Translational inhibition of the restriction factor APOBEC3G (A3G) by the HIV-1 Vif protein : role of a uORF in the 5'-UTR of A3G mRNA and identification of cellular factors

Tanja Seissler

▶ **To cite this version:**

Tanja Seissler. Translational inhibition of the restriction factor APOBEC3G (A3G) by the HIV-1 Vif protein : role of a uORF in the 5'-UTR of A3G mRNA and identification of cellular factors. Virology. Université de Strasbourg, 2019. English. NNT : 2019STRAJ028 . tel-03270780

HAL Id: tel-03270780
https://theses.hal.science/tel-03270780

Submitted on 25 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Strasbourg

UNIVERSITÉ DE STRASBOURG

École Doctorale des Sciences de la Vie et de la Santé STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

CNRS UPR 9002 - Architecture et Réactivité de l'ARN

THÈSE présentée par :

Tanja SEISSLER

soutenue le : **13 septembre 2019**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Biologie moléculaire, cellulaire et Virologie

Inhibition traductionnelle du facteur de restriction APOBEC3G par la protéine Vif du VIH-1

Role d'une uORF dans la 5'-UTR de l'ARNm d'A3G et identification de facteurs cellulaires

THÈSE dirigée par :

Dr PAILLART Jean-Christophe	Directeur de recherches, Institut de biologie moléculaire et cellulaire, Strasbourg

RAPPORTEURS :

Dr GALLOIS-MONTBRUN Sarah	Chargée de recherches, Institut Cochin, Paris
Dr GOUJON Caroline	Chargée de recherches, Institut de recherche en infectiologie, Montpellier

AUTRES MEMBRES DU JURY :

Prof. GILMER David	Professeur, Institut de biologie moléculaire des plantes
Dr RICCI Emiliano	Chargé de recherches, Laboratoire de biologie et modélisation de la cellule, Lyon

University of Strasbourg
Doctoral School of Life and Health Sciences
CNRS UPR 9002, Architecture and Reactivity of RNA

PhD Thesis

Submitted to obtain the degree of Doctor of the University of Strasbourg

Discipline: Life and Health Sciences
Specialty: Molecular Biology, Cellular Biology and Virology

by

Tanja Seissler

Translational inhibition of the Restriction Factor APOBEC3G (A3G) by the HIV-1 Vif protein

Role of a uORF in the 5'-UTR of A3G mRNA and identification of cellular factors

Defended on September 13th, 2019 in front of the examining committee:

Dr. Jean-Christophe Paillart	Thesis Director
Dr. Sarah Gallois-Montbrun	External Reporter
Dr. Caroline Goujon	External Reporter
Prof. David Gilmer	Examiner
Dr. Emiliano Ricci	Examiner

Acknowledgements

J’aimerais tout d’abord remercier mon Jury de thèse, Dr. Sarah Gallois-Montbrun, Dr. Caroline Goujon, Prof. David Gilmer et Dr. Emiliano Ricci pour avoir accepté de juger mon travail et de faire de cette thèse leur lecture de l’été.

Un grand merci aussi à toutes les personnes qui ont contribué à l’avancement de cette thèse. Merci à Philippe, Johana et Lauriane de la plateforme de spectrométrie de masse pour avoir analysé les nombreux échantillons que je vous ai donnés au fil des ans. Merci également à Béatrice pour ton outil d’analyse des données et aussi pour avoir répondu à mes questions de bio-informatique. Merci aussi au Dr. Emiliano Ricci pour nous avoir fourni les Nanoblades pour la délétion de l'uORF dans les cellules H9. Merci à Géraldine Laumond et au Dr. Christiane Moog pour les tests d'infection et merci au Dr. Karim Majzoub pour les plasmides RaPID.

La personne la plus importante pour cette thèse c‘était toi, JC. Tu m’as tout appris. Tu as toujours été disponible pour répondre à mes questions, essayer de résoudre des problèmes et corriger mes écrits en un temps record. Tu m’as remonté le moral maintes fois pour mes manips ratées et tu as fêté avec moi chaque avancée. Je te remercie pour tout ce que tu as fait. Tu es le meilleur !

Fous rires, cris, des gens qui se font frapper par des calendriers, des gants qui fusent dans tous les sens … tout ça c’est le labo 433-435. Et toute cette bonne humeur, c’est grâce à vous, Benjamin, Orian, Erwan, Cédric. Même si les manips ont été difficiles par moment, l’ambiance était toujours au top et je vous remercie pour ça !
Je remercie également tous les autres membres de l'équipe 433-435 : Valérie, Anne, Roland et Serena. Merci de m’avoir accueilli parmi vous si chaleureusement. Merci aussi aux autres personnes dans l'unité pour votre gentillesse et votre accueil.
J'aimerais également remercier les personnes qui sont déjà parties vers d’autres horizons. Camille, merci de m’avoir initié dans la team ApoVif et de m’avoir appris les western blot et compagnie. Elodie, tu m’as accueilli dans notre petit coin du labo 435 et tu m’as gentiment aidé avec tous mes petits problèmes et mes questions au début de ma thèse. Merci d’avoir toujours pris le temps pour moi. Merci aussi à Damien, Maud, Noé et Red pour votre bonne humeur et votre amitié !

Je tiens aussi à remercier tous mes stagiaires. Orian, mon premier bébé stagiaire. Je t'ai utilisé comme robot à minipreps et calculatrice vivante. Avec toute cette expertise que tu as appris au cours de ce stage, je suis sûre que tu vas faire une thèse extra. Melanie, tu as fait preuve d'une surprenante autonomie et efficacité dans tout ce que tu faisais. On a formé une super équipe toutes les deux ! Léa, merci de m'avoir aidé à boucler ce travail en un temps record. J'ai eu de la chance de t'avoir pour m'aider à survivre à ce dernier mois de manips et de rédaction. Et finalement Cédric, l'ultime padawan. Tu n'as pas eu la vie facile, avec une encadrante stressée par sa dernière année. Tu as été autonome très vite et c'est une de tes grandes qualités. Je te transmets le flambeau ApoVif. Je sais que tu vas faire un super travail !

Ich möchte mich auch von ganzem Herzen bei meiner Familie bedanken. Ihr habt mich mein ganzes Studium über unterstützt. Danke, dass ihr immer hinter mir steht und mich anfeuert.

Finalement j'aimerais remercier Alan. Tu as été le plus grand soutien dans ma vie. Rien n'aurait été pareil sans toi. Je ne pourrais jamais te remercier assez.

Table of contents

ACKNOWLEDGEMENTS **2**

TABLE OF CONTENTS **4**

LIST OF ABBREVIATIONS **7**

I INTRODUCTION **11**

1. THE HUMAN IMMUNODEFICIENCY VIRUS 11

1.1 HIV discovery and pandemic *11*

1.2 Origin and evolution *11*

1.3 Natural progression of the HIV infection *13*

1.4 Treatment *15*

1.5 HIV-1 life cycle *16*

1.5.1 Entry 17

1.5.2 Reverse transcription 18

1.5.3 Uncoating, Nuclear Import and Integration 20

1.5.4 Expression of the viral genome and latency 22

1.5.5 Assembly, budding and maturation 23

2. HOST IMMUNE DEFENSES AND VIRAL COUNTER-DEFENSES 25

2.1 The immune system *25*

2.2 Restriction Factors *28*

2.3 APOBEC *50*

2.4 Vif *54*

3. TRANSLATION 56

3.1 The canonical mechanism of translation in eukaryotes *56*

3.2 Regulation of translation initiation *59*

3.2.1 RNA binding proteins 59

3.2.2 Non canonical translation initiation 60

3.2.3 Upstream open reading frames 61

3.2.4 RNA granules 64

3.3 Translational regulation of HIV-1 ..... *65*

II OBJECTIVES OF THIS THESIS **67**

III. IMPORTANCE OF A CONSERVED UORF IN THE 5'-UTR OF A3G AND A3F MRNA **69**

1. A CONSERVED UORF IN APOBEC3G MRNA IMPACTS ITS OWN TRANSLATION AND IS ESSENTIAL TO ITS TRANSLATIONAL INHIBITION BY THE HIV-1 VIF PROTEIN 69

2. ROLE OF THE A3G UORF IN VIRAL INFECTION 107

2.1 Introduction *107*

2.2 Material and methods *107*

2.2.1 Generation of a ΔuORF cell line using nanoblades 107

2.2.2 Screening of the ΔuORF deletion by RT-PCR on A3G mRNA and PCR on genomic DNA 108

2.2.3 Infection of cells 109

2.2.4 Western blotting 110

2.2.5 RT-qPCR 111

2.3 Results *111*

2.3.1 Deletion of the A3G uORF in an H9 T-cell line 111

2.3.2 Effect of the uORF on viral production and infectivity 113

2.4 Discussion *114*

IV. IDENTIFICATION OF PROTEINS INTERACTING WITH THE 5'-UTR OF A3G MRNA **116**

1. INTRODUCTION 116

2. MATERIAL AND METHODS 117

2.1 Plasmids *117*

2.2 Preparation of cells *117*
2.3 Pull-down of biotinylated proteins *118*
2.4 Identification of proteins by mass spectrometry *118*
2.5 Bioinformatic analysis of identified proteins *119*
2.6 Co-immunoprecipitation (without crosslinking) of Vif with its potential cellular partners *119*
2.7 Co-immunoprecipitation (with crosslinking) of Vif with its potential cellular partners *120*
2.8 Western blotting *121*
3. RESULTS 121
3. DISCUSSION 125

