB	C304
VectaShield antifade mounting medium	Vector Laboratories	H-1000-10
X-tremeGENE 9 DNA transfection reagent	Sigma Aldrich	6365787001
Zeocin	Invitrogen	R25001

2.1.2. Consumables

Table 2.2. Consumables used in this study

REAGENT	SUPPLIER	IDENTIFIER
0.45 µM filters	Thermo Fisher	15216869
0.45 µM PVDF Stericup-HV filters	Millipore	MPSCHVU02RE
High precision cover slips	Marienfeld	0107052
MaXtract high density tubes	Qiagen	129056

2.1.3. Machines

Table 2.3. Machines used in this study

MACHINE NAME	SUPPLIER/IDENTIFIER
API DeltaVision Elite widefield fluorescence microscope	Oxford Micron
Countess II FL Automated Cell Counter	Thermo Fisher/TF-CACC2FL
LI-COR Odyssey Fc imaging system	Li-Cor
Nanodrop 2000 Spectrophotometer	Thermo Fisher/ND-2000
Trans-Blot Turbo Transfer system	Bio-Rad/1704150

2.1.4. Buffers and solutions

Table 2.4. Buffers used in this study

BUFFER	COMPOSITION
10% sucrose cushion	50 mM Tris-HCl pH 7.4, 100 mM NaCl, 0.5 mM EDTA, 10% sucrose
2x HBS pH 7.09	50 mM HEPES, 1.5 mM Na_2HPO_4, 280 mM NaCl, 10 mM KCl, 6 mM sucrose
Halo dilution buffer	10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA
Halo elution buffer	10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA
Halo wash buffer	10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05 % Nonidet P40 Substitute, 0.5 mM EDTA
Hybridisation buffer	2X SSC, 10% Deionized Formamide, 10% Dextran Sulphate, DEPC water
PK-SDS buffer	10 mM Tris-HCl pH 7.4, 100 mM NaCl, 1 mM EDTA, 0.2% SDS
Pre-hybridisation buffer	2X SSC, 10% Deionized Formamide, DEPC water
RIPA lysis buffer	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% Triton-X 100, 0.5 mM EDTA, 25 U Bensonase, 1x AEBSF
TAE buffer	40 mM Tris-HCl, 1 mM EDTA-Na, 20 mM acetic acid, pH 8.5
Wash buffer 1	50 mM Tris-HCl pH 7.5, 150mM NaCl, 1% Triton-X, 0.1% SDS, 0.5% Sodium Deoxycholate, 1x AEBSF
Wash buffer 2	50 mM Tris pH 7.5, 150mM NaCl, 0.2% IGEPAL, 0.5 mM EDTA

2.1.5. Commercial kits

Table 2.5. Commercial kits used in this study

KIT	SUPPLIER	IDENTIFIER

Clarity ECL kit	Bio-Rad	11705061
HiScribe T7 ARCA mRNA kit	NEB	E2065S
Luna Universal One-step RT-qPCR kit	NEB	E3005S
Monarch Genomic DNA Purification kit	NEB	T3010S
Monarch Total RNA Miniprep kit	NEB	T2010
Pierce 660nm Protein Assay kit	Thermo Fisher	22662
PureLink HiPure Plasmid Filter Maxiprep kit	Thermo Fisher	K210016
QIAquick Gel Extraction kit	Qiagen	28706X4
QIAquick PCR Purification kit	Qiagen	28104
QIAprep Spin Miniprep kit	Qiagen	27104
Qubit RNA High Sensitivity Assay kit	Thermo Fisher	Q32855
RNA 6000 Pico kit	Agilent	5067-1513
RNeasy Mini kit	Qiagen	74004
SilverQuest Silver Staining kit	Thermo Fisher	LC6070
SuperScript IV Reverse Transcriptase	Thermo Fisher	18090010
Zymo RNA Clean & Concentrator	Zymo	R1019

2.1.6. Enzymes

Table 2.6. Enzymes used in this study

ENZYME	SUPPLIER	IDENTIFIER
BamHI-HF	NEB	R3136
EcoRI		R0101
HpaI		R0105
KpnI		R0142
NotI		R0189

2.1.7. Antibodies and dyes

Table 2.7. Antibodies and dyes used in this study

	PRIMARY ANTIBODY		SECONDARY ANTIBODY	
ANTIBODY/DYE	SUPPLIER/IDENTIFIER	SPECIES	SUPPLIER/IDENTIFIER	SPECIES
ß-actin	Sigma/A2228	Mouse	Li-Cor/926-32210	Goat anti-mouse
ERGIC3	Sigma Aldrich/SAB4200585	Rabbit	Li-Cor/926-32211	Goat anti-rabbit
ELMO1	Cell Signalling/D4K2E	Rabbit	Li-Cor/926-32211	Goat anti-rabbit
GFP	Chromotek/3H9	Rat	Li-Cor/926-32219	Goat anti-rat

GFP booster	Chromotek gb2AF488	Alpaca	NA	
Halo ligand (dye-labelled)	Janelia Fluor/GA1120	NA		
Myc	Cell Signalling/2278S	Rabbit	Li-Cor/926-32211	Goat anti-rabbit
p24	Center for AIDS Reagents/ARP3279	Human	Li-Cor/926-32232	Goat anti-human
PURA	Abcam/ab79936	Rabbit	Li-Cor/926-32211	Goat anti-rabbit
PURB	Proteintech/18128-1-AP	Rabbit	Li-Cor/926-32211	Goat anti-rabbit

2.1.8. Plasmids

Table 2.8. Plasmids used in this study

PLASMID	BACKBONE	SOURCE
pCDNA5-FRT-TO	pcDNA5-FRT-TO	Thermo Fisher
pCDNA5-FRT-TO-ELMO1-GFP	pcDNA5-FRT-TO	Made in Castello lab
pCDNA5-FRT-TO-ERGIC3-GFP		
pCDNA5-FRT-TO-PURA-GFP		
pCDNA5-FRT-TO-PURB-GFP		
pHEF-VSV-G	pHEF-VSV-G	Sattentau lab
pNL4-3.Luc.R-E	pNL4-3.Luc.R-E	Centre for AIDS Reagents #2128
pNL4.3-R-E-T2A-Nef	pNL4-3.Luc.R-E	Made in Castello lab
pNL4.3-R-E-T2A-Nef-ΔRev	pNL4.3-R-E-T2A-Nef	
pNL4.3-R-E-T2A-Nef-Rev-Flag	pNL4.3-R-E-T2A-Nef	
pNL4.3-R-E-T2A-Nef-Rev-Halo	pNL4.3-R-E-T2A-Nef	
pNL4.3-R-E-T2A-Nef-Rev-Spot	pNL4.3-R-E-T2A-Nef	
pOG44 Flp-Recombinase Expression vector	pOG44	Thermo Fisher

2.1.9. Oligonucleotides and primers

Table 2.9. Oligonucleotides and primers used in this study

NAME	SEQUENCE
ELMO1_F	AAGCTTACCATGCCGCCACCCGCGGAC
ELMO1_R	CTCGAGTCAGTTACAGTCATAGACGAAGTCATAGTTGCTGGGC
ERGIC3_F	AAGCTTACCATGGAGGCGCTGGGGAAGCTG
ERGIC3_R	CTCGAGCTACGTTGTCTTCCCTAGATCAATTTTCTTCTGG
T5974C_Primer_A	gggcaggagtggaagccataataagaattctgcaac

T5974C_Primer_B	cgtcccagataagtgctaaggatccgttcac
T5974C_Primer_C	gcatctcctacggcaggaagaagcggagacag
T5974C_Primer_D	ctgtctccgcttcttcctgccgtaggagatgc
T6041A_Primer_C	gtcagactcatcaagcttctctaacaaagcagtaagtagtacatg
T6041A_Primer_D	catgtactacttactgctttgttagagaagcttgatgagtctgac

2.1.10. Cell lines

Parental cell lines

Table 2.10. Parental cell lines used in this study

CELL LINE	CELL TYPE	SOURCE
HEK293	Human embryonic kidney cells	Castello lab
HEK293T	Human embryonic kidney cells	
HEK293 Flp-In T-REx	Human embryonic kidney cells	
Hela Flp-In T-REx	Hela cells	
Hela Kyoto	Human cervical cells	
SupT1	Human T lymphoblasts	

Genetically modified cell lines

Table 2.11. Genetically modified cell lines used in this study

CELL LINE	CELL TYPE
HEK293 Flp-In T-Rex PURA-GFP	HEK293 Flp-In T-REx
HEK293 Flp-In T-REx PURB-GFP	HEK293 Flp-In T-REx
Hela Flp-In T-REx ELMO1-GFP	Hela Flp-In T-REx
Hela Flp-In T-REx ERGIC3-GFP	Hela Flp-In T-REx

3. Methods

3.1. Cell biology

3.1.1. Maintaining cells

All cells were maintained at 37ºC with 5% CO_2 under sterile conditions. Adherent and suspension cell lines were cultured in DMEM and RPMI 1640 media respectively, both with 10% FBS and 1x PS. All HEK293 Flp-In T-REx and Hela Flp-In T-REx cell lines were cultured in DMEM with 10% FBS, 1x P/S, 5 µg/ml blasticidin S and 150 µg/ml zeocin. Cells were passaged regularly to avoid over-confluency by removing media by washing with PBS and incubating briefly in trypsin EDTA solution (0.5 ml per T25 flask, 1 ml per T75 flask, 2 ml per T175 flask) in a 37ºC incubator followed by resuspension in the relevant medium. When required, cells were counted using trypan blue stain (0.4%) and the countess II FL automated cell counter (both Invitrogen).

3.1.2. Stable transfection of Flp-In T-Rex cells

Hela Flp-In T-Rex cells were seeded to be 80% confluent on the day of transfection. X-tremeGENE 9 DNA transfection reagent was diluted in opti-MEM reduced serum media at a ratio of 6 µl reagent:100 µl media and 0.5 µg of both the relevant pcDNA5-FRT-TO plasmid (*e.g.,* pcDNA5-FRT-TO-ELMO1-GFP) and pOG44 were incubated in this mix for 15 min at room temperature. Transfection mix was added to cells to grow for 24 h before selection with media supplemented with 150 µg/ml hygromycin B and 5 µg/ml blasticidin S. Induction of the relevant RBP-GFP fusion protein was achieved by induction with 1 µg/ml doxycycline.

3.1.3. Generating and titrating pseudotyped virus

Pseudotyping was carried out using the protocol by Kutner *et al* (197). Briefly, HEK293T cells were seeded to 30-40% confluency and transfected with 2M $CaCl_2$, 84 µg of the pHEF-VSVG plasmid and 252 µg of the relevant pNL4-3 HIV-1 plasmid in 2x HBS pH 7.09 buffer. Media was exchanged 1 day after transfection and the resulting supernatant centrifuged at 500xg for 10 min and passed through a 0.45 µm filter to remove cell debris. Pseudotyped virus was precipitated by the addition of 50% PEG 6000, 4M NaCl and PBS. After 1.5 h of incubation at 4ºC, the solution was centrifuged at 4500 xg for 30 min and supernatant decanted. The resulting pellet was resuspended in 50 mM Tris-HCl pH 7.4, divided into 20-100 µl aliquots and stored at -80 ºC. Pseudotyped viruses were freeze-thawed once only. To titrate virus, HEK293T cells were seeded at 50,000 cells per well in a 96-well plate and transduced with virus in triplet. Virus was serially diluted at a ratio of 1:10 in PBS up to a dilution of 10^{-9} and 20 µl used to trans wells in triplicate. At 48 hours post-transduction, mCherry-fluorescent cells were counted in wells under a brightfield microscope to ascertain the viral titre.

3.2. Molecular biology techniques

3.2.1. Cloning Rev-tagged HIV-1 replicons

The pre-existing pNL4-3-R-E-T2A-Nef plasmid produced by the Castello lab was made as follows. A sequence encoding the end of *env* followed by a linker, mCherry, T2A self-cleaving peptide and the beginning of Nef protein was synthesized using the gene art gene synthesis service (Thermo Fisher Scientific), and cloned between the BamHI and XhoI restriction sites of pNL4-3.Luc.R-E- plasmid, which is defective for *vpr* and *envelope*. Rev-tagged plasmids were made as follows. Insert plasmids containing an ampicillin resistance site were designed with the sequence: a BamHI

site, 173 bp of the NL4-3 Rev CTD, a flexible glycine-serine linker (TCGGCCGGAGGA), the relevant tag (Halo or Spot), a stop codon and a HpaI site. To generate pNL4.3-R-E-T2A-Nef replicons with different tags on the C-terminal of Rev, the backbone of pNL4.3-R-E-T2A-Nef and the insert plasmid of interest were digested with BamHI-HF and HpaI enzymes in the following reactions for 1 h at 37ºC (**Table 3.1**). pNL4.3-R-E-T2A-Nef and insert vector digests were diluted in 6x purple loading buffer and run on 0.5% and 2.5% agarose gels respectively (agarose dissolved in TAE buffer) with 1:20000 midori green DNA stain and 2-log DNA ladder.

Table 3.1: Reaction mix for restricting HIV-1 vector

COMPONENT	VOLUME
Enzymes	20 units each
DNA	4.5 µg
H_2O	To 20 µl
CutSmart buffer (10x)	2 µl (1x)

Gels were run in 1x TAE buffer at 150 V and visualised with a UV transilluminator at 600 nm. Fragment bands were gel extracted using the QIAquick gel extraction kit according to instructions. Recovered DNA was quantified on nanodrop spectrophotometer and used for subsequent ligation. Ligation was carried out using a 3:1 molecular insert:backbone ratio at 16ºC for 1 h according to **Table 3.2**. Ligated plasmids were sequenced (SourceBioscience) for correct insert and insert orientation.

Table 3.2: Reaction mix for ligation

COMPONENT	VOLUME
Backbone DNA	0.02 pmol
Insert DNA	0.06 pmol
T4 DNA ligase	1 µl
H_2O	To 20 µl
T4 DNA ligase buffer (10x)	2 µl (1x)

3.2.2. Fusion PCR to generate ΔRev replicon

The pNL4-3-R-E-T2A-Nef-ΔRev plasmid was made by introducing 2 point mutations in the Rev ORF of the pNL4-3-R-E-T2A-Nef plasmid. T5974C and T6041A mutations change the start methionine residue into a threonine residue and introduce a stop codon in place of tyrosine residue respectively. Firstly, pNL4-3-R-E-T2A-Nef was restricted with BamHI-HF and EcoRI enzymes following **Table 3.1** to generate a 2.7 kb template containing the Rev ORF. Digests were extracted from agarose gels using the QIAquick gel extraction kit as before. Recovered DNA was quantified on nanodrop spectrophotometer and used for subsequent polymerase chain reactions (PCRs), which were all performed according to **Tables 3.3** and **3.4**. To generate the T5974C point mutation, PCR was performed at a T_m of 72ºC firstly with primers T5974C_Primer_A and T5974C_Primer_D and next with primers T5974C_Primer_B and T5974C_Primer_C. PCR mixtures were then run on 1% agarose gels and extracted, isolating 2 fragments encoding the point mutation. PCR was carried out without primers or DNA to anneal extracted fragments, using 7 cycles at a T_m of 70ºC. After 7 cycles, 1.3 µl of T5974C_Primer_A and T5974C_Primer_B were added and remaining cycles completed to produce a contiguous fragment harbouring the T5974C mutation. PCR mixtures were gel-purified and ligated into the restricted pNL4-3-T2A-Nef backbone. This process was then repeated with primers targeting the T6041A point mutation.

Tables 3.3, 3.4: PCR thermocycles and reagent mixtures used to generate fusion PCR products.

STEP	TEMPERATURE (ºC)	TIME (min)	CYCLES
Initial denature	98	00:30	
Denature	98	00:10	25x
Anneal	Variable	00:30	
Elongation	72	00:25/kb	
Final extension	72	10:00	
Hold	4	Hold	

REAGENT	VOLUME (µl)
10 mM DTPs	0.4
Phusion polymerase	0.2
DNA template	5 ng
5x HF Phusion polymerase buffer	4
Forward primer (10mM)	1
Reverse primer (10mM)	1
Water	To 20 µl

3.2.3. Plasmid preparation

Constructs previously validated by sequencing were transformed into stable competent *E. coli* cells according to the manufacturer's instructions. In brief, 50 µl competent cells were thawed on ice and 100-500 pg of the plasmid of interest added for 30 min on ice. Cells were heat shocked at 45ºC for 42 seconds and incubated on ice for a further 5 min. SOC media was added to 900 µl and the cells incubated at 37ºC at 1100 rpm for 30-60 min. Resulting reactions were diluted variably and 100 µl of the diluted mix was spread onto selective LB agar containing 100 µg/ml of ampicillin. These were incubated overnight at 37ºC. The next day, 3-5 colonies were transferred into 5 or 250 ml LB containing 50 µg/ml ampicillin for mini and maxi preps respectively and incubated overnight at 37ºC with shaking at 225 rpm. DNA was harvested from cultures according to the purelink hipure plasmid filter maxiprep kit or QIAprep spin miniprep kit and quantified using a nanodrop. All plasmid stocks were regularly sequenced to maintain sequence integrity.

3.2.4. X-tremeGENE 9 Transfection of HIV-1 plasmids

To transfect viral plasmids into cells for FISH, IF or Western blotting, 0.1 x 10^6 Hela cells were seeded per well in 6 well plates. After 24 h, 100 µl opti-MEM, 6 µl X-

tremeGENE 9 DNA transfection reagent (Sigma Aldrich) and 2 µg relevant DNA were incubated at room temperature for 15 min and added to cells.

3.2.5. Western blotting and silver staining

Cells to be lysed for Western blotting were washed with PBS and incubated at 4ºC in RIPA lysis buffer for 30 min before being lifted with a cell scraper and centrifuged at 500 xg for 5 min. Supernatant was decanted and mixed with NuPAGE 4x LDS sample buffer and 10 mM DTT before heating to 95ºC for 5 min. Proteins were separated on 15% polyacrylamide gels. If silver staining was performed, the silverquest staining kit was followed. If Western blotting was performed, gels were transferred to nitrocellulose proteins using the transblot turbo transfer system. Membranes were blocked in 5% milk powder in PBS for 1 h at room temperature or overnight at 4ºC. Blots were washed 0.1% PBS with 0.1% tween20 (PBST) and incubated with primary antibodies in 3% milk powder in PBST overnight at 4ºC or for 1 h at room temperature. Following further washing with PBST, blots were incubated with secondary antibodies in 5% milk powder in PBST for 1 h at room temperature in foil. If chemiluminescent secondary antibodies were used, the clarity ECL Kit was followed and exposed at the relevant wavelengths using the LI-COR odyssey Fc imaging system.

3.2.6. Immunofluorescence

Hela cells were seeded on high precision cover slips, prewashed for 30 min in 1 M hydrochloric acid and sterilised with 70% ethanol in wells of a 6-well plate. Cells were transduced, transfected or induced with doxycycline at the relevant time point. At the relevant time point, cells were washed with PBS and fixed with 4% paraformaldehyde (PFA) for 10 min at room temperature. Cells were washed with 3 times for 5 min with

PBS while gently rocking and permeabilised with PBS + 0.1% triton-X (PBSTX). Cells were washed with PBSTX three times for 5 min with gentle rocking. Primary antibodies were diluted 1:100 in PBSTX and dropped onto parafilm on a flat surface. Coverslips with attached cells were incubated face-down in PBSTX/antibody mix for 1 h at room temperature. Coverslips were replaced in 6 well plates and washed 3 times with PBSTX before being incubated in secondary antibodies diluted to 1:500 in the same manner. If FISH was being carried, 1:1000 ribolock RNase inhibitor was included. At this point, samples could be prepared for FISH. Otherwise, cells were incubated with DAPI stain in PBSTX for 5 min and washed with PBS and water before being mounted on to glass slides using antifade mounting medium. Slides were stored at 4ºC and imaged on an API deltavision elite widefield fluorescence microscope using a 100X oil UPlanSApo objective (1.4 NA) and deconvolved with softWoRx v6.5.2 (GE Healthcare).

3.2.7. Single molecule fluorescence in situ hybridisation

Cells intended for FISH were seeded, harvested and fixed in the same manner as IF. After necessary IF antibody staining was complete, a secondary PFA fixing step was carried out. Coverslips were then washed in PBSTX, PBSTX + 1x saline-sodium citrate solution (SSC), PBSTX + 2x SSC and finally twice for 10 min at 37ºC in pre-hybridisation buffer. Coverslips were then incubated in 125 nM stellaris FISH probes (LGC Biosearch Technologies) diluted in hybridisation buffer overnight in a wet chamber at 37ºC. After 24 h, slides were washed twice for 10 min at 37ºC in pre-hybridisation buffer. Finally, slides were and incubated for 5 min at 37°C with 2 µg/ml DAPI in pre-hybridization buffer. Slide mounting was performed as for IF.

3.2.8. Cloning RBP-GFP plasmids

Hela cells were seeded in a 6-well plate to 80% confluency and lifted with a cell scraper. Total RNA was then obtained using the RNeasy mini kit with on column DNase digestion using the RNase-free DNase set, according to the manufacturer's protocol. Reverse transcription of extracted RNA was carried out using the superscript IV reverse transcriptase (Thermo Fisher). PCR was carried out using primers designed to amplify ORFs of ERGIC3 and ELMO1 (*e.g.,* ELMO1_F) and flanking KpnI (5') and NotI (3') restriction sites, with the following thermocycle and reagents:

Tables 3.5, 3.6: Reaction mix for amplifying RBP ORFs by PCR

STEP	TEMPERATURE (ºC)	TIME (min)	CYCLES
Initial denature	98	01:00	
Denature	98	00:30	25x
Anneal	72	00:30	
Elongation	72	01:30	
Final extension	72	5:00	
Hold	4	Hold	

REAGENT	VOLUME (µl)
10 mM DTPs	0.4
Phusion polymerase	0.2
DNA template (RT mix)	0.5 µl
5x HF Phusion polymerase buffer	4
Forward primer (10mM)	1
Reverse primer (10mM)	1
Water	To 20 µl

20 µl of PCR products were mixed with 1x µl purple loading gel and run on 1.5% agarose gels alongside 10 ul of 1 kb ladder. Gels were exposed at 600 nm and bands of expected size were extracted using the QIAquick gel extraction kit. Fragments, and

the pcDNA5 FRT/TO vector were digested using the enzymes KpnI and NotI according to **Table 3.1** and ligated according to **Table 3.2**.

3.3. Proteomic Analysis

3.3.1. Tagged Rev immunoprecipitations

SupT1 cells were seeded at 60% confluency and transduced with pseudotyped HIV-$1_{Rev\text{-}Flag}$, HIV-$1_{Rev\text{-}Spot}$ or HIV-$1_{Rev\text{-}Halo}$ alongside a negative control dish infected with pseudotyped HIV-$1_{mCherry\text{-}Nef}$. At 48 hpi, cells were imaged under a brightfield microscope to confirm ~80%-90% infection as measured by the proportion of mCherry-fluorescent cells. Cells were centrifuged at 400 xg for 5 min at room temperature and resuspended in RIPA lysis buffer. After incubation for 30 min on ice, cells were span down at 500xg for 5 min and the supernatant saved. Lysates were incubated with washed pierce control agarose resin for 30 min at 4°C. Lysates were then centrifuged 2500 xg for 2 min and supernatant decanted. Anti-Flag M2 and Spot-trap were washed with RIPA wash buffer. For Halo-trap agarose, Halo dilution buffer was used to wash beads. Beads were incubated with lysates for 1 h at 4°C. To remove contaminants, anti-Flag M2 and Spot-trap beads were washed with wash buffer 2 and Halo-trap agarose was washed with Halo wash buffer. To elute, Spot-trap and anti-Flag M2 beads were incubated with 10 µl Spot peptide or 10 µl 3x Flag peptide respectively with 20 µl wash buffer 2 for 1 h on ice. Halo-trap agarose were eluted by pipetting beads in pH 2.5 Halo elution buffer for 1 min at room temperature before centrifuging for 2 min at 2500 xg. The beads/buffer mix was neutralised with 1 M tris pH 10.4. For optimising IPs, elutions were repeated and pooled.

3.3.2. PURA-GFP and PURB-GFP immunoprecipitations

PURA-GFP, PURB-GFP and GFP HEK293T Flp-In T-Rex cells were seeded in two 10 cm plates and transduced with pseudotyped HIV-$1_{mCherry\text{-}Nef}$ viral particles to produce ~90-100% infected cells or left uninfected as determined by mCherry fluorescence. Overexpression was induced using 1 µg/ml doxycycline for 24 h. Lysis was performed at 48 hpi as above and lysates were incubated with GFP trap beads for 2 h and washed to remove contaminants. GFP fusion proteins were eluted identically to Halo-trap agarose. Two rounds of elution were pooled to increase the elution efficiency.

3.3.3. Testing bensonase RNA digestion

To measure RNA concentration in Flag IP eluates, the Flag IP was repeated with 2 HIV-$1_{Rev\text{-}Flag}$-infected lysate samples. Bensonase was removed from lysis for 1 sample as a control. The eluates resulting from the Flag IP of these lysates were incubated with 20 µl proteinase K and 180 µl PK-SDS buffer for 1 h at 50°C at 1100 rpm. Next, 250 µl Phenol:Chloroform:Isoamyl alcohol pH 6.6 was incubated with samples for 10 min at 37°C, 1100 rpm. MaXtract high density tubes were spun at 15000 xg at room temperature for 5 min. Samples were transferred to maXtract tubes and spun again. The aqueous layer containing RNA was transferred to a new tube and RNA eluted using the RNA clean & concentration kit following kit instructions. RNA concentration was assessed by running 1 µl on a bioanalyser using an RNA 5000 pico kit, following the manufacturer's protocol.

3.3.4. Mass spectrometry

Sample preparation, LC-MS/MS analysis, protein identification and data processing were performed by Honglin Chen and Shabaz Mohammed.

Sample preparation

IP eluates were processed and digested using filter-aided sample preparation protocol (FASP) (198) with custom adjustments. IP eluates were denatured in 200 µl of urea buffer (8 M urea, 100 mM AmBic) by pipetting. Denatured samples were subjected to reduction and alkylation in 1 solution environment by addition of 10 mM TCEP and 50 mM CAA, and incubate in dark for 30 min. VIVACON500 30 kDa filters were washed with 200 µl of 0.1 % TFA, 50 % ACN, and centrifuged at 14,000 xg for 15 min. Alkylated samples were loaded onto 30 kDa filters and centrifuged at 14,000 xg for 15 min. Detergent contaminants were removed by repeating the following washing cycles for 3 rounds: addition of 200 µl urea buffer, gentle pipetting, and centrifuge at 14,000 xg for 15 min. Detergent-removed samples were washed twice with 200 µl 1 M urea in 25 mM AmBic, and centrifuge at 14,000 xg for 5 min. On-filter digestion were performed by addition of 200 µl trypsin endoproteinase mixture (250 ng worth of trypsin, 1 M urea, 25 mM AmBic, 1 mM $CaCl_2$), and 37 C^o incubation overnight. Peptides in digested samples were collected in eluates after centrifuge the filters at 14,000 xg for 15 min. The filters were washed with first 150 µl 0.1 % TFA, centrifuged at 14,000 xg for 15 min, and then 150 µl 50 % ACN in 0.1 % TFA, centrifuged at 14,000 xg for 15min. Flow-through fractions in two washes were combined with previous eluates to maximise peptide recovery rate. Combined flow-through were dried using speedvac. dried peptides were reconstituted with 45 µl MS loading buffer (5 % DMSO, 5 % FA) for LC-MS/MS analysis.

LC-MS/MS

FASP processed samples were analysed using ultimate 3000 ultra-HPLC system (Thermo Fisher) and electrosprayed into a QExactive mass spectrometer (Thermo Fischer). Injected samples were first trapped on C18 PepMap100 pre-column (300 µm inner diameter x 5 mm, 100 Å, Thermo Fisher) in solvent A (0.1% FA). Peptides were separated on an in-house packed analytical column (75 µm inner diameter x 50 cm packed with ReproSil-Pur 120 C18-AQ, 1.9 µm, 120 Å, Dr. Maisch GmbH) using a linear 15-35% (vol/vol) ACN gradient (2 h for interactome capture samples) and a flow rate of 200 nl/min. Full-scan mass spectra were acquired in the Orbitrap (scan range 350-1500 m/z, resolution 70000, AGC target 3 x 106, maximum injection time 50 ms) in a data-dependent mode. Subsequently, the 10 most intense peaks were selected for higher energy collisional dissociation fragmentation at 30% of normalised collision energy. Higher-energy collisional dissociation fragmentation spectra were also acquired in the orbitrap (resolution 17500, AGC target 5104, maximum injection time 120 ms) with first fixed mass at 180 m/z.

Protein identification and data processing

Raw data were processed by maxquant (199) (version 1.6.3.4) for peptide protein identification, using the integrated andromeda search engine. For Rev-Flag analysis, the mass spectra were searched against human database (UniProt, version 01/16) and HIV-1 (NL4-3) viral proteins. The Flag-Myc-tagged Rev protein sequence was included in the Flag IP. For PURA-GFP/PURB-GFP analysis, the mass spectra were searched against HIV-1 (NL4-3) viral proteins were with Nef-T2A-mCherry protein sequence. The following search parameters were used: trypsin was chosen with full

tryptic specificity and maximal two missed cleavage sites. Fixed modifications were set to carbamidomethyl (C) and the variable modifications to acetylation (protein N-term, *i.e.,* only the amino terminus of the protein) and oxidation (M). Match between runs was applied. Default parameters were used for all others. False discovery rate (FDR) cut-off for peptide identification was set to 0.01. The protein group file of maxquant search result was imported in RStudio for further processing. Intensities from all 3 replicates were extracted and log-2 transformed. Proteins that had less than 2 valid intensity measurements in at least one condition were removed prior to downstream analysis. For Rev-Flag analysis, missing values were imputed with minimal deterministic value method (200) using 1% quantile of global intensities. For PURA-GFP/PURB-GFP analysis, for proteins that only have valid quantification in 1 condition, the missing values in the opposite condition were imputed with minimal deterministic value method (200) using 1% quantile of global intensities. Batch effects were assessed using principle component analysis. Linear modelling and Bayesian-model-based moderated T-test was performed using package *limma* (201), with replicate number incorporated in modelling as co-variable to account for batch effects. P values obtained in T-test were adjusted to account for effect of large dataset using Benjamini-Hochburg method.

3.4. Analysis

3.4.1. Mapping and comparing gene IDs

For all datasets analyses, including mass spectrometry data, genes were mapped to Hugo gene nomenclature committee IDs (HGNC) using R package BiomaRt (202, 203). Any genes that could not be mapped were manually curated by adding manually or removing if that gene was a pseudogene or it's HGNC entry had been withdrawn.

Viral proteins were removed and remaining IDs compared in RStudio. For HIV-1 analysis, the HIV-1 NCBI database was downloaded from the NCBI web server and parsed to gene IDs in the same manner. Upset plots and euler diagrams were made using the R packages upsetR (204) and venneuler respectively.

3.4.2. GO terms/STRING analysis

To generate viral GO term plots, the R package annotation dbi (205) was queried with HGNC IDs and resulting GO terms filtered by the following GO categories: 'viral', 'immune', 'infection', 'pathogen', 'immune cell' and 'immune molecule.' General GO terms were extracted using the tool EnrichR (206). Spiderweb string diagrams were downloaded from the string-db webserver (207); string analyses were based on textmining, experiments and database with a confidence of at least 0.4.

3.4.3. RNA-binding and RBD analysis

To determine the percentage of proteins from a given dataset that were RNA-binding, proteins were considered RNA-binding if they were identified in at least 3 studies listed in the manually curated datasets available from the EMBL RBPbase dataset maintained by the Hentze lab. To generate lollipop charts of RBDs, domains were extracted by firstly selecting PFAM IDs for genes using the R package org.Hs.eg.db (208) and annotation dbi (205) packages. These IDs were matched to domain names using a look-up table of proteins domains generated using the PFAM.db package (209). Finally, domains were considered 'classical' based on PFAM domains (**Table 3.7**). Specifically, proteins containing the following domains were considered classical: RRM, KH, dsRNA, Piwi, DEAD (DDX and DHX), Pumilio, CSD, ZnF-CCCH, and the

curated helicase domains AAA11, AAA12 and AAA19 matching the analyses previously carried out in (188).

Table 3.7: RNA-binding domains classified as 'classical' in domain analysis

PFAM ID	Domin name	Domain description
PF13086	AAA_11	AAA domain
PF13087	AAA_12	AAA domain
PF13245	AAA_19	AAA domain
PF00313	CSD	'Cold-shock' DNA-binding domain
PF00270	DEAD	DEAD/DEAH box helicase
PF00035	dsrm	Double-stranded RNA binding motif
PF00013	KH_1	KH domain
PF07650	KH_2	KH domain
PF13083	KH_4	KH domain
PF13184	KH_5	NusA-like KH domain
PF15985	KH_6	KH domain
PF17214	KH_7	KH domain
PF05172	Nup35_RRM	Nup53/35/40-type RNA recognition motif
PF14605	Nup35_RRM_2	Nup53/35/40-type RNA recognition motif
PF02171	Piwi	Piwi domain
PF00806	PUF	Pumilio-family RNA binding repeat
PF10378	RRM	Putative RRM domain
PF00076	RRM_1	RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)
PF04059	RRM_2	RNA recognition motif 2
PF08777	RRM_3	RNA binding motif
PF10598	RRM_4	RNA recognition motif of the spliceosomal PrP8
PF13893	RRM_5	RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)
PF16367	RRM_7	RNA recognition motif
PF00642	zf-CCCH	Zinc finger C-x8-C-x5-C-x3-H type (and similar)
PF14608	zf-CCCH_2	RNA-binding, Nab2-type zinc finger
PF15663	zf-CCCH_3	Zinc-finger containing family

3.4.4. Drug interaction analyses

The drug-gene interaction database was downloaded from the DGIdb web server (210). Drugs were filtered by those with inhibitory effects and matched to drug names using the 'dbdataset' available from the interstellar consultation services github.

4. Establishing a system for researching the HIV-1 Rev protein under conditions approximating physiological infection

4.1. Introduction

Rev-mediated RNA export is a complex process, requiring nuclear/cytoplasmic shuttling, interactions with RNA and multiple conformational changes to engage or disengage RNP complexes (211). Though Rev must interact with host cell proteins to accomplish these steps, few interactors beyond its shuttling partner CRM1 were known for many decades. Similarly, despite posited links to splicing (126, 127), stabilisation of RNA (107, 109), translation (134, 135) and genome encapsidation (146, 147), the roles of Rev outside of RNA export remain unclear.

Rev is poorly characterised due to technical challenges (211). These technical caveats have forced researchers to employ a variety of non-physiological approaches. For example, Rev is expressed at low levels in HIV-1 infected cells, making it difficult to visualise or immunoprecipitate (IP) (212). To overcome this difficulty, Rev is often overexpressed to non-physiological levels, which can lead to aberrant localisation and artificial interactions. In addition, the lack of efficient commercial antibodies against Rev makes it difficult to analyse by Western blotting or IP. This could be rectified by the fusion of Rev to a well-characterised tag which has effective and available reagents. However, the Rev ORF is entirely overlapped by other viral genes, making it difficult to genetically manipulate without nullifying intersecting HIV-1 genes. Instead, peptides of Rev such as the RNA-binding ARM (126) are often ectopically expressed, which can lead to non-physiological folding, functions and interactions. The cell systems typically used to study Rev also present limitations. CD4+ cells such as T lymphocytes are the natural target for HIV infection, but are difficult to transfect, and consequently CD4(-) cell lines are used. Different cell systems harbour varying transcriptomes and proteomes, which may affect the interaction network established by Rev (213).

The dependence of singly-spliced vRNAs on Rev also appears to vary between cell lines (136). Therefore, using a variety of CD4(-) cells may uncover Rev interactions that do not occur in lymphocytes. A common alterative to studying Rev in CD4(+) cell lines is using uninfected cells via plasmid transfection. As infection is known to profoundly remodel the host transcriptome and proteome (214-217), including production of high levels of viral proteins and RNA, uninfected cellular environments do not resemble physiological infection conditions. An overlooked consideration in Rev research is the co-expression of the HIV-1 gRNA. Rev multimerization on the RRE is a pre-requisite to Rev-vRNA export (82, 87), and therefore the localisation, interactions and functions of Rev are drastically altered without it. Many studies compensate by expressing a synthetic RRE RNA. However, the full-length HIV-1 gRNA contains many internal binding sites and signals that alter viral protein and nucleic acid behaviour. For example, INS sequences in HIV-1 gRNA require binding of Rev to be translated. Therefore, smaller synthetic RNAs lacking these signals may experience export, transcription or translation at varying rates.

Because of the technical challenges and compensations made when studying Rev, research into Rev interactions and functions is limited and often discordant (211). In an optimal system, full-length Rev would be expressed at physiological levels from a complete HIV-1 gRNA in infected T CD4+ lymphocytic cell lines. In this chapter, I detail the generation of such HIV-1 replicons and validate the activity of the produced tagged Rev with several orthogonal techniques.

4.2. Designing HIV-1 replicons expressing tagged Rev

To uncover Rev interactions using conditions nearer to physiological infection than have often been used in Rev research, a HIV-1 replicon encoding full-length Rev from

a complete pNL4-3 HIV-1 genome was engineered. To achieve this, a novel replicon created by the Castello lab was used as a backbone for cloning. pNL4-3-R-E-T2A-Nef, henceforth referred to as 'HIV-$1_{mCherry-Nef}$' (**Figure 4.1, A**), encodes the full HIV-1 NL4-3 coding sequence in a plasmid carrying an ampicillin-resistance gene for selection. mCherry is fused to the C-terminal of *nef,* separated by a self-cleaving T2A element. This allows mCherry fluorescence to be used as proxy for HIV-1 replication without disrupting Nef activity. To ensure that viruses would be limited to single-round replication and therefore safe for use in a category 2 laboratory, *vpr* and *env* genes were nullified with frameshift mutations. Vpr is an accessory viral factor which is dispensable for viral replication. Envelope is unnecessary when a replicon is expressed via transfection and can be replaced by trans-complementation with a plasmid encoding envelope or a foreign glycoprotein such as vestitular stomatitis virus glycoprotein (VSVG) to produce single-round viral particles. Importantly, the nullification of *env* allows Rev to be tagged on the CTD without interruption or frameshifting effects of other viral ORFs.

To generate a toolkit of replicons suited to different techniques for studying Rev with different approaches, several tags with unique advantages were selected for cloning. pNL4-3-R-E-T2A-Nef-Rev-Flag-3xMyc (HIV-$1_{Rev-Flag}$) was generated by the fusion of a flag and three myc tags to the Rev CTD (**Figure 4.1, B**). Flag and myc tags comprise short 8 and 10 aa peptide sequences (DYKDDDDK and EQKLISEEDL respectively) and were selected as both have well-established commercial antibodies for IP and Western blotting. The small ~4 kDa size and disordered nature of these combined tags, together with the fact that the Rev CTD is highly disordered, makes them unlikely to perturb Rev folding. pNL4-3-R-E-T2A-Nef-Rev-Spot (HIV-$1_{Rev-Spot}$) was generated by the fusion of a Spot tag to the Rev CTD. Spot tag is a 12 aa,

uncharged sequence also unlikely to perturb Rev folding. The tag is bound by a specific alpaca nanobody with a Kd of 0.7 nM. Nanobodies are recombinant antibody fragments consisting of the antigen-binding domain of the heavy chain and exhibit highly specific antigen binding (218). Spot tag was therefore selected for use in high-affinity IPs. pNL4-3-R-E-T2A-Nef-Rev-Halo (HIV-$1_{Rev\text{-}Halo}$) was generated by the fusion of a Halo tag to the Rev CTD. Halo tag is a 33 kDa tag derived from a Haloalkane dehalogenase of bacteria *Rhodococcus rhodochrous* and forms covalent bonds with reactive chloroalkane-based ligands (219).

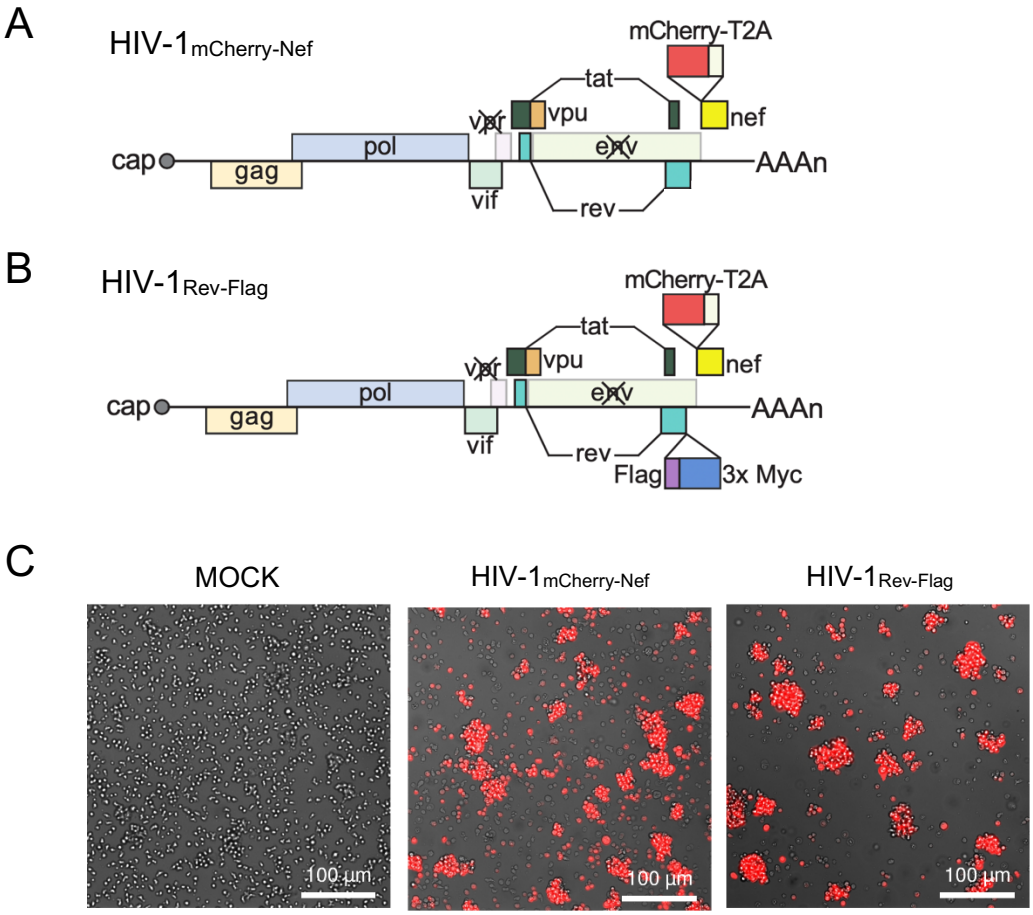


Figure 4.1: Novel pNL4-3 replicons were engineered to express mCherry and tagged Rev proteins. (**A**) HIV-$1_{mCherry\text{-}Nef}$ and (**B**) HIV-$1_{Rev\text{-}Flag}$ replicons comprise the coding region of pNL4-3 HIV-1 gRNA and express untagged and Flag-3xMyc-tagged Rev respectively. mCherry, fused to a self-cleaving T2A element, is encoded at the N-terminal of the *nef* gene. Both *env* and *vpr* are nullified by frameshift mutations, as signified by crosses through gene names. (**C**) Brightfield images of SupT1 lymphocytic

cell lines mock-transduced and transduced with HIV-$1_{mCherry\text{-}Nef}$ and HIV-$1_{Rev\text{-}Flag}$ replicons, as labelled. mCherry-fluorescent cells are HIV-1 positive, courtesy of mCherry expression from HIV-1 replicons (scale bars = 10 μM).

Formation of covalent binding enables stringent washing conditions in IPs. In addition, Halo ligands can be attached to several functional groups for different applications, including IP and dyes for immunofluorescence (IF). Dye-coupled Halo ligands render Halo a powerful tool in microscopy experiments (220). Each Rev-tagged replicon was cloned by restricting HIV-$1_{mCherry\text{-}Nef}$ at BamHI and HpaI sites, therefore removing the last 178 bp of the Rev CTD. A unique insert was then ligated using these restriction sites, encoding the last 178 bp of the Rev CTD fused to the appropriate tag, separated by a flexible glycine-serine linker. As a result, all replicons harbour the same backbone containing mCherry-T2A-*nef* and express mCherry (**Figure 4.1, C**). In addition, a replicon unable to express Rev was generated as a negative control for tagged-Rev replicons. pNL4-3-R-E-T2A-Nef-ΔRev (HIV-$1_{\Delta Rev}$) was generated by 2 rounds of fusion PCR using primers to 1) amplify the Rev ORF and 2) introduce 2 point mutations at specific sites in this ORF. These mutations were T5974C, which mutates the start methionine residue into a threonine and T6041A, which replaces a tyrosine with a stop codon downstream in the Rev sequence. Both were selected to recreate a previously developed ΔRev mutant (148, 221, 222).

My aim was to use transduction instead of transfection using the newly-generated tagged Rev replicons, as transfection efficiency of lymphocytes is generally poor and transduction with low MOI approximates physiological HIV-1 infection to a greater extent. To achieve a high transduction rate, I took advantage of the nullified *env* gene in replicons through trans-complementation with VSVG. The co-expression of a replicon lacking an envelope protein with a foreign glycoprotein able to bind to host cell receptors is henceforth referred to as pseudotyping. Pseudotyping with VSVG

can markedly increase transduction efficiency (197). I transfected HEK293T cells with each HIV-1 replicon and a plasmid expressing VSVG. After 48 h, cell supernatant containing viral particles coated with VSVG were concentrated using PEG precipitation and stored at -80ºC. Transduction with pseudotyped viral particles led to 100% of cells expressing mCherry. Given these high infection rates, pseudotyped HIV-1 particles are a suitable approach with which to study the Rev interactome.

4.3. Tagging of Rev does not affect HIV-1 gene expression

To determine whether the tagging of Rev affects HIV-1 replication, cells were transduced with pseudotyped, single-round viruses or transfected with plasmids encoding the different replicons and lysed 24, 48 and 72 hours later. For HIV-$1_{mCherry-Nef}$ and HIV-$1_{Rev-Flag}$ replicons, lymphocytic SupT1 cell lines were transduced with pseudotyped viral particles and lysates immunoblotted against HIV-1 CA (p24) and β-actin as a loading control (**Figure 4.2**). In the infected cell, p24/CA typically exists as a precursor polyprotein (Gag/p55 – MA/CA/NC) that is processed by HIV-1 PR into MA, CA and NC to generate mature viral particles. Gag polyprotein is synthesised from the unspliced gRNA and therefore is reliant on Rev-mediated RNA export. The p24 antibody detects both the precursor polyprotein (Gag/p55) and all CA-containing proteolytic fragments (225). Importantly, Gag/CA (p55/p24) expression was almost indistinguishable between Flag-tagged (HIV-$1_{Rev-Flag}$) and untagged Rev (HIV-$1_{mCherry-Nef}$)-transduced cell lysates at each time point. No bands were detectable in mock lanes, demonstrating the specificity of the antibody. To further validate that Rev is critical for Gag/CA expression, I used HIV-$1_{\Delta Rev}$. As Rev-mediated RNA export is required for the expression of structural viral proteins which form viral particles, HIV-$1_{\Delta Rev}$ can not be pseudotyped. Therefore, for comparing HIV-$1_{Rev-Halo,}$ HIV-$1_{\Delta Rev}$ and

HIV-$1_{Rev\text{-}Spot}$ replicons, I employed transfection with X-tremeGENE 9 DNA transfection reagent. In addition, experimental conditions for downstream validation IF experiments were compatible with adherent cells, and I thus used HEK293T instead of lymphocytic lines. Similarly, no Gag/CA bands were detectable when Rev was nullified in the replicon (HIV-$1_{\Delta Rev}$) (**Figure 4.2, B**). These results confirmed that Rev is critical for Gag expression and that the tagging of Rev does not affect HIV-1 gene expression. Nef is a HIV-1 protein expressed from fully spliced vRNA which does not require Rev-mediated export for translation. Blotting against Nef showed similar expression in both HIV-$1_{\Delta Rev}$ and HIV-$1_{mCherry\text{-}Nef}$transfected lysates (**Figure 4.2, B,** third panel). Conversely, no detectable Nef was seen in the mock-transfected lane. Taken together, these data suggest that Rev is not required for the expression of accessory viral proteins such as Nef but is crucial for expression of Gag/p55 from unspliced HIV-1 gRNA. Moreover, tagged Rev can mediate Gag gene expression at the same level as untagged Rev.

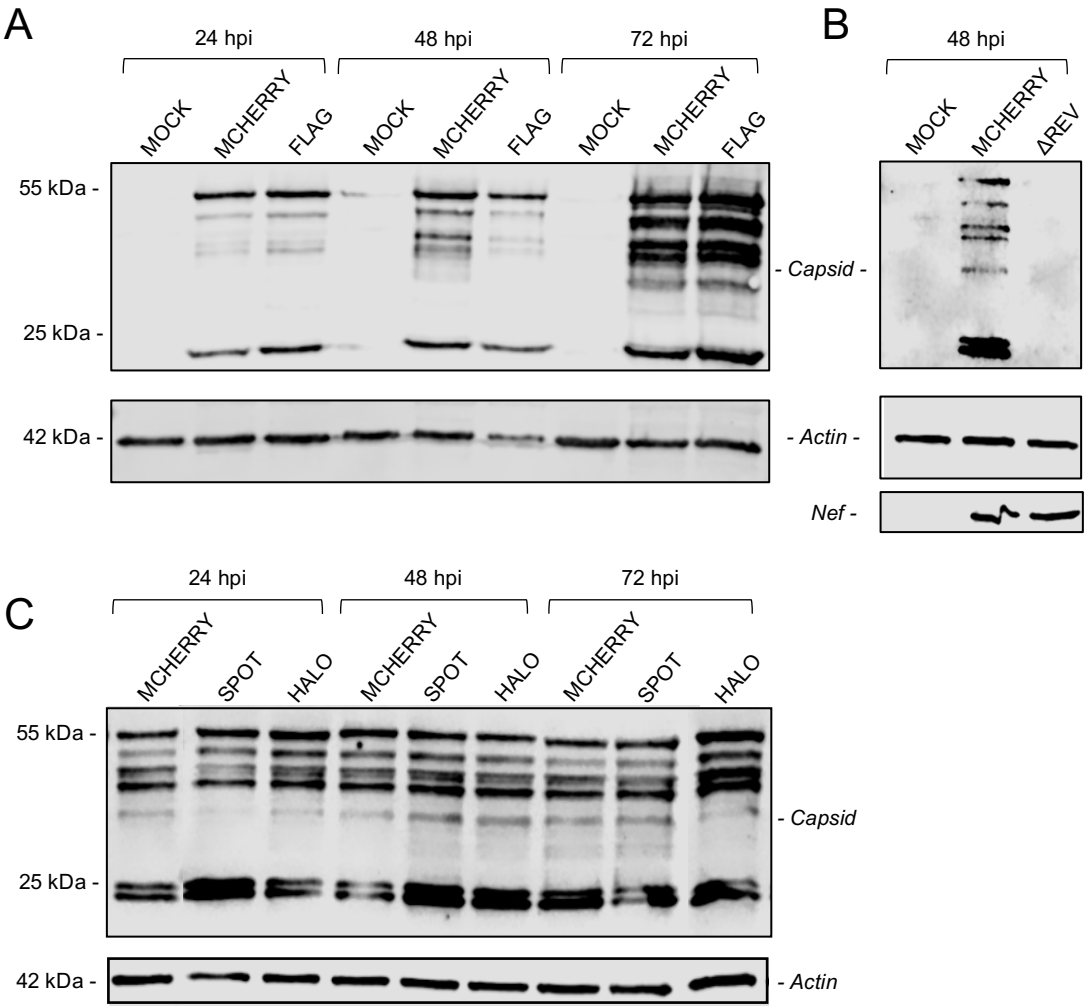


Figure 4.2: Tagged Rev HIV-1 replicons lead to normal Gag/CA expression. (**A**) SupT1 cells were mock transduced and transduced with pseudotyped HIV-$1_{mCherry\text{-}Nef}$ and HIV-$1_{Rev\text{-}Flag.}$ Cells were lysed at 24, 48 and 72 hpi and immunoblotted against p24 (CA, Capsid) and Actin. CA expression is indistinguishable between lysates transduced with untagged and tagged Rev HIV-1 replicons. (**B**) As in (A), with HEK293T cells transfected instead with HIV-$1_{\Delta Rev,}$ HIV-$1_{Nef\text{-}mCherry}$ and mock transfected. (**C**) As in (A) with HEK293T cells transfected with HIV-$1_{mCherry\text{-}Nef}$, HIV-$1_{Rev\text{-}Spot}$ and HIV-$1_{Rev\text{-}Halo}$ replicons (**C**).

4.4. Assessing subcellular localisation of tagged Rev

While Rev participates in nucleocytoplasmic shuttling, it also displays a well-documented tendency to localise in the nucleolus (64, 91, 226-230). Though the function behind this localisation is undetermined, Rev mutants that do not localise in the nucleolus are impaired in their ability to export vRNA (227, 228), suggesting it is

critical to Rev function. To ensure that canonical nucleolar localisation of Rev is preserved in its tagged form, Hela cells were transduced with HIV-$1_{Rev\text{-}Halo}$ and HIV-$1_{mCherry\text{-}Nef}$ as a negative control, fixed at 48 hpt and stained with dye-labelled Halo ligands, alongside chromatin-binding DAPI to label the nucleus. IF showed that Rev-Halo strongly recapitulates the well-established nucleolar localisation of Rev (**Figure 4.3**).

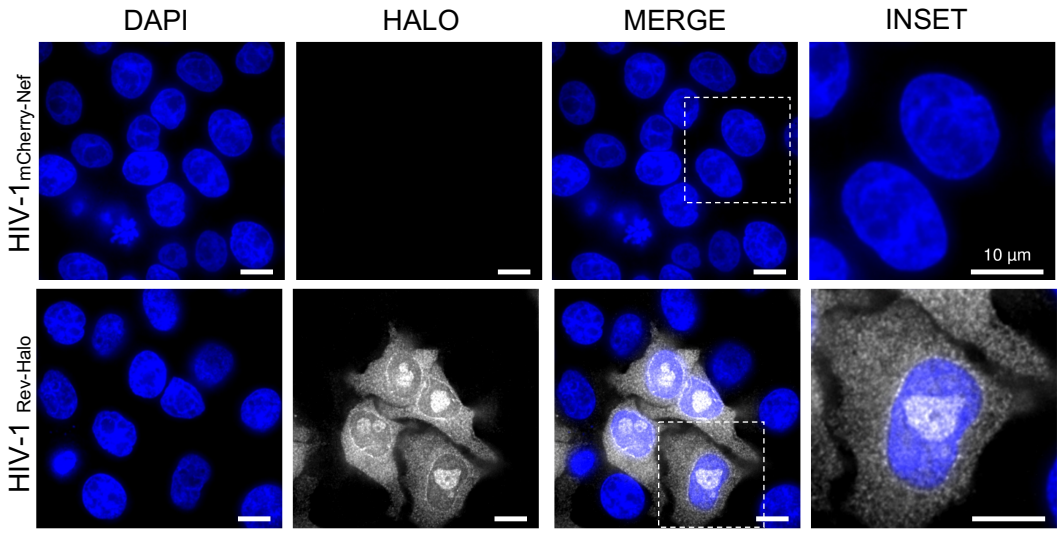


Figure 4.3: Rev-Halo recapitulates the known nucleolar localisation of Rev and highlights spots at the nuclear membrane. Hela cells transduced with HIV-$1_{mCherry\text{-}Nef}$ and HIV-$1_{Rev\text{-}Halo}$ replicons were fixed with paraformaldehyde and stained with dye-labelled Halo ligands (seen as white foci) and DAPI at 48 hpi. Fixed samples were imaged with fluorescence microscopy on a 100x lens (scale bars = 10 µM). Dotted white box indicates area of enlarged inset region.

HIV-$1_{mCherry\text{-}Nef}$ transduced cells showed no Halo signal, confirming that this pattern relates specifically to the Rev-Halo protein. Interestingly, Rev-Halo also appears to localise strongly at the nuclear envelope with a punctuate distribution that evokes nuclear pores. To my knowledge, this pattern has not been seen in previous Rev IF studies and likely reflects interactions between Rev and the NPC, as Rev translocates across the nuclear membrane. This novel localisation pattern suggests that expressing Rev-Halo from a full-length replicon in concert with Halo-reactive dyes is a more

sensitive and specific approach than those previously used. In addition, this IF data confirms that tagged Rev localises as expected, even when a relatively large tag such as Halo is used.

4.5. Tagged Rev can effectively export HIV-1 gRNA

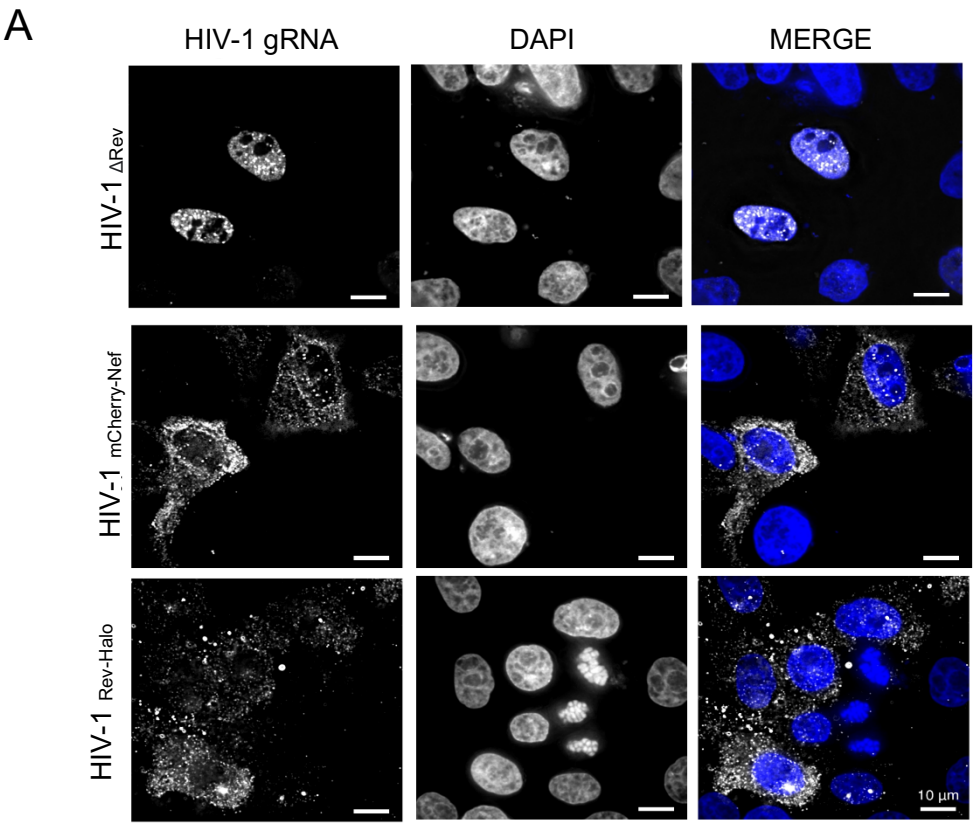
Rev not only functions as an essential RNA export factor, but also blocks the cellular TAP/NXF1 export pathway (106). This may be to shift the stoichiometry of exported HIV-1 RNAs towards the unspliced and partially spliced viral transcripts, allowing balanced expression of viral proteins. To directly compare the export pattern of unspliced HIV-1 RNAs by tagged and untagged Rev, single molecule fluorescence *in situ* hybridisation (smFISH) was used. In this technique, dye-labelled probes are hybridised on to target RNAs in fixed cells and imaged using fluorescence microscopy. To carry out smFISH, I seeded Hela cells on washed cover slips and transfected them with HIV-$1_{mCherry\text{-}Nef}$, HIV-$1_{\Delta Rev}$ and HIV-$1_{Rev\text{-}Halo}$. At 48 hpi, cells were fixed with paraformaldehyde and smFISH was performed with probes targeting the major 5' intron that is only present in the HIV-1 gRNA. In cells transfected with both HIV-$1_{mCherry\text{-}Nef}$ and HIV-$1_{Rev\text{-}Halo}$, dye-labelled HIV-1 RNAs (visible as white foci) were largely cytoplasmic, which suggests active Rev-mediated export from the nucleus (**Figure 4.4, A**). In both cases I observed similar patterns of RNA distribution, with a high degree of clustering around the nuclear envelope and dense clustering at plasma membrane sites, where virus is likely assembling. On the other hand, cells transfected with HIV-$1_{\Delta Rev}$ exhibited almost exclusively nuclear gRNA, demonstrating that when Rev is not functional, there is ineffective nuclear export of gRNA. Together, my results strongly support that tagged Rev is functional and mediates the export of under-spliced HIV-1 RNAs to the cytoplasm in an analogous manner to untagged Rev. It is probable

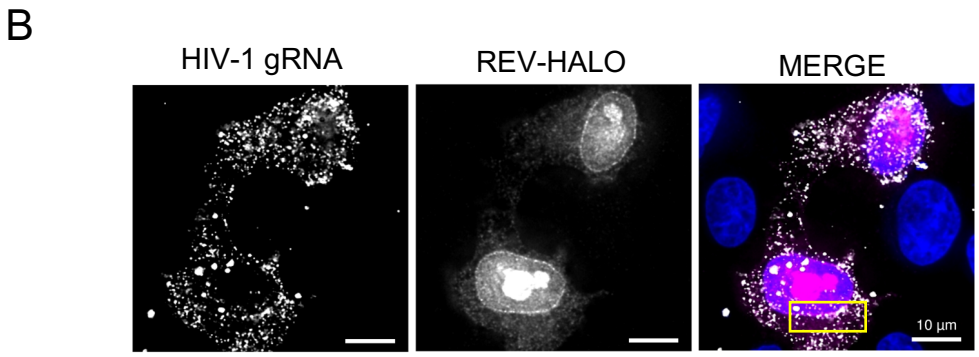
that the much smaller sizes of Flag-3xMyc and Spot tags lead to a similar accumulation of HIV-1 gRNA.

To gauge the degree of co-localisation between Rev-Halo and HIV-1 gRNA, Hela cells were transfected with HIV-$1_{mCherry\text{-}Nef}$ and HIV-$1_{Rev\text{-}Halo,}$ fixed and used for both smFISH and IF at 48 hpt (**Figure 4.4, B**). While the strongest Rev-Halo signal at 48 hpt was nucleolar as expected, it also concentrated at a speckled pattern resembling NPCs. Small portions of the HIV-1 gRNA appear to occupy a similar space in nuclear pore regions (**Figure 4.4, B**, yellow box). This localisation likely reflects Rev-mediated RNA export across the nuclear pore. The low levels of Rev-Halo in the cytoplasm relative to the nucleus is expected, as Rev is thought to be imported into the nucleus upon releasing vRNA in the cytoplasm. Both Rev-Halo and HIV-1 gRNA were indeed spread ubiquitously across the cytoplasm. However, as cytoplasmic Rev-Halo was far less concentrated than nucleolar or NPC-associated Rev-Halo, it is hard to visualise both simultaneously. Overexposing Rev-Halo signal to counteract this reveals very similar patterns of localisation compared to HIV-1 gRNA in the cytoplasm (**Figure 4.4, C**). Co-localisation of Rev and gRNA is expected, as the predominant role of Rev is believed to be RNA export. However, as Rev-Halo signal is ubiquitous, it is hard to determine if individual protein/RNA molecules are interacting or if the signals from closely localised molecules are simply overlapping. Surprisingly, Rev-Halo can be seen even at the plasma membrane close to viral gRNA, despite Rev not being known to become packaged into virions (25, 231, 232).

This data shows that tagged Rev can recognise and export vRNAs. Using my novel replicons, I can visualise Rev-mediated nuclear export of vRNAs during HIV-1 replication and similar localization patterns of Rev and gRNA in the cytoplasm. In

upcoming chapters, I will exploit these Rev-tagged replicons to investigate the interactions of Rev and host cell proteins during infection.





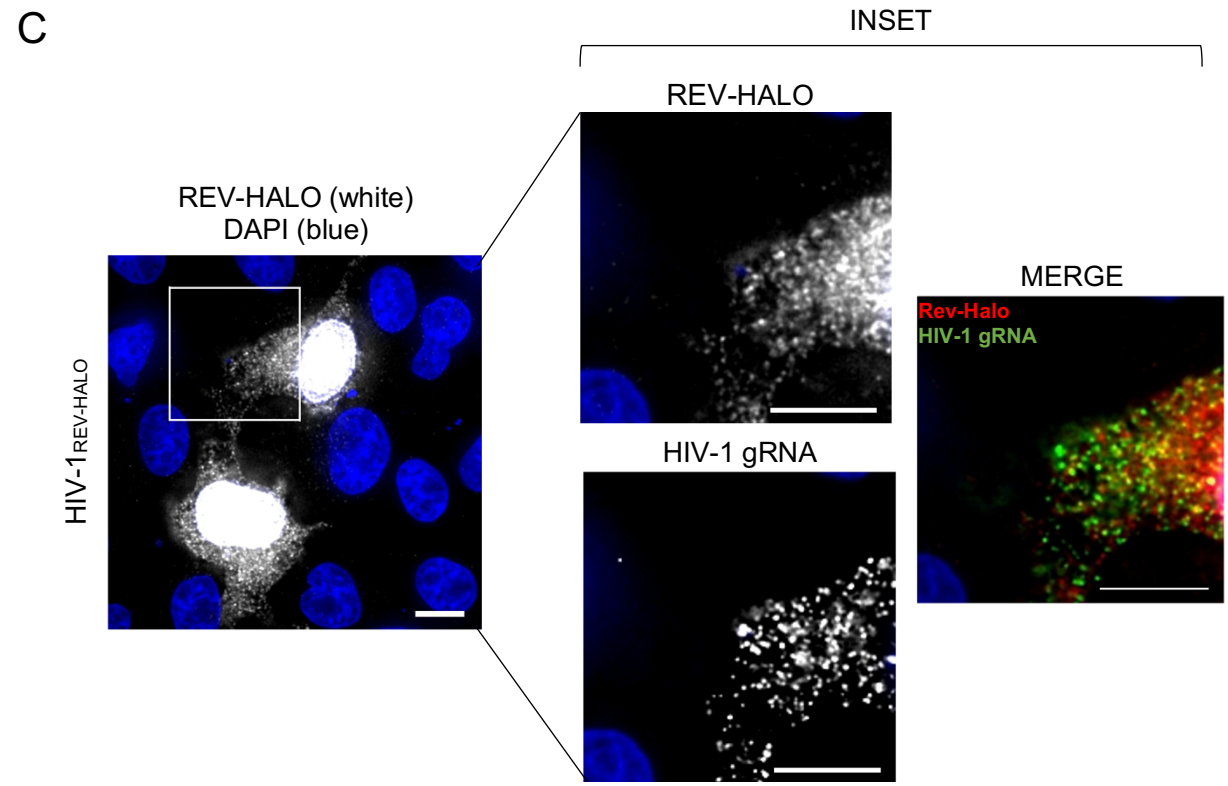


Figure 4.4: Immunofluorescence and fluorescence *in situ* hybridisation of Rev-Halo and HIV-1 gRNA during HIV-1 replication. (**A**) Hela cells transfected with HIV-$1_{mCherry\text{-}Nef,}$ HIV-$1_{\Delta Rev}$ or HIV-$1_{Rev\text{-}Halo}$ replicons were fixed at 48 hpt and labelled with dye-coupled oligonucleotide probes complementary to the 5' intron of HIV-1 gRNA. This revealed white foci corresponding to unspliced HIV-1 gRNA. (**B**) As in HIV-$1_{Rev\text{-}Halo}$-transfected cells in (A), with the additional staining of Rev-Halo using dye-coupled Halo ligands. (**C**) Overexposing signal from HIV-1 gRNA-complementary oligonucleotide probes (green foci in merged image) and Rev-Halo ligands (red foci in merged image) in cells transfected with HIV-$1_{Rev\text{-}Halo}$ at 48 hpt reveals that Rev-Halo and HIV-1 gRNA both localise partially in the cytoplasm. Cells were imaged by fluorescence microscopy on a 100x lens (scale bars = 10 μM). White box indicates area that is enlarged in inset images.

4.6. Discussion

4.6.1. Future Rev studies should utilise near-physiological systems

Rev is an essential HIV-1 protein; nullifying Rev results in HIV-1 failure at expressing the late viral genes (110). Despite its significance in the viral infection cycle, Rev's functions remain poorly understood because of a variety of technical challenges. The

compromises made to overcome these challenges in the study of Rev have resulted in contrasting and often contradictory data. For example, Rev has been reported to assemble on the RRE as a dimer in some studies (63, 71, 72, 76, 86, 87) and a monomer in others (70, 73, 77, 78, 81-85). One study suggested that Rev may not be necessary for the cytoplasmic accumulation of vRNAs (136), directly contradicting many other previous findings (233). To unify published data, the conditions used to study Rev must be standardised and aim to recapitulate physiological infection.

In this chapter, I have described the development of a novel system to study Rev in conditions nearer to physiological infection than previously used. A particular challenge of studying Rev is the inability to tag it and express it to physiological levels from its cognate vRNA. I have designed and engineered Rev-tagged replicons, which express Rev and vRNAs to physiological levels and can be pseudotyped with VSVG to produce singe-round infectious particles. One technical challenge of HIV-1 research is the difficulty of transfecting lymphocytic cell lines. By implementing VSVG pseudotyping, I can achieve high rates of HIV-positive cells in CD4+ T lymphocytic lines such as SupT1. This is particularly important in the case of Rev, which exhibits low abundancy in the infected cell, making it difficult to detect with conventional proteomics approaches or to gain sufficient depth in next generation sequencing experiments. Moreover, expression of Rev in cells does not guarantee that the protein is functional or engaging with the correct pathways. Conversely, I have shown through several orthogonal approaches that tagged Rev expressed from HIV-1 replicons is a) able to export RNAs from the nucleus, b) consequently allows translation of essential viral proteins and c) displays a localisation pattern that is compatible with its function. While this system approximates physiological infection to a greater extent than many of the systems conventionally used to study Rev, there are limitations that future work

should address. Firstly, the use of lymphocytic cell lines is an improvement on the use of adherent CD4- cells lines. However, cancer cell lines also differ biologically from primary cells which are natively infected by HIV-1 (234). Future works could improve my system through the use of primary cell lines. The SupT1 cell line is a T-lymphoblastoid cell line expressing the CD4 receptor which has previously been used extensively for HIV-1 research, though expresses a very low level of the CCR5 receptor (223). However, as the NL4-3 variant of HIV-1 which is encoded in HIV-1 replicons is a CXCR4 strain (224), this will not affect transduction experiments, though it is notable that these experiments could be repeated in different cell lines using different viral strains in future work. Moreover, the replicons I have engineered are derived from NL4-3. pNL4-3 is a recombinant proviral clone generated from DNA from HIV-1 NY5 and LAV strains (235). This clone is therefore not representative of all HIV-1 strains and thus, biological differences in replication may not be accounted for in the use of these replicons. To address this, future studies could repeat this work in additional proviral replicons derived from other clinical isolates to give a fuller picture of Rev functions and interactions during infection. Finally, while the transduction of cells with VSVG-pseudotyped viral particles achieves a high rate of HIV-1-positive cells, it does not mirror the biological methods through which capsid core deposition occurs naturally. Binding of envelope proteins to CD4+ and chemokine co-receptors on immune cells permits fusion at the plasma membrane. While HIV-1 entry has been found by occur by endocytosis, these viral cores become degraded and do not lead to successful infection (236). VSVG-mediated virions enter host cells following internalization into endosomes, and thereafter escape from endosomes (237). The biological differences between these entry mechanisms result in capsid cores becoming deposited at different locations in the cell (endosomal network versus the

plasma membrane) and therefore being trafficked to the nucleus and experiencing early infection steps differently. For example, VSVG-mediated endocytosis was found to circumvent interactions with the cortical actin network (237, 238). Additionally, Nef enhances viral infectivity by influencing an as-yet unknown early-phase facet of HIV-1 infection, perhaps uncoating; this enhancement is lost when VSVG-mediated fusion is used (237). While I have validated that downstream steps of HIV-1 replication, such as translation, occur under my experimental approach, future works could improve upon this system by using fully infectious viral particles to infect CD4+ lymphocytes with envelope(+) HIV-1.

4.6.2. Rev-Halo is a powerful tool to study Rev localisation during infection

Expression of Rev-Halo to physiological levels reveals novel Rev localisations at the plasma membrane and nuclear envelope

Previous Rev IF experiments showed Rev localisation in the nucleolus and/or ubiquitously in the cytoplasm (64, 226-230). However, in my system Rev-Halo also concentrated at the nuclear envelope and plasma membrane at 48 hpt. This novel localisation is a consequence of the powerful specificity and sensitivity of the Halo system. The covalent bond formed between the reactive Halo tag and dye-coupled ligands means that Halo molecules are labelled strongly and specifically with a very bright fluorophore that cannot be lost during the multiple stringent washing steps of smFISH/IF protocols. The novel localisations of Rev-Halo are also likely a consequence of native Rev functions, since previous IF studies ectopically overexpressed Rev without gRNA (64, 227, 229, 230). It has been shown that binding to RNA occludes the Rev NLS, facilitating CRM1 recruitment and nuclear export (63). Likewise, recruitment of importins to Rev in the cytoplasm only occurs when Rev is

not bound to RNA (63). Therefore, binding and releasing vRNA is a critical component of the Rev shuttling cycle. Without vRNA, Rev may be retained in the nucleolus in an intrinsically disordered state, leading to an artificial localisation pattern. One study that performed Rev IF in the context of infection used Hela cells stably expressing Rev-deficient HIV-1_{HXB2} provirus (228), thus including the viral gRNA. Rev-eGFP was ectopically overexpressed in this cell line. Interestingly, the resulting IF shows exclusively nucleolar Rev-eGFP signal, and the vRNA was not imaged. It is probable that Rev overexpression alters Rev/vRNA stoichiometry, leading to the accumulation of non-functional Rev in the nucleolus. This would impair the visualisation of cytoplasmic and nuclear pore-associated Rev-eGFP signal. Our newly observed protein pattern may thus be a consequence of using experimental conditions which closely maintain physiological levels of vRNA and viral proteins. Additionally, many previous IF studies that image a tagged Rev fusion protein do not carry out sufficient controls to ensure that large tags such as eGFP affect the behaviour of Rev under a given experimental system. Indeed, I created a replicon that expressed GFP-tagged Rev, but found that Rev shed this large tag, demonstrating the need for stringent control experiments. I have shown that my HIV-1 replicons encoding tagged Rev do not perturb the RNA export functions of Rev and do not alter the classical nucleolar localisation. In addition, Western blotting against CA protein demonstrates that tagged Rev can sustain the normal expression of vital structural HIV-1 proteins throughout infection.

Concentration of Rev at the nuclear pore is not unexpected, although this is the first time it has been observed to my knowledge. The mechanism for Rev translocation is poorly defined, though Rev must interact with NPC components in order to shuttle between the nucleus and cytoplasm. Indeed, Rev is able to interact with proteins

containing phenylalanine-glycine repeats, a motif conserved in nucleoporins (239). However, research into these interactions is extremely limited and potential binding partners of Rev at NPCs are unknown. Interestingly, HIV-1 infection causes downregulation and remodelling of NPC components (101). In particular, the proteome of the nuclear envelope in HIV-infected versus uninfected cells showed a marked decrease in proteins involved in nucleocytoplasmic transport (101). This downregulation may hinder cellular export pathways, shifting the predominant export pathway in the cell towards Rev/CRM1-mediated translocation. In a similar manner to Rev blocking NXF1/TAP-mediated export (106), this may ultimately increase the levels of under-spliced vRNAs which are exported. The high concentration of Rev-Halo at NPCs suggests that Rev may participate in this NPC remodelling. More mysterious is the localisation of Rev-Halo at the plasma membrane. Rev is thought to be absent in virions. Localisation of Rev at sites distal to the nucleus may simply be a consequence of the ubiquitous distribution of Rev across the cytoplasm. However, since excess Rev localises in the nucleolus, this would suggest that Rev protein that accumulates towards the plasma membrane does so because of specific interactions in the cell. In smFISH experiments, HIV-1 gRNA is spread similarly to Rev-Halo across the cytoplasm and at the plasma membrane. It is also known to participate in viral/host cell RNP complexes in the cytoplasm. Therefore, a pool of Rev may accompany gRNA to distal cytoplasmic sites, likely to participate in vRNA metabolism following RNA export. In upcoming chapters, I will exploit my validated tagged Rev replicons to probe these interactions during HIV-1 infection in CD4(+) T lymphocytic lines.

HIV-1$_{\text{Rev-Tag}}$ as a tool to study the Rev interactions in the nucleolus

As previously seen, Rev-Halo is concentrated in the nucleus and particularly the nucleolus, even after the majority of vRNA has been exported to the cytoplasm. The strong localisation of Rev in the nucleolus has long been documented (64, 91, 226-229), yet is currently unexplained. This nucleolar residence may arise from masking of the NES, as mutation of the NES constrains Rev to the nucleolus. Moreover, Behrens *et al* have shown that deliberate masking and unmasking of the NLS is able to alter this phenotype (240). Though the function behind this localisation is undetermined, Rev mutants that do not localise in the nucleolus are impaired in their ability to export vRNA, resulting in impaired vRNA packaging (227, 228). This suggests that it is critical to the Rev functional cycle and downstream Rev functions. There are multiple studies which indicate that the nucleolus may act as a 'meeting ground' for important Rev interactors. For example, Rev and CRM1 overexpressed separately yield nucleolar and nuclear envelope localisation, respectively (177). However, when both are overexpressed together, CRM1 mobilises to the nucleolus in a Rev-dependent manner. FRET measurements suggest that Rev dimerization occurs in the nucleolus, which may be important for binding to the RRE (241). Another interaction that may occur in the nucleolus is the binding of Rev to ribosomal (r)RNA or ribosomal transcription components. Rev relocalises from the nucleolus to the cytoplasm upon treatment with RNA polymerase I and III inhibitors, but is unaffected by inhibitors of RNA polymerase I (242), suggesting that pre-rRNA transcription is a determinant of Rev nucleolar localisation (242, 243). Indeed, Rev is localised in dense fibrillar and granular compartments of the nucleus, which are important for rRNA transcription (244). While the interactions of nucleolar Rev are unclear, HIV-$1_{\text{Rev-Halo}}$ represents a specific and powerful tool that could be used to determine precise localisation of Rev during infection. The difficulty in defining precise co-localisation of Rev-Halo/HIV-1

gRNA could, in future work, be overcome using centroid analysis. Centroid analysis can determine the location of a single RNA foci within a few nm and has been employed in techniques such as fluorescence imaging (245). This technique can determine the center of individual fluorescent particles to a precision about an order of magnitude greater than the microscope resolution, where the center is less challenging to pinpoint than the diffraction-limited edges of an individual foci. Centroid analysis therefore allows a more precise understanding of whether 2 molecules are co-localised or simply close to each other.