V. DEVELOPMENT OF A PROTOCOL TO IDENTIFY PROTEINS ASSOCIATED WITH THE FULL-LENGTH A3G MRNA 129
1. STUDY OF RNA-PROTEIN COMPLEXES: STATE OF THE ART 129
2. MATERIAL AND METHODS 131
2.1 Pull-down of A3G mRNA using complementary, biotinylated oligonucleotides *131*
2.1.1 Design of oligos for pull-down 131
2.1.2 Preparation of cells 132
2.1.3 Cell lysis in a syringe 132
2.1.4 Pull-down 132
2.1.5 Analysis of oligo retention 133
2.1.6 RT-PCR 133
2.1.7 Quantification of bands on agarose gel 134
2.1.8 *In vitro* transcription of A3G mRNA 134
2.1.9 Preparation of A3G mRNA-associated proteins for mass spectrometry 135
2.1.10 Mass spectrometry 135
2.2 Pull-down and specific elution of A3G mRNA using complementary, desthiobiotinylated oligonucleotides *136*
2.3 Pull-down of A3G mRNA using complementary, biotinylated oligonucleotides and elution by RNase H *136*
2.3.1 Digestion of *in vitro* transcribed A3G mRNA by RNase H followed by migration on a denaturing gel 136
2.3.2 Pull-down followed by elution with RNase H 137
2.3.3 RT-qPCR 137
2.4 Pull-down of A3G mRNA using a complementary, biotinylated oligo and elution by a competitor oligo *138*
2.5 Pull-down of cellular proteins on an in vitro transcribed, biotinylated A3G mRNA *139*
2.5.1 *In vitro* transcription of biotinylated, capped, poly-adenylated A3G mRNA 139
2.5.2 Preparation of cellular lysates using a nitrogen cell bomb 139
2.5.3 Pull-down of cellular proteins on biotinylated A3G mRNA 140
3. RESULTS 140
3.1 Pull-down of A3G mRNA using complementary, biotinylated oligonucleotides *140*
3.1.1 Experimental setup 140
3.1.2 Binding of the biotinylated oligonucleotides to different types of beads 141
3.1.3 Retention of the A3G mRNA on magnetic beads in presence and absence of complementary oligonucleotides 141
3.1.4 Tests to eliminate aspecifically bound RNAs from the magnetic beads 145
3.1.5 Identification of proteins associated with A3G mRNA by mass spectrometry 146
3.2 Pull-down and specific elution of A3G mRNA using complementary, desthiobiotinylated oligonucleotides *148*
3.2.1 Experimental setup 149
3.2.2 Tests 150
3.3 Pull-down of A3G mRNA using complementary, biotinylated oligonucleotides and elution by RNAse H *150*
3.3.1 Experimental setup 150
3.3.2 Tests 151
3.4 Pull-down of A3G mRNA using a complementary, biotinylated oligo and elution by a competitor oligo *152*
3.4.1 Experimental strategy 152

3.4.2 Tests 153
3.5 Pull-down of cellular proteins on an in vitro transcribed, biotinylated A3G mRNA 155
3.5.1 Experimental strategy 155
3.5.2 Retention of biotinylated A3G mRNA on streptavidin-coated beads 155
4. DISCUSSION 156

VI IDENTIFICATION OF CELLULAR FACTORS IMPLICATED IN APOBEC3G TRANSLATIONAL INHIBITION BY THE HIV-1 VIF PROTEIN 160

VII DISCUSSION 194

VIII CONCLUSION AND PERSPECTIVES 198

IX. SUMMARY IN FRENCH / RESUME EN FRANÇAIS 200

X BIBLIOGRAPHY 219

XI APPENDICES 229

List of abbreviations

A	
A1CF	APOBEC1 complementation factor
aa-tRNA	Aminoacyl-tRNA
ABCE1	ATP-binding cassette sub-family E member 1
ADAT2	Adenosine deaminase tRNA specific 2
AID	Activation-induced cytidine deaminase
AIDS	Acquired immunodeficiency syndrome
ALLN	Acetyl-leucine-leucine-norleucine (proteasome inhibitor)
AMV	Avian myeloblastosis virus
ApoB	Apolipoprotein B
APOBEC	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like
ATF4	Activating transcription factor 4
AZT	Azidothymidine
B	
BCR	B-cell receptor
BirA	Bifunctional ligase/repressor BirA
BSA	Bovine serum albumin
C	
CA	Capsid protein
Cas9	CRISPR-associated protein 9
cat-1	Cationic amino acid transporter 1
CBC	Cap-binding complex
CBF-β	Core binding factor β
CBP	Cap-binding protein
CCR4	Carbon catabolite repressor protein 4
CCR5	C-C chemokine receptor type 5
CD4, 8	Cluster of differentiation 4, 8
CDK7	Cyclin dependent kinase
cDNA	Complementary DNA
cGas	Cyclic GMP-AMP synthase
cpa1	Carboxypeptidase A1
CPSF6	Cleavage and polyadenylation specific factor 6
cpz	Chimpanzee
cpz	Chimpanzee
CRF	Circulating recombinant form
CRISPR	Clustered regularly interspaced short palindromic repeats
Crm1	Chromosomal maintenance 1
CTP	Cytidine triphosphate
Cul5	Cullin 5
CXCR4	C-X-C chemokine receptor type 4
CycT1	Cyclin T1
CYFIP1	Cytoplasmic FMR1 interacting protein 1
CypA	Cyclophilin A
D	
DAMP	Damage-associated molecular pattern
DAVID	Database for annotation, visualization and integrated discovery
DC-SIGN	Dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin
DCP1, 2	mRNA decapping enzyme 1, 2
DDX3, 9	DEAD box protein 3
DEAE	Diethylaminoethanol
DENR	Density regulated re-initiation and release factor
DHX29	DExH-Box helicase 29
DIS	Dimerization initiation site
DMEM	Dulbecco's modified Eagle medium
DNA	Deoxyribonucleic acid
DNase	Deoxiribonuclease
dPBS	Dulbecco's phosphate buffered saline
DSIF	DRB sensitivity inducing factor
DSP	Dithiobis(succinimidyl propionate)
dsRBD	Double-stranded RNA binding domain
DTT	Dithiothreitol
E	
EDTA	Ethylenediaminetetraacetic acid
eEF	Eukaryotic elongation factor
eIF	Eukaryotic initiation factor
eIF4E-BP	eIF4E binding protein
EJC	Exon-junction complex
ELISA	Enzyme-linked immunosorbent assay
EloB	Elongin B
EloC	Elongin C
Env	Envelope protein
EPRS	Glutamyl-Prolyl-tRNA synthetase

ER	Endoplasmic reticulum
eRF	Eukaryotic release factor
ESCRT	Endosomal sorting complexes required for transport

F

FDR	False discovery rate
FMRP	Fragile X mental retardation protein

G

G2BP	Ras-GAP SH3 domain binding protein
GAIT	IFN-γ-activated inhibitor of translation
GALT	Gut-associated lymphatic tissue
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
gor	Gorilla
gp	Glycoprotein
gRNA	Genomic RNA

H

HAART	Highly active antiretroviral therapy
HEK	Human embryo kidney
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen
hnRNP	Heterogeneous nuclear RNP

I

IFI16	Gamma-interferon-inducible protein 16
IFN	Interferon
IFNAR	Interferon-α/β receptor
IKK	IκB kinase
IN	Integrase
IRES	Internal ribosome entry site
IRF3,7	Interferon regulatory factors
IRP	Iron regulatory protein
ISG	Interferon stimulated gene
ISRE	Interferon-stimulated responde element
IκB	Inhibitor of NFκB

J

JAK1	Janus Kinase 1

K

KH	K-homology domain

L

L13a	60S ribosomal protein L13a
LARP1	La-related protein 1
LC	Liquid chromatography
LEDGF	Lens epithelium-derived growth factor
lox	Lysyl oxidase
LTR	Long terminal repeat

M

MA	Matrix protein
map1B	Microtubule associated protein 1B
mAUG	Main ORF start codon
MAVS	Mitochondrial antiviral-signaling protein
MCT-1	Monocarboxylate transporter 1
MDA5	Melanoma differentiation-associated protein 5
MHC	Major histocompatibility complex
MOI	Multiplicity of infection
mORF	Main ORF
mRNA	Messenger RNA
MS/MS	Tandem mass spectrometry
msl2	Male-specific lethal 2
MYD88	Myeloid differentiation primary response gene 88

N

NC	Nucleocapsid protein
NELF	Negative elongation factor
NES	Nuclear export signal
NF-kB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NFAT	Nuclear factor of activated T-cells
NK-cell	Natural killer cells
NLS	Nuclear localization signal
NMD	Nonsense mediated decay
NNRTI	Non-nucleoside reverse transcriptase inhibitors
NRTI	Nucleoside reverse transcriptase inhibitors
NSAP	NS-1 associated protein-1
nts	Nucleotides
NUP153	Nucleoporin 153
NUP358	Nucleoporin 358

O

oligo	Oligonucleotide
ORF	Open reading frame

P

p-TEFb	Positive transcription elongation factor
PABP	Poly-A binding protein
PAGE	Polyacrylamide gel electrophoresis
PAMP	Pathogen-associated molecular pattern
PAR-)CLIP	(Photoactivable Ribonucleoside-Enhanced) Crosslinking and immunoprecipitation
PBS	Primer binding site OR Phosphate buffered saline
PC	Principal component
PCA	Principal component analysis
PIC	Preintegration complex
PIP2	Phosphatidylinositol 4,5-bisphosphate
PKR	Protein kinase R
PNA	Peptide nucleic acid
ppt	Poly-pyrimidine-tract
PRR	Pattern recognition receptor
PTC	Peptidyl transferase centre
PVDF	Polyvinylidene fluoride

R

Ran	Ras-related nuclear protein
RaPID	RNA-protein interaction detection
RBM47	RNA binding motif protein 47
RBP	RNA binding protein
Rbx2	RING-box protein 2
RIG-I	Retinoic acid-inducible gene I
RLR	RIG-I-like receptor
RNA	Ribonucleic acid
RNase	Ribonuclease
RNP	Ribonucleoproteins
ROS	Reactive oxygen species
RPMI	Roswell Park Memorial Institute medium
RRE	Rev response element
RRM	RNA recognition motif
RT	Reverse transcriptase
RT-PCR	Reverse transcription followed by polymerase chain reaction
RTC	Reverse transcription complex
RUNX	Runt-related transcription factor

S

SCF	Skp, Cullin, F-box containing
SDS	Sodium dodecyl sulfate
SEC	Super-elongation complex
sgRNA	Single guide RNA
SIV	Simian immunodeficiency virus
SL	Stem loop
sm	Sooty mangabey
SMG1	Serine/threonine-protein kinase SMG1
SNP	Single nucleotide polymorphism
snRNP	Small nuclear RNP
SP1	Specificity protein 1
SpC	Spectral count
SPINK1	Serine peptidase inhibitor, kazal type 1
SR	Serine and arginine-rich
STAT	Signal transducer and activator of transcription
STING	Stimulator of interferon genes
SURF	SMG1-UPF1-eRFs complex
SXL	Sex-lethal

T

Tad	tRNA-specific adenosine deaminase
Taq	Thermus aquaticus
TAR	Trans-activation response
TBK1	Serine/threonine-protein kinase TBK1
TC	Ternary complex
TCR	T-cell receptor
TFIID	Transcription factor II D
TFIIH	Transcription factor II H
TIA-1	T-cell-restricted intracellular antigen-1
TIAR	TIA-1-related protein
TLR	Toll-like receptor
TNPO3	Transportin 3
TOP	Terminal oligopyrimidine tract
Tpr	Translocated promoter region protein
tRNA	Transfer RNA
Tsg101	Tumor susceptibility gene 101
TYK2	Tyrosine kinase 2

U

uAUG	Upstream ORF start codon
UNG2	Uracil-DNA glycosylase
UNG2	Uracil-DNA glycosylase
uORF	Upstream ORF

	UPF1, 2, 3	Regulator of nonsense transcripts 1, 2, 3
	URF	Unique recombinant form
	UTP	Uridine triphosphate
	UTR	Untranslated region
	UV	Ultraviolet light
V		
	VLP	Virus-like particle
	Vpr	Viral protein R
	VPS4	Vacuolar protein sorting-associated protein 4
W		
	WT	Wild-type
X		
	XRN1	5'-3' exoribonuclease 1
...		
	Ψ	Packaging signal

Introduction

I INTRODUCTION

1. The Human immunodeficiency virus

1.1 HIV discovery and pandemic

The human immunodeficiency virus (HIV) has first been characterized in 1983 as the etiological agent of the acquired human immunodeficiency syndrome (AIDS) (17). This disease first came to global attention in the early 1980s in the United States, when several patients presented with a severely compromised immune system with very low $CD4^+$ T-cell counts, giving rise to rare opportunistic infections and cancers (17, 18, 73, 184, 217). In the following years, the disease has rapidly developed into a global pandemic with a great number of patients all over the world. Thanks to the discovery of the first antiretroviral treatment in 1987 (64) and the establishment of a highly efficient tri-therapy in 1996 (44, 81, 136, 194), the life expectancy of patients has nowadays greatly improved, however, there still is neither vaccine nor cure (18, 184). Therefore, HIV still represents a global health issue today with 37.9 million infected people in 2018, a number that is continuously growing with 1.7 million new infections every year. Since the beginning of the pandemic, 32 million people have already died from AIDS and in 2018, 0.8 million deaths could still be counted. While the number of new infections, as well as the number of deaths per year has considerably declined since the peak of the pandemic between 1996 and 2004, the issue still maintains a considerable extent today (226).

1.2 Origin and evolution

While HIV has only been discovered in 1983, it has been introduced into the human population much sooner by cross-species transmission of the simian immunodeficiency virus (SIV) from monkeys in Africa. Molecular clock studies suggest that the first cross-species transmission event has occurred around 1908 (165, 222).

Monkeys harbor a great number of different SIVs with a specific virus for each monkey species. SIVs are thought to have co-evolved with monkeys for around 30,000 years. In general, these SIVs don't cause any disease in the infected monkeys with the only potential effect being a

slightly reduced lifespan and overall fitness (177, 204). Cross-species transmission may occur from pet monkeys but also through hunting of monkeys as bushmeat. Cross-species transmission as well as rapid diversification of the virus is possible thanks to the ability of the virus to rapidly evolve. This is due to the high error rate of the viral reverse transcriptase, frequent recombination and rapid turnover of viral particles (165, 191, 224).

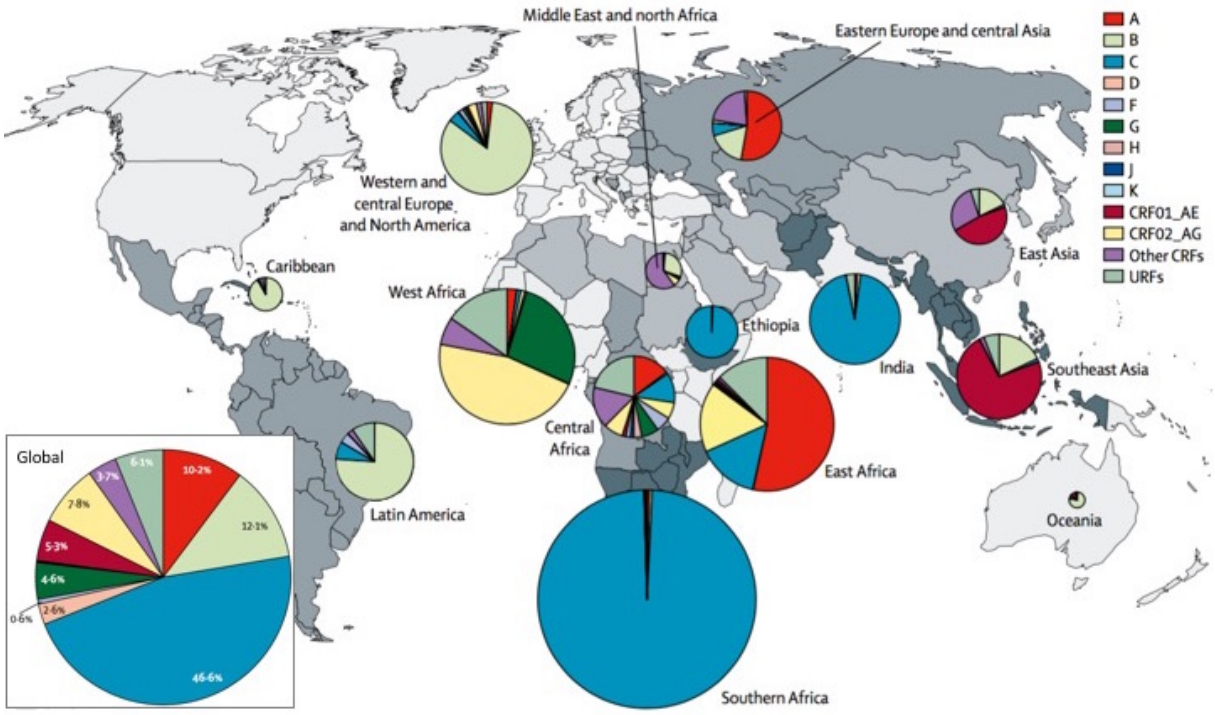


Figure 1: Global distribution of HIV-1 subtypes in the period from 2010 to 2015. *The number of people living with a given subtype, circulating recombinant form (CRF) or unique recombinant form (URF) of HIV-1 in the different regions of the world (each region being colored with different shades of grey) as well as their global distribution from 2010 to 2015 are represented as pie charts. The color-code for different subtypes in indicated on the right (adapted from 78)*

There has been a total of two distinct cross-species transmissions of SIV from the chimpanzee *Pan troglodytes troglodytes* (SIVcpz) to humans. The first one gave rise to HIV-1 group M, which is at the origin of today's pandemic, responsible for over 98 % of total HIV infections. The second transmission of SIVcpz resulted in HIV-1 group N. Two other cross-species transmission events occurred between humans and the SIVgor from gorillas (*Gorilla gorilla*), giving rise to HIV-1 group O and P. Finally, a total of 9 transmissions of SIVsm from sooty mangabeys (*Cercocebus atys*) to humans have occurred and given rise to HIV-2 groups A-I (165, 191, 222).

While HIV-1 group M has spread all over the world and has infected millions of people, HIV-1 group N, O and P, as well as HIV-2 have mostly remained confined to West and Central Africa with a limited number of patients. HIV-1 group M has diversified into subtypes A-D, F-H and J-

K. Subtype C is the most virulent one causing the largest number of infections, mostly in Sub-Saharan Africa. However, subtype B is predominant in western and central Europe as well as the Americas and has been the most studied (Fig. 1). In addition to the already immense variability of HIV due to different types, groups and subtypes as well as considerable genetic variability within one given subtype (8-17 % variation), further complexity is added by recombinant forms which present a mosaic structure combining two or more different subtypes. These viruses are named circulating recombinant forms (CRF) when characterized in at least 3 unrelated individuals (85, 165, 222).

1.3 Natural progression of the HIV infection

HIV is contained in blood, cervicovaginal secretions, rectal secretions and breast milk of an untreated infected person. HIV can thereby be transmitted through sexual intercourse, parenterally for example by transfusion or needlestick or from the mother to the child, either *in utero* through the placenta, intrapartum or by breastfeeding (91, 205).

Once the virus has reached the site of primary infection, most of the time the genital or the intestinal tract of the new host, it crosses the mucosal barrier in a couple of hours. During the first 3 days, HIV first establishes a local infection in the underlying tissue, then disseminates to draining lymph nodes (91, 103, 205). The main target cells of HIV are $CD4^+$ T-cells. While the virus cannot infect dendritic cells productively, it can attach to the surface of these cells or be captured in intracellular vesicles, which helps in the transport of the virus to lymph nodes and its presentation to susceptible $CD4^+$ T-cells (91, 142). These early stages of the infection are qualified as the eclipse phase, during which the virus cannot be detected in the blood (Fig. 2). However, the virus already begins building a latent reservoir at this early time point. Starting around day 7 to 21 of the infection, HIV disseminates systemically and becomes detectable in the blood (142, 205). During this second phase of HIV infection, qualified as the acute phase, HIV replicates exponentially and reaches a peak of very high titer of approximately 10^6-10^7 copies of viral RNA per ml of blood (Fig. 2). The acute phase can sometimes also be accompanied by flu-like symptoms (53, 162). After around 4 weeks starts the asymptomatic chronic phase of HIV infection, during which the viral load decreases to a steady state and the $CD4^+$ T-cell count, which had dramatically dropped during the acute phase, partially recovers (Fig. 2). The chronic phase can last for a period of several years where the virus steadily multiplies at low levels (53, 162, 184, 224).