5. Revealing the landscape of Rev/host cell protein-protein interactions in HIV-1 infected T CD4+ lymphocytic cells

5.1. Introduction

HIV-1 cannot encode all the machinery required for its replication and spread. Thus, it is heavily reliant on host cell resources (246). One prevalent strategy that viruses use to hijack cellular resources is to express viral proteins that interact with and recruit key cellular factors (247). By knowing which cellular proteins Rev interacts with, it is thus possible to obtain deeper mechanistic insights into its regulatory roles. The most common approach for discovering protein-protein interactions (PPIs) involves IP followed by mass spectrometry analysis. This approach has been employed to reveal the interactomes of HIV-1 proteins (248-252), including a few studies focusing on Rev (249, 252-254). Disappointingly, there is no overlap between the three available Rev PPI studies (**Figure 5.1**), which is probably due to technical limitations discussed in chapter 4 and the non-physiological approaches used to overcome these limitations (**Table 5.1**). For example, each study used a different tagged Rev fusion protein (249, 252, 253). IP efficiency depends on strength and selectivity of antibody/epitope binding and efficacy of elution. As my work in this chapter shows, fusion proteins harbouring different tags are captured, cleaned and eluted from commercial affinity systems with varying efficiencies. Thus, different IP approaches are likely to enrich Rev-interacting proteins to different extents. Moreover, none of these studies demonstrated that activity of their fusion proteins was not altered by the tag (249, 252, 253). As Rev is highly disordered and folds upon interaction with RNA, tagging may alter this folding or reduce the accessibility of the ARM motif that is critical for RNA binding (255). Each Rev IP study also uses different cellular systems, one of them using cell extracts instead of whole cells. As transcriptomes and proteomes vary between cell lines (213) and lysis can lead to non-physiological interactions due to the mixing of intracellular compartments, each study may purify Rev PPIs which are distinct, explaining the poor

agreement between studies. Importantly, both Arizala and Naji *et al* purified Rev in CD4(-) cell lines, which are not natively infected by HIV-1 and may lead to the identification of interactions that do not occur in natural infection. Only Arizala *et al* used cells lines stably expressing a ΔRev replicon (**Table 5.1**). Although this does not fully recapitulate infection conditions, it is a more suitable scenario to study Rev PPIs.

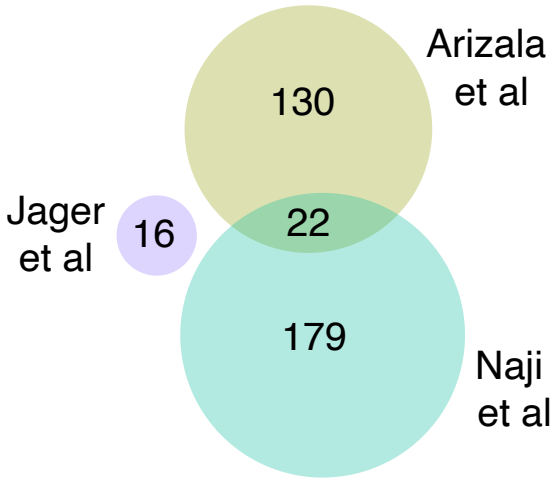


Figure 5.1: 3-way Venn diagram illustrating the overlap between Rev interactome studies. There are 3 publicly available datasets generated from studies which used IP and mass spectrometry to investigate the Rev interactome. Overlap between identified Rev interactors in these studies is poor, as seen by the lack of any 3-way overlap. Proteins exhibiting >0% binding to wildtype Rev from Arizala *et al* were considered. Proteins identified in both Jurkat and HEK293 from Jager *et al* were considered. For Naji, the top 250 Rev interactors were considered.

Table 5.1. Experimental specifications of the 3 publicly available Rev IP interactome studies

STUDY	REV EXPRESSION	RNASE	CELL LINE
Arizala *et al* (2019) (253)	Plasmid encoding Rev-Flag transfected in Hela cells expressing a stably-integrated ΔRev HIV- 1 provirus	No	HLfB
Naji *et al* (2012) *(249)*	Purified MBP-Rev incubated in nuclear or cytosolic cellular extracts (+/- purified RRE, +/- RNase)	Yes/No	Hela
Jager *et al* (2011) (252)	Rev-2xStreptavidin-3xFlag transfected in HEK293 or induced from stably transfected Jurkat cell lines	No	Jurkat and HEK293

Due to these diverging experimental systems, individual Rev PPIs are often contradicted between studies. For example, eukaryotic initiation factor (eIF)5A and Sam68/KHDRBS1 are 2 proteins that are heavily studied as putative Rev interactors and, indeed, have been referred to as a 'Rev co-factor' and 'Rev homologue' respectively (256-258). Yet, evidence of their interactions with Rev remains controversial. For example, eIF5A can bind specifically to the Rev NES (259, 260) and hampers nuclear export of Rev-CRM1 complexes when impaired (259-261). However, several groups subsequently failed to confirm the existence of an interaction between eIF5A and Rev (63, 262). Sam68 is an RBP often linked to Rev, despite their direct binding being demonstrated solely through co-IPs when both proteins are overexpressed in CD4(-) cell lines (263). For example, Sam68 can bind a synthetic RRE-RNA when overexpressed (229), but has not been shown to bind the RRE of a full HIV-1 RNA expressed to physiological levels. Other evidence linking Sam68 and Rev is correlative rather than direct. For example, overexpression of Rev in 293T cells transfected with Rev(-) proviral DNA increased free CA levels 25 to 30-fold, and further addition of a Sam68 variant increased this to 120-fold (229). Indirect evidence identified in non-physiological cell lines means that many reported Rev interactors still require strict validation using standardised, near-physiological approaches. In this chapter, I detail the establishment and optimisation of an approach that better approximates physiological infection and use it to unveil the interactome of Rev in HIV-1-transduced CD4+ lymphocytic lines for the first time.

5.2. Tagged Rev is efficiently purified using a Flag IP

To find the most efficient system to IP Rev, I took advantage of the toolkit of tagged-Rev HIV-1 replicons constructed and validated previously and performed IPs of each

tagged Rev protein. In brief, SupT1 CD4(+) lymphocytic cells were transduced with pseudotyped HIV-$1_{Rev\text{-}Flag}$, HIV-$1_{Rev\text{-}Halo}$ or HIV-$1_{Rev\text{-}Spot}$ viral particles in parallel with pseudotyped HIV-$1_{mCherry\text{-}Nef}$ particles as a negative control. The volume of pseudotyped virus was adjusted to produce ~80% infected cells at 48 hpt, as indicated by mCherry fluorescence (data not shown). This allowed me to achieve a high rate of HIV-1-positive cells while reducing superinfection, in which multiple virions infect a single cell and cause rapid infection progression and cell death. While it is likely that cells under these conditions could indeed by infected by multiple virions, a linear relationship between fluorophores and MOI has been demonstrated (264), suggesting that at least some of the cells in this system avoid superinfection. Regardless, a high MOI is needed to overcome low Rev expression levels during infection. Cells were lysed at 48 hpt, as abundance of Rev has been shown to peak at this time (265). Typically, all early phases of the HIV-1 infection cycle (*i.e.,* membrane fusion, reverse transcription and integration) are completed after 12 h (265). Therefore, for cells undergoing non-latent infection, integration and expression of vRNAs, as well as plasma membrane assembly of virions, occurs by 48 hpt (266). This allows potential host-virus PPIs from every step of the HIV-1 replication cycle to be represented across the asynchronously infected population of cells. Lysis was performed in the presence of bensonase, an endonuclease able to degrade all forms of RNA and DNA. Bensonase is required to remove RNA-mediated interactions and ensure only direct PPIs are isolated. This step is critical to elucidate Rev PPIs, as Rev binds to and oligomerises on vRNA and can co-precipitate proteins through RNA bridges. Lysates were then subjected to IP with anti-Flag beads, following a protocol previously established in the lab, or with Halo- or Spot-trap beads according to manufacturer's instructions. Briefly, beads were incubated with lysates for 1 h at 4ºC and washed 3

times to remove contaminating proteins. Composition of wash buffers included physiological levels of salt (150 mM NaCl). Finally, purified Rev/host cell protein complexes were released with competing Spot and 3x Flag peptides for 1 h at 4ºC. Rev-Halo was eluted by mixing beads in a low pH solution to denature bonds between bead antibodies and their epitopes. It is note-worthy that Halo-trap agarose consist of agarose beads covalently bound to anti-Halo nanobodies rather than Halo ligands and therefore do not form covalent bonds with their substrates.

To determine whether tagged Rev proteins were successful isolated, I analysed input and eluate samples of each IP by silver staining (**Figure 5.2, A-C**). These silver staining analyses revealed high band densities at all molecular weights in input samples, reflecting the whole cell proteome as expected. Conversely, eluate lanes had only few bands, the most prominent at 50, 25 and 18 kDa in HIV-$1_{Rev\text{-}Halo}$, HIV-$1_{Rev\text{-}Flag}$ and HIV-$1_{Rev\text{-}Spot}$ silver stains respectively (**Figure 5.2, A-C,** marked '*****'). These bands correspond to Rev-tagged proteins and were absent in HIV-$1_{mCherry\text{-}Nef}$ lanes. Thus, silver stains showed that each tag system could specifically purify tagged Rev. This also confirmed that Rev was expressed in infected cells to sufficient levels for deep PPI analyses. Rev-Halo elution was inefficient even after implementation of several elution rounds (data not shown). Halo-trap agarose binds Halo protein very strongly at a Kd of 2 nM, 3 times stronger than the interaction between Spot-trap beads/Spot peptide (220). This can explain the poor elution efficiency of Rev-Halo. Boiling Halo-trap agarose at 95ºC rendered a stronger IP band which was confirmed by Western blotting to be Rev, but this approach also eluted many contaminating proteins (data not shown). On the other hand, Rev-Spot and Rev-Flag IP eluates were consistently efficient and produced strong, singular bands based on silver staining (**Figure 5.2, B-C**).

To determine the performance of these IPs to a greater extent, single replicates of both HIV-$1_{Rev\text{-}Flag}$ and HIV-$1_{Rev\text{-}Spot}$ IPs with corresponding HIV-$1_{mCherry\text{-}Nef}$ negative control samples were analysed by label-free mass spectrometry analysis (**Figure 5.2, D-E**). Rev-Flag and Rev-Spot were the most enriched proteins in respective IPs (marked on plots), with the highest protein intensities and fold changes. This indicates that both IPs isolated tagged Rev proteins efficiently. However, most proteins in the HIV-$1_{Rev\text{-}Spot}$ IP exhibited a fold change close to 0 (**Figure 5.2, E**, red box), indicating that they bound to Spot-trap beads and to Spot-trap beads/Rev-Spot complexes with similar affinity. Thus, binding of these proteins was non-specific, signifying that they were contaminants. Moreover, there were a limited number of proteins with a fold change >2 in both IPs, indicating a poor enrichment of proteins in HIV-$1_{Rev\text{-}Spot}$/HIV-$1_{Rev\text{-}Flag}$ eluates over HIV-$1_{mCherry\text{-}Nef}$ eluates. Finally, both IPs exhibited high numbers of proteins with a negative fold change (<0) (**Figure 5.2, D-E, grey dots**). These proteins are enriched in negative control eluates over HIV-$1_{Rev\text{-}Spot}$ and HIV-$1_{Rev\text{-}Flag}$ eluates and are contaminants. Hence, these scatter plots showed that although both IPs could efficiently pull-down tagged Rev, there was a high volume of non-specific proteins binding to Spot-trap and anti-Flag beads and a poor enrichment of Rev interactors. Particularly, Rev-Spot underperformed with regards to isolating potential Rev-interacting proteins (fold change >2) when compared to Rev-Flag. Therefore, Rev-Flag was selected for further optimisation.

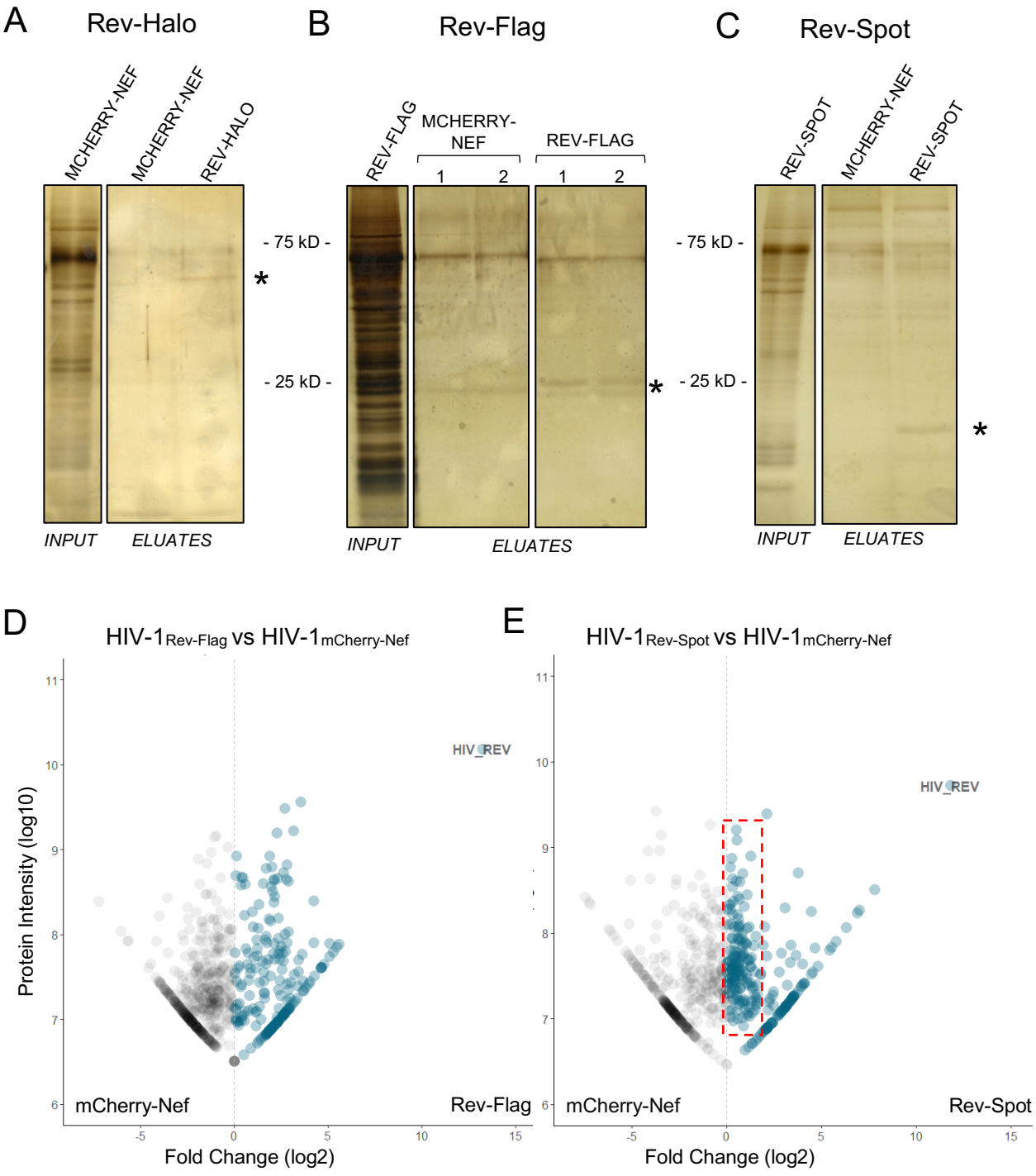


Figure 5.2: Immunoprecipitation of tagged Rev using different tags and their commercial antibodies. Lymphocytes were infected with HIV-1 replicons expressing Rev-Halo (**A**), Rev-Flag-3xMyc (**B**) and Rev-Spot (**C**) and untagged Rev protein (HIV-$1_{mCherry-Nef)}$ as a negative control. Infected cells were lysed and captured using anti-Flag or Spot-trap beads, or Halo-trap agarose. Inputs of whole cell lysates and eluates were analysed by silver staining. '*' marks expected molecular weight of the relevant Rev-tag protein. (**D-E**) Single-replicate eluates of Rev-Spot and Rev-Flag IPs were analysed by mass spectrometry. X axes plot the fold change in protein intensity in

Rev-Flag/mCherry-Nef eluates (log2) and Y axes plot the combined protein intensities (log10). Red box indicates host cell proteins interacting with Rev-Spot clustering close to 0 fold change, indicating that they may be false positive interactors.

To improve the purity and efficiency of the Rev-Flag IP, 2 steps were undertaken. Firstly, a new washing buffer was formulated that maintained physiological salt levels (150 mM NaCl), no denaturing agents and a single non-ionic detergent (0.2% IGEPAL), since excessive detergents and denaturing agents can disrupt native PPIs. Secondly, the number of washing steps was carefully optimised by carrying out Rev-Flag IPs with 2, 4, 6 and 8 washes. To ascertain the purity and efficiency of these IPs, I analysed input and eluate samples by silver staining (**Figure 5.3, A**). This analysis showed that the Rev-Flag IP band became more prominent as the number of washes increased, since increasing washes efficiently removed background contaminants without disturbing Rev-Flag binding to anti-Flag beads. In addition, IPs performed with 6 and 8 washes revealed bands in HIV-$1_{Rev\text{-}Flag}$ lanes that did not correspond to Rev-Flag and which were absent from HIV-$1_{mCherry\text{-}Nef}$ lanes. Thus, these bands may well represent interacting cellular proteins. These results confirmed that the increase in the number of washes and the modification of the IP buffer helped to increase the abundance of Rev/interactor complexes in eluates while removing contaminant proteins. Six washes were selected for a large-scale IP experiment, as these conditions represent an optimal equilibrium between efficiency and specificity. Western blotting analysis revealed that Gag/CA expression was identical in HIV-$1_{Rev\text{-}Flag}$ and HIV-$1_{mCherry\text{-}Nef}$ infected cells (*i.e.,* inputs) (**Figure 5.3, B**). Immunoblotting with anti-myc antibody showed strong bands matching Rev-Flag-3xMyc in inputs of HIV-$1_{Rev\text{-}Flag}$, but not in HIV-$1_{mCherry\text{-}Nef}$ inputs, confirming the specificity of the anti-myc antibody. This myc band appeared even stronger in Rev-Flag eluates and was absent

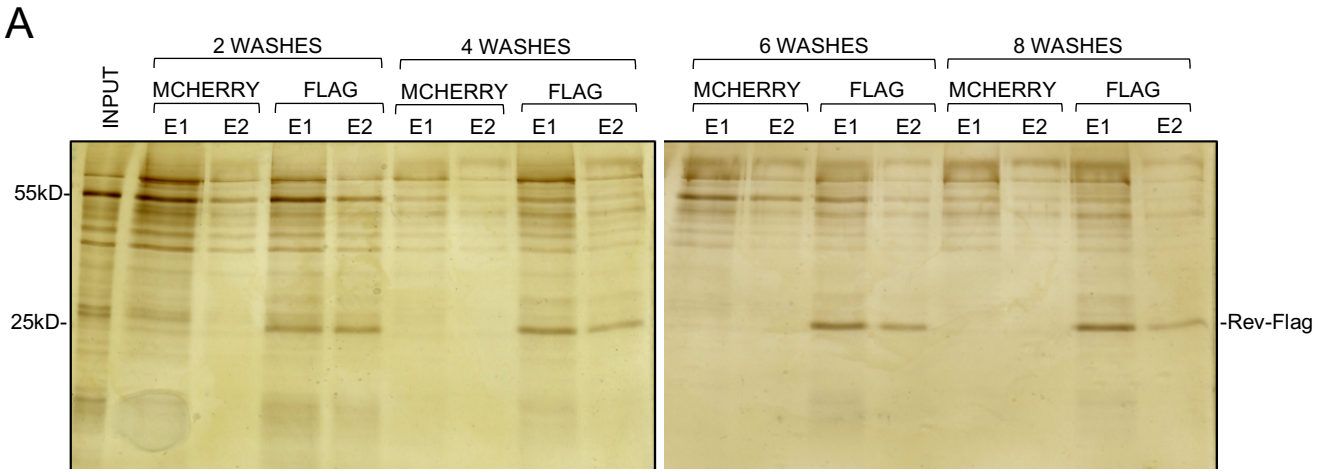
in HIV-$1_{mCherry\text{-}Nef}$ eluates, which demonstrated that elution of Flag beads was highly efficient and specific.

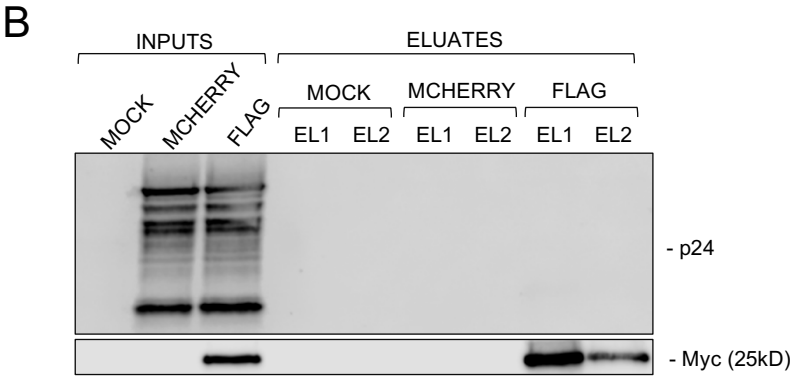
To ensure that bensonase was effectively degrading nucleic acids during lysis, the Flag IP was repeated with 2 HIV-$1_{Rev\text{-}Flag}$-infected samples, which were lysed with or without bensonase. Following IP, RNA was extracted from eluates and resulting RNA levels measured using chip-based capillary electrophoresis (bioanalyser, Agilent) (**Figure 5.3, C**). The bioanalyser assay showed that without bensonase, Rev-Flag IP eluates contained 1,864 pg/µl of RNA, ranging from ~20 to ~2000 nt. The addition of bensonase almost entirely eliminated detectable RNA, reducing total RNA in the eluate to 22 pg/µl. This confirms that bensonase eliminates almost all detectable RNA in IP eluates and severely diminishes the possibility of detecting PPIs via RNA bridges. It should be noted that interactions that require an RNA bridge, such as those that nucleate on vRNA, may be missed with the use of bensonase. However, these PPIs may have been identified in prior Rev interactome studies which lacked RNase treatment; in this way, my interactome identifies a specific and possibly novel subset of Rev interactors.

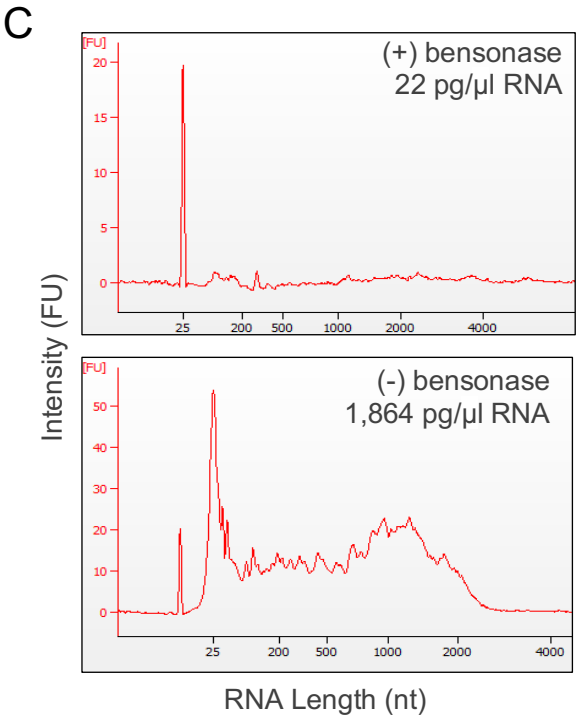
Finally, to assess the quality of the optimised Rev-Flag IP, a single-replicate IP was performed with optimised washing conditions and analysed by label-free proteomics in collaboration with Honglin Chen and Professor Shabaz Mohammed. A scatter plot of the results was made and compared to the previous single-replicate IP results (**Figure 5.3, D**). The optimised IP showed a drastic reduction in the volume of contaminant proteins (those with fold change less than or close to 0). There was also a strong increase in the number of proteins with a high log2 fold change. Indeed, the straight diagonal line on the right-hand side of the plot is an artefact from intensity imputation in the control, as these proteins are only detected in HIV-$1_{Rev\text{-}Flag}$ eluates

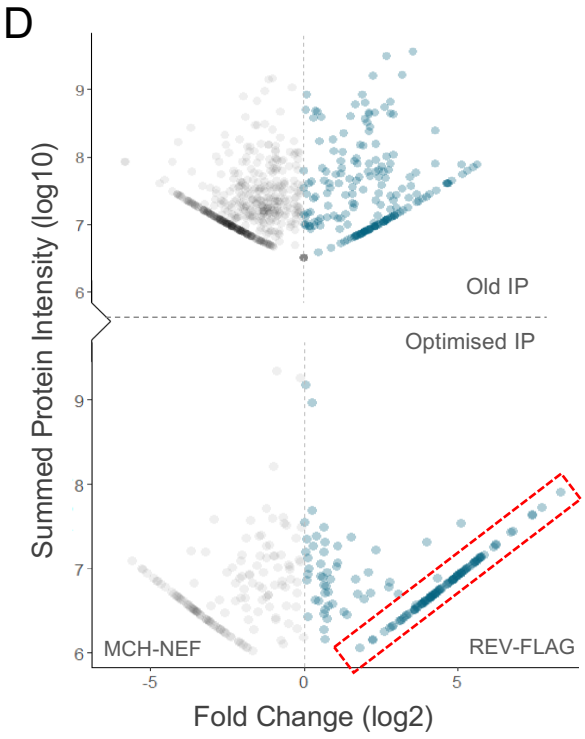
(**Figure 5.3, D, red box**). The number of proteins only detected in Rev-Flag samples is substantially higher with the optimised conditions, suggesting an enhanced detection of Rev interactors. Therefore, my optimised Rev-Flag IP allowed me to identify direct Rev interactors by reducing the enriched background noise. Importantly, this IP was performed in lymphocytic cell lines transduced with pseudotyped HIV-1, which expresses viral proteins and vRNA to physiological levels, recapitulating natural infection conditions to a greater extent than has conventionally been used in Rev research.

(below) Figure 5.3: Rev-Flag is efficiently purified using an optimised Flag IP under conditions approximating physiological HIV-1 infection. Lymphocytes transduced with HIV-$1_{mCherry\text{-}Nef}$ and HIV-$1_{Rev\text{-}Flag}$ were lysed in the presence of bensonase and subjected to IP with anti-Flag beads. Inputs and eluates generated in IPs that used different numbers of washes with an optimised wash buffer were analysed by silver staining (**A**) and Western blotted against p24/CA and myc (**B**). (**C**) RNA extraction was performed on Rev-Flag IP eluates lysed with or without bensonase and the amount of RNA analysed using a bioanalyser. The sharp peak at 25 nt is a control carried out for each bioanalyser run and does not reflect eluate RNA. (**D**) Single-replicate proteomics analysis of HIV-$1_{Rev\text{-}Flag/mCherry\text{-}Nef}$ IP eluates was carried out for IPs using both the old (top panel) and the optimised (lower panel) wash buffer conditions and compared by scatter plots. X axes plot the fold change in protein intensity in Rev-Flag/mCherry-Nef eluates (log2) and Y axes plot the protein intensity summed from both eluates (log10). Individual dots are proteins, with grey indicating a foldchange < 0 and blue indicating a fold-change > 0. Rev box outlines the increased number of proteins detected solely in Rev-Flag eluates (**D**).


A
2 WASHES
4 WASHES
6 WASHES
8 WASHES
INPUT
MCHERRY
FLAG
E1
E2
55kD-
25kD-
-Rev-Flag


B
INPUTS
ELUATES
MOCK
MCHERRY
FLAG
EL1
EL2
- p24
- Myc (25kD)


C
(+) bensonase
22 pg/µl RNA
(-) bensonase
1,864 pg/µl RNA
Intensity (FU)
RNA Length (nt)


D
Old IP
Optimised IP
MCH-NEF
REV-FLAG
Summed Protein Intensity (log10)
Fold Change (log2)

5.3. The Rev interactome comprises 282 host cell proteins

Once the Rev-Flag IP was established and optimised, 3 replicates were performed and analysed by label-free mass spectrometry (LC-MS and initial analyses performed by Honglin Chen). Quantitative proteomics revealed 282 host cell proteins significantly enriched in HIV-$1_{Rev\text{-}Flag}$ over HIV-$1_{mCherry\text{-}Nef}$ samples (**Figure 5.4, A**, **Supplementary Table 1**). Of these proteins, 141 exhibited an adjusted p value < 0.01 (marked in red) and the other 141 exhibited an adjusted p value between 0.01 and 0.1 (marked in orange). The number of contaminants specifically enriched in HIV-$1_{mCherry\text{-}Nef}$ eluates over HIV-$1_{Rev\text{-}Flag}$ eluates is remarkably low, further supporting the high quality of the IP. Rev-Flag protein exhibited the highest fold change and the lowest adjusted p value, denoting its strong and specific enrichment by the Flag IP. To determine the reproducibility of individual IP samples, I plotted the correlation of protein intensities across all samples (**Figure 5.4, B**) and calculated r correlation coefficients. This yielded an average r value of ~0.8, showing that samples heavily correlate with one another and that IPs are highly reproducible.

If my system truly allows me to identify true Rev PPIs, I expect to enrich many known Rev interactors and proteins involved in HIV-1 infection. To investigate this, I compared Rev interactors to those listed in the national center for biotechnology information (NCBI) HIV-1 database (267), which catalogues human proteins which have been linked to HIV-1. I classified proteins in my Rev interaction dataset using this resource as (a) proteins that were known to interact with Rev, (b) proteins that have been linked to HIV-1 previously or (c) novel proteins not previously related to HIV-1 or Rev (**Figure 5.4, C**). Despite the heterogeneous quality of the known Rev interactome, 25% of the proteins in my dataset have previously been reported to interact with Rev

directly. Moreover, many of these proteins are well-established Rev interactors which have been linked to various steps of HIV-1 infection. For example, factors that are essential for nuclear import of Rev, such as NPM1/B23 (254, 268-270) and NAP1L1 (254) are identified. Moreover, several DEAD/H box helicases are putatively reported to interact with Rev to promote RNA packaging into virions and Rev-mediated RNA export (258), and I identified here DDX24 (149), DDX21 (271) and DDX9 (272). Additionally, 42% of the proteins have been linked to HIV-1 previously but were not known to interact directly with Rev. Finally, ~30% of the Rev interactome comprises proteins that have not previously been linked to Rev or HIV-1. This strong enrichment of both HIV-1-linked proteins and Rev-interacting proteins (almost 70% of Rev PPIs) suggests that that Rev interactions I have uncovered reflect true interactions that occur during HIV-1 infection.

PPIs identified by previous Rev interactome studies do not overlap (**Figure 5.1**). To determine if proteins in my dataset have been reported in these studies, I generated an upset plot using the R package UpsetR (204) (**Figure 5.4, D**), which shows degree of intersection between datasets. When comparing to previous interactomes, I find that a striking 43% of the Rev interactome has been identified in at least 1 previous interactome. In this way, my Rev dataset appears to corroborate many factors that were incidentally identified, adding confidence to those interactions and highlighting host cell proteins whose Rev-binding activity is conserved across different cell lines.

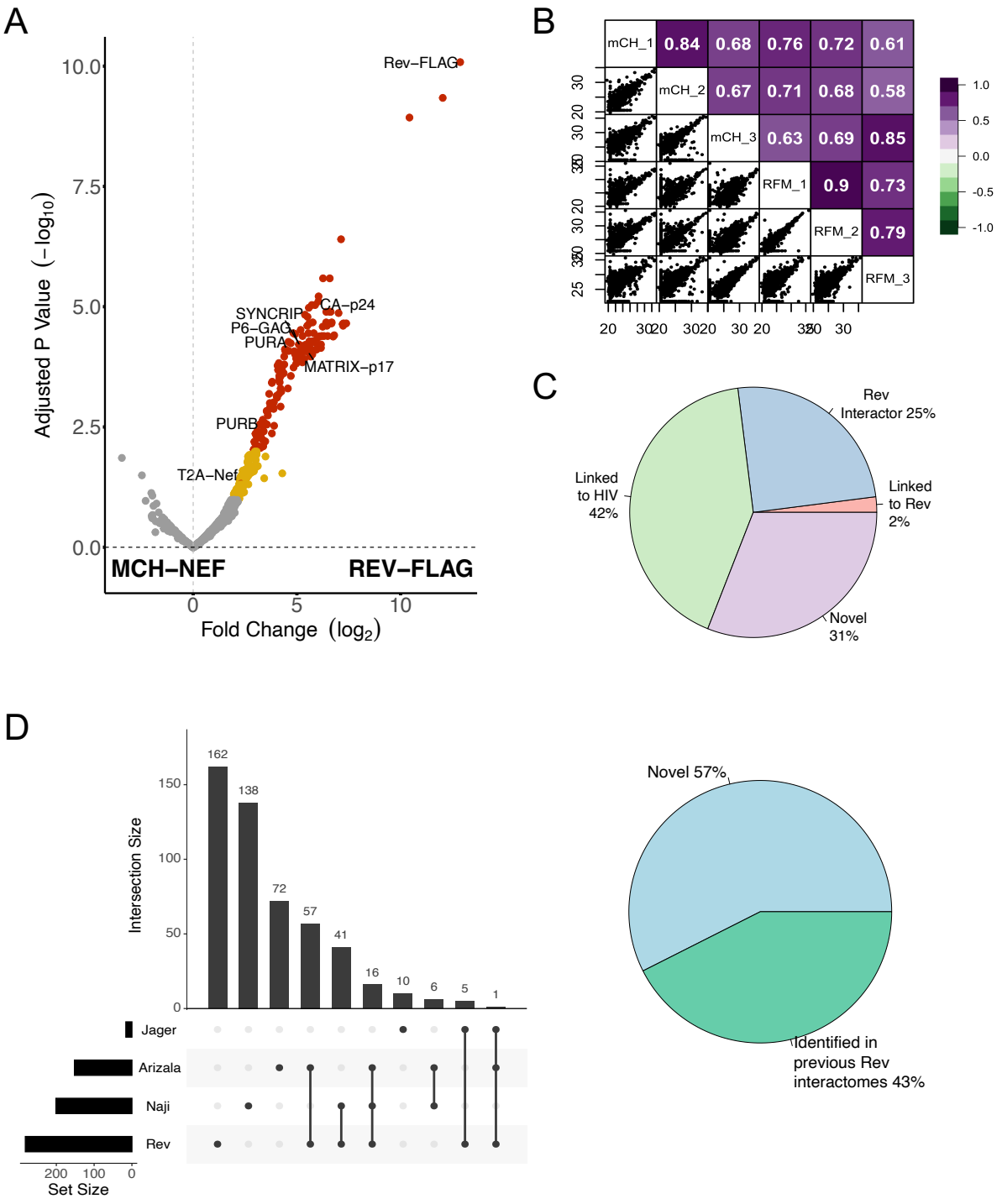


Figure 5.4: The Rev interactome is enriched in known HIV-1 interactors. (**A**) IP using anti-Flag beads was carried out in triplicate in HIV-$1_{Rev\text{-}Flag}$ versus HIV-$1_{mCherry\text{-}Nef}$ transduced lysates and plotted as a volcano plot, showing log2 fold change of protein intensities in HIV-$1_{Rev\text{-}Flag}$ versus HIV-$1_{mCherry\text{-}Nef}$ eluates on the X axis and adjusted -log10 p value on the Y axis. Individual dots correspond to proteins: grey dots indicate non-significant proteins, yellow dots indicate a p value > 0.01 and < 0.1 (significant) and red dots indicate a p value < 0.01 (highly significant). This revealed 282 significant host cell proteins that interact with Rev-Flag. (**B**) Left-hand side: Scatter plots where dots indicate proteins and X and Y axes plot protein intensity in individual

mCherry-Nef and Rev-Flag eluate samples. Right-hand side: the correlation coefficient of the plot on the left-hand side as indicated by labelled diagonal boxes. 0 = no correlation, -1 = strong negative correlation, 1 = strong positive correlation. (**C**) Pie-chart: annotating the 282 significant Rev interactors using the NCBI HIV database with indicated categories revealed a strong enrichment in known Rev interactors and proteins involved in HIV-1 infection. (**D**) Upset plot/pie-chart: comparison of Rev interactors identified by previous Rev IP studies with my Rev interactome shows that 43% of my dataset has been identified in at least 1 previous Rev IP study, despite their lack of overlap. Vertical black bars plot the number of proteins in the intersection indicated directly below, visualised as vertical black lollipop bars joining different studies. Horizontal black bars indicate total number of proteins in each study.

5.4. Rev interacts with RBPs involved in splicing and translation

To globally identify functional networks enriched in the Rev interactome, I used the STRING protein-protein association database (207) (**Figure 5.5**). This analysis revealed large clusters of proteins that participate in RNP complexes which interact with and/or modify RNA. There is a large cluster of proteins involved in ribosome genesis and function. Indeed, almost 40 proteins from the 60S ribosomal subunit and 19 from the 40S ribosomal subunit are present in the Rev interactome, suggesting that Rev recruits ribosomal proteins for translation of viral mRNAs, in agreement with previous studies (136, 137). Several large complexes are present in the dataset, including the 80S ribosome, despite efficient RNase treatment, suggesting that Rev IP enriches for both proteins that interact directly with Rev and proteins that bind indirectly via complex PPI networks. In addition to ribosomal proteins, I identified Rev-Flag-interacting proteins involved in initiation of translation, such as eIF4E, which is responsible for the recognition of the cap in mRNAs, eIF1, eIF5A and eIF6. These eIF proteins are all novel Rev interactors. STRING analysis, which charts known and precited PPI networks based on protein participation in common functions, complexes, subcellular organelle localisations etc., also revealed an array of splicing-related proteins, including 16 heterogeneous nuclear ribonucleoproteins (HNRNPs), many of

which are known to regulate alternative splicing of HIV-1 genes (273). Several HNRNPs have been shown to interact with Rev and promote HIV-1 infection (274). Moreover, serine/arginine-rich splicing factor (SRSF) proteins are involved in HIV-1 splicing and viral transcription (275) and many, including SRSF2, 3, 5, 6, 10, SR repetitive matrix 1 (SRRM1) and SR protein phosphatase PP1-alpha catalytic subunit (PPP1CA), are found in the Rev interactome. Finally, STRING analysis showed an enrichment of proteins that participate in mRNA surveillance pathways. For example, there are proteins involved in polyadenylation and cleavage of pre-mRNA (CPSF1, CSTF3) and the nonsense-mediated decay (NMD) pathway, which is responsible for degrading mRNAs with premature stop codons (UPF1). These data show a striking enrichment in RBPs with regulatory roles in RNA metabolism. Previous research has posited roles for Rev in RNA splicing, stabilisation, transcription and translation but no molecular mechanisms for these Rev-mediated effects have yet been discovered. The Rev-Flag interactome reveals vital PPIs which may underpin these enigmatic Rev functions. In this way, I have established a database of proteins for future works to examine to elucidate the roles of Rev in HIV-1 infection.