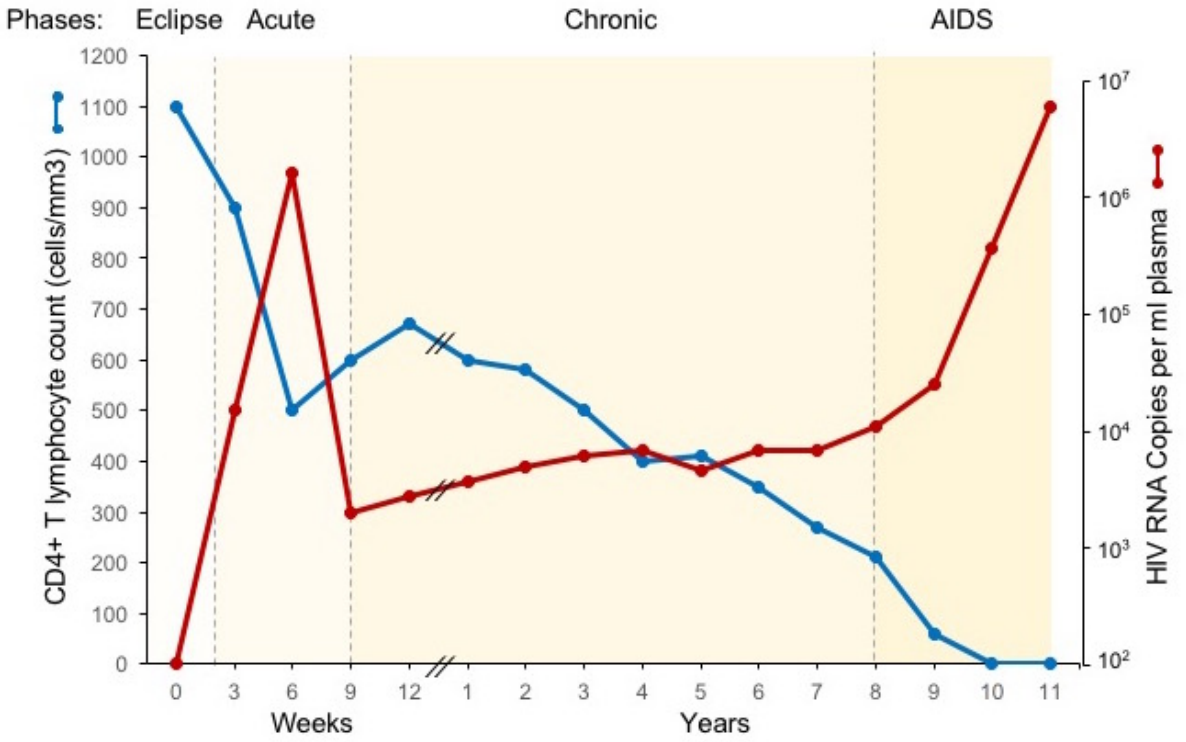


***Figure 2: Natural progression of the HIV-1 infection**. The viral load is represented in red, measured by the number of viral RNA copies in the blood. The number of CD4$^+$ T-cells in the blood is represented in blue. The viral infection plays out over the course of a decade and can be separated into 4 distinct phases: the eclipse phase, acute infection, chronic infection and the AIDS phase (Adapted from 56).*

At early stages of the infection, HIV is usually R5 T-cell tropic, which means that the virus infects cells presenting high levels of the CD4 receptor as well as the CCR5 co-receptor at their surface. This is the case for most memory T-cells and activated T-cells. While activated T-cells are the main target cells of HIV and allow efficient viral replication, memory T-cells are quiescent cells with a low gene expression in which HIV establishes the main latent reservoir. As a result of HIV infection, the available pool of CD4$^+$ CCR5$^+$ T-cells rapidly decreases and the virus evolves in order to broaden its tropism. R4 T-cell tropic viral variants arise which use CXCR4 as a coreceptor instead of CCR5. This allows latent infection of naive T-cell subsets. In parallel the virus also becomes M-tropic, which means that it becomes competent for infection of cells expressing low levels of CD4, like macrophages. While infected T-cells rapidly die due to cytopathic effects as well as the cytotoxic immune response, infected macrophages can subsist for a long time and constitute a consistent source of viral production (19, 43, 47, 181).

During the phase of chronic infection, the CD4$^+$ T-cell count decreases slowly but steadily over the years (Fig. 2). This is due to exhaustion of the immune system, which is caused on the one hand by its constitutive activation due to chronic HIV replication. Especially replication of HIV

in the gut-associated lymphoid tissue (GALT) leads to continuous inflammation accompanied by a steady depletion of $CD4^+$ T-cells. In addition to the loss of T-cells, their homeostatic renewal is exhausted over time, finally leading to a drop in $CD4^+$ T-cells (53, 162). The drop below the threshold of 200 $CD4^+$ T-cells per ml of blood characterizes the onset of the AIDS phase (Fig. 2), where the patient becomes vulnerable to opportunistic infections and cancers, ultimately resulting in death (53, 162, 184, 224).

This natural progression of the disease is determined by multiple factors, depending both on the virus and on the patient. The different variants of HIV behave differently in the infection (224). While HIV-1 group M viruses have efficiently spread all over the world and generally lead to infections resembling what has been described above, HIV-2 for example is much less virulent, less transmissible and harbors lower viral loads (165). During the transmission of the virus from an infected individual to a noninfected person, several bottlenecks contribute to the selection of a single viral variant which is called the transmitted virus or founder virus. On the side of the donor, a large variety of different viral variants circulate in the blood. Only a certain number is capable of penetrating the transmission fluid. On the side of the recipient, the virus is first selected by the mucosal barrier. Then, only the highest-fitness virus with best infectiousness, better interferon resistance and large burst size can establish a successful infection in the underlying tissue (103). It has been found that the transmitted/founder virus is almost always an R5-tropic virus. This particularity leads to almost full protection of people carrying the *CCR5-Δ32* deletion. Once the infection is established, most patients ultimately progress to the AIDS stage when not treated. Only a very small portion of patients (<1 %) maintain very low to undetectable levels of virus even in the absence of treatment. These patients are qualified as elite controllers. The better control of viral replication in these patients has often been associated with particular alleles of HLA (especially HLA-B5701), but also with an efficient and specific $CD8^+$ cytotoxic T-cell response. Overall, people who manage to maintain a lower viral load during the chronic phase of the infection progress slower to AIDS than people with an overall higher steady state viral load (53, 103, 224).

1.4 Treatment

Since the discovery of the virus in 1983 (17) and the development of the very first antiretroviral treatment in 1987 (64), considerable progress has been made in terms of treatment of HIV. This has led to a significant improvement of the lifespan of infected individuals to near-normal levels. While the first treatment was given on its own and rapidly gave rise to resistant viral

variants, since 1996 three different antiretroviral agents are combined into a highly active antiretroviral therapy (HAART) also referred to as tri-therapy, which considerably increases the genetic barrier for viral escape mutations (11, 18, 44, 81, 136).

The different classes of antiretrovirals act at the different stages of the viral life cycle:

(I) The first antiretroviral agent to be approved for treatment was 3'-azido-2',3'-dideoxythymidine (AZT) which is a nucleoside reverse transcriptase inhibitor (NRTI). These nucleotide analogues are incorporated into viral DNA by the viral reverse transcriptase (RT) and terminate the elongation of the DNA synthesis.

(II) Non-nucleotide reverse transcriptase inhibitors (NNRTI), like for example efavirenz, also inhibit the viral reverse transcriptase but instead of being incorporated into viral DNA, they bind to an allosteric site of the viral RT enzyme and thereby inhibit its functioning.

(III) Fusion inhibitors, like for example enfuvirtide, which interact with the gp41 moiety of the viral Env protein, inhibit the entry of the virus into the cell.

(IV) Integrase inhibitors, like Raltegravir, interact with the viral integrase enzyme and inhibit the strand-transfer reaction which allows the virus to integrate into the cellular genome.

(V) Protease inhibitors, like darunavir, can either be nonhydrolyzable peptidomimetic analogues which directly interact with the active site of the viral protease or they can be nonpeptidic inhibitors.

(VI) Co-receptor inhibitors, like maraviroc, are up to now the only antiretroviral agents which target the host rather than the virus itself. These molecules interact with the viral coreceptor CCR5 and inhibit its interaction with the viral envelope (11, 144).

While these drugs are very efficient in inhibiting viral replication and have considerably improved patient lives, they don't allow a cure of the disease as HIV establishes latent reservoirs which are long-lived and lead to rebound of viral replication as soon as the treatment is interrupted (138, 199). Furthermore, despite multiple trials, no efficient vaccine against HIV infection has yet been found (59).

1.5 HIV-1 life cycle

HIV-1 is an enveloped virus of 80-150 nm of diameter. Approximately 7-14 trimers of the viral envelope protein (Env) are present as spikes on the surface of the particle (Fig. 3A). The inside of the envelope is covered by matrix proteins (MA). The conical capsid formed by viral capsid proteins (CA) is present in the center of the particle (Fig. 3A). This structure contains a dimer of the viral gRNA which is associated with nucleocapsid proteins (NC) as well as the viral reverse transcriptase (RT) and integrase (IN) enzymes. Certain viral accessory proteins,

including Vif and Vpr, as well as several cellular proteins like APOBEC3G, cyclophilin A and lysyl-tRNA-synthetase are also encapsidated into HIV-1 particles (32, 66, 221, 236).

The replication cycle of HIV-1 begins with the entry of the virus into the cytoplasm of the target cell, where the viral RNA genome is reverse transcribed in order to form a double stranded DNA. This DNA is then integrated into the cellular genome as a so-called provirus which allows production of viral proteins and RNAs. These components assemble at the plasma membrane and bud in order to form new viral particles (65). In the following chapter, each of these steps will be described individually. For an overview over the entire viral life cycle, please refer to pages 31-32.

1.5.1 Entry

Env is composed of gp120 and gp41 which form a heterodimer. The gp120 is composed of five variable regions (V1-V5) which are exposed on the surface of the protein and protect five conserved regions (C1-C5) underneath. The gp41 is composed of a fusion peptide and two helices. On average 7-14 trimers of these heterodimers (gp120-gp41) can be found on the surface of each viral particle. In these so-called spikes, gp41 is at the center of the three gp120 proteins (61).

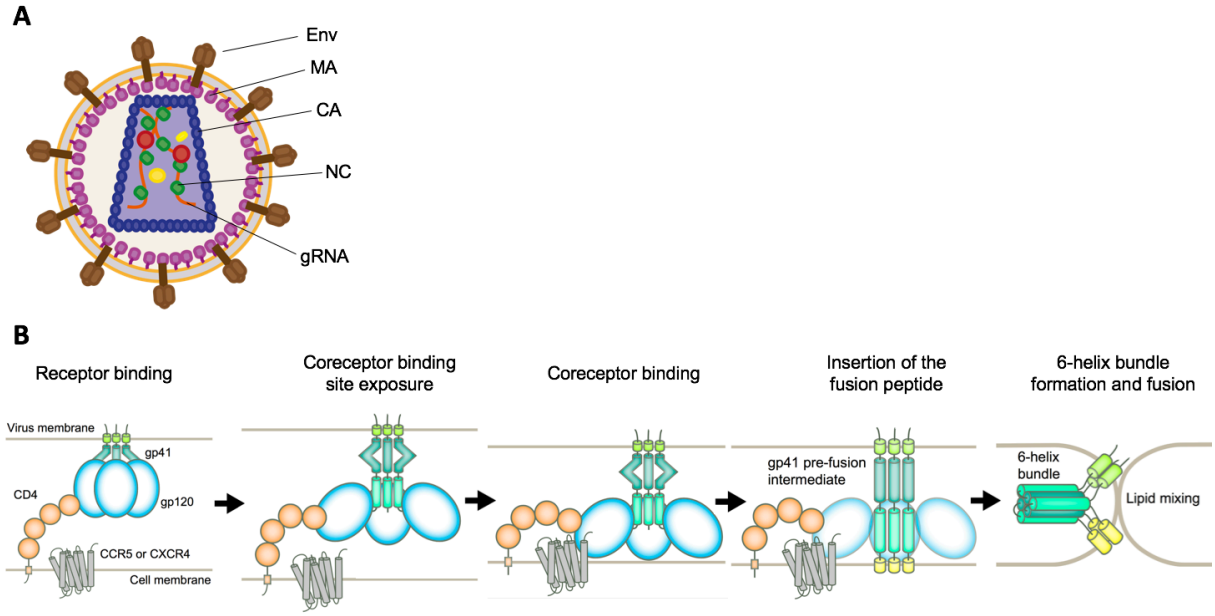


Figure 3: HIV-1 particle and attachment of the viral Env on cellular receptors followed by fusion. (A) *Schematic representation of the mature viral particle. The principal components are indicated.* ***(B)*** *The HIV Env protein is composed of gp120 and gp41. gp120 binds to its receptor CD4. This induces a conformational change which exposes the co-receptor binding site. Co-receptor binding leads to the exposure and activation of gp41. The fusion peptide of gp41 is inserted into the plasma membrane. Then gp41 forms a 6-helix bundle which stimulates fusion of the two membranes (adapted from 55).*

The main receptor of HIV is CD4, a transmembrane protein which is expressed mainly on T-cells and monocytes and is composed of four immunoglubulin-like domains (D1-D4). In addition to CD4, one of two possible coreceptors is required to allow HIV infection of target cells. The possible coreceptors are CCR5 and CXCR4, two chemokine receptors composed of seven transmembrane domains.

The first step of infection of a new target cell by HIV is a relatively nonspecific attachment on the surface of the cell through interaction of Env with heparan sulfate proteoglycans, α4β7 integrins or DC-SIGN. This helps bringing the viral particle into close proximity with the cell surface but is not absolutely required to achieve infection. However, the attachment promotes interaction of Env with its receptor CD4. This interaction is mediated by a region at the base of the variable loops V1, V2 and V5 of gp120, which binds to the most distal domain of CD4 (D1). The binding of gp120 with its receptor induces a structural rearrangement of the spike, that leads to the opening of the trimer and grants accessibility to the V3 loop of gp120 and to gp41, which were hidden at the center of the complex before (Fig. 3B). Exposure of the V3 loop allows its interaction with the coreceptor followed by another conformational change which leads to activation of gp41. The fusion peptide domain of gp41 can then insert into the cellular plasma membrane. The helices of the gp41 trimer fold back on themselves to form a six-helix bundle. This brings the cellular plasma membrane and the viral membrane into close proximity of each other and finally induces their fusion (Fig. 3B). The fusion pore then allows release of the viral core into the cytoplasm of the target cell (28, 61, 229).

1.5.2 Reverse transcription

Directly after entry of the viral core into the cytoplasm of the new host cell, reverse transcription is initiated in order to form the double-stranded proviral DNA.

The viral reverse transcriptase (RT) is a heterodimeric protein, that consists of the p66 and p51 moieties. The p66 comprises a domain capable to polymerize DNA from an RNA or a DNA template, as well as an RNase H domain, which exclusively degrades RNA in an RNA-DNA duplex. The p51 is identical to p61 except that it lacks the RNase H domain. In the heterodimer, p61 is thought to be catalytically active, while p51 only seems to play a structural role.

The first step of reverse transcription is the formation of a single-stranded negative-sense DNA intermediate, using the positive-sense viral gRNA as a template. The reverse transcription begins at a region near the 5'-end of the gRNA, called the PBS, where a $tRNA^{Lys3}$ is annealed with 18 nucleotides at its 3'-extremity. The viral RT starts adding nucleotides to the 3'-end of this tRNA until it reaches a strong stop at the 5'-end of the viral gRNA (Fig. 4). During this

process, the viral gRNA is degraded by the RNase H activity of the viral RT, leaving the newly formed negative sense DNA single-stranded. The 5'-end sequence of the viral gRNA is called R region and is repeated at the 3'-end of the viral gRNA. This allows what is called the first strand-transfer, where the newly synthetized (-) strong-stop DNA strand hybridizes with the 3'-end of the viral gRNA (Fig. 4). RT can then continue all the way along the viral gRNA, degrading the RNA along on its way. Only certain pyrimidine-rich regions, called poly-pyrimidine-tracts (ppt) resist the RNase H activity and serve as primers for synthesis of the positive-sense DNA. RT primarily uses a ppt near the 5' end of the (-) DNA to initiate synthesis of the (+) DNA, copying the 5'-extremity of (-) DNA as well as the 18 nucleotides at the 3'-extremity of $tRNA^{Lys3}$ (Fig. 4). At this stage, $tRNA^{Lys3}$ is degraded by the RNase H activity of viral RT, however a single r-adenine is maintained at the 5'-end of (-) DNA. The sequence corresponding to the $tRNA^{Lys3}$ can then hybridize with the PBS at the 3'-extremity of the (-) DNA in what is called the second strand-transfer (Fig. 4). RT can then extend both strands to form a complete double-stranded DNA. Initiation of (+) strand synthesis from a central ppt leads to the formation of a flap, which is later repaired (93, 95, 137, 145).

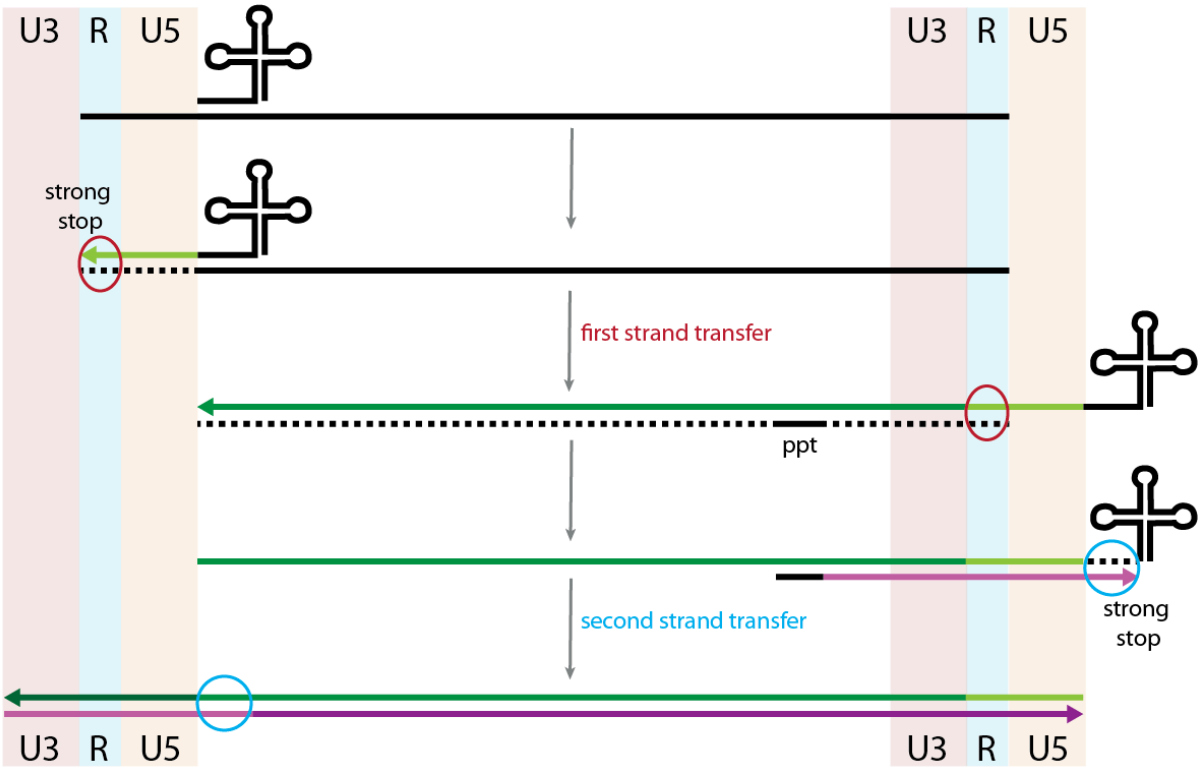


Figure 4: Reverse transcription of the HIV genomic RNA. *The viral reverse transcriptase uses the $tRNA^{Lys3}$ as a primer for synthesis of the (-) DNA strand (green) and poly-pyrimidine-tracts (ppt or $_c$ppt) for synthesis of the (+) DNA strand (violet). Dashed lines indicate degraded RNA due to RNase H activity. The R region is repeated on the 5'- and 3'-extremity of the viral gRNA and allows the first strand transfer (red circles). The region hybridizing to the $tRNA^{Lys3}$ allows the second strand transfer (blue circles). The resulting proviral DNA contains LTRs composed of the U3, R and U5 regions at both ends and a central flap ($_c$flap) due to the $_c$ppt which is later repaired.*

While it is possible to form the entire viral DNA from a single gRNA molecule, the virus actually very often makes use of both gRNA molecules inside the viral particle by jumping from one to the other leading to recombination. This contributes to the formation of mosaic viruses that combine sequences from different viral strains, like it is the case for CRFs (93, 145).
The reverse transcription complex (RTC) formed by the viral core at this stage of the viral life cycle contains multiple viral proteins that influence reverse transcription, apart from the RT itself. Notably, MA, CA, NC, IN and Vpr are present within the RTC (93). CA protects the RTC from cellular antiviral sensors in the cytoplasm while still allowing import of the necessary nucleotides for reverse transcription through pores formed at the center of each CA-hexamer (100). NC possesses a nucleic acid chaperone activity, that facilitates reverse transcription. Vpr interacts with cellular UNG2 (Uracil-DNA glycosylase), implicated in DNA-repair mechanisms (93).