A

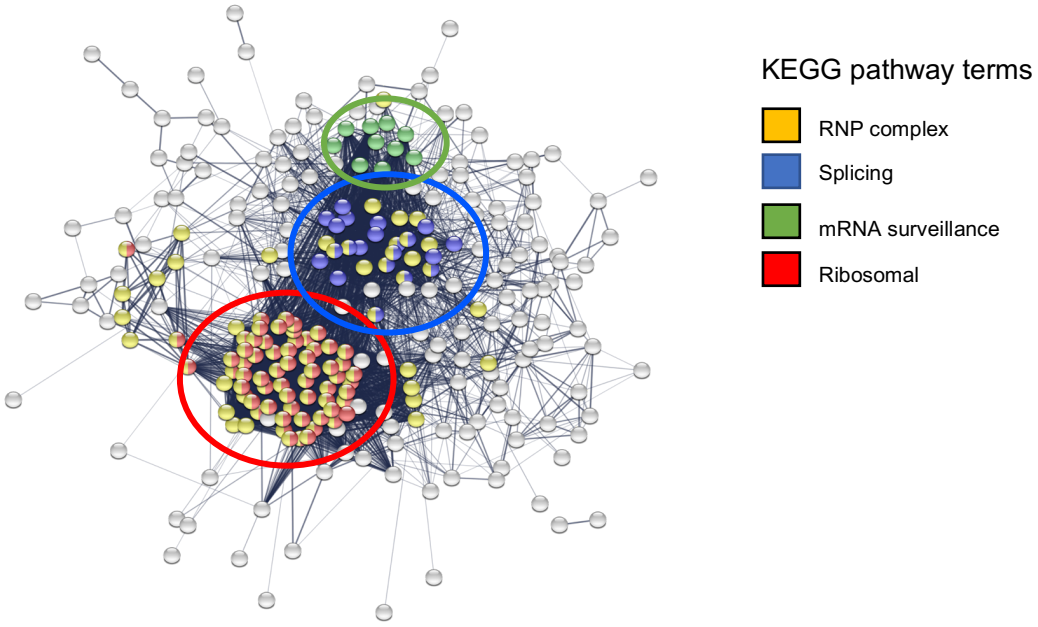


Figure 5.5: The Rev interactome is enriched in RBPs that are involved in vRNA metabolism. The Rev interactome was annotated using STRING software, which charts known and precited PPI networks. Here, prediction is based on Kyoto encyclopaedia of genes and genomes (KEGG) pathway terms. These terms refer to molecular interactions underpinning systemic functions of the cell and the organism. Individual nodes represent proteins and links between them represent common functionality, as indicated by the colour of the node, which is described in the legend. Large circles indicate clusters of proteins identified to participate in a common physiological function.

To ascertain functions underpinning Rev PPIs, GO terms of Rev interactors were analysed using the gene list enrichment analysis tool enrichR (206). The top 8 most enriched terms were plotted using log2 odds ratio (**Figure 5.6**). Bars were coloured based on p value, with deeper purple indicating smaller p values. The most enriched 'biological process' term was '*SRP-dependent co-translation protein targeting to membrane*,' which refers to the signal recognition particle (SRP) (**Figure 5.6, A**). The SRP is a cytosolic complex that recruits the ribosome/nascent peptide chain complex to the endoplasmic reticulum. Interestingly, I identified that 5/6 of the proteins which form the SRP interacted with Rev-Flag (SRP54, SRP72, SRP19, SRP9 and SRP68). Rev may co-opt this complex to sites of viral translation in the cytoplasm to ensure swift processing of viral proteins that are modified in the secretory pathway, such as envelope glycoproteins SU and TM. The term '*viral transcription*' was also enriched, along with terms related to cellular protein synthesis, such as '*gene expression*,' '*translation*,' and '*cellular macromolecule biosynthetic process*.' These GO terms, together with enriched STRING annotations, suggest that Rev may regulate HIV-1 RNA expression at every step, by recruiting host cell proteins important for transcription, splicing and translation. 'Cellular component' GO terms revealed an enrichment in both cytoplasmic and nuclear bodies (**Figure 5.6, B**), which is expected given the nucleocytoplasmic shuttling of Rev and demonstrated through localisation

studies of Rev-Halo detailed in chapter 4. The nucleolus was also a strongly enriched term, which also agrees with the observed nucleolar localisation of Rev-Halo. The most enriched ‘molecular function’ GO term was ‘RNA binding’ (**Figure 5.6, C**). Similarly, all other significant ‘molecular function’ terms denoted interactions with nucleic acids, including ‘dsRNA binding.’ This stark enrichment of RNA- and DNA-binding proteins suggests that many components of the Rev interactome are specifically recruited by Rev to regulate the HIV-1 vRNA. Terms such as ‘dsRNA binding’ and ‘snRNA binding’ illuminate that Rev may contact RBPs to modulate host cell nucleic acids in addition to vRNA. To my knowledge, the relationship between Rev and host cell nucleic acids is entirely unstudied, though HIV-1 infection has been found to increase global intron retention in host cell RNAs (133). Therefore, Rev may promote the stability and/or export of intron-containing host cell RNAs, perhaps to dysregulate gene expression during HIV-1 infection.

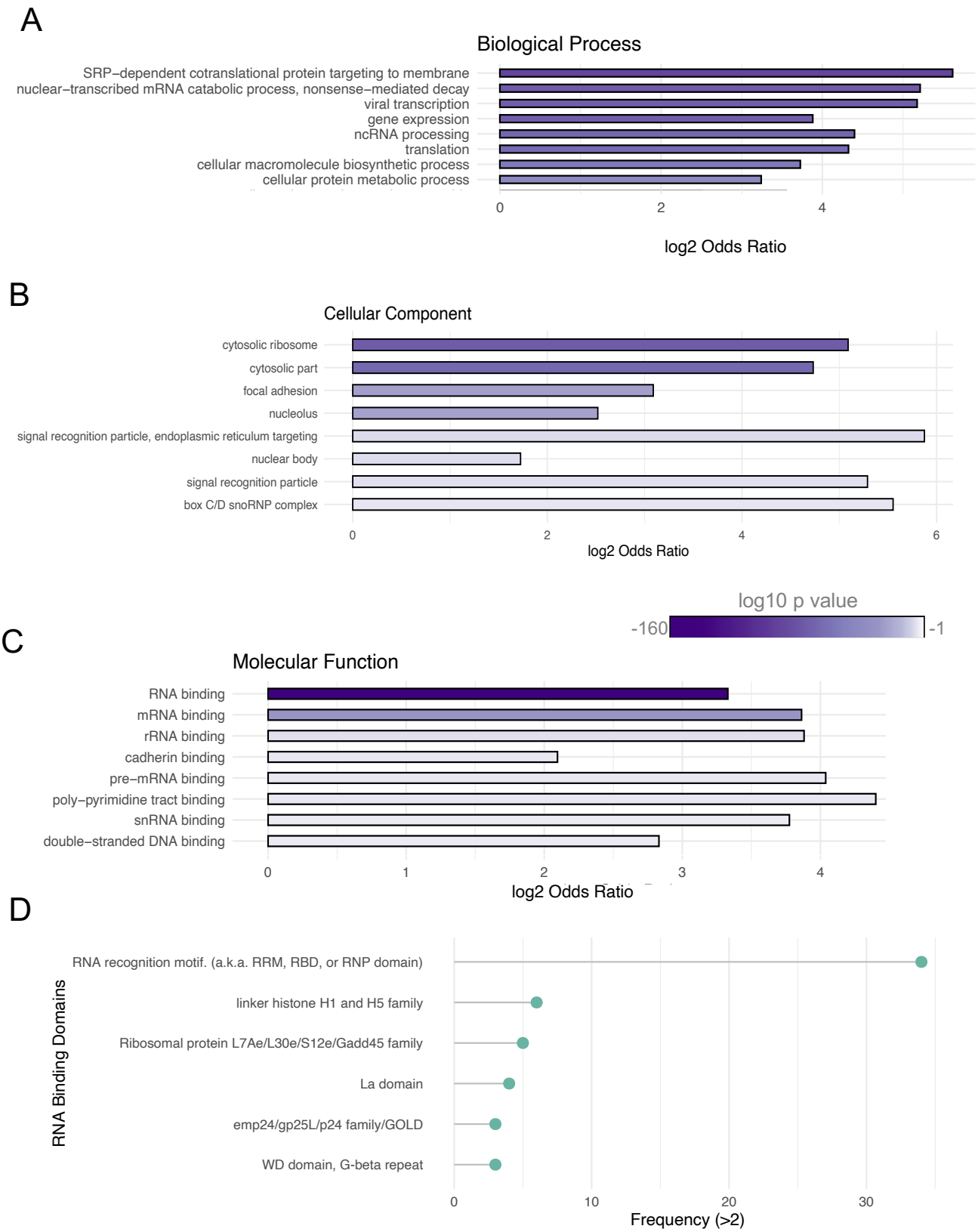


Figure 5.6: Rev interactors are mostly RBPs. (**A-C**) Bar plots showing the enriched gene ontology terms of the Rev interactome, which label groups of proteins according to their functional characteristics. Bars plot the odds ratio (log2) on the Y axis and are coloured by p value (log10), which deeper purple indicating increased significance. (**D**)

I used the R package PFAM.db (209), a database of known and predicted protein domains, to obtain protein domain information from all Rev-interacting proteins. I then plotted the 6 domains with the highest frequency as a lollipop chart. X axis plots the number of proteins with the domain indicated on the Y axis.

To ascertain the number of Rev-interacting proteins that interact with RNA, I annotated my dataset using RBPbase, an online database of experimentally identified RBPs. I labelled proteins as RBPs if they had been reported to interact with RNA in at least 3 independent studies in RBPbase. Under this stringent criteria, a remarkable 81% of Rev interactors were identified as RBPs. This striking enrichment further supports the hypothesis that Rev engages mainly with cellular RBPs, probably on HIV-1 RNAs. To gain further insights into the composition of the Rev interactome, I used PFAM to classify and predict the domain architecture of Rev interactors and plotted the domain frequency as a lollipop chart (**Figure 5.6, C**). This analysis revealed surprisingly small and disparate frequencies of enriched protein domains. Indeed, only 6 domains were common to more than 2 proteins. The RNA-recognition motif (RRM) was the most highly represented domain and mapped to 34 proteins. The RRM is a biochemically well-characterised RBD and consists of ~90 amino acids with a very well-established mode of ssRNA binding through biochemical and structural studies (276). Previous examination of the human RBPome (186) and virus-linked RBPs (277) revealed that the RRM is the most predominant RBD. This data further supports the notion that these proteins are recruited by Rev to bind HIV-1 ssRNAs. The other domains included the La domain, found in RBPs which bind the 3' untranslated region of nascent RNA polymerase III transcripts in the nucleus to assist with folding (278). RBPs wielding the La domain also use it to bind mRNAs in the cytoplasm at specific 5' oligopyrimidine motifs to modulate translation (279). Interestingly, many proteins labelled as RBPs lacked domain information. Thus, Rev interacts with many RBPs employing unknown

domains to bind RNA. I predict that these proteins may harbour IDRs, which mediate RNA-binding activity with highly flexible interfaces, despite lacking a consistent tertiary structure (280). The high proportion of IDRs in the human proteome has only begun to be elucidated in recent years, and proteins containing IDRs can be challenging to identify due to the lack of structure in these regions. As an illustrative example, RNA binding motif protein 14 (RBM14) is an RBP which is required for non-homologous end joining (281) and was found in the Rev-Flag interactome. While it is recognised by the PFAM database as containing RRM domains, recent work identified that RBM14 employs an intrinsically disordered prion-like domain to create a liquid environment with which to bind damaged DNA (281). Indeed, annotation of the Rev-Flag interactome using IUPred2 revealed that most proteins that interacted with Rev-Flag (59%) contained at least 20% sequence disorder, supporting this hypothesis (data not shown). The high percentage of RBPs in the Rev interactome suggests that Rev may unite host cell RBPs with vRNAs, to promote RNA metabolism steps or to 'clothe' the RNA and protect it from host cell immune sensors.

From these data, I can conclude that Rev interacts with many host cell RBPs that modulate the splicing, stability and translation of host RNAs and which may be co-opted by Rev to control the fate of vRNAs.

5.5. Testing the co-localisation of Rev-Halo and Rev interactors

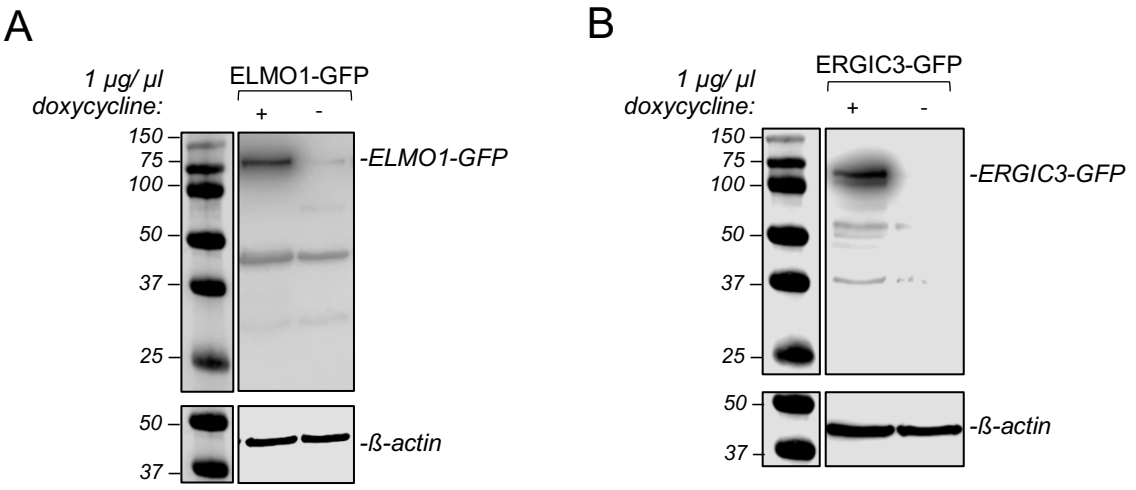
To validate some of the PPIs that I uncovered using the Rev-Flag IP, and to investigate their functional relationship, 2 proteins were selected for further analyses. These proteins were chosen based on high fold change values and low adjusted p values, indicating that they interacted with Rev strongly and specifically in all 3 IP replicates. Both proteins are novel Rev interactors with limited or no research into their roles in

HIV-1 infection. Engulfment and cell motility protein 1 (ELMO1) is the most enriched protein in the Rev interactome after Rev itself. It may be important for immunity, as it is involved in mediating the cytoskeletal rearrangements which occur during phagocytosis, in cell migration and has been posited to act as a cytoplasmic regulator of specific neutrophil receptors (282). ELMO1 has not previously been linked to Rev but was found to participate in a complex with host cell protein DOCK2 and the HIV-1 Nef protein in T cells, which promotes T cell activation (283). Endoplasmic reticulum-Golgi intermediate compartment protein 3 (ERGIC3) is a transmembrane protein that localises in the secretory pathway. While its roles in the cell are poorly characterised, ERGIC3 is thought to facilitate transport between the endoplasmic reticulum and Golgi apparatus. Its sole link to HIV-1 is an identified interaction with envelope (248, 252), which is predicted to occur in the secretory pathway.

To ratify that these proteins interact with Rev during HIV-1 infection using an orthogonal approach, I firstly generated GFP-tagged ERGIC3 and ELMO1 constructs and transfected them into Hela Flp-In T-Rex cells to establish stable and inducible RBP-GFP cell lines. To ensure that expression of ERGIC3-GFP and ELMO1-GFP could be effectively induced in these cell lines, I immunoblotted induced and uninduced cell lysates against GFP and ß-Actin as a loading control (**Figure 5.7, A-B**). Immunoblotting against GFP revealed strong bands in doxycycline induced lanes corresponding to the molecular weight of each RBP + 27 kDa of GFP. On the other hand, either no bands or very faint bands were detected in uninduced lanes, confirming that expression of ELMO-GFP and ERGIC3-GFP proteins could be induced with doxycycline and that minimal RBP-GFP protein expression occurred without doxycycline induction.

Next, to investigate the degree of colocalization between ERGIC3-GFP and ELMO1-GFP with Rev during HIV-1 replication, I transfected inducible cell lines with either HIV-$1_{Rev\text{-}Halo}$ or mock-transfected cells as a control for 48 h and induced RBP-GFP expression with 1 µg/ml doxycycline for 24 h. Cells were fixed and subjected to IF with dye-labelled Halo ligands and a GFP booster dye (**Figure 5.7, C**). In mock-transfected cells, ELMO1-GFP localised in the cytoplasm in a pattern resembling the endoplasmic reticulum. Previous IF of ELMO1 exhibited similar perinuclear cytoplasmic localisation (284-286), indicating that the fusion of GFP to ELMO1 does not perturb its native localisation. In HIV-$1_{Rev\text{-}Halo}$ transfected samples, ELMO1-GFP maintained this cytoplasmic localisation, but became concentrated at regions resembling the Golgi apparatus (**Figure 5.7, C**, red box). Indeed, ELMO1 is known to co-localise with Golgi markers in *Arabidopsis thaliana* (284-286). This shift to the Golgi apparatus during HIV-1 replication indicates that ELMO1 may be involved in ER-Golgi stress response pathways which are induced by HIV-1 infection (287). Where there was no detectable Rev-Halo signal in mock-transfected cell lines, demonstrating the specificity of dye-labelled Halo ligands, Rev-Halo colocalised strongly with ELMO1-GFP in HIV-$1_{Rev\text{-}Halo}$ transfected cells. In these cell lines, Rev-Halo is cytoplasmic and similarly concentrated at perinuclear regions resembling the Golgi apparatus. Together, these results indicate that ELMO1 overexpression may redistribute Rev from the nucleolus to the cytoplasm during HIV-1 infection, possibly sequestering Rev. This close localization supports the idea of a strong interaction between ELMO1 and Rev-Flag that I previously identified by affinity purification and suggests that ELMO1 could play an antiviral role in HIV-1 infection. In mock-transfected cells, ERGIC3-GFP exhibited cytoplasmic localisation, similarly concentrating at structures resembling the endoplasmic reticulum. This is expected, as ERGIC3 is posited to participate in the

secretory pathway and suggests that the native localisation of ERGIC3 is maintained in the GFP-tagged form. Remarkably, while the localisation of ERGIC3-GFP is maintained in cells transfected with HIV-$1_{Rev\text{-}Halo}$, Rev-Halo signal was consistently decreased in cells expressing ERGIC3-GFP. This can be seen in **Figure 5.7**, **C** which shows that cells strongly expressing ERGIC3-GFP exhibit decreased Rev-Halo signal. Conversely, in cells exhibiting no detectable ERGIC3-GFP signal, Rev-Halo signal is strong and recapitulates the nucleolar/nuclear membrane localisation previously seen. This implies ERGIC3 may suppress expression of Rev-Halo, potentially acting in an antiviral role. This may be through direct interaction with Rev, or as a consequence of a general effect on HIV-1 infection mediated by ERGIC3-GFP overexpression. Further experiments should aim to characterise the effect of these proteins on HIV-1 replication.



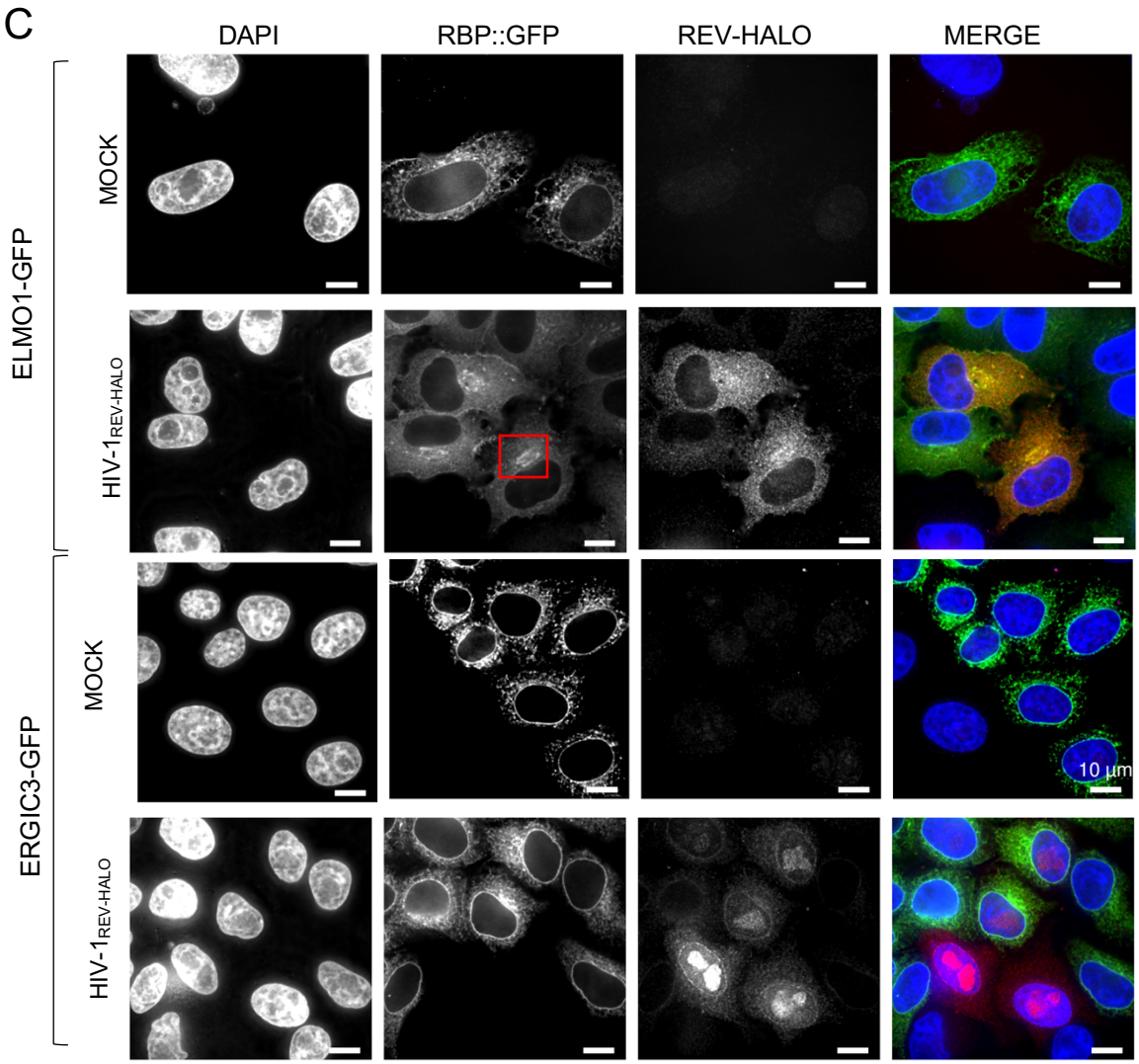


Figure 5.7: Rev-Halo colocalises with ELMO1-GFP and is downregulated by ERGIC3-GFP. (**A-B**) To test the colocalization of Rev with host cell proteins identified by affinity purification, I generated Hela cells stably expressing inducible ELMO1-GFP and ERGIC3-GFP. I tested inducible protein expression by inducing expression with doxycycline or leaving cells uninduced and immunoblotting lysates against GFP and ß-actin. Left panel is a protein molecular weight ladder. (**C**) These cells were then induced and mock transfected or transfected with HIV-$1_{Rev-Halo}$ and subjected to immunofluorescence with dye-labelled Halo ligands, DAPI and GFP booster at 48 hours post-transfection. Scale bars = 10µM.

5.6. Discussion

5.6.1. My Rev interactome addresses limitations of prior Rev IP studies

Previous research into the interactions of Rev typically utilised artificial approaches, culminating in contrasting data. To readdress this question using conditions nearer to

physiological infection than used in prior Rev research, I generated a panel of VSVG-pseudotyped HIV-1 particles expressing different tagged Rev proteins, which permitted a high rate of HIV-1 positive cells in SupT1 CD4(+) lymphocytes. These particles were generated from novel HIV-1 replicons which expressed distinct tagged Rev fusion proteins and vRNAs to physiological levels. Furthermore, I lysed transduced lymphocytes in the presence of bensonase to minimise the capture of RNA/DNA-mediated PPIs, a control that was lacking in previous Rev PPI experiments (228, 252). The approach that I used to uncover Rev PPIs culminated in a strong and specific pull-down. This strength is reflected in the efficient capture of the Rev-Flag protein, which exhibited the lowest adjusted p value and the highest fold change across IP eluates, as revealed by quantitative proteomics. The specificity of this IP is evident in the low number of proteins enriched in the negative control, non-specifically binding anti-Flag beads. Ultimately, this approach allowed me to unveil Rev PPIs in a system closely recapitulating natural HIV-1 infection, which is confirmed by the high number of proteins identified which were previously known to interact with Rev and/or to be involved in HIV-1 infection (67%).

5.6.2. The shuttling of Rev and expression of vRNA is integral to Rev interactions

My Rev-Flag interactome can help to identify true PPIs previously identified using different approaches. For example, previous Rev IP studies did not overlap, whereas the addition of my interactome defined 120 proteins identified in at least 2 studies, adding confidence to them as potential Rev interactors and possibly revealing PPIs which are conserved across different cell lines. My dataset overlaps to different extents with previous Rev IP studies, which is probably a result of specific conditions that alter Rev PPIs. For example, my dataset overlaps the most with Arizala *et al*, where almost

50% of the PPIs they report are also present in my interactome. This is compared to an overlap of 28% and 38% of the PPIs reported by Naji and Jager *et al* respectively. I hypothesise that the conservation between my interactome and Arizala *et al* is due to the compounded effect of both a) the cell systems used and b) co-expression of the vRNA. Firstly, GO term and STRING analysis of my dataset reveals a strong enrichment of proteins from both cytoplasmic and nuclear RNP complexes such as the spliceosome and the ribosome. The ability of Rev to shuttle between distinct cellular compartments permits Rev to interact with these complexes. Furthermore, Rev may redistribute host cell proteins that it interacts with by shuttling, possibly recruiting them to cytoplasmic or nuclear complexes in this way. Thus, assaying Rev in the absence of these subcellular structures may alter specific complexes that Rev interacts with. Whereas both Arizala and Jager *et al* studied Rev in full cell lines, Naji *et al* overexpressed Rev in cytoplasmic or nuclear cell extracts, removing these distinct compartments and altering Rev PPIs. This may explain the reduced overlap between mine and Naji *et al* datasets. Secondly, the interaction of Rev with the vRNA is critical to Rev localisation. In the nucleus, Rev binding to the vRNA displaces nuclear import factor B23 and masks the NLS, promoting vRNA/Rev nuclear export (270). Similarly, importin-mediated translocation of Rev into the nucleus occurs exclusively after Rev has dissociated from vRNA in the cytoplasm (63). In this way, binding to vRNA regulates the shuttling of Rev. Moreover, as discussed in chapter 4, overexpression of Rev without similar levels of its vRNA substrate may culminate in excessive, unbound Rev localising non-physiologically (for example, high levels of Rev in the nucleolus relative to the cytoplasm seen in previous IF studies) and thus permitting Rev to form PPIs that do not occur in natural infection. In addition, HIV-1 RNA may act as a nucleation point for RBP/Rev interactions. Indeed, HIV-1 gRNA contains several

cis-acting sequences that recruit viral and cellular RBPs (288-290) and analysis of Rev-Flag PPIs shows a strong enrichment in RBPs and proteins linked to RNA modification. Thus, Rev may “meet” interaction partners along the vRNA. Consequently, the absence of vRNA would limit the formation of these interactions. Indeed, all 3 former Rev IP studies overexpress Rev. However, while Arizala *et al* utilise a cell line stably expressing HIV-1 RNAs, both Naji and Jager *et al* isolated tagged Rev in uninfected cells lacking vRNA, with Naji *et al* overexpressing only a purified, synthetic RRE RNA. This may explain the stronger commonality between my dataset and that of Arizala *et al*. These comparisons suggest that both the shuttling of Rev between subcellular compartments and the presence of vRNAs is integral to Rev PPIs in physiological HIV-1 infection.

5.6.2. The Rev interactome: implications for Rev function

My dataset has revealed over 80 proteins which had not previously been linked to HIV-1, in addition to over 120 novel Rev-interacting proteins. These PPIs can contribute significantly to our knowledge of Rev functions, which I discuss below.

Transcription

Despite no known role for Rev in transcription, GO term analysis of my dataset reveals a strong enrichment of the term ‘*viral transcription.*’ A major component in the transcription of vRNA is positive transcription elongation factor b (P-TEFb), a heterodimeric kinase consisting of CDK9 and Cyclin T1 that releases RNA polymerase II into active elongation (291). P-TEFb is usually kept sequestered in an inactive state by the 7SK snRNP. Displacement of P-TEFb and recruitment to the HIV-1 LTR is triggered by viral Tat protein, which binds the 5' TAR element (292) and recruits P-

TEFb to RNA polymerase II, which it phosphorylates, stimulating transcription (292). I identify that Tat, Cyclin T1 (CCNT1) and important members of the 7SK snRNP (LARP7, MECPE) interacted with Rev-Flag, in addition to RNA polymerase II subunit B (POLR2B). Interestingly, knockdown of either MEPCE or LARP7 leads to destabilization of 7SK *in vivo,* releasing P-TEFb (293). The interaction of these proteins with Rev suggests that Rev may help displace P-TEFb from the 7SK snRNP and recruit it to vRNA, illuminating a novel function for Rev in infection.

Splicing

HIV-1 infection results in an increased proportion of introns within cellular RNAs in primary T cells (133). Rev may underpin this effect, as it has been shown to block recruitment of the U4/U5/U6 small nuclear (sn)RNP complex to immature spliceosome complexes on pre-mRNAs, preventing complete spliceosome assembly (126). While the mechanism behind Rev-mediated blockage is unknown, some of the proteins comprising the U4/U5/U6 snRNP complex interacted with Rev-Flag (NHP2L1, PRPF8, PRPF19, EFTUD2, SNRPA1) (294). These proteins may be targeted by Rev to prevent their participation in the snRNP complex, preventing full spliceosome assembly. Interestingly, I also identified an interaction between Rev-Flag and PHAX, an export factor which complexes with CRM1 and RanGTP to export U snRNA from the nucleus (295). PHAX promotes U snRNA export but is not required for general CRM1-mediated export (295). Thus, Rev may bind PHAX to prevent redistribution of snRNAs to the cytoplasm, similarly blocking spliceosome assembly. The co-opting of CRM1 by Rev for RNA export, in addition to the enriched GO term '*snRNA binding',* further supports this idea. Another explanation is that PHAX may play a role in the

Rev-mediated export cycle, joining the CRM1/RanGTP/Rev/vRNA complex to promote nuclear export.

In addition to preventing assembly of the spliceosome, Rev also interacts with host cell factors that enhance or downregulate splicing. Principally, I identified Rev-Flag interacting with many SR proteins and HNRNPs (**Table 5.2**). These proteins bind to exonic splicing enhancer/silencer (ESE, ESS) and intronic splicing silencer (ISS) motifs in RNA transcripts to modify the efficiency of nearby splice sites. The ability of Rev to mediate vRNA export is impaired by altering splice site efficiencies in the HIV-1 genome (296-298). Therefore, Rev probably interacts with these proteins to ensure an optimal balance of splice site usage to enable efficient vRNA production and export. Several HNRNPs bind with varying specificity along the HIV-1 genome to regulate splicing (299). For example, crosslinking and IP (CLIP) experiments revealed that HNRNPH1 binds to a small number of specific sites on HIV-1 RNA, while HNRNPA1 and HNRNPA2B1 bind frequently throughout the entire genome, regulating splicing globally (273). This regulation is important for efficient HIV-1 infection. Knocking down HNRNPA1, HNRNPA2B1, HNRNPD and HNRNPH1 led to an increased use of acceptor sites involved in the removal of the major 5' intron in HIV-1 gRNA, impairing infection (249, 273). I hypothesise that Rev recruits HNRNPs to downregulate splicing, ultimately creating a bigger pool of intron-containing vRNAs to export. HNRNPs also contribute to RNA trafficking, stability, export and translation (124, 300, 301), so it is challenging to pin down the effects of these interactions without further experiments.

Table 5.2. HNRNP and SR proteins identified in the Rev-Flag interactome. Proteins coloured in blue are novel Rev interactors.

HNRNP PROTEINS			SR PROTEINS		
hnRNPA0	hnRNPA1	hnRNPA2B1	SRRM1	SRSF10	SRSF2
hnRNPA3	hnRNPAB	hnRNPC	SRSF3	SRSF5	SRSF6

hnRNPD	hnRNPDL	hnRNPF		PPP1CA	STRAP
hnRNPH3	hnRNPL	hnRNPM			
hnRNPR	hnRNPU	hnRNPUL1			
hnRNPQ	hnRNPUL2				

SR proteins bind to ESE sites to promote host cell splicing and are also vital regulators of HIV-1 RNA splicing. The overexpression of SRSF proteins 1, 2, 5, 6 or 7 led to a substantial impairment in HIV-1 virus production and infectivity by decreasing packaging of envelope and the abundance of gRNA in virions (19, 130-132, 302). I hypothesise that Rev both recruits and sequesters these modulators to regulate HIV-1 RNA splicing in a temporal manner. For example, a splice site important for expression of envelope, Tat and Rev proteins is enhanced by the binding of SRSF1 (ASF/SF2) and SRSF5 (SRp40) to a nearby ESE motif in vRNA (17). These SR proteins bind the ESE and recruit the U1 snRNP to initialise splicing (17). My work shows that SRSF5 interacted strongly with Rev-Flag. Thus, in early-phase infection, Rev may recruit SRSF5 to this ESE site to induce splicing, allowing Rev to regulate the production of its own RNA. Marchand *et al* investigated the proteome of a region of HIV-1 gRNA containing splice site A7, which is important for production of Rev and Tat pre-mRNAs. Of 42 proteins identified to bind this vRNA, I find 15 in the Rev interactome (36%), confirming that Rev interacts with RBPs that may affect splicing and stability of HIV-1 gRNA. As a threshold level of Rev protein is needed to switch from early to late phase infection, to allow Rev-mediated vRNA export (303), this self-regulation may increase concentration of Rev in the cell, triggering the switch to late-phase infection. During late-phase infection, Rev may then sequester SR proteins to prevent splicing and promote the synthesis of intron-containing vRNAs.

Stabilisation

Rev has been posited to stabilise RNA. Evidence for this largely comes from the RNA export function of Rev, which overcomes the nuclear retention of vRNAs triggered by their INS sites. Rev-dependent mRNAs harbour these INS sites and are AU-rich, containing the AU-rich element (ARE) AUUUA pentanucleotide (117). These vRNAs are therefore vulnerable to host cell pathways which degrade AU-rich transcripts. I hypothesise that Rev stabilises RNAs by sequestering decay pathway proteins. For example, ARE-medicated decay (AMD) is a pathway which recognises and degrades transcripts with AU-rich 3' LTRs. I identified both INS-binding proteins (PABPC1 (122), HNRNPA1 (304)) and key mediators of the AMD (AUF1/HNRNPD, PABPC1 (305), EXOSC8 (306), CPSF1) in the Rev interactome, supporting this hypothesis. NMD is similar quality control pathway which degrades mRNAs harbouring premature stop codons (307). These transcripts are recognised by helicase UPF1 and cofactors UPF2 and UPF3. Recent work has shown that UPF1 participates in the Rev/CRM1 export complex (100). The co-opting of UPF1 by Rev may serve to prevent its roles in NMD, preventing degradation of intron-containing vRNAs. Interestingly, I also identified novel interactions between Rev-Flag and 3 of the 10 core proteins of the exosome (EXOSC2, EXOSC4, EXOSC8). The exosome is a multi-subunit exonuclease complex responsible for degrading RNAs (308) from the 3' end and contributes to several decay pathways, including NMD and AMD (the exosome directly recruits AU-rich RNAs for AMD) (308). It is also recruited by antiviral protein ZC3HAV1 to vRNAs, which I also found interacted with Rev-Flag (309). While EXOSC2 inhibits HIV-1 replication in lymphocytes (310), neither EXOSC4 or EXOSC8 have otherwise been linked to HIV-1 or Rev. Thus, my interactome reveals novel interactions which may identify new mechanisms for how Rev stabilises vRNAs.

Translation

Evidence that Rev promotes translation of vRNAs remains controversial. However, GO term and STRING analysis of Rev-Flag PPIs showed that a large cluster of translation-related proteins interacted with Rev-Flag. I find almost 40 proteins comprising the 60S ribosomal subunit and 19 comprising the 40S subunit in the Rev interactome. This suggests that Rev may directly recruit the ribosome to vRNAs for efficient translation. Indeed, I identified Rev-Flag interacting with crucial initiation factors responsible for recruiting the ribosome to RNA in a cap-dependent fashion, including critical cap-binding protein eIF4E, eIF1, eIF6 and PABPC1 and PABPC4 (311). Rev-Flag also interacted with RNA helicases SSB, RNA Helicase A/DDX9, DDX1 and DDX24, which each unwind the highly-structured 5' end of HIV-1 RNA to permit ribosome binding (258, 272, 312).

Rev PPIs may bias the type of translation that occurs for specific vRNAs. Late-phase viral translation is important for the synthesis of Gag (313) and is mainly cap-independent, being initiated through an internal ribosome entry site (IRES). IRESs are common in retroviruses (314) and form a tertiary fold which acts as a scaffold for RBPs, which bind and recruit translation factors (315). Although the mechanism by which the HIV-1 IRES functions is unknown, RBPs that are required for this process interacted with Rev-Flag, including RPS25 (316), hnRNPA1 (317), SSB (312) and Staufen1 (318). This suggests that Rev may recruit these RBPs to vRNA to facilitate IRES-mediated translation. Interestingly, HIV-1 PR cleaves eIF4GI and PABP to prevent cap- and poly(A)-dependent translation, shifting ribosomes to sites of IRES-dependent translation (319-321). Rev-Flag interacted with PABP and eIF4E and may therefore enact this role in HIV-1 infection to increase a bias toward IRES-mediated

translation, or may bring these factors to the RNA to promote cap-dependent translation.

I identified that all of the core SNORD proteins interacted with Rev-Flag (NHP2L1, NOP56, NOP58 and FBL). The SNORD complex is an RNP of host cell proteins and small nucleolar (sno)RNAs containing C and/or D box motifs (RUGAUGA and CUGA respectively). These snoRNAs base-pair with rRNAs and act as a scaffold which allows host cell protein FBL to catalyse 2′-O-methylation of target rRNAs. This methylation modulates the ability of ribosomes to translate mRNAs (322). SNORD complexes have been shown to be important for infection of multiple viruses through siRNA knockdown experiments (323). The interaction of Rev with SNORD proteins suggests that Rev participates in this complex to increase the translation competency of ribosomes at early stages in infection.

5.6.3. Future work

Uncovering the Rev interactome has revealed Rev PPIs which imply several novel functions for Rev during infection, in addition to corroborating prior research ascertained in non-physiological conditions. Furthermore, it has exposed a strong relationship between Rev and host cell RBPs and revealed many novel Rev interactors. Some of these interactions may be crucial to Rev function in HIV-1 infection, potentially representing antiviral or proviral proteins with potential as therapeutic targets. In this way, I have generated a valuable resource which future research can build on to uncover the functions of Rev in HIV-1 infection. While my dataset has revealed many interesting Rev PPIs, further experiments should be performed to confirm Rev PPIs using orthogonal methods. To do this, I generated inducible GFP-tagged interactor cell lines, as performed for ERGIC3 and ELMO1. I

induced expression of fusion proteins in these cells, transfected them with HIV-$1_{Rev\text{-}Halo}$ and performed IF to directly visualise the degree of colocalisation between Rev-Halo and host cell proteins during HIV-1 infection. Further work should explore whether these interactions have pro or antiviral effects. One way to investigate this is to transduce induced and uninduced ERGIC3-GFP and ELMO1-GFP cell lines with HIV-$1_{mCherry\text{-}Nef}$ and monitor the kinetics of HIV-1 replication over time. This has previously been carried out in the Castello lab using a plate reader which maintains cells in an environment supplied with 5% CO_2 at 37ºC and can monitor mCherry fluorescence over 72 hours, as a proxy of HIV-1 replication. This experiment would indicate whether ERGIC3 and ELMO1 impair or promote HIV-1 replication. While I began this work, time restraints prevented me from acquiring all necessary replicates. I could also perform inverse IPs, capturing GFP-tagged Rev interactors using GFP-trap beads and immunoblotting against Flag for Rev-Flag-3xMyc. To instead validate the Rev interactome in a high-throughput manner, I generated a HIV-1 replicon encoding Rev tagged with APEX2 (HIV-$1_{Rev\text{-}APEX2}$) in an analogous way to the other replicons. APEX2 is ascorbate peroxidase which biotinylates nearby proteins (those within a few nm) upon catalysis with hydrogen peroxide. Biotinylated proteins can be then purified using streptavidin beads and analysed using mass spectrometry, essentially corroborating the interactome with similar quantitative power through a separate experimental approach. While I cloned and validated this replicon as previously, time constraints restricted me from fully optimising the proximity-ligation assay. Future work on this project could continue this process, validating many Rev PPIs in a high-throughput manner.

6. The Rev interactome overlaps heavily with the interactome of encapsidated HIV-1 RNA

6.1. Introduction

Rev is not known to be encapsulated in HIV-1 virions and localises primarily in the nucleolus (64, 91, 226-230). However, as discussed in chapter 1, there is sparse evidence that Rev promotes packaging into virions by an unknown mechanism (147, 148, 324). To determine whether Rev influences virion packaging, the Rev-Flag interactome dataset can be a powerful resource to learn about Rev's physiological functions. For example, if Rev interacts with proteins that are ultimately incorporated into the viral particles. However, overlap between proteome studies of HIV-1 particles is poor (23, 25, 231, 232, 325) (**Figure 6.1, A**) due to extensive methodological differences in virus isolation, producer cell lines, proteomic depth, data analysis and essential controls. An important caveat of these experiments is that a cytoplasmic portion is captured during viral particle assembly, which can cause 'passive protein travellers' to be carried in virions. A second limitation is that HIV-1 particles are similar in density to microvesicles and exosomes (130-150 nm (326-329)), and these often contaminate viral particle fractions. Purification methods to remove these contaminants have been used across studies, such as sucrose density centrifugation and CD45 depletion, but they lacked a standardised control with which to compare the proteomic results (330, 331). The use of different cell lines between studies also alters the isolated HIV-1 particle proteome, as virion assembly differs between cell type (23). Ultimately, this poor overlap suggests that these datasets do not accurately capture the physiological HIV-1 particle proteome and cannot be used for a valid comparison with the Rev-Flag interactome.

To address this question, Manuel Garcia-Moreno developed 'in virion RNA-interactome capture' (ivRIC) in the Castello lab (**Figure 6.1, B**). Instead of capturing the proteins 'loaded' into virions, ivRIC identifies the proteins bound directly to HIV-1

gRNA inside particles. While 'binding' does not always mean 'function', it does reflect protein activity, unlike virion proteome analysis. In brief, SupT1 lymphocytes were infected with fully infectious (multiple round) HIV-$1_{mCherry\text{-}Nef}$ particles, or mock-infected as a control. Particles in the supernatant were purified by ultracentrifugation on a 10% sucrose cushion and irradiated with 150 mJ/cm^2 UV light at 254 nm, or left unirradiated to test UV dependency. UV induces free radicals at the nucleotide bases that can form 'zero distance' RNA-to-protein crosslinks. As HIV-1 gRNA is polyadenylated, crosslinked RNA-protein complexes were captured with oligo(dT) beads under denaturing conditions. Importantly, RT-qPCR using HIV-specific primers revealed that ~95% of captured RNA was viral (data not shown), strongly diminishing the presence of proteins via contaminant RNA. Finally, eluates were digested with RNase to liberate bound proteins, which were identified by label-free mass spectrometry. Moderated t-tests were performed on two sets of conditions. Firstly, **eluates** from **HIV-1** infected and **crosslinked** particles (***E.HIV.CL***) were compared to those from **crosslinked, mock-infected** particles (***E.MOCK.CL***), identifying factors enriched inside HIV-1 virions relative to mock supernatant. Secondly, E.HIV.CL samples were compared to **eluates** from **infected** and **non-crosslinked** particles (***E.HIV.NCL***), enabling me to distinguish between proteins specifically bound to HIV-1 gRNA and those surviving the purification procedure in UV-independent fashion. Quantitative proteomics revealed 103 host cell proteins significantly enriched in both comparisons, using a threshold adjusted p value < 0.1 (**Figure 6.1, C-D)**. Thus, ivRIC allowed the Castello lab to reveal proteins directly bound to HIV-1 gRNA inside viral particles generated from CD4+ T cells, with very low risk of contamination. This dataset is a highly useful resource, which can be compared to the Rev-Flag interactome to investigate the question "does Rev influence packaging of proteins into HIV-1 particles?"

In this chapter, I will investigate this question using data mining and meta-analysis of my Rev-Flag interactome, the ivRIC dataset and previously published data. I will also experimentally investigate proteins which Rev interacts with to package, and which may serve as critical controllers of HIV-1 infection. Ultimately, this chapter will shed light on the interactions of Rev, its role as a regulator of virion packaging and the proviral functions of packaged host cell proteins.

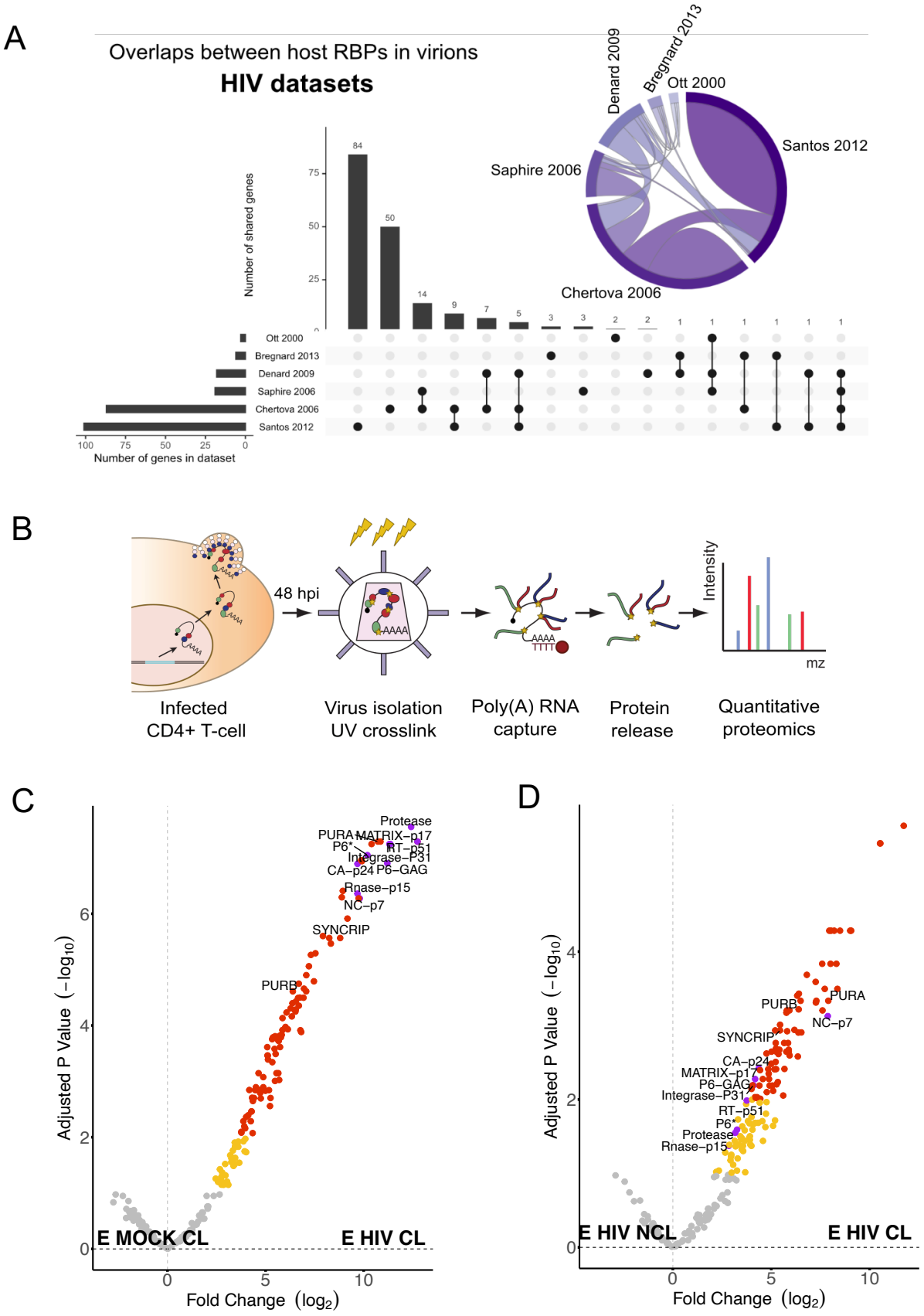


Figure 6.1: Uncovering host cell proteins encapsidated in HIV-1 virions. (**A**) Previous studies that have attempted to uncover the HIV-1 particle proteome utilised a variety of methods with differing controls, cell types and methods and overlap poorly.

This upset plot charts the intersection between the indicated datasets (vertical black bars plot number of common proteins, horizontal black bars plot total number of proteins in the indicated study and lollipop bars indicate which studies are included in each intersection. Inset chord diagram plots this information, which each purple thread indicating 1 protein found in the studies it connects to. (**B**) The Castello lab developed ivRIC, a method to identify host cell proteins bound directly to HIV-1 gRNA inside virions to address this question. This schematic illustrates the main steps of the protocol. (**C-D**) Volcano plots of ivRIC proteomics (n = 3). X axis plots fold-change of protein intensity in (C) HIV-1 infected and crosslinked eluates versus mock-infected crosslinked eluates and (D) HIV-1 infected and crosslinked eluates versus HIV-1 infected and non-crosslinked eluates (log2). Y axis plots the adjusted p value across 3 replicates. Individual dots indicate proteins. There were 103 proteins significantly enriched in both cross-linked, infected particle eluates versus mock-infected or non-crosslinked particle eluates. Proteins with a log2 adjusted p value < 0.01 are marked in red, and those between 0.1 and 0.01 are marked in in yellow. Grey dots are non-significant.

6.2. A substantial proportion of Rev interactors are encapsidated into HIV-1 virions

To investigate the role of Rev in encapsidation, I compared the Rev interactome to the in-virion HIV-1 RNA interactome (ivRBPome) revealed by ivRIC. This comparison exposed a strong overlap: I identified 71 proteins common to both datasets, meaning that almost 70% of the ivRBPome also interacted significantly with Rev-Flag in my IP study (**Figure 6.2, A**). This surprising discovery suggested that Rev may interact with host cell proteins that become packaged into HIV-1 virions. Therefore, it is highly likely that Rev plays a role in specifying the fate of the gRNA that is encapsidated. To gauge the correlation between these protein networks, I plotted the fold change of proteins in Rev-Flag IP eluates (*i.e.,* HIV-$1_{\text{Rev-Flag}}$/HIV-$1_{\text{mCherry-Nef}}$) against the fold change of proteins in each ivRIC eluate condition (E.HIV.CL/E.MOCK.CL and E.HIV.CL/E.HIV.NCL) and calculated respective correlation coefficients (**Figure 6.2, B**). This analysis yielded coefficient values of 0.69 and 0.61 when comparing against E.HIV.CL/E.MOCK.CL and E.HIV.CL/E.HIV.NCL eluates respectively, indicating

strong correlation between the Rev-Flag interactome and ivRBPome. Thus, proteins which prominently interacted with Rev-Flag did so with the encapsidated HIV-1 gRNA as well.

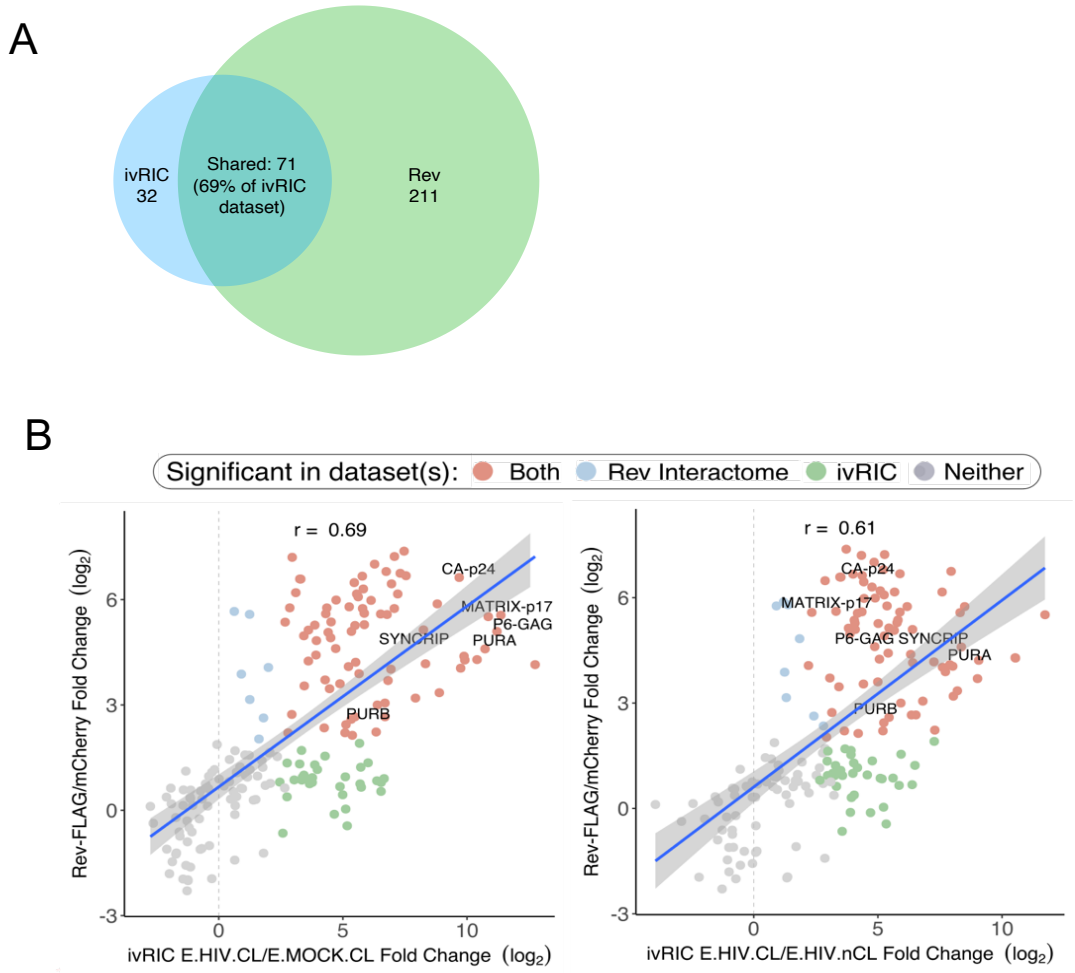


Figure 6.2: Comparison of the HIV-1 Rev interactome and ivRBPome. (**A**) Euler plot (blue circle = significant ivRIC proteins, green circle = significant Rev-Flag interactors) showing intersection between the ivRBPome and Rev-Flag interactome. 69% of proteins identified to significantly interact with HIV-1 gRNA in the viral capsid core were also identified by my Rev interactome. (**B**) Correlation plots of protein intensity fold change in HIV-$1_{Rev\text{-}Flag/mCherry\text{-}Nef}$ against E.HIV.CL/E.HIV.NCL and E.HIV.CL/E.MOCK.CL eluates were generated and revealed strong correlation coefficients (r = 0.69 and 0.61 respectively). Individual dots represent proteins coloured by significance. Blue/grey line indicates correlation line of best fit.

To gauge a deeper understanding of the functions of these proteins, I mined the GO terms of the 71 shared Rev-Flag interactors/ivRBPome proteins using the gene list

enrichment analysis tool EnrichR (206) as before (**Figure 6.3**). Interestingly, the 'cellular component' GO terms '*nucleolar*', '*nuclear matrix'* and '*nuclear periphery*' were enriched. This suggests that proteins to be packaged are defined in the nucleus of the infected cell. Perhaps Rev couriers these nuclear/nucleolar proteins to the cytoplasm during its RNA nuclear export cycle and its interactions with the RNA metabolic machinery. The term '*cytosolic part'* was strongly enriched and refers to areas of the cytoplasm which contain particulate matter, such as protein complexes. This likely reflects the presence of proteins that participate in cytoplasmic RNPs, such as ribosomal proteins, and suggests that encapsidated proteins may reach the distal plasma membrane through participation in cytoplasmic viral complexes. Finally, 'molecular function' GO terms revealed a striking enrichment in the term '*RNA binding*'. This validates the ability of the ivRIC method to recover bona fide RBPs. Moreover, there was an enrichment in the more specific term '*poly-pyrimidine tract binding.'* Indeed, the 71 shared proteins exhibit an abundance of proteins which bind to poly(A) tail of mRNAs and to AU-rich regions.

The strong correlation between the Rev-Flag interactome and the ivRBPome suggests that Rev somehow interacts with host cell proteins to fate their packaging into virions. How Rev can promote encapsidation at the plasma membrane while localising primarily in nuclear and perinuclear regions is unclear. I hypothesise that Rev binds host cell RBPs in the nucleus and recruits them to unspliced vRNAs during its nuclear export cycle. This would result in the assembly of an early vRNP containing host cell RBPs bound to vRNA. The prevalence of RBPs in both datasets, in addition to the enriched term '*polypyrimidine binding'* supports this, suggesting that many Rev interactors and encapsidated proteins may bind the AU-rich HIV-1 genome. I hypothesise that this vRNP is exported to the cytoplasm by the Rev/CRM1 export

pathway and becomes incorporated into cytoplasmic viral complexes. Indeed, relocalisation of host RBPs to viral complexes in the cytoplasm has been demonstrated during infection cycle of other viruses (191), and several cytoplasmic complexes form during HIV-1 infection which act as viral assembly sites (discussed below). Finally, these complexes are likely targeted to the plasma membrane by (a) signals in the HIV-1 RNA such as the 5' packaging *psi* signal and (b) by host cell proteins which mediate packaging. This model envisions Rev acting as a molecular spiderweb, bringing together essential host cell proteins and viral components to optimise the proteome of progeny virions.

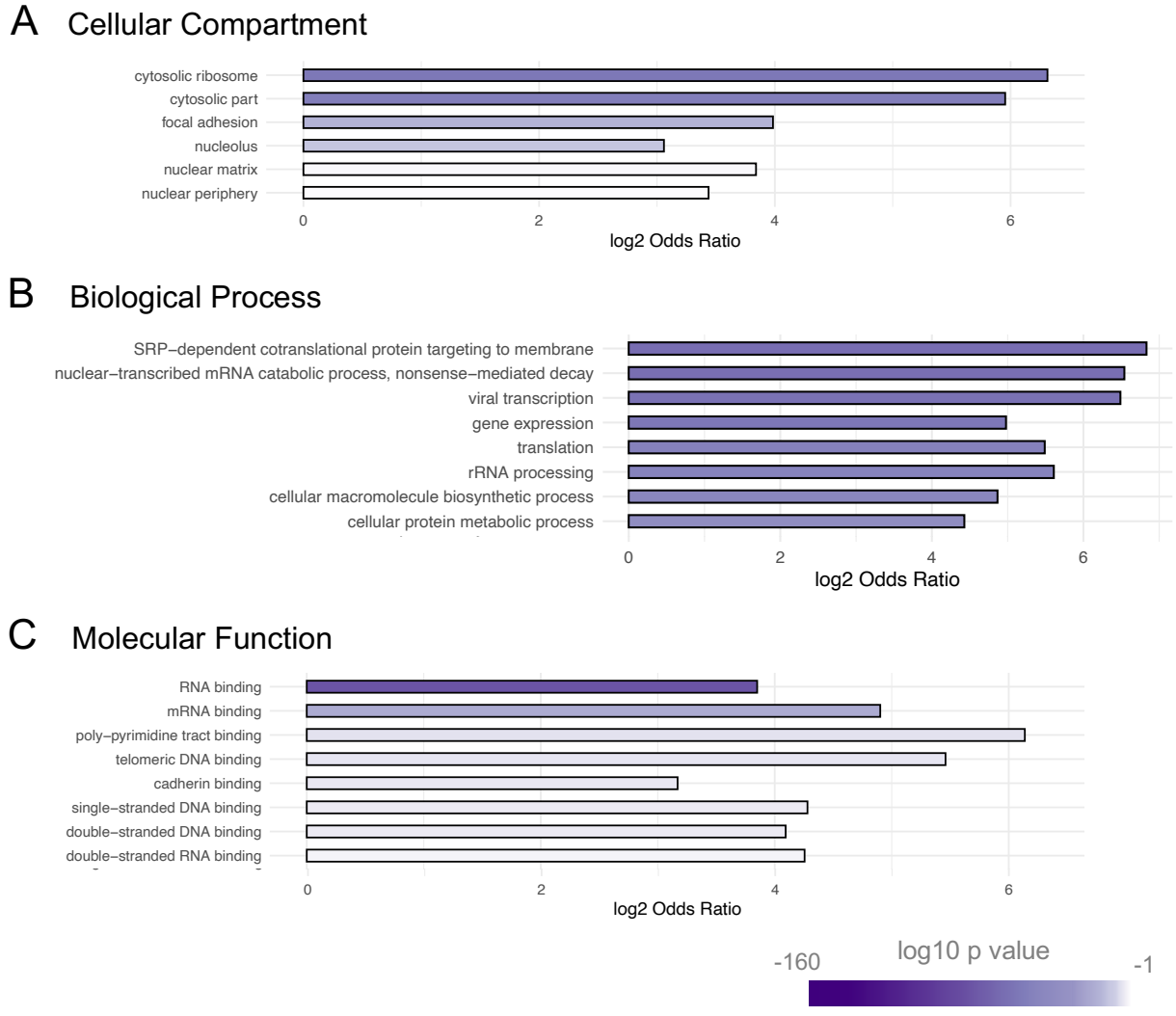


Figure 6.3: Proteins common to the Rev interactome and ivRBPome are RBPs that modify RNA and emerge from the nucleus. Bar plots showing the enriched gene ontology terms of proteins common to both the Rev interactome and ivRBPome.

These terms label groups of proteins according to their functional characteristics. Bars plot the odds ratio (log2) on the Y axis and are coloured by p value (log10), which deeper purple indicating increased significance. Bars are coloured by significance (**A-C**).

6.3. Rev delivers host cell proteins to viral complexes in the cytoplasm

Next, I investigated whether Rev promotes encapsidation by delivering host cell protein/vRNA RNPs to viral complexes in the cytoplasm. To achieve this, I compared Rev-Flag and ivRBPome datasets to the proteome of a Staufen1 particle and the Gag interactome. Staufen1 is an dsRBP that regulates RNA metabolism and is involved in RNA transport in neurons (332, 333). During HIV-1 infection, Staufen1 binds Gag and HIV-1 gRNA to form distinct cytosolic RNPs and drives their incorporation into assembling particles (334, 335). Gag participates in Staufen1 complexes and mediates the downstream steps necessary for viral packaging and assembly at the plasma membrane (46, 336). Thus, Staufen1 and/or Gag drive the packaging of HIV-1 proteins and RNAs into virions. Host cell proteins that are encapsidated will likely engage with these complexes. For Staufen1 analyses, I consulted a dataset generated by Milev *et al,* who isolated Staufen1 complexes in either mock infected or pNL4-3 HIV-1 infected Jurkat cells at 40 hpi, using tandem affinity purification (316). For Gag analyses, I utilised a proteome published by Engeland *et al,* which identified Gag interactors through 6 AP screens using 3 different IP methods, probing 293T cells transfected with a plasmid expressing Gag-GFP (226).

Remarkably, this comparison revealed almost 40 proteins found in all 4 datasets (**Figure 6.4, Table 6.1**). This indicates that many encapsidated proteins both interact with Rev and are part of cytoplasmic HIV-1 RNP complexes prior to packaging. Indeed, there were 75 proteins common to at least 3 of the 4 datasets, revealing a large pool of host cell proteins which follow a common interaction pathway during HIV-

1 infection. In particular, 60 of the 71 shared ivRBPome/Rev-Flag interactome proteins were also isolated in the Gag proteome. Interestingly, there was a large cohort of 53 proteins which interacted with Rev-Flag and Gag only. These proteins may be necessary for steps of the viral replication cycle prior to viral particle assembly and may leave the complex before arriving to the plasma membrane.

This striking commonality supports the idea that Rev assembles an early vRNP in the nuclei of infected cells which is absorbed into cytoplasmic HIV-1 complexes and trafficked to the assembling virion. Finally, these analyses unveil a common pathway that host cell proteins engage with to reach the assembling virion. In this pathway, HIV-1 proteins (Rev, Gag) may create a network of interactions with host cell proteins to shuttle specific RBPs from the nucleus to the assembling virion, beginning with Rev-mediated nuclear export and continuing with Staufen1-Gag-gRNA complexes in the cytoplasm.

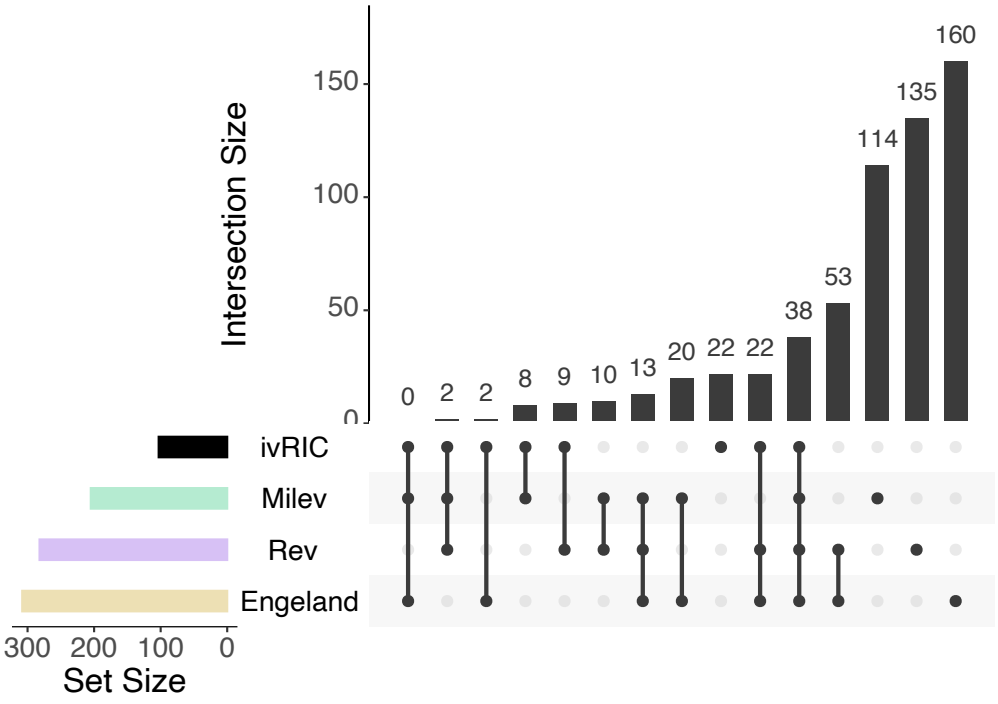


Figure 6.4: There is a high degree of conservation between proteins of the Rev, Staufen1, Gag and ivRIC interactomes. An upset plot comparing the overlap between proteins found in the Staufen1 particle proteome, the ivRBPome, the Rev-Flag interactome and the Gag proteome. Vertical black bars plot the number of proteins in the intersection indicated directly below, visualised as vertical black lollipop

bars joining different studies. Horizontal black bars indicate total number of proteins in each study.

Table 6.1. The 38 proteins shared between Gag, ivRIC, Staufen1 and Rev interactomes.

Miscellaneous	HNRNP	60S ribosomal		40S ribosomal	Helicases
IGF2BP3	HNRNPA1	RPL10	RPL5	RPS2	DHX9
ILF3	HNRNPA2B1	RPL13	RPL6	RPS4X	MOV10
NCL	HNRNPA3	RPL14	RPL7	RPS6	UPF1
PABPC1	HNRNPAB	RPL18	RPL7A	RPS8	
PABPC4	HNRNPD	RPL19	RPL8	RPS9	
RBMX	HNRNPM	RPL24			
YBX1	HNRNPR	RPL29			
	HNRNPU	RPL3			
	HNRNPQ/SYNCRIP	RPL4			

6.4. Revealing the interactions of proteins encapsidated in HIV-1 virions

6.4.1. Introduction

Recent discoveries indicate that the HIV-1 CA core disassembles inside the nucleus of the infected cell (29, 32, 36, 37). Thus, viral and cellular proteins required for early infection steps, such as reverse transcription and integration, are likely pre-packaged with the HIV-1 gRNA in the producer cell.

To examine the possibility that encapsidated proteins mediate early-phase HIV-1 infection, purine-rich ssDNA-binding proteins alpha and beta (PURA and PURB) were selected for further analysis. Both proteins interacted strongly with Rev-Flag, were identified in the ivRBPome and were found in the Gag interactome by Engeland *et al*. PURA binds to purine-rich sequences on several gene promoter regions to regulate gene expression (337-339) and interacts with cell cycle factors cyclin T1/CDK9 and CDK2 to control the cell cycle (340). Conversely, the physiological roles of PURB are poorly defined. PURA may enact proviral roles in HIV-1 infection, as it

colocalises with Rev (341) and inhibits viral replication when depleted with siRNAs (342). However, the exact role of PURA in HIV-1 infection remains unclear. PURB has no known roles in HIV-1 infection. Experiments in the Castello lab demonstrated that HIV-1 particles produced from PURA or PURB knock-out SupT1 cells exhibit low infectivity (data not shown). Thus, depletion of PURA and PURB in HIV-1 particles strongly impairs infection, signifying that these proteins are required for the initial steps of HIV-1 infection. Based on this data, other host cell proteins targeted for packaging through interaction with Rev may be packaged due to proviral roles in HIV-1 infection.

6.4.2. Revealing the interactomes of encapsidated proteins during HIV-1 replication

To elucidate the roles of PURA and PURB in HIV-1 infection, I investigated their interactions during HIV-1 replication. To achieve this, I used HEK293T Flp-In T-Rex cell lines stably expressing inducible PURA-GFP, PURB-GFP or free GFP protein, generated by Manuel Garcia-Moreno. Briefly, expression of these proteins was induced with 1 µg/ml doxycycline for 24 h and cells were either transduced with pseudotyped HIV-$1_{mCherry\text{-}Nef}$ viral particles or mock-transduced as a control. At 48 hpt, cells were lysed in the presence of bensonase to degrade nucleic acids, incubated with GFP trap beads for 2 h and washed to remove contaminants. Finally, interacting complexes were eluted using a low pH solution. Two rounds of elution were pooled to increase the volume of eluted proteins. To assess the efficiency of each IP, I performed silver staining and Western blotting (**Figure 6.5, A**). Silver staining showed expected size-matched bands at ~75 kDa, 74 kDa and 27 kDa in PURA-GFP, PURB-GFP and free GFP eluate lanes respectively, indicating successful isolation of target proteins. Densities of these bands were similar in both HIV-$1_{mCherry\text{-}Nef}$ and mock-transduced conditions for all eluates, suggesting that expression levels of PURA and PURB are

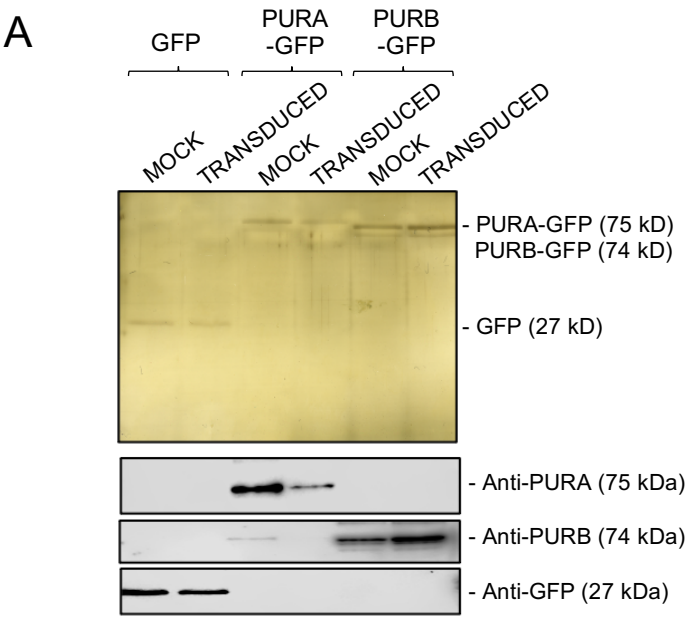
not significantly altered by HIV-1 infection. Western blotting of eluates using antibodies against GFP, PURA and PURB revealed strong and specific bands in each eluate lane, matching band sizes observed in the silver staining and further confirming successful elution of target proteins. These data confirmed that PURA-GFP and PURB-GFP could be captured using GFP trap beads and productively eluted under my IP conditions.

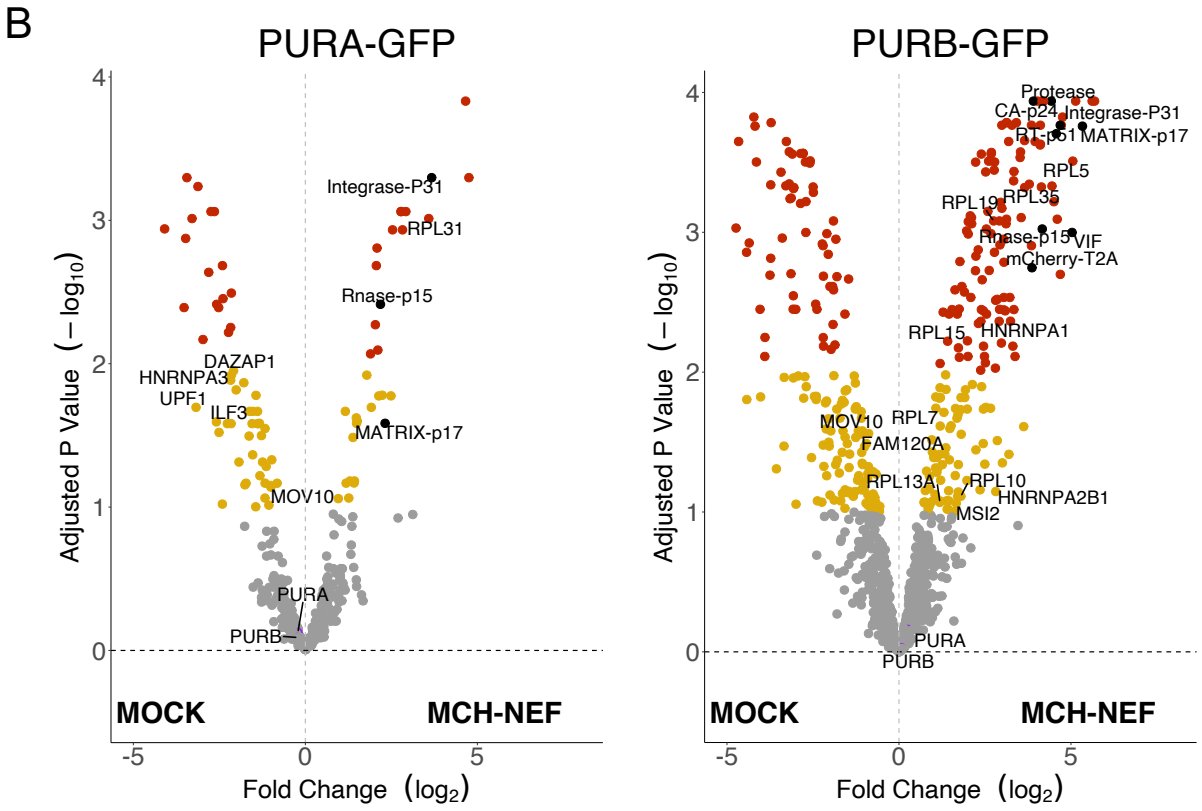
Next, I performed 3 IP replicates which were analysed by label-free mass spectrometry. Firstly, protein intensity ratios between (A) PURA-GFP versus GFP eluates and (B) PURB-GFP vs GFP eluates of the same transduction condition were calculated. Proteins that did not significantly interact with the relevant PUR protein compared to the GFP control (those exhibiting an adjusted p value > 0.05) were removed, filtering contaminants binding non-specifically to GFP. Next, protein intensity ratios between (A) transduced PURA-GFP versus mock-transduced PURA-GFP eluates and (B) transduced PURB-GFP versus mock-transduced PURB-GFP eluates were calculated. This allowed me to identify proteins whose interaction with PURA-GFP or/and PURB-GFP was specifically altered by HIV-1 infection and thus whose interactions may underpin the proviral functions of PURA and PURB. Significance of all ratios was estimated using a moderated t-test (LC-MS and initial analyses performed by Honglin Chen). Protein sequences of *Homo sapiens* PURA and PURB were searched for during MS analyses.

Quantitative proteomics revealed 83 and 337 host cell proteins which displayed significant differential interactions with PURA-GFP and PURB-GFP respectively between HIV-1 transduced versus mock transduced conditions (**Figure 6.5, B**) (**Supplementary Table 2, Supplementary Table 3**). This large difference in interactome sizes is surprising, given that PURA and PURB have a sequence

conservation of 70% (343). This difference is likely technical, probably due to lower levels of PURA-GFP expressed following doxycycline induction (**Figure 6.5, A**). Both proteins exhibited HIV-1 transduced/mock transduced fold changes close to 0, indicating that their expression levels are not significantly altered during HIV-1 replication, as seen in silver staining. Moreover, both proteins exhibited similar intensities across replicates and group closely on both volcano plots. This is probably due to PURA and PURB forming a heterodimer *in vivo* and suggests that this heterodimer persists during HIV-1 replication (344).

Interestingly, PURA-GFP and PURB-GFP both interacted significantly with many viral proteins during HIV-1 replication (**Figure 6.5, C**). In particular, while PURB is only known to interact with Rev and has no known role in HIV-1 infection, I identified extensive novel interactions between PURB-GFP and HIV-1 proteins, including PR, INT, CA and MA (both derived from Gag/Gag-Pol), Vif and RT. These viral proteins are present in HIV-1 particles and thus, these interactions with Gag are likely to occur at virus assembly sites or formed inside viral particles. Strong interactions between PURB-GFP and HIV-1 structural and enzymatic proteins indicates that PURB may assist in the assembly of viral particles or the very first steps of infection in the new cell. Surprisingly, I did not detect Rev in PURA/PURB IPs, despite the presence of both PUR proteins in the Rev-Flag interactome. This may be because of the small size and low abundance of Rev in infected cells that makes it challenging to identify. As Rev is the most enriched protein in Rev IPs, this maximises the identification chances, even if a limited number of peptides are detected. On the other hand, IP of PURA- or PURB-GFP may purify Rev, together with a complex interactome, at a reduced stoichiometry, making it challenging to identify Rev.