1.5.3 Uncoating, Nuclear Import and Integration

During reverse transcription, the viral RTC is transported through the cytoplasm from the site of viral entry towards the nucleus using dynein-dependent transport along the microtubule network. Once reverse transcription is achieved, the so called pre-integration complex is transported into the nucleus through the nuclear pore, where viral DNA is integrated into the cellular genome (35, 232).
At some stage along these steps the viral genome has to be uncoated. While the core needs to stay assembled in order to protect the viral genome from detection by antiviral sensors in the cytoplasm, nuclear import of an intact core (50-60 nm width) seems to be impossible due to size constraints imposed by the nuclear pore (upper limit of approximately 40 nm). The cellular cyclophilin A (CypA) plays an important role in the process. CypA has been shown to interact with the viral capsid, stabilizing the viral core and thereby regulating the early steps of the viral life cycle. Several mechanisms have been proposed for uncoating. While dynein and kinesin-dependent transport of the viral core seems to favour uncoating, the main driving force for uncoating is thought to be the reverse transcription process itself. The transformation of the flexible viral gRNA into a relatively rigid double-stranded DNA leads to the build-up of pressure inside the core and this pressure ultimately induces disassembly. Nevertheless, a certain amount of capsid proteins and other viral factors present in the viral core stay associated with the pre-integration complex up until after the import into the nucleus, where they play important regulatory roles for integration into cellular DNA (4, 35, 174, 232).
While MA, IN and also Vpr contain nuclear localization signals (NLS) that might play a role in nuclear import of the pre-integration complex, CA seems to be the major player in this process,

interacting with many essential cellular factors. It has been shown that the cellular factors NUP358 (Nucleoporin 358), CPSF6 (Cleavage and polyadenylation specific factor 6), TNPO3 (Transportin 3) and NUP153 are necessary for nuclear import. The viral pre-integration complex is thought to dock to the nuclear pore by CA interaction with NUP358, a filamentous protein which is positioned at the cytoplasmic side of the nuclear pore. Then, the interaction of the pre-integration complex with CPSF6 and TNPO3 allows shuttling through the pore to the nuclear side, where NUP153 is present and helps to achieve the process (88, 137, 232).
Nuclear import is directly followed by, and to a certain extent coupled with integration of the viral DNA into the cellular genome. This is influenced by NUP153, which recruits active chromatin to the periphery of nuclear pores through binding of Tpr (Translocated promoter region protein). CPSF6 and LEDGF (Lens epithelium-derived growth factor) target the pre-integration complex to the body of actively RNA-Pol-II-transcribed genes in proximity of the nuclear pore (128, 137, 230).

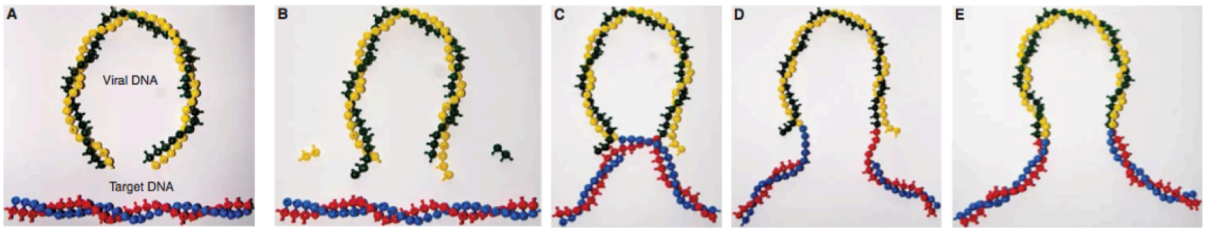


Figure 5: Integration of proviral DNA into the cellular genome. (A) *The proviral DNA (yellow+green) and the cellular DNA (red and blue).* ***(B)*** *The viral DNA is activated by 3'-end processing, which generates reactive 3'-OH residues.* ***(C)*** *The 3'-extremity of each strand attacks a target sequence in the cellular DNA, around 5 nts apart.* ***(D)*** *This generates a duplicated, 5 nts region in the cellular DNA as well as a 2 nts overhang of viral 5'-extremities due to previous 3'-end processing.* ***(E)*** *The cellular machinery repairs the gap which results in the integrated provirus (adapted from 45).*

The core of the pre-integration complex is formed by an intasome composed of an IN tetramer, one dimer of IN being associated with each end of the viral DNA. The first step of integration is 3'-end processing, during which IN cleaves the 3'-end of each DNA strand at a CA dinucleotide in order to form a hydroxyl extremity (Fig. 5A-B). This hydroxyl group is then used in a second step for a nucleophilic attack on the cellular DNA, called strand-transfer (Fig. 5C). The integration of both extremities of the viral DNA is targeted towards a sequence, 5 base pairs apart, which leads to a duplication of 5 nucleotides of the target sequence (Fig. 5D). The integration intermediate that is formed comprises a single-stranded region and an overhang, that are repaired by the cellular repair machinery (Fig. 5E) (49, 128).

1.5.4 Expression of the viral genome and latency

The virus is integrated into the cellular genome as a so-called provirus with a length of approximately 9.7 kb flanked by two long terminal repeats (LTRs) which each contain the U3, R and U5 regulatory regions (Fig. 6). U3 contains the promoter of transcription composed of one TATA box, 3 SP1 (specificity protein 1) binding sites as well as 2 enhancer regions (Fig. 6). Binding of the transcription factors NFκB (Nuclear factor kappa-light-chain-enhancer of activated B cells) or NFAT (Nuclear factor of activated T-cells) to the enhancer regions as well as recruitment of SP1 proteins to their binding sites stimulates transcription initiation and facilitates recruitment of the TFIID (Transcription factor II D) complex to the promoter through interaction of its TBP (TATA binding protein) subunit with the TATA box. This stimulates the recruitment of additional transcription factors, notably TFIIH, as well as RNA pol II. The CDK7 (Cyclin dependent kinase 7) subunit of TFIIH phosphorylates the C-terminal repeat domains of RNA pol II at the Ser5 position, which allows transcription initiation. The association of NELF (Negative elongation factor) and DSIF (DRB sensitivity inducing factor) causes stalling of RNA pol II after transcription of the first elements of the viral RNA which include the TAR (Trans-activation response) region. The TAR region allows recruitment of Tat to the transcription complex. Tat itself recruits the transcriptional stimulator p-TEFb (positive transcription elongation factor) which is composed of CDK9 and CycT1 (Cyclin T1) and is sequestrated in an inactive state in a 7SK-snRNP. CDK9 then phosphorylates NELF, DSIF as well as the Ser2 position of the C-terminal repeat domain of RNA pol II, leading to dissociation of NELF and recruitment of additional stimulatory complexes like SEC (super-elongation complex). This results in resuming of transcriptional elongation (104, 138, 154, 216).

Naive as well as memory $CD4^+$ T-cells are in a quiescent state where the overall transcriptional activity is low. Transcriptional activators like NFAT and NFκB are sequestered in the cytoplasm and thereby can't access the viral promoter. Furthermore, the chromatin is in a condensed state with several epigenetic markers, like deacylation of histones as well as methylation of histones and DNA, which repress transcriptional activation. In these cells, the HIV provirus enters an inactive state defined as latency. Activation of the host cell for example by stimulation of its TCR (T-cell receptor) can lead to reactivation of viral transcription (8, 138, 187).
The entire unspliced transcript of around 9 kb represents the viral gRNA and also serves as an mRNA for translation of the Gag and GagPol polyproteins. However, this RNA can be incompletely spliced to generate mRNAs of Vif, Vpr, Env and Vpu or completely spliced to generate mRNAs of Tat, Rev and Nef. In total, alternative splicing through the 4 different splice

donor and 8 acceptor sites (Fig. 6) generates over 100 different viral transcripts. Splicing is regulated by binding of SR proteins (Serine and arginine-rich proteins) to splicing enhancer sites and binding of hnRNPs (heterogenous nuclear RNPs) to splicing silencer regions (104, 156, 182, 201).

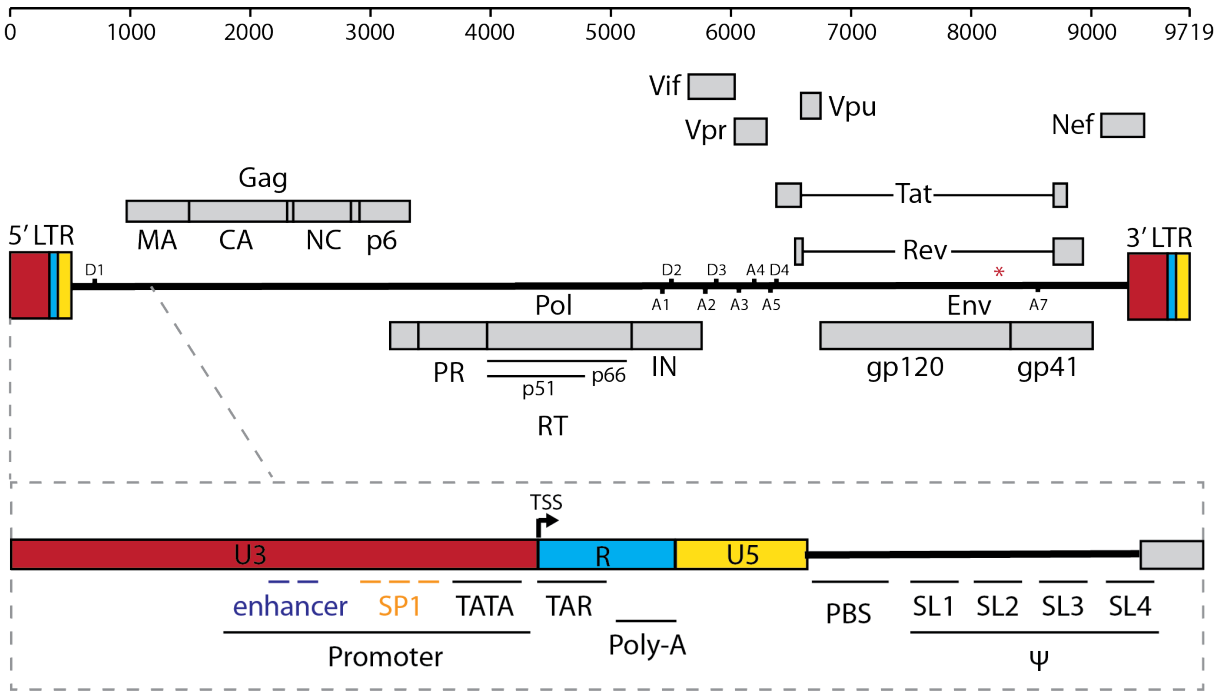


Figure 6: Organization of the HIV-1 genome. *The positions are indicated in nucleotides on the top. Protein coding sequences are indicated as grey boxes. The regulatory regions which constitute the long terminal repeats are indicated in color. The main splicing donor site positions (D) and acceptor sites (A) are indicated and numbered. The position of the RRE is indicated with a red star. The transcription promoter with enhancer sequences, SP1 binding sites and a TATA box are situated in the U3 region. The R region starts with the transcription start site (TSS) and contains the TAR and Poly-A signal. Between U5 and the Gag coding sequence are the PBS as well as the packaging signal Ψ (adapted from 59, 96, 123).*