A
GFP
PURA -GFP
PURB -GFP
MOCK
TRANSDUCED
MOCK
TRANSDUCED
MOCK
TRANSDUCED
- PURA-GFP (75 kD)
PURB-GFP (74 kD)
- GFP (27 kD)
- Anti-PURA (75 kDa)
- Anti-PURB (74 kDa)
- Anti-GFP (27 kDa)


B
PURA-GFP
PURB-GFP
Adjusted P Value (−log10)
Fold Change (log2)
MOCK
MCH-NEF
Integrase-P31
RPL31
Rnase-p15
DAZAP1
HNRNPA3
UPF1
ILF3
MATRIX-p17
MOV10
PURA
PURB
Protease
CA-p24
Integrase-P31
RT-p51
MATRIX-p17
RPL5
RPL19
RPL35
Rnase-p15
VIF
mCherry-T2A
RPL15
HNRNPA1
MOV10
RPL7
FAM120A
RPL13A
RPL10
HNRNPA2B1
MSI2
PURA
PURB

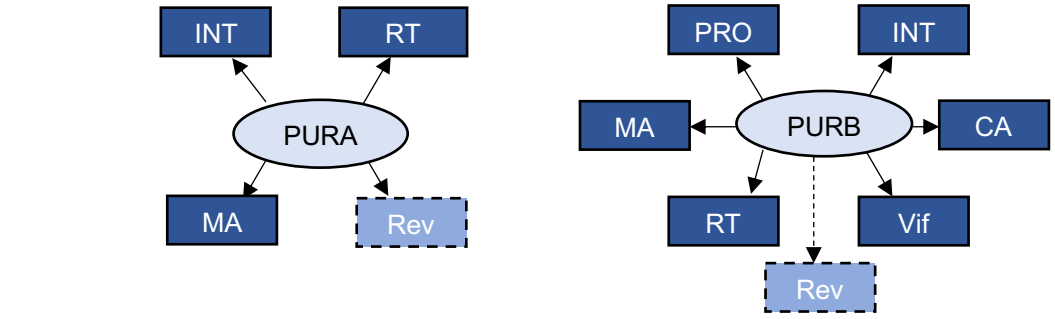

C
INT
RT
PURA
MA
Rev
PRO
INT
MA
PURB
CA
RT
Vif
Rev

Figure 6.5: PURA-GFP and PURB-GFP interactomes revealed during HIV-1 replication. (A) The interactions of PURA-GFP and PURB-GFP were investigated in lymphocytic cell lines transduced with HIV-$1_{mCherry-Nef}$ or mock transduced at 48 hpi using GFP-trap beads. IPs were carried out in triplicate and compared to IP eluates of Jurkat cells expressing free GFP by silver staining and Western blotting. (**B**) Volcano plots of PUR protein IPs, showing log2 fold change of protein intensities in HIV-$1_{mCherry-Nef}$ versus mock infected eluates on the X axis and adjusted -log10 p value on the Y axis (n=3). Individual dots correspond to proteins: grey dots indicate non-significant proteins; yellow dots indicate a p value > 0.01 and < 0.1 (significant) and red dots indicate a p value < 0.01 (highly significant). (**C**) The viral proteins identified to bind significantly to PURA- and PURB-GFP by quantitative proteomics. Rev is shown in a dashed box, as it was identified to bind both PUR proteins in the Rev-Flag interactome rather than through PUR protein IPs.

To discover the interactions of PURA and PURB that may underpin regulatory roles in HIV-1 infection, I filtered significant interactions by those with a fold change >1 in transduced versus mock conditions. These are proteins whose interaction with PURA-GFP or PURB-GFP is substantially increased by HIV-1 replication, and which may therefore be required for efficient infection. Both proteins interacted significantly with DNA topoisomerase II alpha (TOP2A), an enzyme that relieves torsional stress of DNA during transcription. TOP2A is a critical regulator of reverse transcription during HIV-1 infection. When TOP2A is downregulated, HIV-1 reverse transcription is completely inhibited and levels of Gag protein severely diminished as a result (345). Thus, PURA and PURB interacting with TOP2A may facilitate its role in reverse transcription, ensuring efficient vDNA synthesis during early infection. Similarly, PURA may be a necessary stimulant of viral transcription, as it interacted strongly with POLR2A, the largest subunit of RNA polymerase II. RNA polymerase II is required to synthesise eukaryotic mRNA and is bound by HIV-1 Tat to enhance vRNA transcription. Knock-down of POLR2A strongly inhibits the early stages of HIV-1 infection (346). Indeed, PURA has been previously reported to aid assembly of viral Tat protein on to promoters to stimulate RNA polymerase II activity (337). Interestingly, both proteins

interacted with many members of protein decay machineries. For example, PURB-GFP interacted significantly with a group of 6 proteasome proteins (PSMD14, PSMC6, PSMA5, PSMD11, PSMC4, PSMD1). The eukaryotic proteasome degrades misfolded proteins by proteolysis (347) and is hijacked by HIV-1 proteins which promote ubiquitination and proteasomal degradation of antiviral proteins to avoid host cell defenses (347). PURA-GFP interacted strongly with EBR4, an E3 ubiquitin-protein ligase which marks proteins with destabilising N-terminal residues for proteasomal degradation. Furthermore, PURA-GFP interacted with both exosome proteins EXOSC4 and ZCCHC8, the protein component of the nuclear exosome targeting (NEXT) complex which targets the exosome to its substrates. These interactions suggest that both PUR proteins have wide-spread roles in RNA metabolism.

If the interactions of PURA and PURB are important for infection, I expect them to interact with proteins known to influence HIV-1 infection. To investigate this, I compared significant PURA-GFP and PURB-GFP interactors to the NCBI HIV-1 database as before (**Figure 6.6, A-B**). This comparison revealed that around half of PURA- and PURB-GFP interactors have been previously linked to HIV-1 infection (55% and 46% respectively). These include interactors which engage in complexes with HIV-1 proteins to downregulate antiviral factors. For example, PURA-GFP bound strongly to gene expression regulator RUNX1. RUNX1 forms a transcription complex with cofactor CBF-β, which is required to generate APOBEC3 proteins (348). APOBEC3 proteins are a family of DNA cytidine deaminases that inhibit HIV-1 infection by hypermutating the viral genome (349). CBF-β/RUNX1 complexes are disrupted by the formation of a Vif/CBF-β heterodimer during HIV-1 infection, reducing the production of APOBEC3 proteins and permitting the assembling and packaging of infective progeny virions (348). Similarly, PURB-GFP interacted with DCAF1, the

substrate recognition component of E3 ubiquitin ligase complexes responsible for downregulating proteins through poly-ubiquitination. Vpr redirects host cell proteins to DCAF1 to promote their downregulation, achieving G2/M cell cycle arrest through this mechanism (350). In this way, I identified both PUR proteins interacting with proteins that are downregulated or mediate downregulation during infection: perhaps they play a role in creating an optimal environment for HIV-1 replication through the degradation of antiviral factors. If PURA and PURB participate in early nuclear vRNPs assembled by Rev, I expect that both proteins interact with known Rev interactors. Interestingly, only 17% and 9% of significant PURA-GFP and PURB-GFP interactors were previously known to interact with Rev respectively. I found a similarly modest overlap when comparing significant PUR-GFP interactors to Rev-Flag interactors (data not shown). While this suggests that PURA and PURB do not contact an identical network of proteins as Rev during HIV-1 infection, I identified interactions between PUR proteins and some of the highest fold change interactors in the Rev-Flag interactome. For example, the 2 proteins which exhibited the lowest adjusted p value and highest fold change in the Rev interactome were ERGIC3 and ELMO1. PURA-GFP interacted with DOCK2, a guanine nucleotide exchange factor that activates the small G protein Rac, which regulates cell movement and cytoskeletal rearrangement. DOCK2 forms a complex with ELMO1, which is required for the DOCK2-mediated Rac activation in lymphocytic cell lines. Similarly, I identified an interaction between PURB-GFP and ERGIC3. These interactions, along with the presence of PURA and PURB in the Rev-Flag interactome, serve as evidence that both PUR proteins likely participate in Rev-assembled complexes.

Work in the Castello lab revealed that HIV-1 gene expression is diminished by depletion of PURA or PURB. If specific interactions of these proteins underpin their

proviral functions, disrupting these interactions serves as a potential therapeutic target against HIV-1 replication. In this way, the interactomes of PURA-GFP and PURB-GFP can be used as resources to identify candidate proteins to target. To explore this, I examined the drug-gene interaction database (210). This database is a curated list of the genes which are affected by currently available compounds. I filtered by compounds which have inhibitory effects on gene expression and next by resulting compounds that affect significant interactors of PURA-GFP or PURB-GFP. While the list of drugs corresponding to PURA-GFP interactors was modest, I identified over 30 available compounds that are inhibitory to PURB-GFP interactors (**Figure 6.6, C**). Many of these compounds have already been found to have anti-retroviral effects against HIV-1 replication (351-355).

While more extensive analysis should be carried out on these datasets, my work, together with research from the Castello lab, has revealed that host cell proteins which are packaged into viral particles are essential for ensuring efficient infection in the next cell. Revealing the interaction networks of two such proteins, PURA and PURB, has provided a valuable resource which can be mined to gain insight into the proviral roles of these proteins and to reveal therapeutic targets to restrict HIV-1 replication. Furthermore, this work supports the notion that Rev is a vital HIV-1 protein that plays an instrumental role in directing necessary proviral proteins to the assembling virion.

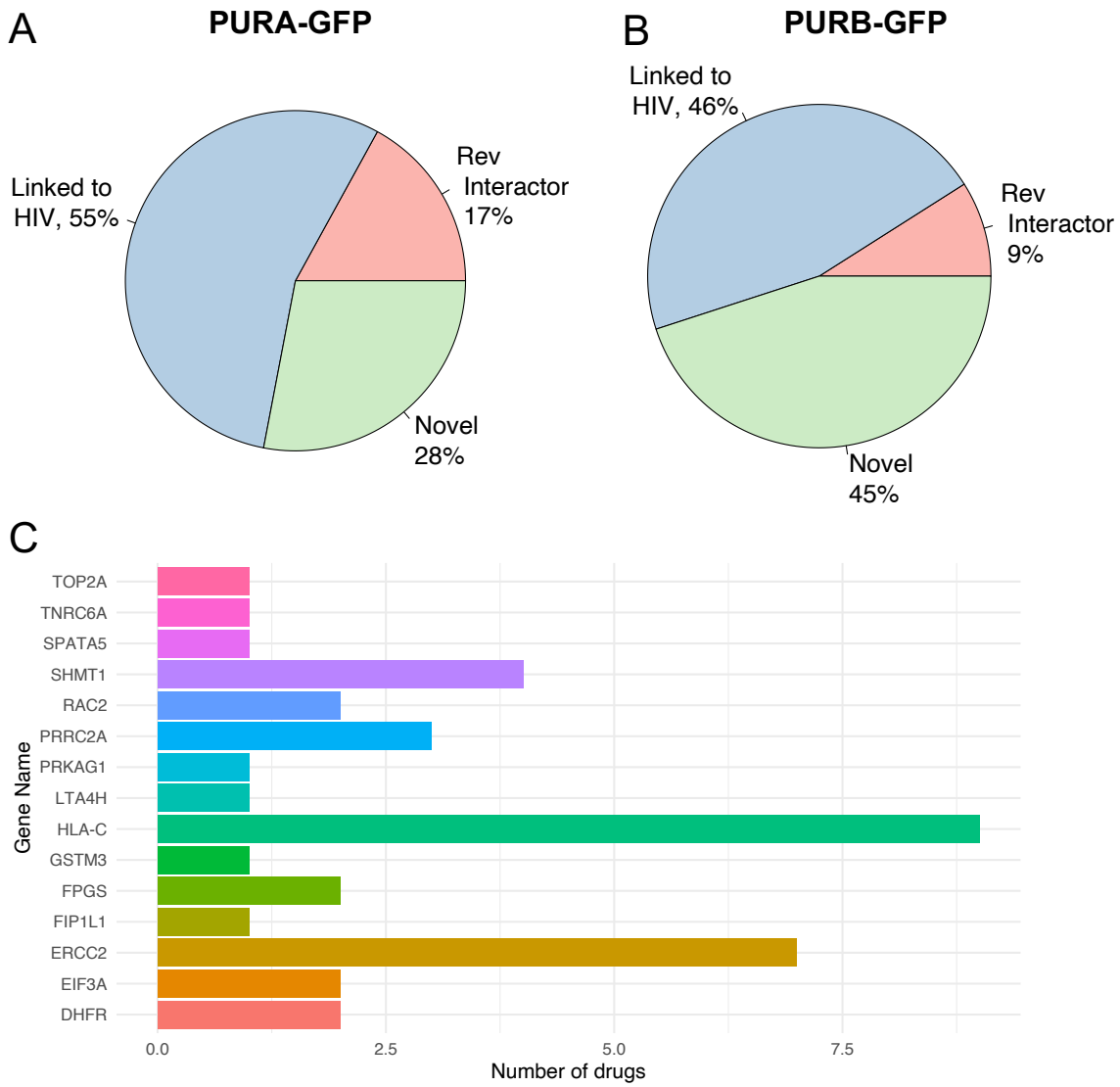


Figure 6.6: PURA-GFP and PURB-GFP interactions may underpin proviral effects. (**A-B**) I filtered the NCBI HIV-1 database by significant PURA-GFP and PURB-GFP interactors in the indicated categories of the pie chart. (**C**). I used the drug-gene interaction database to identify the number of inhibitory drugs against significant PURB-GFP interactors and found over 30 drugs, many of which are inhibitory against HIV-1 replication. X axis plots the number of compounds available that target the gene indicated on the Y axis.

6.5. Discussion

6.5.1. Rev acts as a molecular spiderweb to catch nuclear proteins for encapsidation

In this chapter, I examined the possibility that Rev promotes the packaging of host cell proteins into viral particles by comparing the ivRBPome to the Rev-Flag interactome. I discovered that almost 70% of encapsidated host cell proteins identified using ivRIC were also present in the Rev-Flag interactome, supporting prior work that Rev influences packaging (324, 356). Through meta-analyses of these large datasets, I have presented a model which begins with Rev-mediated recruitment of host cell RBPs to vRNAs in the nuclei of infected cells. This recruitment is supported by the enrichment of RBPs in the Rev-Flag interactome, which comprised over 80% RBPs. Furthermore, the idea that viral complexes assemble in the nucleus during early HIV-1 infection is supported by new work by Tuffy *et al,* in which the co-localisation of HIV-1 Gag, Rev and unspliced vRNA in discrete nuclear foci was revealed using FISH/IF (357). The number of these foci increased in the presence of CRM1 inhibitor leptomycin B, indicating that nuclear export of these vRNPs is dependent on the Rev export pathway (357). The model continues with Rev-mediated export of vRNPs to the cytoplasm. GO terms of shared ivRBPome/Rev-Flag interactors revealed an enrichment in nuclear proteins, supporting the idea that Rev exports these proteins concomitant with vRNAs to allow their encapsidation downstream. While Rev is often relegated to the role of RNA export factor, several host cell proteins have been identified to participate in the Rev/HIV-1 gRNA export complex (98, 99, 358-360), suggesting Rev may export host cell proteins too. Future work should focus on challenging the assumption that Rev solely transports vRNAs. To examine which nuclear proteins Rev exports during HIV-1 infection, I propose similar experiments to those performed by Tuffy *et al*. The Rev-Flag interactome can be filtered by interactors

which exhibit nuclear localisation, using the Human Cell Atlas, and the localisation of these proteins could be examined during HIV-1 infection. Leptomycin B can be applied to identify whether these proteins exhibit different localisation patterns when Rev is unable to complete its export cycle. This would indicate whether these proteins are exported from the nucleus by Rev during HIV-1 infection. Following nuclear export, I hypothesise that RBPs bound to vRNA become incorporated into viral assembly complexes. Two proteins that form vRNPs during HIV-1 infection are Staufen1 and Gag, which both bind vRNA and drive packaging. To examine this, I compared Rev and ivRIC datasets to proteomes of Gag and Staufen1 and discovered a conserved pool of almost 40 proteins common to all 4 datasets. Indeed, the recruitment of RBPs to viral particles by means of binding encapsidated RNA is not a new phenomenon. APOBEC3G piggybacks into HIV-1 virions by binding 7SL, the RNA component of the SRP complex, which is encapsidated into viral particles at high volumes (361). In this model, Rev brings together viral factors required for packaging, such as Gag and the vRNA, with host cell proteins that require packaging.

6.5.2. Rev promotes encapsidation of host cell proteins that are critical proviral factors

Recent evidence suggests that the HIV-1 CA core breaks down inside the nucleus (32). Thus, host cell proteins required for viral processes between virion maturation and CA disassembly/early-phase infection steps in the nucleus must be pre-packaged inside the CA core. Accordingly, the proteome of HIV-1 virions can elucidate which host cell proteins are vital for HIV-1 infection. The Castello lab used ivRIC to identify host cell proteins interacting directly with HIV-1 gRNA inside virions and discovered 103 packaged proteins, many of which have well characterised roles in infection. For

example, RNA helicases UPF1 and DHX9/RNA helicase A are known to associate with the Gag protein to stimulate its translation and facilitate viral reverse transcription respectively (23, 362, 363). To examine the hypothesis that encapsidated proteins are vital for HIV-1 infection, the Castello lab tested the infectivity of HIV-1 particles produced from SupT1 PURA and PURB knock-out cell lines. These particles were almost entirely uninfective compared to particles produced from wildtype cells, confirming that PURA and PURB are essential for the early phases of HIV-1 infection. I continued this work by investigating PURA and PURB interactions during HIV-1 infection in CD4+ lymphocytic cell lines, to gain a deeper insight into PURA/PURB proviral functions. This revealed a broad network of interactions between PURA, PURB and HIV-1 proteins packaged into the virion core, which facilitate early-phase infection steps such as reverse transcription and integration. Further, it revealed interactions with other host cell proteins which are essential for early-phase reactions such as RNA pol II, which is required for efficient viral transcription. These findings signify that host cell proteins which Rev may convey to assembling virions play crucial roles in early-phase HIV-1 infection. Thus, the ivRBPome represents a resource for locating proviral host cell proteins whose functions or interactions can be targeted by therapeutics to counteract viral infection. With more time, other ivRIC/Rev-Flag proteins could be tested in the context of knock-out cell lines, as performed for PURA and PURB, allowing us to identify further interactions between host and viral proteins that permit infection.

6.5.3. Staufen1 and Rev: assembling efficient sites of translation

Rev has been posited to increase the translation of vRNAs by increasing the association of underspliced vRNA with polysomes in an RRE-dependent manner

(135). I hypothesise that this effect may be mediated through interaction with Staufen1. Staufen1 RNPs promote efficient translation of vRNAs by disassembling stress granules (SGs), which are translationally silent cytoplasmic RNPs. HIV-1 infection prevents SG formation, promoting the assembly of Staufen1 particles instead. In Staufen1 knock-out cells, SGs could not be efficiently disassembled by HIV-1 infection and vRNA instead accumulated in SGs and exhibited impaired translation. Expression of Staufen1 rescued vRNA trafficking and translation, demonstrating that Staufen1 forms complexes that are rich in translation machinery and facilitate optimal vRNA assembly and protein synthesis. In accordance with my model, I estimate that Rev interacts with Staufen1 in the nucleus and recruits vRNAs and translational proteins to promote the early formation of translation-rich RNPs. This hypothesis is supported by my work in this chapter, which uncovered a strong overlap between Staufen1 and Rev interactomes during HIV-1 infection, both of which are rich in translation machinery. My hypothesis is further supported by work on Staufen1 and the HERV-K Rec protein. The HERV-K genome encodes the Rec protein, which binds to a Rec-Response Element (RcRE) on HERV-K RNAs to export them to the cytoplasm, in an analogous manner to Rev. Rev can also bind the RcRE, which forms the same topological structure as the RRE (364), and export HERV-K RNAs, contributing to their elevated expression levels during physiological HIV-1 infection (365). Conservation between Rec and Rev suggests that proviral interactions between Staufen1 and Rev are ancient and vital for infection. Interestingly, IF studies revealed strict co-localisation between Staufen1 and Rec in the nucleoli and cytoplasm of 293T cells (365). Moreover, overexpression of Staufen1 caused a 5-fold increase in the expression of unspliced RcRE-containing HERV-K vRNAs in the presence of Rec (365). I

hypothesise that Staufen1 interacts in a conserved manner with Rev, promoting Rev-mediated vRNA export and participating in a translation-rich complex.

7. General Discussion and future directions

7.1. Revisiting Rev functions in a more physiological system

Rev was discovered in 1986 by Sodroski *et al,* who transfected B lymphocytes with a HIV-1 provirus and measured levels of viral and cellular proteins and RNAs at 48 h-post transfection. Mutations in the *rev* gene of this provirus, which was not known to be a discrete ORF, caused significant depletion of Gag. This revealed a novel gene that was required for viral protein expression (49). In the decades since this discovery, research into Rev has continued. However, non-physiological approaches have been used to compensate for the technical difficulties of studying Rev (detailed in chapter 4), culminating in contrasting data. Indeed, while Sodroski *et al* reported that mutation in the *rev* ORF reduced Gag protein levels, they found that it did not affect levels of unspliced vRNAs, contrary to many subsequent studies (49). Much of the contrasting data concerning Rev has yet to be revisited using more physiological approaches.

In this work, I developed and optimised a system to study Rev in conditions nearer to physiological infection than used in prior research, using HIV-1 tagged Rev replicons. I exploited these tools to revisit the functions and interactions of Rev which were uncovered in less physiologically relevant conditions, and garnered unexpected and novel data. For example, while Rev is reported to interact with the CRM1/RanGTP export complex, research into its other interactions have focused on a limited number of proteins (reviewed in (211)) such as eIF5A and Sam68, whose interactions with Rev remain controversial. Uncovering the Rev-Flag interactome revealed novel interactions between Rev and over 200 host cell proteins and unveiled interactions between Rev and a network of host cellular RBPs with prominent roles in RNA metabolism. I also employed this system to visualise Rev-Halo localisation *in vivo* using a Halo-tag that is compatible with single molecule analyses. While previous Rev

IF studies often identified either strictly nucleolar or cytoplasmic localisation (64, 91, 226-230), the expression of Rev-Halo from its cognate HIV-1 gRNA revealed a strong localisation of the Rev-Halo fusion protein at the nuclear pore. Rev must interact with the nuclear pore during RNA export, though I would expect this interaction to be transient. However, concomitant visualisation of genomic vRNA and Rev-Halo showed high levels of Rev-Halo in the NPC during late-phase viral replication, when most of the vRNAs in the cell were cytoplasmic. Thus, my data suggests that large pools of Rev protein maintain interactions with the NPC during infection even beyond RNA export. HIV-1 infection is known to remodel the nuclear pore through depletion of nucleoporins (101): Rev may contribute to this remodelling during HIV-1 infection. Moreover, concomitant visualisation of genomic vRNA and Rev-Halo revealed a lower degree of co-localisation than would be expected if the only function of Rev is to regulate the HIV-1 RNA metabolic cycle, as is often presumed in the literature. This suggests that Rev performs functions beyond RNA export and that Rev-Halo foci without co-localising HIV-1 gRNA likely represent Rev proteins that are engaging with host cell nucleic acids and proteins.

I can also employ my experimental approach to test Rev functions not yet investigated. For example, I could analyse Rev-Halo localisation using super-resolution microscopy, which has not previously been performed to my knowledge. This could inform whether Rev escorts genomic HIV-1 RNA to the assembling virion at the plasma membrane, for example, or reveal unexplored interactions between Rev and core or peripheral components of the nuclear pore during RNA export. Another poorly investigated question that could be explored using Rev-tagged replicons is whether Rev interacts with cellular nucleic acids during HIV-1 infection. Rev uses an ARM to bind gRNA, and similar motifs are found in several host RNA and DNA-binding

proteins (366). Thus, Rev may also interplay with cellular RNA and/or DNA, which has never been investigated to my knowledge. Using the tagged Rev HIV-1 replicons and individual-nt resolution (i)CLIP experiments, I could identify which cellular RNAs are bound by Rev. Indeed, I began implementing this method by applying UV crosslinking to SupT1 CD4+ lymphocytic cells transduced with HIV-$1_{Rev\text{-}Flag}$ and subjecting lysates to the Flag IP I optimised in chapter 5. While I was unable to complete this work due to time restraints and the COVID19 lockdown, my preliminary data showed RNAs that co-precipitated with Rev in HIV-1 replication. This work demonstrates that complex questions can be addressed using the new experimental approaches I have developed, and hypotheses regarding Rev functions can be revisited under experimental conditions better approximating physiological conditions than used previously. Importantly, the new data I have uncovered stresses the importance of revisiting previous research using modern technology and more optimal, physiological experimental approaches.

7.2. Rev-mediated export may dictate protein packaging fate

In this work, I discovered that Rev-Flag interacted with almost 70% of the host cell proteins identified to bind HIV-1 gRNA inside virions using ivRIC. Furthermore, many of these shared proteins are common to the interactomes of Staufen1 and Gag, which mediate viral assembly vRNPs and drive packaging. From these data, I have formed a hypothesis wherein Rev acts as a molecular spiderweb to recruit proviral host RBPs that are required during the metabolic cycle of the vRNA, including its assembly into virions. I envision a model whereby Rev binds to RBP complexes in the nuclei of infected cells and recruits them to the vRNA by simultaneously binding the RRE. This assembles an early vRNP which, following nuclear export, becomes integrated into

Gag RNP complexes in the cytoplasm, that, after potential remodelling, are ultimately packaged in HIV-1 virions.

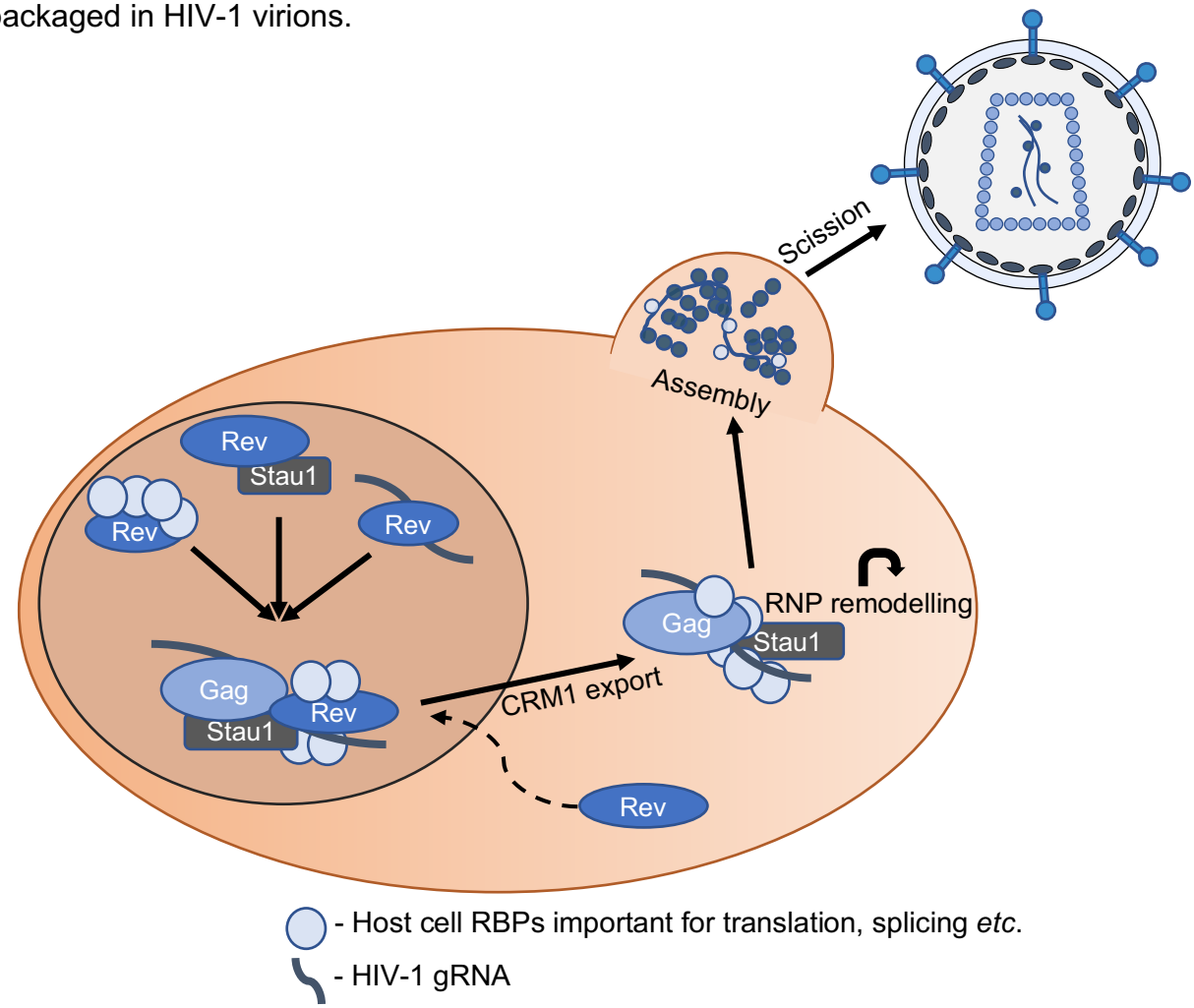


Figure 7.1: Rev acts as a molecular spiderweb during HIV-1 infection. Based on the work in this thesis, I hypothesise that Rev acts as a molecular spiderweb, assembling an early vRNP in the nucleus of infected cells which is rich in RBPs necessary for successful infection. This model begins with Rev forming interactions with viral and cellular RBPs in the nucleus (Staufen1, Gag) and bringing them into contact with vRNAs during export. Rev then disassembles from this vRNA by masking of the NES in the cytoplasm and continues its export cycle in the nucleus. The vRNP then remodels during translation and transit in the cytoplasm. RBPs recruited to the vRNA by Rev are encapsidated into virions through Gag/vRNA association.

This idea is supported by the observation that a Rev-deficient, codon-optimized HIV-1 RNA construct (HIV-$1_{(-)Rev}$), which does not require Rev for nuclear export, was packaged into virions 100-fold less efficiently in the absence of Rev (148). These results suggest that Rev is required for packaging in addition to nuclear export, as

vRNA that reaches the cytoplasm cannot be efficiently packaged without Rev. Similarly, tethering TAP protein to HIV-$1_{(-)Rev}$ permitted HIV-1 gRNA to reach the cytoplasm at similar levels in presence or absence of Rev (148). However, levels of RNA packaged into virions increased strongly when Rev was present (148, 367). How can TAP/NXF1 or Rev/CRM1 pathways successfully export vRNA to the cytoplasm (148, 367), but those vRNAs only achieve encapsidation in the presence of Rev? I hypothesise that Rev mediates the assembly of a vRNP complex containing specific host cell RBPs required for downstream infection steps such as translation and encapsidation. My work in chapter 5 uncovered that Rev forms interactions with proteins involved in virtually every step of RNA metabolism, including splicing, decay and translation. Thus, Rev may recruit host cell proteins on to vRNA which are important for downstream steps, such as translation and packaging. Previous research supports the hypothesis that vRNA must form specific interactions in the nucleus in order to be efficiently packaged. For example, Moore *et al* demonstrated that 2 copies of the HIV-1 gRNA must be exported by the same export pathway (*e.g.,* Rev/CRM1 or TAP/NXF1) to be efficiently co-packaged together in a virion (367). Similarly, HIV-1 Gag RNA exported by the Hepatitis B virus post-transcriptional regulatory element caused a defect in plasma membrane assembly, which was ameliorated by expression of Gag in a Rev-dependent manner (368). The idea that Rev initiates vital interactions in the nucleus is already well-evidenced. The co-localisation of HIV-1 Gag, Rev and unspliced vRNA in discrete nuclear foci has been demonstrated and has been proven to be dependent on Rev-mediated export (357). Additionally, when Rev and CRM1 are overexpressed separately, they yielded nucleolar and nuclear envelope localization, respectively (177). However, when both were overexpressed together, CRM1 mobilized to the nucleolus in a Rev-dependent manner. These results confirm

that Rev mediates important interactions in the nucleus that influence the downstream fate of HIV-1 RNA.

7.3. Summary and future perspectives

To summarise, this work details an experimental approach to study the HIV-1 Rev protein under conditions nearer to physiological infection than previously used and unveils the Rev protein-protein interactome using this approach. While these insights provide novel conclusions regarding Rev/host cell PPIs, future works should aim to validate the Rev PPIs revealed here using orthogonal methods. For example, a proximity ligase assay could be performed using the HIV-$1_{Rev\text{-}APEX2}$ replicon that I generated and validated (discussed in chapter 5) (369). This would identify Rev PPIs in a high-throughput manner, under highly similar conditions (*e.g.,* CD4+ lymphocytic cells transduced with pseudotyped Rev-tagged replicons) but using a different experimental approach. Following validation experiments, future works should address the mechanistic details behind the interactions of Rev and how these interactions promote or hinder HIV-1 replication. For example, I identified interactions between Rev and host cell proteins ERGIC3 and ELMO1. Through IF experiments, I have observed that ERGIC3 may play an antiviral role in infection, potentially downregulating Rev in the cell. On the other hand, I showed that ELMO1 co-localised strongly with Rev-Halo, causing an increase in cytoplasmic Rev-Halo levels. We similarly determined that PURA and PURB are required for early HIV-1 infection steps. How do these host cell proteins promote or restrict HIV-1 infection? Experiments can be performed to identify which stage(s) of HIV-1 replication these proteins effect and how. For example, to test whether reverse transcription is effected by these proteins, RT-qPCR can be performed on RNA extracted from HIV-1 infected cells, using primers

that target different regions of the viral genome (370). Viral genome integration can be tested using *Alu-gag* PCR, which quantifies the level of integrated HIV DNA in HIV-1 infected cells (371). Viral transcription can be tested using uridine analog 4-thiouridine (4sU) labelling. 4sU becomes integrated in nascently transcribed RNAs, metabolically labelling them, and these RNAs can be biotinylated via 4sU sulfylhydryl groups and isolated using streptavidin-coated beads. The levels of vRNA transcribed in HIV-1 infected cells can then be quantified (372). Translation of viral proteins can be tested by Western blotting against viral proteins in HIV-1 infected cells, in a similar manner to the blots I presented in chapter 4. These experiments can be carried out using tools I have generated and validated in this work. For example, performing these experiments in wildtype versus the PURA/PURB/ERGIC3/ELMO1 fusion GFP overexpression cell lines we have generated. Moreover, knock down cell lines or inducible shRNA cell lines can be generated and similarly used to test these experiments. An excellent tool to study the requirement of Rev in HIV-1 infection is the HIV-$1_{\Delta Rev}$ replicon I have generated and validated. The localisation and expression levels of Rev interactors can be tested in both HIV-$1_{mCherry\text{-}Nef}$ and HIV-$1_{\Delta Rev}$ transduced cells to immediately reveal whether Rev alters the localisation of these proteins during HIV-1 replication.

In addition to deeper research into Rev interactions, the molecular spiderweb hypothesis should also be tested. I firstly recommend that IF experiments combined with super-resolution microscopy be carried out. For example, the localisation of Rev interactors (identified using the Rev-Flag interactome) can be visualised with commercial antibodies in cells infected with HIV-$1_{Rev\text{-}Halo}$, with or without the RNA export inhibitor leptomycin B. The visualisation of both Rev and Rev interactors using fluorescence microscopy would indicate whether these proteins co-localise in the

nucleus and if their localisations change in a Rev/CRM1-dependent manner, similar to work previously described by Tuffy *et al (357)*.

Appendices

Supplementary table 1 – The Rev-Flag interactome: fold change, gene name, average expression and adjusted p value for Rev-Flag/mCherry-Nef t-test proteomics comparison.

Supplementary table 2 – PURA-GFP interactome: gene name, fold change, average expression, p value and adjusted p value for PURA-GFP transduced vs PURA-GFP mock-transduced t-test proteomics comparison.

Supplementary table 3 – PURB-GFP interactome: gene name, fold change, average expression, p value and adjusted p value for PURA-GFP transduced vs PURA-GFP mock-transduced t-test proteomics comparison.

Bibliography

1. (WHO) WHO. Fact Sheet, World AIDS Day. 2020.
2. Laskey SB, Siliciano RF. A mechanistic theory to explain the efficacy of antiretroviral therapy. Nature Reviews Microbiology. 2014;12:772.
3. Moreno M, Caballero E, Mateus RM, Samba F, Gil E, Falcó V. HIV drug resistance in Africa: an emerging problem that deserves urgent attention. AIDS. 2017;31(11).
4. Deeks SG. Antiretroviral treatment of HIV infected adults. BMJ. 2006;332(7556):1489.
5. Mertz JA, Lozano MM, Dudley JP. Rev and Rex proteins of human complex retroviruses function with the MMTV Rem-responsive element. Retrovirology. 2009;6(1):10.
6. Watts JM, Dang KK, Gorelick RJ, Leonard CW, Bess Jr JW, Swanstrom R, et al. Architecture and secondary structure of an entire HIV-1 RNA genome. Nature. 2009;460(7256):711-6.
7. Martin Stoltzfus C. Chapter 1 Regulation of HIV-1 Alternative RNA Splicing and Its Role in Virus Replication. Advances in Virus Research. 74: Academic Press; 2009. p. 1-40.
8. Ocwieja KE, Sherrill-Mix S, Mukherjee R, Custers-Allen R, David P, Brown M, et al. Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule enrichment and long-read sequencing. Nucleic acids research. 2012;40(20):10345-55.
9. Benko DM, Schwartz S, Pavlakis GN, Felber BK. A novel human immunodeficiency virus type 1 protein, tev, shares sequences with tat, env, and rev proteins. Journal of virology. 1990;64(6):2505-18.
10. Carrera C, Pinilla M, Pérez-Álvarez L, Thomson MM. Identification of Unusual and Novel HIV Type 1 Spliced Transcripts Generated in Vivo. AIDS Research and Human Retroviruses. 2010;26(7):815-20.
11. Lützelberger M, Reinert LS, Das AT, Berkhout B, Kjems J. A Novel Splice Donor Site in the gag-pol Gene Is Required for HIV-1 RNA Stability*. Journal of Biological Chemistry. 2006;281(27):18644-51.
12. O'Reilly MM, McNally MT, Beemon KL. Two Strong 5′ Splice Sites and Competing, Suboptimal 3′ Splice Sites Involved in Alternative Splicing of Human Immunodeficiency Virus Type 1 RNA. Virology. 1995;213(2):373-85.
13. Si Z, Amendt BA, Stoltzfus CM. 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. 1997;25(4):861-7.
14. Staffa A, Cochrane A. The tat/rev intron of human immunodeficiency virus type 1 is inefficiently spliced because of suboptimal signals in the 3' splice site. Journal of virology. 1994;68(5):3071-9.
15. Zahler AM, Damgaard CK, Kjems J, Caputi M. 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. 2004;279(11):10077-84.
16. Caputi M, Mayeda A, Krainer AR, Zahler AM. hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. The EMBO journal. 1999;18(14):4060-7.
17. Caputi M, Freund M, Kammler S, Asang C, Schaal H. A Bidirectional SF2/ASF- and SRp40-Dependent Splicing Enhancer Regulates Human Immunodeficiency Virus Type 1 rev, env, vpu, and nef Gene Expression. Journal of Virology. 2004;78(12):6517.