While entirely spliced transcripts are exported into the cytoplasm using the regular Tap/p15 export machinery of most cellular mRNAs, incompletely and unspliced viral transcripts are bound by the viral Rev protein through specific recognition of the RRE (Rev response element), positioned in the Env coding region (Fig. 6). Rev binds to this element, multimerizes and recruits Crm1 (Chromosomal maintenance 1) through its nuclear export signal (NES). Crm1 in association with the GTP-bound Ran protein (ras-related nuclear protein) then mediates export of these RNAs through the nuclear pore (63, 104, 175, 182).

1.5.5 Assembly, budding and maturation

In order to form new virions, the viral proteins have to assemble at the plasma membrane. Env is a membrane protein that is co-translationally inserted into the membrane of the endoplasmic

reticulum (ER) followed by its trafficking to the plasma membrane. On its way, it is maturated through glycosylation and furin-mediated cleavage into gp120 and gp41. Gag as well as Gag-Pol are expressed in the cytoplasm and recruit the viral gRNA through interaction of the two CCHC zinc fingers of the NC domain. The gRNA is specifically selected for packaging from the pool of available cellular and spliced viral RNAs through a yet unknown mechanism which seems mediated by interaction of Gag with the stem-loop 1 of the packaging region (Ψ, Fig. 6) of the RNA and specificity is significantly influenced by the p6 domain of Gag (58, 131, 146, 221).

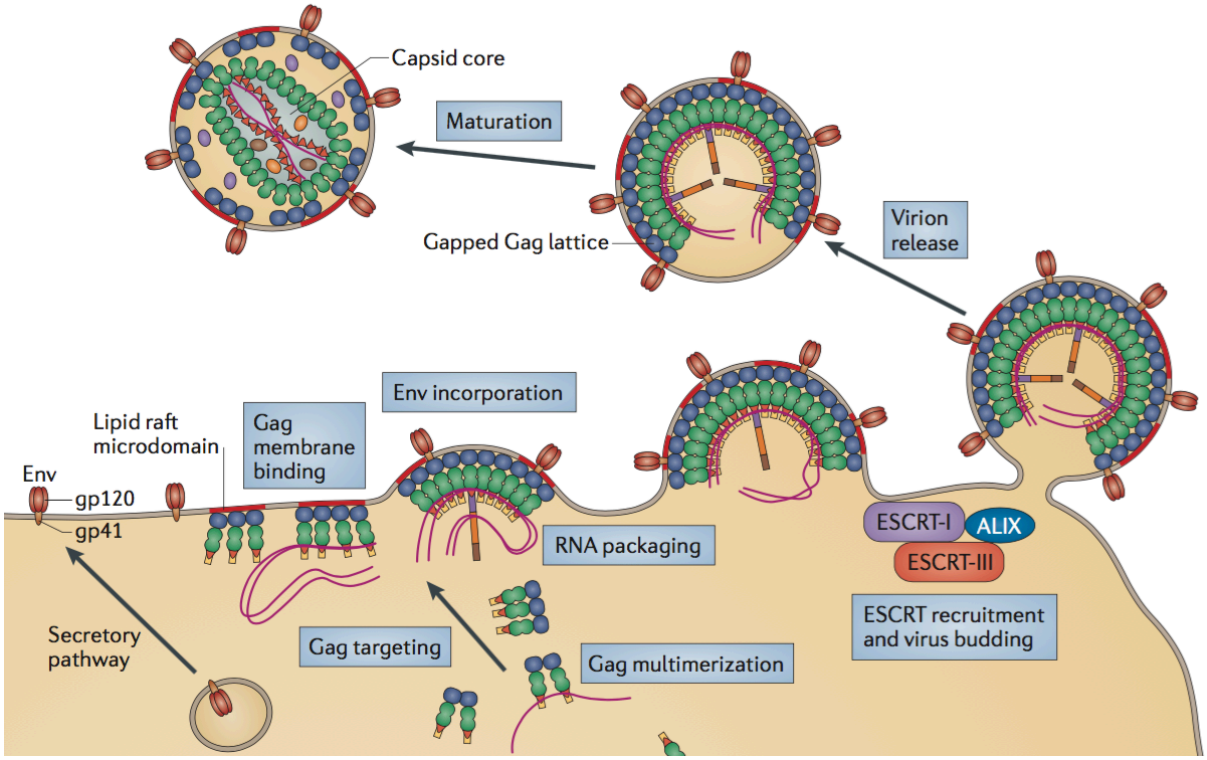


***Figure 7: HIV-1 virion assembly and maturation**. The Gag polyprotein consists of an MA domain (blue), a CA domain (green), an NC domain (red) and a p6 domain (yellow). It also exists in fusion with Pol, which is composed of the protease (violet), the reverse transcriptase (orange) and the integrase (brown). Gag recruits gRNA and Env to lipid rafts in the plasma membrane, where it multimerizes to form a budding viral particle. The p6 domain of Gag mediates recruitment of the ESCRT machinery which allows scission of the new viral particle from the cell. The viral polyproteins are then maturated by the viral protease (adapted from 60).*

HIV gRNAs form dimers through the dimer initiation site (DIS), a palindromic sequence in the Ψ region which forms a kissing-loop complex (161, 213). Gag is targeted to the membrane through its MA domain which contains many basic residues that interact preferentially with the PIP2 (Phosphatidylinositol 4,5-bisphosphate), cholesterols and sphingomyelins of lipid rafts (Fig. 7). Once at the plasma membrane, a myristic acid added post-translationally on the MA domain is inserted into the lipid bilayer in order to anchor Gag. MA also interacts with the cytoplasmic tail of gp41, which recruits Env to the forming particle. Gag multimerizes through

interactions in the CA domain and the forming particle starts budding (Fig. 7). The p6 domain mediates recruitment of the ESCRT (Endosomal sorting complex required for transport) machinery through direct interaction with the TSG101 (Tumor susceptibility gene 101) subunit of ESCRTI (Fig. 7). The subsequent recruitment of ESCRTIII and VPS4 (Vacuolar protein sorting-associated protein 4) allows scission of the new virion from the plasma membrane. During or shortly after budding, the maturation of the new particle begins through cleavage of the Gag and Gag-Pol polyproteins by the viral protease. Liberation of the different domains allows capsid formation, stabilisation of the gRNA dimer through formation of additional interactions and structural changes that render Env fusion-competent. The mature particle has a diameter of 80-150 nm and can disseminate to infect new target cells (Fig. 7) (56, 66, 96, 221).

2. Host immune defenses and viral counter-defenses

2.1 The immune system

The human immune system is characterized by an innate and an adaptive immune response, which work in concert to eliminate invading viruses. HIV infects immune cells and deeply disregulates the entire system in order to escape immune defenses (7, 50, 167).

As described above, the viral infection begins by a local amplification of HIV in $CD4^+$ cells present at the site of primary infection. This local infection triggers production of inflammatory cytokines which lead to recruitment of immune effector cells and to activation of dendritic cells. Dendritic cells patrol the entire organism in the search for pathogens. Upon encounter of an infection, they take up antigens and present them on major histocompatibility complex (MHC) class I and II molecules on their surface. This is followed by activation and maturation of the dendritic cells and their migration to lymph nodes, where they present captured antigens to $CD4^+$ T-cells. In the case of an HIV infection, the virus takes advantage of dendritic cell migration to disseminate to lymph nodes and infect $CD4^+$ cells. When a naive T-cell specifically recognizes the MHC-II-associated proteins on a dendritic cell through their TCR, it becomes activated and undergoes clonal expansion (2, 50, 167).
$CD4^+$ T-cells represent a central hub in the adaptive immune response due to their important role in activation of $CD8^+$ cytotoxic T-cells as well as antibody-producing B-cells.

When B-cells are activated, they undergo class-switch recombination and maturation of their B-cell receptor (BCR) by hypermutation. These processes are stimulated by the cytosine deaminase AID (activation-induced cytidine deaminase). B-cells with a BCR that efficiently recognizes the viral antigen undergo clonal expansion and secrete a soluble version of their BCR in the form of antibodies. Antibodies bind to viral particles and can inhibit viral attachment and entry into target cells (50). In the case of an HIV-infection, the antibody response is often poorly efficient because produced antibodies rarely achieve broad specificity, easily allowing viral escape by mutations in the targeted antigens (167).

The cytotoxic immune response seems more efficient and is thought to be in part responsible of the decline in viral titer following the acute infection. Cytotoxic $CD8^+$ T-cells recognize infected cells thanks to the viral antigens presented on the MHC-I, which is ubiquitously expressed on all cell types. $CD8^+$ T-cells then secrete cytotoxic molecules like perforin and granzyme in order to kill the target cell. Some viruses downregulate MHC-I from the cell surface in order to evade $CD8^+$ T-cell-mediated cytotoxicity, however, cells lacking MHC-I are recognized and killed by natural killer cells (NK-cells) (2, 7, 50).

The innate immune response is essential to activate and direct adaptive immune responses. Indeed, the two are connected by dendritic cells which are differentially activated by the cytokine environment present at the site of infection (2, 50, 186).