18. Dowling D, Nasr-Esfahani S, Tan CH, O'Brien K, Howard JL, Jans DA, et al. HIV-1 infection induces changes in expression of cellular splicing factors that regulate alternative viral splicing and virus production in macrophages. Retrovirology. 2008;5:18-.
19. Erkelenz S, Hillebrand F, Widera M, Theiss S, Fayyaz A, Degrandi D, et al. Balanced splicing at the Tat-specific HIV-1 3'ss A3 is critical for HIV-1 replication. Retrovirology. 2015;12(1):29.
20. Sun W, Reyes-Serratos E, Barilla D, Santos JRL, Bujold M, Graves S, et al. Mathematical determination of the HIV-1 matrix shell structure and its impact on the biology of HIV-1. PLOS ONE. 2019;14(11):e0224965.
21. Briggs JAG, Grünewald K, Glass B, Förster F, Kräusslich H-G, Fuller SD. The Mechanism of HIV-1 Core Assembly: Insights from Three-Dimensional Reconstructions of Authentic Virions. Structure. 2006;14(1):15-20.
22. Campbell EM, Hope TJ. HIV-1 capsid: the multifaceted key player in HIV-1 infection. Nature reviews Microbiology. 2015;13(8):471-83.
23. Santos S, Obukhov Y, Nekhai S, Bukrinsky M, Iordanskiy S. Virus-producing cells determine the host protein profiles of HIV-1 virion cores. Retrovirology. 2012;9(1):65.
24. Gurer C, Cimarelli A, Luban J. Specific incorporation of heat shock protein 70 family members into primate lentiviral virions. Journal of virology. 2002;76(9):4666-70.
25. Saphire ACS, Gallay PA, Bark SJ. Proteomic Analysis of Human Immunodeficiency Virus Using Liquid Chromatography/Tandem Mass Spectrometry Effectively Distinguishes Specific Incorporated Host Proteins. Journal of Proteome Research. 2006;5(3):530-8.
26. Chen B. Molecular Mechanism of HIV-1 Entry. Trends in microbiology. 2019;27(10):878-91.
27. Hu W-S, Hughes SH. HIV-1 reverse transcription. Cold Spring Harbor perspectives in medicine. 2012;2(10):a006882.
28. Feng YX, Copeland TD, Henderson LE, Gorelick RJ, Bosche WJ, Levin JG, et al. HIV-1 nucleocapsid protein induces "maturation" of dimeric retroviral RNA in vitro. Proceedings of the National Academy of Sciences of the United States of America. 1996;93(15):7577-81.
29. Blanco-Rodriguez G, Gazi A, Monel B, Frabetti S, Scoca V, Mueller F, et al. Remodeling of the Core Leads HIV-1 Preintegration Complex into the Nucleus of Human Lymphocytes. Journal of Virology. 2020;94(11):e00135-20.
30. Mattei S, Glass B, Hagen WJH, Kräusslich H-G, Briggs JAG. The structure and flexibility of conical HIV-1 capsids determined within intact virions. Science. 2016;354(6318):1434.
31. von Appen A, Kosinski J, Sparks L, Ori A, DiGuilio AL, Vollmer B, et al. In situ structural analysis of the human nuclear pore complex. Nature. 2015;526(7571):140-3.
32. Burdick RC, Li C, Munshi M, Rawson JMO, Nagashima K, Hu W-S, et al. HIV-1 uncoats in the nucleus near sites of integration. Proceedings of the National Academy of Sciences. 2020;117(10):5486.
33. Monroe KM, Yang Z, Johnson JR, Geng X, Doitsh G, Krogan NJ, et al. IFI16 DNA sensor is required for death of lymphoid CD4 T cells abortively infected with HIV. Science (New York, NY). 2014;343(6169):428-32.
34. Dharan A, Talley S, Tripathi A, Mamede JI, Majetschak M, Hope TJ, et al. KIF5B and Nup358 Cooperatively Mediate the Nuclear Import of HIV-1 during Infection. PLOS Pathogens. 2016;12(6):e1005700.
35. Fernandez J, Machado AK, Lyonnais S, Chamontin C, Gärtner K, Léger T, et al. Transportin-1 binds to the HIV-1 capsid via a nuclear localization signal and triggers uncoating. Nature Microbiology. 2019;4(11):1840-50.

36. Bejarano DA, Peng K, Laketa V, Börner K, Jost KL, Lucic B, et al. HIV-1 nuclear import in macrophages is regulated by CPSF6-capsid interactions at the nuclear pore complex. eLife. 2019;8:e41800.
37. Zila V, Müller TG, Laketa V, Müller B, Kräusslich H-G. Analysis of CA Content and CPSF6 Dependence of Early HIV-1 Replication Complexes in SupT1-R5 Cells. mBio. 2019;10(6):e02501-19.
38. Kutluay SB, Bieniasz PD. Analysis of the Initiating Events in HIV-1 Particle Assembly and Genome Packaging. PLOS Pathogens. 2010;6(11):e1001200.
39. Kemler I, Meehan A, Poeschla Eric M. Live-Cell Coimaging of the Genomic RNAs and Gag Proteins of Two Lentiviruses. Journal of Virology. 2010;84(13):6352-66.
40. Poole E, Strappe P, Mok H-P, Hicks R, Lever AML. HIV-1 Gag–RNA Interaction Occurs at a Perinuclear/Centrosomal Site; Analysis by Confocal Microscopy and FRET. Traffic. 2005;6(9):741-55.
41. Rein A. RNA Packaging in HIV. Trends in microbiology. 2019;27(8):715-23.
42. Chukkapalli V, Ono A. Molecular determinants that regulate plasma membrane association of HIV-1 Gag. Journal of molecular biology. 2011;410(4):512-24.
43. Jouvenet N, Simon SM, Bieniasz PD. Imaging the interaction of HIV-1 genomes and Gag during assembly of individual viral particles. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(45):19114-9.
44. Olety B, Ono A. Roles played by acidic lipids in HIV-1 Gag membrane binding. Virus research. 2014;193:108-15.
45. Burniston MT, Cimarelli A, Colgan J, Curtis SP, Luban J. Human immunodeficiency virus type 1 Gag polyprotein multimerization requires the nucleocapsid domain and RNA and is promoted by the capsid-dimer interface and the basic region of matrix protein. Journal of virology. 1999;73(10):8527-40.
46. Sundquist WI, Kräusslich H-G. HIV-1 assembly, budding, and maturation. Cold Spring Harbor perspectives in medicine. 2012;2(7):a006924-a.
47. Gupta S, Bromley J, Saffarian S. High-speed imaging of ESCRT recruitment and dynamics during HIV virus like particle budding. PLOS ONE. 2020;15(9):e0237268.
48. Feinberg MB, Jarrett RF, Aldovini A, Gallo RC, Wong-Staal F. HTLV-III expression and production involve complex regulation at the levels of splicing and translation of viral RNA. Cell. 1986;46(6):807-17.
49. Sodroski J, Goh WC, Rosen C, Dayton A, Terwilliger E, Haseltine W. A second post-transcriptional trans-activator gene required for HTLV-III replication. Nature. 1986;321(6068):412-7.
50. Maria LZ, Thomas JH, Tristram GP, Michael RG. Oligomerization and RNA Binding Domains of the Type 1 Human Immunodeficiency Virus Rev Protein: A Dual Function for an Arginine-Rich Binding Motif. Proceedings of the National Academy of Sciences of the United States of America. 1991;88(17):7734-8.
51. Daugherty MD, Liu B, Frankel AD. Structural basis for cooperative RNA binding and export complex assembly by HIV Rev. Nature Structural &Amp; Molecular Biology. 2010;17:1337.
52. DiMattia Michael A, Watts Norman R, Cheng N, Huang R, Heymann JB, Grimes Jonathan M, et al. The Structure of HIV-1 Rev Filaments Suggests a Bilateral Model for Rev-RRE Assembly. Structure. 2016;24(7):1068-80.

53. Jayaraman B, Crosby DC, Homer C, Ribeiro I, Mavor D, Frankel AD. RNA-directed remodeling of the HIV-1 protein Rev orchestrates assembly of the Rev-Rev response element complex. eLife. 2014;3:e04120-e.
54. DiMattia MA, Watts NR, Stahl SJ, Rader C, Wingfield PT, Stuart DI, et al. Implications of the HIV-1 Rev dimer structure at 3.2 Å resolution for multimeric binding to the Rev response element. Proceedings of the National Academy of Sciences. 2010;107(13):5810.
55. Watts NR, Misra M, Wingfield PT, Stahl SJ, Cheng N, Trus BL, et al. Three-Dimensional Structure of HIV-1 Rev Protein Filaments. Journal of Structural Biology. 1998;121(1):41-52.
56. Fischer U, Meyer S, Teufel M, Heckel C, Lührmann R, Rautmann G. Evidence that HIV-1 Rev directly promotes the nuclear export of unspliced RNA. The EMBO journal. 1994;13(17):4105-12.
57. Mészáros B, Erdós G, Dosztanyi Z. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding Nucleic Acids Research2018. p. W329-W37
58. Katahira J. Nuclear export of messenger RNA. Genes. 2015;6(2):163-84.
59. Luo M-j, Reed R. Splicing is required for rapid and efficient mRNA export in metazoans. Proceedings of the National Academy of Sciences. 1999;96(26):14937.
60. Jacob AG, Smith CWJ. Intron retention as a component of regulated gene expression programs. Human genetics. 2017;136(9):1043-57.
61. Pasquinelli AE, Ernst RK, Lund E, Grimm C, Zapp ML, Rekosh D, et al. The constitutive transport element (CTE) of Mason-Pfizer monkey virus (MPMV) accesses a cellular mRNA export pathway. EMBO J. 1997;16(24):7500-10.
62. Sandri-Goldin RM. Viral regulation of mRNA export. Journal of virology. 2004;78(9):4389-96.
63. Henderson BR, Percipalle P. Interactions between HIV rev and nuclear import and export factors: the rev nuclear localisation signal mediates specific binding to human importin-β11Edited by M. Yaniv. Journal of Molecular Biology. 1997;274(5):693-707.
64. Arnold M, Nath A, Hauber J, Kehlenbach RH. Multiple Importins Function as Nuclear Transport Receptors for the Rev Protein of Human Immunodeficiency Virus Type 1. Journal of Biological Chemistry. 2006;281(30):20883-90.
65. Szebeni A, Mehrotra B, Baumann A, Adam SA, Wingfield PT, Olson MOJ. Nucleolar Protein B23 Stimulates Nuclear Import of the HIV-1 Rev Protein and NLS-Conjugated Albumin. Biochemistry. 1997;36(13):3941-9.
66. Fankhauser C, Izaurralde E, Adachi Y, Wingfield P, Laemmli UK. Specific complex of human immunodeficiency virus type 1 rev and nucleolar B23 proteins: dissociation by the Rev response element. Mol Cell Biol. 1991;11(5):2567-75.
67. Malim MH, Tiley LS, McCarn DF, Rusche JR, Hauber J, Cullen BR. HIV-1 structural gene expression requires binding of the rev trans-activator to its RNA target sequence. Cell. 1990;60(4):675-83.
68. Tiley LS, Malim MH, Tewary HK, Stockley PG, Cullen BR. Identification of a high-affinity RNA-binding site for the human immunodeficiency virus type 1 Rev protein. Proceedings of the National Academy of Sciences. 1992;89(2):758.
69. Battiste JL, Mao H, Rao NS, Tan R, Muhandiram DR, Kay LE, et al. α Helix-RNA Major Groove Recognition in an HIV-1 Rev Peptide-RRE RNA Complex. Science. 1996;273(5281):1547-51.

70. Fisk GJ, Rusche JR, Cook KS, Daly TJ, Hauber J, Usman N. Characterization of HIV-1 REV protein: binding stoichiometry and minimal RNA substrate. Nucleic Acids Research. 1991;19(7):1577-83.
71. Heaphy S, Finch JT, Gait MJ, Karn J, Singh M. Human immunodeficiency virus type 1 regulator of virion expression, rev, forms nucleoprotein filaments after binding to a purine-rich "bubble" located within the rev-responsive region of viral mRNAs. Proceedings of the National Academy of Sciences of the United States of America. 1991;88(16):7366-70.
72. Kjems J, Brown M, Chang DD, Sharp PA. Structural analysis of the interaction between the human immunodeficiency virus Rev protein and the Rev response element. Proceedings of the National Academy of Sciences of the United States of America. 1991;88(3):683-7.
73. Daly TJ, Cook KS, Gray GS, Maione TE, Rusche JR. Specific binding of HIV-1 recombinant Rev protein to the Rev-responsive element in vitro. Nature. 1989;342(6251):816-9.
74. Zapp ML, Green MR. Sequence-specific RNA binding by the HIV-1 Rev protein. Nature. 1989;342(6250):714-6.
75. Karn J, Dingwall C, Finch JT, Heaphy S, Gait MJ. RNA binding by the tat and rev proteins of HIV-1. Biochimie. 1991;73(1):9-16.
76. Heaphy S, Dingwall C, Ernberg I, Gait MJ, Green SM, Kern J, et al. HIV-1 regulator of virion expression (Rev) protein binds to an RNA stem-loop structure located within the Rev response element region. Cell. 1990;60(4):685-93.
77. Cole JL, Gehman JD, Shafer JA, Kuo LC. Solution oligomerization of the rev protein of HIV-1: Implications for function. Biochemistry. 1993;32(44):11769-75.
78. Pond SJK, Ridgeway WK, Robertson R, Wang J, Millar DP. HIV-1 Rev protein assembles on viral RNA one molecule at a time. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(5):1404-8.
79. Jain C, Belasco JG. A Structural Model for the HIV-1 Rev–RRE Complex Deduced from Altered-Specificity Rev Variants Isolated by a Rapid Genetic Strategy. Cell. 1996;87(1):115-25.
80. Fang X, Wang J, O'Carroll Ina P, Mitchell M, Zuo X, Wang Y, et al. An Unusual Topological Structure of the HIV-1 Rev Response Element. Cell. 2013;155(3):594-605.
81. Daly TJ, Doten RC, Rennert P, Auer M, Jaksche H, Donner A, et al. Biochemical characterization of binding of multiple HIV-1 Rev monomeric proteins to the Rev responsive element. Biochemistry. 1993;32(39):10497-505.
82. Malim MH, Cullen BR. HIV-1 structural gene expression requires the binding of multiple Rev monomers to the viral RRE: Implications for HIV-1 latency. Cell. 1991;65(2):241-8.
83. Mann DA, Mikaélian I, Zemmel RW, Green SM, Lowe AD, Kimura T, et al. A Molecular Rheostat: Co-operative Rev Binding to Stem I of the Rev-response Element Modulates Human Immunodeficiency Virus Type-1 Late Gene Expression. Journal of Molecular Biology. 1994;241(2):193-207.
84. Robertson-Anderson RM, Wang J, Edgcomb SP, Carmel AB, Williamson JR, Millar DP. Single-molecule studies reveal that DEAD box protein DDX1 promotes oligomerization of HIV-1 Rev on the Rev response element. Journal of molecular biology. 2011;410(5):959-71.
85. Hammond JA, Lamichhane R, Millar DP, Williamson JR. A DEAD-Box Helicase Mediates an RNA Structural Transition in the HIV-1 Rev Response Element. Journal of Molecular Biology. 2017;429(5):697-714.

86. Zapp ML, Hope TJ, Parslow TG, Green MR. Oligomerization and RNA binding domains of the type 1 human immunodeficiency virus Rev protein: a dual function for an arginine-rich binding motif. Proceedings of the National Academy of Sciences of the United States of America. 1991;88(17):7734-8.
87. Olsen HS, Cochrane AW, Dillon PJ, Nalin CM, Rosen CA. Interaction of the human immunodeficiency virus type 1 Rev protein with a structured region in env mRNA is dependent on multimer formation mediated through a basic stretch of amino acids. Genes & Development. 1990;4(8):1357-64.
88. Daugherty MD, D'Orso I, Frankel AD. A solution to limited genomic capacity: using adaptable binding surfaces to assemble the functional HIV Rev oligomer on RNA. Molecular cell. 2008;31(6):824-34.
89. Carroll IP, Thappeta Y, Fan L, Ramirez-Valdez EA, Smith S, Wang Y-X, et al. Contributions of Individual Domains to Function of the HIV-1 Rev Response Element. Journal of Virology. 2017;91(21):e00746-17.
90. Watts NR, Eren E, Zhuang X, Wang Y-X, Steven AC, Wingfield PT. A new HIV-1 Rev structure optimizes interaction with target RNA (RRE) for nuclear export. Journal of Structural Biology. 2018;203(2):102-8.
91. Malim MH, Böhnlein S, Hauber J, Cullen BR. Functional dissection of the HIV-1 Rev trans-activator--derivation of a trans-dominant repressor of Rev function. Cell. 1989;58(1):205-14.
92. Mermer B, Felber BK, Campbell M, Pavlakis GN. Identification of trans-dominant HIV-1 rev protein mutants by direct transfer of bacterially produced proteins into human cells. Nucleic acids research. 1990;18(8):2037-44.
93. Huang XJ, Hope TJ, Bond BL, McDonald D, Grahl K, Parslow TG. Minimal Rev-response element for type 1 human immunodeficiency virus. Journal of virology. 1991;65(4):2131-4.
94. Fischer U, Huber J, Boelens WC, Mattajt LW, Lührmann R. The HIV-1 Rev Activation Domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. Cell. 1995;82(3):475-83.
95. Fornerod M, Ohno M, Yoshida M, Mattaj IW. CRM1 Is an Export Receptor for Leucine-Rich Nuclear Export Signals. Cell. 1997;90(6):1051-60.
96. Booth DS, Cheng Y, Frankel AD. The export receptor Crm1 forms a dimer to promote nuclear export of HIV RNA. eLife. 2014;3:e04121.
97. Güttler T, Madl T, Neumann P, Deichsel D, Corsini L, Monecke T, et al. NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. Nature Structural &Amp; Molecular Biology. 2010;17:1367.
98. Wang Y, Zhang H, Na L, Du C, Zhang Z, Zheng Y-H, et al. ANP32A and ANP32B are key factors in the Rev dependent CRM1 pathway for nuclear export of HIV-1 unspliced mRNA. Journal of Biological Chemistry. 2019.
99. Huang F, Zhang J, Zhang Y, Geng G, Liang J, Li Y, et al. RNA helicase MOV10 functions as a co-factor of HIV-1 Rev to facilitate Rev/RRE-dependent nuclear export of viral mRNAs. Virology. 2015;486:15-26.
100. Ajamian L, Abel K, Rao S, Vyboh K, García-de-Gracia F, Soto-Rifo R, et al. HIV-1 Recruits UPF1 but Excludes UPF2 to Promote Nucleocytoplasmic Export of the Genomic RNA. Biomolecules. 2015;5(4):2808-39.
101. Monette A, Panté N, Mouland AJ. HIV-1 remodels the nuclear pore complex. The Journal of cell biology. 2011;193(4):619-31.

102. Kula A, Guerra J, Knezevich A, Kleva D, Myers MP, Marcello A. Characterization of the HIV-1 RNA associated proteome identifies Matrin 3 as a nuclear cofactor of Rev function. Retrovirology. 2011;8:60-.
103. Hofmann W, Reichart B, Ewald A, Müller E, Schmitt I, Stauber RH, et al. Cofactor requirements for nuclear export of Rev response element (RRE)- and constitutive transport element (CTE)-containing retroviral RNAs. An unexpected role for actin. The Journal of cell biology. 2001;152(5):895-910.
104. Toro-Ascuy D, Rojas-Araya B, García-de-Gracia F, Rojas-Fuentes C, Pereira-Montecinos C, Gaete-Argel A, et al. A Rev-CBP80-eIF4AI complex drives Gag synthesis from the HIV-1 unspliced mRNA. Nucleic acids research. 2018;46(21):11539-52.
105. Edgcomb SP, Carmel AB, Naji S, Ambrus-Aikelin G, Reyes JR, Saphire ACS, et al. DDX1 is an RNA-dependent ATPase involved in HIV-1 Rev function and virus replication. Journal of molecular biology. 2012;415(1):61-74.
106. Taniguchi I, Mabuchi N, Ohno M. HIV-1 Rev protein specifies the viral RNA export pathway by suppressing TAP/NXF1 recruitment. Nucleic Acids Research. 2014;42(10):6645-58.
107. Schwartz S, Felber BK, Pavlakis GN. Distinct RNA sequences in the gag region of human immunodeficiency virus type 1 decrease RNA stability and inhibit expression in the absence of Rev protein. Journal of virology. 1992;66(1):150-9.
108. Cochrane AW, Jones KS, Beidas S, Dillon PJ, Skalka AM, Rosen CA. Identification and characterization of intragenic sequences which repress human immunodeficiency virus structural gene expression. Journal of virology. 1991;65(10):5305-13.
109. Felber BK, Hadzopoulou-Cladaras M, Cladaras C, Copeland T, Pavlakis GN. rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. Proceedings of the National Academy of Sciences of the United States of America. 1989;86(5):1495-9.
110. Hadzopoulou-Cladaras M, Felber BK, Cladaras C, Athanassopoulos A, Tse A, Pavlakis GN. The rev (trs/art) protein of human immunodeficiency virus type 1 affects viral mRNA and protein expression via a cis-acting sequence in the env region. Journal of virology. 1989;63(3):1265-74.
111. Nasioulas G, Zolotukhin AS, Tabernero C, Solomin L, Cunningham CP, Pavlakis GN, et al. Elements distinct from human immunodeficiency virus type 1 splice sites are responsible for the Rev dependence of env mRNA. Journal of virology. 1994;68(5):2986-93.
112. Maldarelli F, Martin MA, Strebel K. Identification of posttranscriptionally active inhibitory sequences in human immunodeficiency virus type 1 RNA: novel level of gene regulation. Journal of virology. 1991;65(11):5732-43.
113. Nguyen K-L, Llano M, Akari H, Miyagi E, Poeschla EM, Strebel K, et al. Codon optimization of the HIV-1 vpu and vif genes stabilizes their mRNA and allows for highly efficient Rev-independent expression. Virology. 2004;319(2):163-75.
114. van der Kuyl AC, Berkhout B. The biased nucleotide composition of the HIV genome: a constant factor in a highly variable virus. Retrovirology. 2012;9(1):92.
115. Haas J, Park E-C, Seed B. Codon usage limitation in the expression of HIV-1 envelope glycoprotein. Current Biology. 1996;6(3):315-24.
116. Schwartz S, Campbell M, Nasioulas G, Harrison J, Felber BK, Pavlakis GN. Mutational inactivation of an inhibitory sequence in human immunodeficiency virus type 1 results in Rev-independent gag expression. Journal of virology. 1992;66(12):7176-82.

117. Schneider R, Campbell M, Nasioulas G, Felber BK, Pavlakis GN. Inactivation of the human immunodeficiency virus type 1 inhibitory elements allows Rev-independent expression of Gag and Gag/protease and particle formation. Journal of virology. 1997;71(7):4892-903.
118. Anson DS, Dunning KR. Codon-optimized reading frames facilitate high-level expression of the HIV-1 minor proteins. Molecular Biotechnology. 2005;31(1):85-8.
119. Ngumbela KC, Ryan KP, Sivamurthy R, Brockman MA, Gandhi RT, Bhardwaj N, et al. Quantitative effect of suboptimal codon usage on translational efficiency of mRNA encoding HIV-1 gag in intact T cells. PloS one. 2008;3(6):e2356-e.
120. Kotsopoulou E, Kim VN, Kingsman AJ, Kingsman SM, Mitrophanous KA. A Rev-independent human immunodeficiency virus type 1 (HIV-1)-based vector that exploits a codon-optimized HIV-1 gag-pol gene. Journal of virology. 2000;74(10):4839-52.
121. Zolotukhin AS, Michalowski D, Bear J, Smulevitch SV, Traish AM, Peng R, et al. PSF acts through the human immunodeficiency virus type 1 mRNA instability elements to regulate virus expression. Molecular and cellular biology. 2003;23(18):6618-30.
122. Afonina E, Neumann M, Pavlakis GN. Preferential Binding of Poly(A)-binding Protein 1 to an Inhibitory RNA Element in the Human Immunodeficiency Virus Type 1 gag mRNA. Journal of Biological Chemistry. 1997;272(4):2307-11.
123. Black AC, Luo J, Watanabe C, Chun S, Bakker A, Fraser JK, et al. Polypyrimidine tract-binding protein and heterogeneous nuclear ribonucleoprotein A1 bind to human T-cell leukemia virus type 2 RNA regulatory elements. Journal of virology. 1995;69(11):6852-8.
124. Najera I, Krieg M, Karn J. Synergistic stimulation of HIV-1 rev-dependent export of unspliced mRNA to the cytoplasm by hnRNP A111Edited by A. R. Fersht. Journal of Molecular Biology. 1999;285(5):1951-64.
125. Valiente-Echeverría F, Vallejos M, Monette A, Pino K, Letelier A, Huidobro-Toro JP, et al. A cis-acting element present within the Gag open reading frame negatively impacts on the activity of the HIV-1 IRES. PloS one. 2013;8(2):e56962-e.
126. Kjems J, Frankel AD, Sharp PA. Specific regulation of mRNA splicing in vitro by a peptide from HIV-1 Rev. Cell. 1991;67(1):169-78.
127. Kjems J, Sharp PA. The basic domain of Rev from human immunodeficiency virus type 1 specifically blocks the entry of U4/U6.U5 small nuclear ribonucleoprotein in spliceosome assembly. Journal of virology. 1993;67(8):4769-76.
128. Xiao H, Wyler E, Milek M, Grewe B, Kirchner P, Ekici A, et al. CRNKL1 is a highly selective regulator of intron-retaining HIV-1 and cellular mRNAs. bioRxiv. 2020:2020.02.04.934927.
129. Powell DM, Amaral MC, Wu JY, Maniatis T, Greene WC. HIV Rev-dependent binding of SF2/ASF to the Rev response element: possible role in Rev-mediated inhibition of HIV RNA splicing. Proceedings of the National Academy of Sciences of the United States of America. 1997;94(3):973-8.
130. Ropers D, Ayadi L, Gattoni R, Jacquenet S, Damier L, Branlant C, et al. 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. 2004;279(29):29963-73.
131. Jablonski JA, Caputi M. Role of cellular RNA processing factors in human immunodeficiency virus type 1 mRNA metabolism, replication, and infectivity. Journal of virology. 2009;83(2):981-92.
132. Jacquenet S, Decimo D, Muriaux D, Darlix J-L. Dual effect of the SR proteins ASF/SF2, SC35 and 9G8 on HIV-1 RNA splicing and virion production. Retrovirology. 2005;2:33-.

133. Sherrill-Mix S, Ocwieja KE, Bushman FD. Gene activity in primary T cells infected with HIV89.6: intron retention and induction of genomic repeats. Retrovirology. 2015;12:79-.
134. Knight DM, Flomerfelt FA, Ghrayeb J. Expression of the art/trs protein of HIV and study of its role in viral envelope synthesis. Science. 1987;236(4803):837.
135. D'Agostino DM, Felber BK, Harrison JE, Pavlakis GN. The Rev protein of human immunodeficiency virus type 1 promotes polysomal association and translation of gag/pol and vpu/env mRNAs. Molecular and cellular biology. 1992;12(3):1375-86.
136. Arrigo SJ, Chen IS. Rev is necessary for translation but not cytoplasmic accumulation of HIV-1 vif, vpr, and env/vpu 2 RNAs. Genes & Development. 1991;5(5):808-19.
137. Perales C, Carrasco L, González ME. Regulation of HIV-1 env mRNA translation by Rev protein. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. 2005;1743(1):169-75.
138. Emerman M, Vazeux R, Peden K. The rev gene product of the human immunodeficiency virus affects envelope-specific RNA localization. Cell. 1989;57(7):1155-65.
139. Yi R, Bogerd HP, Cullen BR. Recruitment of the Crm1 nuclear export factor is sufficient to induce cytoplasmic expression of incompletely spliced human immunodeficiency virus mRNAs. Journal of virology. 2002;76(5):2036-42.
140. Gallego J, Greatorex J, Zhang H, Yang B, Arunachalam S, Fang J, et al. Rev Binds Specifically to a Purine Loop in the SL1 Region of the HIV-1 Leader RNA. Journal of Biological Chemistry. 2003;278(41):40385-91.
141. Groom HCT, Anderson EC, Dangerfield JA, Lever AML. Rev regulates translation of human immunodeficiency virus type 1 RNAs. Journal of General Virology. 2009;90(5):1141.
142. Greatorex JS, Palmer EA, Pomerantz RJ, Dangerfield JA, Lever AML. Mutation of the Rev-binding loop in the human immunodeficiency virus 1 leader causes a replication defect characterized by altered RNA trafficking and packaging. Journal of General Virology. 2006;87(10):3039.
143. Berkowitz RD, Hammarskjöld M-L, Helga-Maria C, Rekosh D, Goff SP. 5’ Regions of HIV-1 RNAs Are Not Sufficient for Encapsidation: Implications for the HIV-1 Packaging Signal. Virology. 1995;212(2):718-23.
144. Richardson JH, Child LA, Lever AM. Packaging of human immunodeficiency virus type 1 RNA requires cis-acting sequences outside the 5' leader region. Journal of Virology. 1993;67(7):3997.
145. Anson DS, Fuller M. Rational development of a HIV-1 gene therapy vector. The Journal of Gene Medicine. 2003;5(10):829-38.
146. Kharytonchyk S, Brown JD, Stilger K, Yasin S, Iyer AS, Collins J, et al. Influence of gag and RRE Sequences on HIV-1 RNA Packaging Signal Structure and Function. Journal of molecular biology. 2018;430(14):2066-79.
147. Brandt S, Blißenbach M, Grewe B, Konietzny R, Grunwald T, Überla K. Rev Proteins of Human and Simian Immunodeficiency Virus Enhance RNA Encapsidation. PLOS Pathogens. 2007;3(4):e54.
148. Blissenbach M, Grewe B, Hoffmann B, Brandt S, Überla K. Nuclear RNA Export and Packaging Functions of HIV-1 Rev Revisited. Journal of Virology. 2010;84(13):6598.
149. Ma J, Rong L, Zhou Y, Roy BB, Lu J, Abrahamyan L, et al. The requirement of the DEAD-box protein DDX24 for the packaging of human immunodeficiency virus type 1 RNA. Virology. 2008;375(1):253-64.
150. Bonyhadi ML, Moss K, Voytovich A, Auten J, Kalfoglou C, Plavec I, et al. RevM10-expressing T cells derived in vivo from transduced human hematopoietic stem-progenitor

cells inhibit human immunodeficiency virus replication. Journal of virology. 1997;71(6):4707-16.
151. Ranga U, Woffendin C, Verma S, Xu L, June CH, Bishop DK, et al. Enhanced T cell engraftment after retroviral delivery of an antiviral gene in HIV-infected individuals. Proceedings of the National Academy of Sciences of the United States of America. 1998;95(3):1201-6.
152. Chan SY, Louie MC, Piccotti JR, Iyer G, Ling X, Yang Z-Y, et al. Genetic Vaccination-Induced Immune Responses to the Human Immunodeficiency Virus Protein Rev: Emergence of the Interleukin 2-Producing Helper T Lymphocyte. Human Gene Therapy. 1998;9(15):2187-96.
153. Bahner I, Sumiyoshi T, Kagoda M, Swartout R, Peterson D, Pepper K, et al. Lentiviral Vector Transduction of a Dominant-negative Rev Gene Into Human CD34⁺ Hematopoietic Progenitor Cells Potently Inhibits Human Immunodeficiency Virus-1 Replication. Molecular Therapy. 2007;15(1):76-85.
154. Ragheb JA, Bressler P, Daucher M, Chiang L, Chuah MKL, Vandendriessche T, et al. Analysis of trans-Dominant Mutants of the HIV Type 1 Rev Protein for Their Ability to Inhibit Rev Function, HIV Type 1 Replication, and Their Use as Anti-HIV Gene Therapeutics. AIDS Research and Human Retroviruses. 1995;11(11):1343-53.
155. Hamm TE, Rekosh D, Hammarskjöld M-L. Selection and Characterization of Human Immunodeficiency Virus Type 1 Mutants That Are Resistant to Inhibition by the Transdominant Negative RevM10 Protein. Journal of Virology. 1999;73(7):5741.
156. Werstuck G, Zapp ML, Green MR. A non-canonical base pair within the human immunodeficiency virus Rev-responsive element is involved in both Rev and small molecule recognition. Chemistry & Biology. 1996;3(2):129-37.
157. Zapp ML, Stern S, Green MR. Small molecules that selectively block RNA binding of HIV-1 rev protein inhibit rev function and viral production. Cell. 1993;74(6):969-78.
158. Matsukura M, Zon G, Shinozuka K, Robert-Guroff M, Shimada T, Stein CA, et al. Regulation of viral expression of human immunodeficiency virus in vitro by an antisense phosphorothioate oligodeoxynucleotide against rev (art/trs) in chronically infected cells. Proceedings of the National Academy of Sciences of the United States of America. 1989;86(11):4244-8.
159. Nakaya T, Iwai S, Fujinaga K, Sato Y, Otsuka E, Ikuta K. Decoy approach using RNA-DNA chimera oligonucleotides to inhibit the regulatory function of human immunodeficiency virus type 1 Rev protein. Antimicrob Agents Chemother. 1997;41(2):319-25.
160. Ratmeyer L, Zapp ML, Green MR, Vinayak R, Kumar A, Boykin DW, et al. Inhibition of HIV-1 Rev−RRE Interaction by Diphenylfuran Derivatives. Biochemistry. 1996;35(42):13689-96.
161. Xiao G, Kumar A, Li K, Rigl CT, Bajic M, Davis TM, et al. Inhibition of the HIV-1 rev–RRE complex formation by unfused aromatic cations. Bioorganic & Medicinal Chemistry. 2001;9(5):1097-113.
162. Good PD, Krikos AJ, Li SXL, Bertrand E, Lee NS, Giver L, et al. Expression of small, therapeutic RNAs in human cell nuclei. Gene Therapy. 1997;4(1):45-54.
163. Dearborn AD, Eren E, Watts NR, Palmer IW, Kaufman JD, Steven AC, et al. Structure of an RNA Aptamer that Can Inhibit HIV-1 by Blocking Rev-Cognate RNA (RRE) Binding and Rev-Rev Association. Structure. 2018;26(9):1187-95.e4.

164. Jin Y, Cowan JA. Targeted Cleavage of HIV Rev Response Element RNA by Metallopeptide Complexes. Journal of the American Chemical Society. 2006;128(2):410-1.
165. Jin Y, Cowan JA. Cellular activity of Rev response element RNA targeting metallopeptides. JBIC Journal of Biological Inorganic Chemistry. 2007;12(5):637-44.
166. Prado S, Beltrán M, Coiras M, Bedoya LM, Alcamí J, Gallego J. Bioavailable inhibitors of HIV-1 RNA biogenesis identified through a Rev-based screen. Biochemical Pharmacology. 2016;107:14-28.
167. Zhang Q, Harada K, Cho HS, Frankel AD, Wemmer DE. Structural characterization of the complex of the Rev response element RNA with a selected peptide. Chemistry & Biology. 2001;8(5):511-20.
168. Mills NL, Daugherty MD, Frankel AD, Guy RK. An α-Helical Peptidomimetic Inhibitor of the HIV-1 Rev−RRE Interaction. Journal of the American Chemical Society. 2006;128(11):3496-7.
169. McColl DJ, Honchell CD, Frankel AD. Structure-based design of an RNA-binding zinc finger. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(17):9521-6.
170. Dai Y, Peralta AN, Wynn JE, Sherpa C, Li H, Verma A, et al. Molecular recognition of a branched peptide with HIV-1 Rev Response Element (RRE) RNA. Bioorganic & Medicinal Chemistry. 2019;27(8):1759-65.
171. Prado S, Beltrán M, Moreno Á, Bedoya LM, Alcamí J, Gallego J. A small-molecule inhibitor of HIV-1 Rev function detected by a diversity screen based on RRE-Rev interference. Biochemical Pharmacology. 2018;156:68-77.
172. Wong RW, Balachandran A, Haaland M, Stoilov P, Cochrane A. Characterization of novel inhibitors of HIV-1 replication that function via alteration of viral RNA processing and rev function. Nucleic acids research. 2013;41(20):9471-83.
173. Shuck-Lee D, Chang H, Sloan EA, Hammarskjold M-L, Rekosh D. Single-nucleotide changes in the HIV Rev-response element mediate resistance to compounds that inhibit Rev function. Journal of virology. 2011;85(8):3940-9.
174. Kudo N, Matsumori N, Taoka H, Fujiwara D, Schreiner EP, Wolff B, et al. Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(16):9112-7.
175. Wolff B, Sanglier J-J, Wang Y. Leptomycin B is an inhibitor of nuclear export: inhibition of nucleo-cytoplasmic translocation of the human immunodeficiency virus type 1 (HIV-1) Rev protein and Rev-dependent mRNA. Chemistry & Biology. 1997;4(2):139-47.
176. Okamura M, Inose H, Masuda S. RNA Export through the NPC in Eukaryotes. Genes. 2015;6(1):124-49.
177. Boons E, Vanstreels E, Jacquemyn M, Nogueira TC, Neggers JE, Vercruysse T, et al. Human Exportin-1 is a Target for Combined Therapy of HIV and AIDS Related Lymphoma. EBioMedicine. 2015;2(9):1102-13.
178. Campos N, Myburgh R, Garcel A, Vautrin A, Lapasset L, Nadal ES, et al. Long lasting control of viral rebound with a new drug ABX464 targeting Rev - mediated viral RNA biogenesis. Retrovirology. 2015;12:30-.
179. Díaz-Muñoz MD, Turner M. Uncovering the Role of RNA-Binding Proteins in Gene Expression in the Immune System. Frontiers in immunology. 2018;9:1094-.
180. Plaschka C, Lin P-C, Nagai K. Structure of a pre-catalytic spliceosome. Nature. 2017;546(7660):617-21.

181. Behrmann E, Loerke J, Budkevich TV, Yamamoto K, Schmidt A, Penczek PA, et al. Structural snapshots of actively translating human ribosomes. Cell. 2015;161(4):845-57.
182. Glisovic T, Bachorik Jl Fau - Yong J, Yong J Fau - Dreyfuss G, Dreyfuss G. RNA-binding proteins and post-transcriptional gene regulation. (0014-5793 (Print)).
183. Castello A, Frese CK, Fischer B, Järvelin AI, Horos R, Alleaume A-M, et al. Identification of RNA-binding domains of RNA-binding proteins in cultured cells on a system-wide scale with RBDmap. Nature Protocols. 2017;12:2447.
184. Hentze MW, Castello A, Schwarzl T, Preiss T. A brave new world of RNA-binding proteins. Nature Reviews Molecular Cell Biology. 2018;19(5):327-41.
185. Castello A, Horos R, Strein C, Fischer B, Eichelbaum K, Steinmetz LM, et al. System-wide identification of RNA-binding proteins by interactome capture. Nature Protocols. 2013;8(3):491-500.
186. Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann Benedikt M, Strein C, et al. Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. Cell. 2012;149(6):1393-406.
187. Castello A, Fischer B, Frese CK, Horos R, Alleaume A-M, Foehr S, et al. Comprehensive Identification of RNA-Binding Domains in Human Cells. Molecular cell. 2016;63(4):696-710.
188. Järvelin AI, Noerenberg M, Davis I, Castello A. The new (dis)order in RNA regulation. Cell communication and signaling : CCS. 2016;14:9-.
189. Castello A, Hentze MW, Preiss T. Metabolic Enzymes Enjoying New Partnerships as RNA-Binding Proteins. Trends in Endocrinology & Metabolism. 2015;26(12):746-57.
190. Iselin L, Palmalux N, Kamel W, Simmonds P, Mohammed S, Castello A. Uncovering viral RNA–host cell interactions on a proteome-wide scale. Trends in Biochemical Sciences. 2021.
191. Garcia-Moreno M, Noerenberg M, Ni S, Järvelin AI, González-Almela E, Lenz CE, et al. System-wide Profiling of RNA-Binding Proteins Uncovers Key Regulators of Virus Infection. Molecular Cell. 2019;74(1):196-211.e11.
192. Kamel W, Noerenberg M, Cerikan B, Chen H, Järvelin AI, Kammoun M, et al. Global analysis of protein-RNA interactions in SARS-CoV-2-infected cells reveals key regulators of infection. Molecular Cell. 2021;81(13):2851-67.e7.
193. Chapman EG, Costantino DA, Rabe JL, Moon SL, Wilusz J, Nix JC, et al. The structural basis of pathogenic subgenomic flavivirus RNA (sfRNA) production. Science (New York, NY). 2014;344(6181):307-10.
194. Gunnery S, Green SR, Mathews MB. Tat-responsive region RNA of human immunodeficiency virus type 1 stimulates protein synthesis in vivo and in vitro: relationship between structure and function. Proceedings of the National Academy of Sciences of the United States of America. 1992;89(23):11557-61.
195. Knoener R, Evans E, 3rd, Becker JT, Scalf M, Benner B, Sherer NM, et al. Identification of host proteins differentially associated with HIV-1 RNA splice variants. eLife. 2021;10:e62470.
196. Knoener RA, Becker JT, Scalf M, Sherer NM, Smith LM. Elucidating the in vivo interactome of HIV-1 RNA by hybridization capture and mass spectrometry. Scientific reports. 2017;7(1):16965-.
197. Kutner RH, Zhang X-Y, Reiser J. Production, concentration and titration of pseudotyped HIV-1-based lentiviral vectors. Nature Protocols. 2009;4:495.

198. Wiśniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. Nature Methods. 2009;6(5):359-62.
199. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature Biotechnology. 2008;26(12):1367-72.
200. Lazwr C. imputeLCMD: A collection of methods for left-censored missing data imputation. 2015.
201. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research. 2015;43(7):e47-e.
202. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nature protocols. 2009;4(8):1184-91.
203. Smedley D, Haider S, Ballester B, Holland R, London D, Thorisson G, et al. BioMart – biological queries made easy. BMC Genomics. 2009;10(1):22.
204. Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics. 2017;33(18):2938-40.
205. Pages H, Carlson M, Falcon S, Li N. *AnnotationDbi: Manipulation of SQLite-based annotations in Bioconductor*. 2020.
206. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC bioinformatics. 2013;14:128-.
207. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic acids research. 2019;47(D1):D607-D13.
208. M C. *org.Hs.eg.db: Genome wide annotation for Human*. 2019.
209. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, et al. The Pfam protein families database. Nucleic acids research. 2012;40(Database issue):D290-D301.
210. Freshour SL, Kiwala S, Cotto KC, Coffman AC, McMichael JF, Song JJ, et al. Integration of the Drug–Gene Interaction Database (DGIdb 4.0) with open crowdsource efforts. Nucleic Acids Research. 2021;49(D1):D1144-D51.
211. Truman CT-S, Järvelin A, Davis I, Castello A. HIV Rev-isited. Open Biology.10(12):200320.
212. Williamson JR. Really exasperating viral protein from HIV. eLife. 2015;4:e05169.
213. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, et al. Alternative isoform regulation in human tissue transcriptomes. Nature. 2008;456(7221):470-6.
214. Garcia-Moreno M, Noerenberg M, Ni S, Järvelin AI, González-Almela E, Lenz C, et al. Understanding RNP remodelling uncovers RBPs functionally required for viral replication. bioRxiv. 2018:350686.
215. Deshiere A, Joly-Beauparlant C, Breton Y, Ouellet M, Raymond F, Lodge R, et al. Global Mapping of the Macrophage-HIV-1 Transcriptome Reveals that Productive Infection Induces Remodeling of Host Cell DNA and Chromatin. Scientific reports. 2017;7(1):5238-.
216. DeBoer J, Wojtkiewicz MS, Haverland N, Li Y, Harwood E, Leshen E, et al. Proteomic profiling of HIV-infected T-cells by SWATH mass spectrometry. Virology. 2018;516:246-57.
217. Naamati A, Williamson JC, Greenwood EJD, Marelli S, Lehner PJ, Matheson NJ. Functional proteomic atlas of HIV infection in primary human CD4+ T cells. eLife. 2019;8:e41431.

218. Hamers-Casterman C, Atarhouch T, Muyldermans S, Robinson G, Hammers C, Songa EB, et al. Naturally occurring antibodies devoid of light chains. Nature. 1993;363(6428):446-8.
219. England CG, Luo H, Cai W. HaloTag technology: a versatile platform for biomedical applications. Bioconjugate chemistry. 2015;26(6):975-86.
220. Los GV, Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, et al. HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. ACS Chemical Biology. 2008;3(6):373-82.
221. Zolotukhin AS, Valentin A, Pavlakis GN, Felber BK. Continuous propagation of RRE(-) and Rev(-)RRE(-) human immunodeficiency virus type 1 molecular clones containing a cis-acting element of simian retrovirus type 1 in human peripheral blood lymphocytes. Journal of virology. 1994;68(12):7944-52.
222. Nappi F, Schneider R, Zolotukhin A, Smulevitch S, Michalowski D, Bear J, et al. Identification of a novel posttranscriptional regulatory element by using a rev- and RRE-mutated human immunodeficiency virus type 1 DNA proviral clone as a molecular trap. Journal of virology. 2001;75(10):4558-69.
223. Dejucq N, Simmons G, Clapham PR. Expanded tropism of primary human immunodeficiency virus type 1 R5 strains to CD4(+) T-cell lines determined by the capacity to exploit low concentrations of CCR5. Journal of virology. 1999;73(9):7842-7.
224. Berkowitz RD, Alexander S, Bare C, Linquist-Stepps V, Bogan M, Moreno ME, et al. CCR5- and CXCR4-utilizing strains of human immunodeficiency virus type 1 exhibit differential tropism and pathogenesis in vivo. Journal of virology. 1998;72(12):10108-17.
225. Erickson-Viitanen S, Manfredi J, Viitanen P, Tribe DE, Tritch R, Hutchison CA, et al. Cleavage of HIV-1 gag Polyprotein Synthesized In Vitro: Sequential Cleavage by the Viral Protease. AIDS Research and Human Retroviruses. 1989;5(6):577-91.
226. Fang J, Kubota S, Yang B, Zhou N, Zhang H, Godbout R, et al. A DEAD box protein facilitates HIV-1 replication as a cellular co-factor of Rev. Virology. 2004;330(2):471-80.
227. Cochrane AW, Perkins A, Rosen CA. Identification of sequences important in the nucleolar localization of human immunodeficiency virus Rev: relevance of nucleolar localization to function. Journal of virology. 1990;64(2):881-5.
228. Arizala JAC, Takahashi M, Burnett JC, Ouellet DL, Li H, Rossi JJ. Nucleolar Localization of HIV-1 Rev Is Required, Yet Insufficient for Production of Infectious Viral Particles. AIDS research and human retroviruses. 2018;34(11):961-81.
229. Reddy TR, Xu W, Mau JKL, Goodwin CD, Suhasini M, Tang H, et al. Inhibition of HIV replication by dominant negative mutants of Sam68, a functional homolog of HIV-1 Rev. Nature Medicine. 1999;5(6):635-42.
230. Zhou X, Luo J, Mills L, Wu S, Pan T, Geng G, et al. DDX5 Facilitates HIV-1 Replication as a Cellular Co-Factor of Rev. PLOS ONE. 2013;8(5):e65040.
231. Chertova E, Chertov O, Coren LV, Roser JD, Trubey CM, Bess JW, Jr., et al. Proteomic and biochemical analysis of purified human immunodeficiency virus type 1 produced from infected monocyte-derived macrophages. Journal of virology. 2006;80(18):9039-52.
232. Ott DE, Coren LV, Kane BP, Busch LK, Johnson DG, Sowder RC, 2nd, et al. Cytoskeletal proteins inside human immunodeficiency virus type 1 virions. Journal of virology. 1996;70(11):7734-43.
233. Fukumori T, Kagawa S, Iida S, Oshima Y, Akari H, Koyama AH, et al. Rev-dependent expression of three species of HIV-1 mRNAs (review). Int J Mol Med. 1999;3(3):297-599.

234. Pastor DM, Poritz LS, Olson TL, Kline CL, Harris LR, Koltun WA, et al. Primary cell lines: false representation or model system? a comparison of four human colorectal tumors and their coordinately established cell lines. International journal of clinical and experimental medicine. 2010;3(1):69-83.
235. Adachi A, Gendelman HE, Koenig S, Folks T, Willey R, Rabson A, et al. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. Journal of virology. 1986;59(2):284-91.
236. Wei BL, Denton PW, O'Neill E, Luo T, Foster JL, Garcia JV. Inhibition of lysosome and proteasome function enhances human immunodeficiency virus type 1 infection. Journal of virology. 2005;79(9):5705-12.
237. Aiken C. 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. 1997;71(8):5871-7.
238. Yu D, Wang W, Yoder A, Spear M, Wu Y. The HIV Envelope but Not VSV Glycoprotein Is Capable of Mediating HIV Latent Infection of Resting CD4 T Cells. PLOS Pathogens. 2009;5(10):e1000633.
239. Stutz F, Izaurralde E, Mattaj IW, Rosbash M. A role for nucleoporin FG repeat domains in export of human immunodeficiency virus type 1 Rev protein and RNA from the nucleus. Molecular and cellular biology. 1996;16(12):7144-50.
240. Behrens RT, Aligeti M, Pocock GM, Higgins CA, Sherer NM. Nuclear Export Signal Masking Regulates HIV-1 Rev Trafficking and Viral RNA Nuclear Export. Journal of virology. 2017;91(3):e02107-16.
241. Daelemans D, Costes S, Cho HE, Erwin-Cohen AR, Lockett S, Pavlakis NG. In Vivo HIV-1 Rev Multimerization in the Nucleolus and Cytoplasm Identified by Fluorescence Resonance Energy Transfer. The Journal of Biological Chemistry2004. p. 50167-75.
242. Dundr M, Leno GH, Hammarskjold ML, Rekosh D, Helga-Maria C, Olson MO. The roles of nucleolar structure and function in the subcellular location of the HIV-1 Rev protein. Journal of Cell Science. 1995;108(8):2811-23.
243. Shi XP, Yin KC, Waxman L. Effects of Inhibitors of RNA and Protein Synthesis on the Subcellular Distribution of the Eukaryotic Translation Initiation Factor, eIF-5A, and the HIV-1 Rev Protein. Neurosignals. 1997;6(3):143-9.
244. Miyazaki Y, Takamatsu T, Nosaka T, Fujita S, Martin TE, Hatanaka M. The Cytotoxicity of Human Immunodeficiency Virus Type 1 Rev: Implications for Its Interaction with the Nucleolar Protein B23. Experimental Cell Research. 1995;219(1):93-101.
245. Selvin PR, Lougheed T, Tonks Hoffman M, Park H, Balci H, Blehm BH, et al. Fluorescence Imaging with One-Nanometer Accuracy (FIONA). Cold Spring Harbor Protocols. 2007;2007(10):pdb.top27.
246. Bolinger C, Boris-Lawrie K. Mechanisms employed by retroviruses to exploit host factors for translational control of a complicated proteome. Retrovirology. 2009;6(1):8.
247. Freed EO. HIV-1 and the host cell: an intimate association. Trends in Microbiology. 2004;12(4):170-7.
248. Luo Y, Jacobs EY, Greco TM, Mohammed KD, Tong T, Keegan S, et al. HIV–host interactome revealed directly from infected cells. Nature Microbiology. 2016;1:16068.
249. Naji S, Ambrus G, Cimermančič P, Reyes JR, Johnson JR, Filbrandt R, et al. Host cell interactome of HIV-1 Rev includes RNA helicases involved in multiple facets of virus production. Molecular & cellular proteomics : MCP. 2012;11(4):M111.015313-M111.

250. Gautier VW, Gu L, O'Donoghue N, Pennington S, Sheehy N, Hall WW. In vitro nuclear interactome of the HIV-1 Tat protein. Retrovirology. 2009;6(1):47.
251. Engeland CE, Brown NP, Börner K, Schümann M, Krause E, Kaderali L, et al. Proteome analysis of the HIV-1 Gag interactome. Virology. 2014;460-461:194-206.
252. Jäger S, Cimermancic P, Gulbahce N, Johnson JR, McGovern KE, Clarke SC, et al. Global landscape of HIV–human protein complexes. Nature. 2011;481:365.
253. Au - Arizala JAC, Au - Chomchan P, Au - Li H, Au - Moore R, Au - Ge H, Au - Ouellet DL, et al. Identification of Nucleolar Factors During HIV-1 Replication Through Rev Immunoprecipitation and Mass Spectrometry. JoVE. 2019(148):e59329.
254. Cochrane A, Murley LL, Gao M, Wong R, Clayton K, Brufatto N, et al. Stable complex formation between HIV Rev and the nucleosome assembly protein, NAP1, affects Rev function. Virology. 2009;388(1):103-11.
255. Chu C-C, Plangger R, Kreutz C, Al-Hashimi HM. Dynamic ensemble of HIV-1 RRE stem IIB reveals non-native conformations that disrupt the Rev binding site. bioRxiv. 2018:498907.
256. Bevec D, Hauber J. Eukaryotic Initiation Factor 5A Activity and HIV-1 Rev Function. Neurosignals. 1997;6(3):124-33.
257. Modem S, Badri KR, Holland TC, Reddy TR. Sam68 is absolutely required for Rev function and HIV-1 production. Nucleic acids research. 2005;33(3):873-9.
258. Modem Suhasini and Thipparthi RR. Cellular Proteins and HIV-1 Rev Function. Current HIV Research. 2009;7(1):91-100.
259. Ruhl M, Himmelspach M, Bahr GM, Hammerschmid F, Jaksche H, Wolff B, et al. Eukaryotic initiation factor 5A is a cellular target of the human immunodeficiency virus type 1 Rev activation domain mediating trans-activation. J Cell Biol. 1993;123(6):1309.
260. Bevec D, Jaksche H, Oft M, Wöhl T, Himmelspach M, Pacher A, et al. Inhibition of HIV-1 Replication in Lymphocytes by Mutants of the Rev Cofactor eIF-5A. Science. 1996;271(5257):1858.
261. Rosorius O, Reichart B, Kratzer F, Heger P, Dabauvalle MC, Hauber J. Nuclear pore localization and nucleocytoplasmic transport of eIF-5A: evidence for direct interaction with the export receptor CRM1. Journal of Cell Science. 1999;112(14):2369.
262. Bogerd H, Fridell RA, Madore SJ, Cullen B. IDENTIFICATION OF A HUMAN COFACTOR FOR THE REV/REX CLASS OF RETROVIRAL REGULATORY PROTEINS. AIDS Research and Human Retroviruses. 1995;11:S86-S.
263. Li J, Liu Y, Park I-W, He JJ. Expression of Exogenous Sam68, the 68-Kilodalton Src-Associated Protein in Mitosis, Is Able To Alleviate Impaired Rev Function in Astrocytes. Journal of Virology. 2002;76(9):4526.
264. Tsai Y-C, Tsai T-H, Chang C-P, Chen S-F, Lee Y-M, Shyue S-K. Linear correlation between average fluorescence intensity of green fluorescent protein and the multiplicity of infection of recombinant adenovirus. Journal of biomedical science. 2015;22(1):31-.
265. Greenwood EJD, Matheson NJ, Wals K, van den Boomen DJH, Antrobus R, Williamson JC, et al. Temporal proteomic analysis of HIV infection reveals remodelling of the host phosphoproteome by lentiviral Vif variants. eLife. 2016;5:e18296.
266. Ivanchenko S, Godinez WJ, Lampe M, Kräusslich H-G, Eils R, Rohr K, et al. Dynamics of HIV-1 Assembly and Release. PLOS Pathogens. 2009;5(11):e1000652.
267. Ako-Adjei D, Fu W, Wallin C, Katz KS, Song G, Darji D, et al. HIV-1, human interaction database: current status and new features. Nucleic acids research. 2015;43(Database issue):D566-D70.

268. Miyazaki Y, Nosaka T, Hatanaka M. The post-transcriptional regulator Rev of HIV: Implications for its interaction with the nucleolar protein B23. Biochimie. 1996;78(11):1081-6.
269. Szebeni A, Mehrotra B, Baumann A, Adam SA, Wingfield PT, Olson MO. Nucleolar protein B23 stimulates nuclear import of the HIV-1 Rev protein and NLS-conjugated albumin. Biochemistry. 1997;36(13):3941-9.
270. Fankhauser C, Izaurralde E, Adachi Y, Wingfield P, Laemmli UK. Specific complex of human immunodeficiency virus type 1 rev and nucleolar B23 proteins: dissociation by the Rev response element. Molecular and cellular biology. 1991;11(5):2567-75.
271. Yasuda-Inoue M, Kuroki M, Ariumi Y. Distinct DDX DEAD-box RNA helicases cooperate to modulate the HIV-1 Rev function. Biochemical and Biophysical Research Communications. 2013;434(4):803-8.
272. Bolinger C, Sharma A, Singh D, Yu L, Boris-Lawrie K. RNA helicase A modulates translation of HIV-1 and infectivity of progeny virions. Nucleic Acids Research. 2010;38(5):1686-96.
273. Kutluay SB, Emery A, Penumutchu SR, Townsend D, Tenneti K, Madison MK, et al. Genome-Wide Analysis of Heterogeneous Nuclear Ribonucleoprotein (hnRNP) Binding to HIV-1 RNA Reveals a Key Role for hnRNP H1 in Alternative Viral mRNA Splicing. Journal of Virology. 2019;93(21):e01048-19.
274. Hadian K, Vincendeau M, Mäusbacher N, Nagel D, Hauck SM, Ueffing M, et al. Identification of a heterogeneous nuclear ribonucleoprotein-recognition region in the HIV Rev protein. The Journal of biological chemistry. 2009;284(48):33384-91.
275. Mahiet C, Swanson CM. Control of HIV-1 gene expression by SR proteins. Biochemical Society Transactions. 2016;44(5):1417-25.
276. Maris C, Dominguez C, Allain FHT. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. The FEBS Journal. 2005;272(9):2118-31.
277. Garcia-Moreno M, Järvelin AI, Castello A. Unconventional RNA-binding proteins step into the virus–host battlefront. WIREs RNA. 2018;9(6):e1498.
278. Alfano C, Sanfelice D, Babon J, Kelly G, Jacks A, Curry S, et al. Structural analysis of cooperative RNA binding by the La motif and central RRM domain of human La protein. Nature Structural & Molecular Biology. 2004;11(4):323-9.
279. Keene JD. Posttranscriptional Generation of Macromolecular Complexes. Molecular Cell. 2003;12(6):1347-9.
280. Cumberworth A, Lamour G, Babu MM, Gsponer J. Promiscuity as a functional trait: intrinsically disordered regions as central players of interactomes. Biochemical Journal. 2013;454(3):361-9.
281. Jang Y, Elsayed Z, Eki R, He S, Du K-P, Abbas T, et al. Intrinsically disordered protein RBM14 plays a role in generation of RNA:DNA hybrids at double-strand break sites. Proceedings of the National Academy of Sciences. 2020;117(10):5329.
282. Arandjelovic S, Perry JSA, Lucas CD, Penberthy KK, Kim T-H, Zhou M, et al. A noncanonical role for the engulfment gene ELMO1 in neutrophils that promotes inflammatory arthritis. Nature Immunology. 2019;20(2):141-51.
283. Janardhan A, Swigut T, Hill B, Myers MP, Skowronski J. HIV-1 Nef binds the DOCK2-ELMO1 complex to activate rac and inhibit lymphocyte chemotaxis. PLoS biology. 2004;2(1):E6-E.

284. Toret CP, Collins C, Nelson WJ. An Elmo-Dock complex locally controls Rho GTPases and actin remodeling during cadherin-mediated adhesion. The Journal of cell biology. 2014;207(5):577-87.
285. Mauldin Joshua P, Lu M, Das S, Park D, Ernst Peter B, Ravichandran Kodi S. A Link between the Cytoplasmic Engulfment Protein Elmo1 and the Mediator Complex Subunit Med31. Current Biology. 2013;23(2):162-7.
286. Kohorn BD, Zorensky FDH, Dexter-Meldrum J, Chabout S, Mouille G, Kohorn S. Mutation of an Arabidopsis Golgi membrane protein ELMO1 reduces cell adhesion. Development. 2021;148(10).
287. Fan Y, He JJ. HIV-1 Tat Induces Unfolded Protein Response and Endoplasmic Reticulum Stress in Astrocytes and Causes Neurotoxicity through Glial Fibrillary Acidic Protein (GFAP) Activation and Aggregation *. Journal of Biological Chemistry. 2016;291(43):22819-29.
288. Martinez-Salas E, Francisco-Velilla R, Fernandez-Chamorro J, Embarek AM. Insights into Structural and Mechanistic Features of Viral IRES Elements. Frontiers in microbiology. 2018;8:2629-.
289. Sylvie Bannwarth and Anne G. HIV-1 TAR RNA: The Target of Molecular Interactions Between the Virus and its Host. Current HIV Research. 2005;3(1):61-71.
290. Keane SC, Heng X, Lu K, Kharytonchyk S, Ramakrishnan V, Carter G, et al. RNA structure. Structure of the HIV-1 RNA packaging signal. Science (New York, NY). 2015;348(6237):917-21.
291. Price DH. P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. Molecular and cellular biology. 2000;20(8):2629-34.
292. Liu R-d, Wu J, Shao R, Xue Y-h. Mechanism and factors that control HIV-1 transcription and latency activation. Journal of Zhejiang University Science B. 2014;15(5):455-65.
293. He N, Jahchan NS, Hong E, Li Q, Bayfield MA, Maraia RJ, et al. A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. Molecular cell. 2008;29(5):588-99.
294. Liu S, Rauhut R, Vornlocher H-P, Lührmann R. The network of protein-protein interactions within the human U4/U6.U5 tri-snRNP. RNA (New York, NY). 2006;12(7):1418-30.
295. Ohno M, Segref A, Bachi A, Wilm M, Mattaj IW. PHAX, a Mediator of U snRNA Nuclear Export Whose Activity Is Regulated by Phosphorylation. Cell. 2000;101(2):187-98.
296. Kammler S, Otte M, Hauber I, Kjems J, Hauber J, Schaal H. The strength of the HIV-1 3' splice sites affects Rev function. Retrovirology. 2006;3(1):89.
297. Pongoski J, Asai K, Cochrane A. Positive and negative modulation of human immunodeficiency virus type 1 Rev function by cis and trans regulators of viral RNA splicing. Journal of virology. 2002;76(10):5108-20.
298. Chang DD, Sharp PA. Regulation by HIV Rev depends upon recognition of splice sites. Cell. 1989;59(5):789-95.
299. Tazi J, Bakkour N, Marchand V, Ayadi L, Aboufirassi A, Branlant C. Alternative splicing: regulation of HIV-1 multiplication as a target for therapeutic action. The FEBS Journal. 2010;277(4):867-76.
300. Lévesque K, Halvorsen M, Abrahamyan L, Chatel-Chaix L, Poupon V, Gordon H, et al. Trafficking of HIV-1 RNA is Mediated by Heterogeneous Nuclear Ribonucleoprotein A2 Expression and Impacts on Viral Assembly. Traffic. 2006;7(9):1177-93.

301. Vincendeau M, Nagel D, Brenke JK, Brack-Werner R, Hadian K. Heterogenous nuclear ribonucleoprotein Q increases protein expression from HIV-1 Rev-dependent transcripts. Virology Journal. 2013;10(1):151.
302. Wong RW, Balachandran A, Ostrowski MA, Cochrane A. Digoxin suppresses HIV-1 replication by altering viral RNA processing. PLoS pathogens. 2013;9(3):e1003241-e.
303. Pomerantz RJ, Seshamma T, Trono D. Efficient replication of human immunodeficiency virus type 1 requires a threshold level of Rev: potential implications for latency. Journal of virology. 1992;66(3):1809-13.
304. Black AC, Luo J, Chun S, Bakker A, Fraser JK, Rosenblatt JD. Specific binding of polypyrimidine tract binding protein and hnRNP A1 to HIV-1 CRS elements. Virus Genes. 1996;12(3):275-85.
305. Chen C-YA, Shyu A-B. AU-rich elements: characterization and importance in mRNA degradation. Trends in Biochemical Sciences. 1995;20(11):465-70.
306. Boczonadi V, Müller JS, Pyle A, Munkley J, Dor T, Quartararo J, et al. EXOSC8 mutations alter mRNA metabolism and cause hypomyelination with spinal muscular atrophy and cerebellar hypoplasia. Nature Communications. 2014;5(1):4287.
307. Hagan KW, Ruiz-Echevarria MJ, Quan Y, Peltz SW. Characterization of cis-acting sequences and decay intermediates involved in nonsense-mediated mRNA turnover. Molecular and cellular biology. 1995;15(2):809-23.
308. Schmid M, Jensen TH. The exosome: a multipurpose RNA-decay machine. Trends in Biochemical Sciences. 2008;33(10):501-10.
309. Guo X, Ma J, Sun J, Gao G. The zinc-finger antiviral protein recruits the RNA processing exosome to degrade the target mRNA. Proceedings of the National Academy of Sciences. 2007;104(1):151.
310. Yeung ML, Houzet L, Yedavalli VSRK, Jeang K-T. A genome-wide short hairpin RNA screening of jurkat T-cells for human proteins contributing to productive HIV-1 replication. The Journal of biological chemistry. 2009;284(29):19463-73.
311. Jackson RJ, Hellen CUT, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. Nature reviews Molecular cell biology. 2010;11(2):113-27.
312. Svitkin YV, Pause A, Sonenberg N. La autoantigen alleviates translational repression by the 5' leader sequence of the human immunodeficiency virus type 1 mRNA. Journal of virology. 1994;68(11):7001-7.
313. Buck CB, Shen X, Egan MA, Pierson TC, Walker CM, Siliciano RF. The human immunodeficiency virus type 1 gag gene encodes an internal ribosome entry site. Journal of virology. 2001;75(1):181-91.
314. Balvay L, Lastra ML, Sargueil B, Darlix J-L, Ohlmann T. Translational control of retroviruses. Nature Reviews Microbiology. 2007;5(2):128-40.
315. Barrera A, Olguín V, Vera-Otarola J, López-Lastra M. Cap-independent translation initiation of the unspliced RNA of retroviruses. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. 2020;1863(9):194583.
316. Carvajal F, Vallejos M, Walters B, Contreras N, Hertz MI, Olivares E, et al. Structural domains within the HIV-1 mRNA and the ribosomal protein S25 influence cap-independent translation initiation. The FEBS journal. 2016;283(13):2508-27.
317. Monette A, Ajamian L, López-Lastra M, Mouland AJ. Human immunodeficiency virus type 1 (HIV-1) induces the cytoplasmic retention of heterogeneous nuclear

ribonucleoprotein A1 by disrupting nuclear import: implications for HIV-1 gene expression. The Journal of biological chemistry. 2009;284(45):31350-62.
318. Dugré-Brisson S, Elvira G, Boulay K, Chatel-Chaix L, Mouland AJ, DesGroseillers L. Interaction of Staufen1 with the 5' end of mRNA facilitates translation of these RNAs. Nucleic acids research. 2005;33(15):4797-812.
319. Alvarez E, Castelló A, Menéndez-Arias L, Carrasco L. HIV protease cleaves poly(A)-binding protein. The Biochemical journal. 2006;396(2):219-26.
320. Castelló A, Franco D, Moral-López P, Berlanga JJ, Alvarez E, Wimmer E, et al. HIV- 1 protease inhibits Cap- and poly(A)-dependent translation upon eIF4GI and PABP cleavage. PloS one. 2009;4(11):e7997-e.
321. Roberts LO, Jopling CL, Jackson RJ, Willis AE. Viral strategies to subvert the mammalian translation machinery. Progress in molecular biology and translational science. 2009;90:313-67.
322. Erales J, Marchand V, Panthu B, Gillot S, Belin S, Ghayad SE, et al. Evidence for rRNA 2'-O-methylation plasticity: Control of intrinsic translational capabilities of human ribosomes. Proceedings of the National Academy of Sciences. 2017;114(49):12934.
323. Murray JL, Sheng J, Rubin DH. A role for H/ACA and C/D small nucleolar RNAs in viral replication. Molecular biotechnology. 2014;56(5):429-37.
324. Cockrell AS, van Praag H, Santistevan N, Ma H, Kafri T. The HIV-1 Rev/RRE system is required for HIV-1 5' UTR cis elements to augment encapsidation of heterologous RNA into HIV-1 viral particles. Retrovirology. 2011;8(1):51.
325. Brégnard C, Zamborlini A, Leduc M, Chafey P, Camoin L, Saïb A, et al. 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. 2013;87(7):3729-40.
326. Floderer C, Masson J-B, Boilley E, Georgeault S, Merida P, El Beheiry M, et al. Single molecule localisation microscopy reveals how HIV-1 Gag proteins sense membrane virus assembly sites in living host CD4 T cells. Scientific Reports. 2018;8(1):16283.
327. Brennan K, Martin K, FitzGerald SP, O'Sullivan J, Wu Y, Blanco A, et al. A comparison of methods for the isolation and separation of extracellular vesicles from protein and lipid particles in human serum. Scientific Reports. 2020;10(1):1039.
328. van der Pol E, Coumans FAW, Grootemaat AE, Gardiner C, Sargent IL, Harrison P, et al. Particle size distribution of exosomes and microvesicles determined by transmission electron microscopy, flow cytometry, nanoparticle tracking analysis, and resistive pulse sensing. Journal of Thrombosis and Haemostasis. 2014;12(7):1182-92.
329. Gluschankof P, Mondor I, Gelderblom HR, Sattentau QJ. Cell Membrane Vesicles Are a Major Contaminant of Gradient-Enriched Human Immunodeficiency Virus Type-1 Preparations. Virology. 1997;230(1):125-33.
330. Cantin R, Diou J, Bélanger D, Tremblay AM, Gilbert C. Discrimination between exosomes and HIV-1: Purification of both vesicles from cell-free supernatants. Journal of Immunological Methods. 2008;338(1):21-30.
331. Steppert P, Burgstaller D, Klausberger M, Berger E, Aguilar PP, Schneider TA, et al. Purification of HIV-1 gag virus-like particles and separation of other extracellular particles. Journal of Chromatography A. 2016;1455:93-101.
332. Broadus J, Fuerstenberg S, Doe CQ. Staufen-dependent localization of prospero mRNA contributes to neuroblast daughter-cell fate. Nature. 1998;391(6669):792-5.
333. Tang SJ, Meulemans D, Vazquez L, Colaco N, Schuman E. A Role for a Rat Homolog of Staufen in the Transport of RNA to Neuronal Dendrites. Neuron. 2001;32(3):463-75.

334. Mouland AJ, Mercier J, Luo M, Bernier L, DesGroseillers L, Cohen EA. The double-stranded RNA-binding protein Staufen is incorporated in human immunodeficiency virus type 1: evidence for a role in genomic RNA encapsidation. Journal of virology. 2000;74(12):5441-51.
335. Kozak SL, Marin M, Rose KM, Bystrom C, Kabat D. The Anti-HIV-1 Editing Enzyme APOBEC3G Binds HIV-1 RNA and Messenger RNAs That Shuttle between Polysomes and Stress Granules *. Journal of Biological Chemistry. 2006;281(39):29105-19.
336. Boeras I, Song Z, Moran A, Franklin J, Brown WC, Johnson M, et al. DHX9/RHA Binding to the PBS-Segment of the Genomic RNA during HIV-1 Assembly Bolsters Virion Infectivity. Journal of molecular biology. 2016;428(11):2418-29.
337. Darbinian N, Sawaya BE, Khalili K, Jaffe N, Wortman B, Giordano A, et al. Functional interaction between cyclin T1/cdk9 and Purα determines the level of TNFα promoter activation by Tat in glial cells. Journal of Neuroimmunology. 2001;121(1):3-11.
338. Thatikunta P, Sawaya BE, Denisova L, Cole C, Yusibova G, Johnson EM, et al. Identification of a cellular protein that binds to tat-responsive element of TGFβ-1 promoter in glial cells. Journal of Cellular Biochemistry. 1997;67(4):466-77.
339. Zambrano N, De Renzis S, Minopoli G, Faraonio R, Donini V, Scaloni A, et al. DNA-binding protein Pur alpha and transcription factor YY1 function as transcription activators of the neuron-specific FE65 gene promoter. The Biochemical journal. 1997;328 (Pt 1)(Pt 1):293-300.
340. White MK, Johnson EM, Khalili K. Multiple roles for Puralpha in cellular and viral regulation. Cell cycle (Georgetown, Tex). 2009;8(3):1-7.
341. Kaminski R, Darbinian N, Sawaya BE, Slonina D, Amini S, Johnson EM, et al. Puralpha as a cellular co-factor of Rev/RRE-mediated expression of HIV-1 intron-containing mRNA. Journal of cellular biochemistry. 2008;103(4):1231-45.
342. Brass AL, Dykxhoorn DM, Benita Y, Yan N, Engelman A, Xavier RJ, et al. Identification of Host Proteins Required for HIV Infection Through a Functional Genomic Screen. Science. 2008;319(5865):921.
343. Molitor L, Bacher S, Burczyk S, Niessing D. The Molecular Function of PURA and Its Implications in Neurological Diseases. Frontiers in genetics. 2021;12:638217-.
344. Kelm RJ, Jr., Cogan JG, Elder PK, Strauch AR, Getz MJ. Molecular Interactions between Single-stranded DNA-binding Proteins Associated with an Essential MCAT Element in the Mouse Smooth Muscle α-Actin Promoter *. Journal of Biological Chemistry. 1999;274(20):14238-45.
345. Lokeswara Balakrishna S, Satyanarayana N, Kondapi AK. Involvement of human topoisomerase II isoforms in HIV-1 reverse transcription. Archives of Biochemistry and Biophysics. 2013;532(2):91-102.
346. König R, Zhou Y, Elleder D, Diamond TL, Bonamy GMC, Irelan JT, et al. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. Cell. 2008;135(1):49-60.
347. Lata S, Mishra R, Banerjea AC. Proteasomal Degradation Machinery: Favorite Target of HIV-1 Proteins. Frontiers in microbiology. 2018;9:2738-.
348. Anderson BD, Harris RS. Transcriptional regulation of APOBEC3 antiviral immunity through the CBF-β/RUNX axis. Science advances. 2015;1(8):e1500296-e.
349. Malim MH. Natural resistance to HIV infection: The Vif–APOBEC interaction. Comptes Rendus Biologies. 2006;329(11):871-5.

350. Zhou X, Monnie C, DeLucia M, Ahn J. HIV-1 Vpr activates host CRL4-DCAF1 E3 ligase to degrade histone deacetylase SIRT7. Virology Journal. 2021;18(1):48.
351. Ryang J, Yan Y, Song Y, Liu F, Ng TB. Anti-HIV, antitumor and immunomodulatory activities of paclitaxel from fermentation broth using molecular imprinting technique. AMB Express. 2019;9(1):194-.
352. Georgiou NA, van der Bruggen T, Healy DMC, van Tienen C, de Bie P, Oudshoorn M, et al. Bleomycin has antiviral properties against drug-resistant HIV strains and sensitises virus to currently used antiviral agents. International Journal of Antimicrobial Agents. 2006;27(1):63-8.
353. Einstein MH, Ndlovu N, Lee J, Stier EA, Kotzen J, Garg M, et al. Cisplatin and radiation therapy in HIV-positive women with locally advanced cervical cancer in sub-Saharan Africa: A phase II study of the AIDS malignancy consortium. Gynecologic oncology. 2019;153(1):20-5.
354. Whiskey E, O'Flynn D, Taylor D. Clozapine, HIV and neutropenia: a case report. Therapeutic advances in psychopharmacology. 2018;8(12):365-9.
355. Hosseinipour MC, Kang M, Krown SE, Bukuru A, Umbleja T, Martin JN, et al. As-Needed Vs Immediate Etoposide Chemotherapy in Combination With Antiretroviral Therapy for Mild-to-Moderate AIDS-Associated Kaposi Sarcoma in Resource-Limited Settings: A5264/AMC-067 Randomized Clinical Trial. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2018;67(2):251-60.
356. Grewe B, Ehrhardt K, Hoffmann B, Blissenbach M, Brandt S, Uberla K. The HIV-1 Rev protein enhances encapsidation of unspliced and spliced, RRE-containing lentiviral vector RNA. PloS one. 2012;7(11):e48688-e.
357. Tuffy KM, Maldonado RJK, Chang J, Rosenfeld P, Cochrane A, Parent LJ. HIV-1 Gag Forms Ribonucleoprotein Complexes with Unspliced Viral RNA at Transcription Sites. Viruses. 2020;12(11):1281.
358. Ajamian L, Abel K, Rao S, Vyboh K, García-de-Gracia F, Soto-Rifo R, et al. HIV-1 Recruits UPF1 but Excludes UPF2 to Promote Nucleocytoplasmic Export of the Genomic RNA. Biomolecules. 2015;5(4).
359. Yedavalli VSRK, Neuveut C, Chi Y-h, Kleiman L, Jeang K-T. Requirement of DDX3 DEAD Box RNA Helicase for HIV-1 Rev-RRE Export Function. Cell. 2004;119(3):381-92.
360. Farjot G, Sergeant A, Mikaélian I. A New Nucleoporin-like Protein Interacts with Both HIV-1 Rev Nuclear Export Signal and CRM-1 *. Journal of Biological Chemistry. 1999;274(24):17309-17.
361. Wang T, Tian C, Zhang W, Luo K, Sarkis PTN, Yu L, et al. 7SL RNA mediates virion packaging of the antiviral cytidine deaminase APOBEC3G. Journal of virology. 2007;81(23):13112-24.
362. Ajamian L, Abrahamyan L, Milev M, Ivanov PV, Kulozik AE, Gehring NH, et al. Unexpected roles for UPF1 in HIV-1 RNA metabolism and translation. RNA. 2008;14(5):914-27.
363. Roy BB, Hu J, Guo X, Russell RS, Guo F, Kleiman L, et al. Association of RNA Helicase A with Human Immunodeficiency Virus Type 1 Particles *. Journal of Biological Chemistry. 2006;281(18):12625-35.
364. O'Carroll IP, Fan L, Kroupa T, McShane EK, Theodore C, Yates EA, et al. Structural Mimicry Drives HIV-1 Rev-Mediated HERV-K Expression. Journal of Molecular Biology. 2020;432(24):166711.

365. Hanke K, Hohn O, Liedgens L, Fiddeke K, Wamara J, Kurth R, et al. Staufen-1 interacts with the human endogenous retrovirus family HERV-K(HML-2) rec and gag proteins and increases virion production. Journal of virology. 2013;87(20):11019-30.
366. Tan R, Frankel AD. Structural variety of arginine-rich RNA-binding peptides. Proceedings of the National Academy of Sciences. 1995;92(12):5282.
367. Moore MD, Nikolaitchik OA, Chen J, Hammarskjöld M-L, Rekosh D, Hu W-S. Probing the HIV-1 genomic RNA trafficking pathway and dimerization by genetic recombination and single virion analyses. PLoS pathogens. 2009;5(10):e1000627-e.
368. Jin J, Sturgeon T, Weisz OA, Mothes W, Montelaro RC. HIV-1 matrix dependent membrane targeting is regulated by Gag mRNA trafficking. PloS one. 2009;4(8):e6551-e.
369. Ritchie C, Cylinder I, Platt EJ, Barklis E. Analysis of HIV-1 Gag protein interactions via biotin ligase tagging. Journal of virology. 2015;89(7):3988-4001.
370. Zerbato JM, Khoury G, Zhao W, Gartner MJ, Pascoe RD, Rhodes A, et al. Multiply spliced HIV RNA is a predictive measure of virus production ex vivo and in vivo following reversal of HIV latency. EBioMedicine. 2021;65.
371. Liszewski MK, Yu JJ, O'Doherty U. Detecting HIV-1 integration by repetitive-sampling Alu-gag PCR. Methods (San Diego, Calif). 2009;47(4):254-60.
372. Garibaldi A, Carranza F, Hertel KJ. Isolation of Newly Transcribed RNA Using the Metabolic Label 4-Thiouridine. Methods in molecular biology (Clifton, NJ). 2017;1648:169-76.