The initial sensing of the pathogen is the first step in the innate immune response. This sensing is mediated by pattern recognition receptors (PRR) which are ubiquitously expressed and specifically recognize different types of pathogen- and damage-associated molecular patterns (PAMPS and DAMPS, respectively). Amongst these PRRs are for example membrane-associated receptors like TLRs (Toll-like receptors), cytoplasmic receptors like RLRs (RIG-I-like receptors) and cytoplasmic DNA sensors. TLR-7 and 8 recognize for example viral single-stranded RNA in endosomes (Fig. 8). RIG-I (Retinoic acid-inducible gene I) and MDA-5 (Melanoma differentiation-associated protein 5) are RLRs and recognize double-stranded RNA in the cytoplasm. IFI16 (Gamma-interferon-inducible protein 16) and cGas (Cyclic GMP-AMP synthase) are DNA sensors, which can sense HIV DNA in the cytoplasm (Fig. 8). DAMPS include for example reactive oxygen species (ROS) and extracellular ATP which signalize cell damage (3, 50, 186, 225).

Upon sensing of an invading pathogen, a signaling cascade is triggered which leads to activation of NFκB, IRF3 and IRF7 (Interferon regulatory factors 3, 7), transcription factors that drive expression of antiviral effector proteins, interferons (IFN) and inflammatory cytokines (Fig. 8) (3, 176).

Type I IFNs play an important role in antiviral immune defenses. They are secreted and bind to cells through the IFNAR receptor (Interferon-α/β receptor) in an autocrine, paracrine and endocrine manner. This leads to activation of a JAK/STAT signaling cascade, resulting in expression of interferon-stimulated genes (ISGs; Fig. 8). ISGs are important mediators of the innate immune response, their expression in the cell leading to an environment that suppresses viral infection. Amongst ISGs are for example cytokines, antigen presenting molecules as well as restriction factors (42, 57, 197).

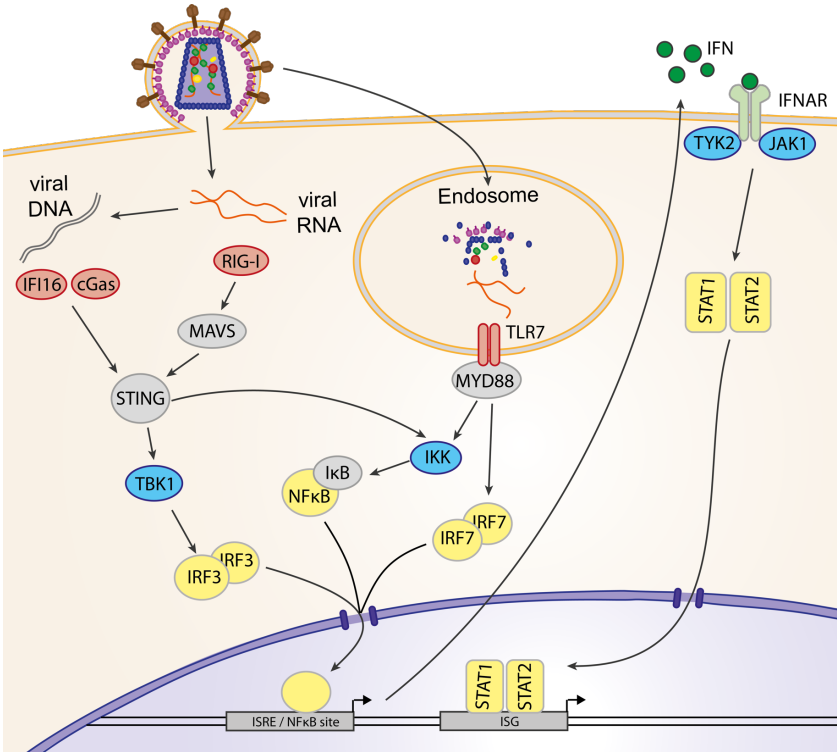


Figure 8: Detection and signaling of HIV infection by the innate immune system. *In the cytoplasm, viral RNA can be detected by RIG-I, which binds to MAVS and activates STING. Viral cDNA can be detected by IFI16 or cGAS, which directly activate STING. STING can activate TBK1 and IKK. TBK1 phosphorylates IRF3, which is activated and induces transcription of different genes. In endosomes, viral RNA can be recognized by TLR7, which recruits MYD88. This leads to activation of IKK. IKK phosphorylates IκB, which leads to detachment of IκB from NFκB. NFκB and IRF7, which is also activated by MYD88, can then translocate into the nucleus and induce transcription. IRF3, NFκB and IRF7 induce for example production of type I IFN, which is secreted and binds to its receptor IFNAR. This activates JAK1 and TYK2 which phosphorylate STAT1 and 2. STAT1 and 2 then translocate into the nucleus and induce transcription of ISGs. PRRs are represented in red, kinases in blue and transcription factors in yellow (adapted from 51, 209). TYK2: Tyrosine kinase 2; JAK1: Janus kinase 1; STAT: Signal transducer and activator of transcription; STING: Stimulator of interferon genes; IκB: Inhibitor of NFκB; IKK: IκB kinase; MYD88: Myeloid differentiation primary response gene 88; ISRE: Interferon-stimulated response element; IRF: Interferon regulatory factor.*

2.2 Restriction Factors

Restriction factors are amongst the first effectors of innate immunity against viral pathogens. Many different restriction factors have been described that inhibit the different steps in the life cycle of HIV and more are continuously being discovered. Discovery of restriction factors has mostly been performed by comparison of expression profiles of permissive cells, in which viral replication can take place even in the absence of certain viral accessory proteins, and non-permissive cells where viral replication is restricted. These non-permissive cells intrinsically express restriction factors which immediately and directly inhibit viral replication (25, 51, 71, 133). Moreover, restriction factors are also ISGs and can therefore be induced by innate immune signalling pathways. Restriction factors are not only essential effectors of innate immunity, but can also participate in the regulation of the immune response, for example by stimulating immune signalling and favouring recognition of infected cells (25, 45, 197). In summary, restriction factors apply a strong block of viral replication, which can in certain cases be part of the species barrier to infection (52, 191). The study of restriction factors is a central topic in HIV research, because it can not only contribute to the understanding of essential viral replication processes, but the enhancement of this preexisting restriction in cells could also be a strategy for novel antiviral drugs (102, 171, 192).

Viruses are under strong evolutive pressure to counteract restriction factors in order to be able to efficiently replicate in non-permissive cells. This counteraction can for example be mediated by viral auxiliary proteins (133, 190, 212, 220). Restriction and counteraction both often rely on the ubiquitin-proteasome system. This is further detailed in the following review.

viruses

MDPI

Review

Hijacking of the Ubiquitin/Proteasome Pathway by the HIV Auxiliary Proteins

Tanja Seissler, Roland Marquet and Jean-Christophe Paillart *

Université de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, UPR 9002, IBMC-15 rue René Descartes, F-67000 Strasbourg, France; t.seissler@ibmc-cnrs.unistra.fr (T.S.); r.marquet@ibmc-cnrs.unistra.fr (R.M.)
* Correspondence: jc.paillart@ibmc-cnrs.unistra.fr; Tel.: +33-388-41-71-35; Fax: +33-388-60-22-18

Received: 5 October 2017; Accepted: 30 October 2017; Published: 31 October 2017

Abstract: The ubiquitin-proteasome system (UPS) ensures regulation of the protein pool in the cell by ubiquitination of proteins followed by their degradation by the proteasome. It plays a central role in the cell under normal physiological conditions as well as during viral infections. On the one hand, the UPS can be used by the cell to degrade viral proteins, thereby restricting the viral infection. On the other hand, it can also be subverted by the virus to its own advantage, notably to induce degradation of cellular restriction factors. This makes the UPS a central player in viral restriction and counter-restriction. In this respect, the human immunodeficiency viruses (HIV-1 and 2) represent excellent examples. Indeed, many steps of the HIV life cycle are restricted by cellular proteins, some of which are themselves components of the UPS. However, HIV itself hijacks the UPS to mediate defense against several cellular restriction factors. For example, the HIV auxiliary proteins Vif, Vpx and Vpu counteract specific restriction factors by the recruitment of cellular UPS components. In this review, we describe the interplay between HIV and the UPS to illustrate its role in the restriction of viral infections and its hijacking by viral proteins for counter-restriction.

Keywords: HIV; ubiquitin; proteasome; restriction factors; TRIM5α; March8; APOBEC; SAMHD1; BST2/Tetherin

1. Introduction

The human cell is in a continuous arms race with various viruses. This has led to the coevolution of cellular restriction factors on the one hand and viral proteins for counter-defense on the other hand. Restriction factors are generally induced as a result of an interferon response—they use unique mechanisms to impair specific steps of the replication cycle and they exhibit a dominant and autonomous effect. In this continuous fight, the ubiquitin-proteasome system (UPS) plays a central role on the cellular as well as on the viral side. The cell expresses restriction factors, some of which are themselves components of the UPS, targeting viral proteins for degradation and thereby inhibiting some crucial steps of the viral life cycle. However, viruses have evolved to use the UPS to their own benefits, subverting components of the UPS to degrade restriction factors, thereby protecting themselves from the cellular defense machinery to allow their dissemination. In this review, we will describe the mechanisms by which the human immunodeficiency viruses (HIV-1 and 2) use and subvert the UPS in the continuous battle between cellular defense and viral counter-defense.

2. The Ubiquitin-Proteasome System

The ubiquitin-proteasome system (UPS) is an important pathway in the cell, ensuring regulation of the protein pool in the cytoplasm as well as in the nucleus. The UPS is constituted by three main components: the proteasome holoenzymes, several ubiquitin ligases and a large variety of deubiquitinating enzymes (DUBs) [1]. Ubiquitin is a ubiquitously expressed and well-conserved

eukaryotic peptide of 76 amino acids, which can be conjugated to proteins, mainly on their lysine residues. The addition of a single ubiquitin or small ubiquitin chains is involved in many regulatory functions, whereas poly-ubiquitination at lysine 48 (K48), corresponding to the addition of chains exceeding four ubiquitins, serves as a signal for degradation [2–5]. Ubiquitination is dependent on the ubiquitin machinery: the ubiquitin-activating enzyme E1 first forms a high-energy thiol-ester link with ubiquitin in an ATP dependent manner; ubiquitin is then transferred onto a thiol group of the ubiquitin-conjugating enzyme E2; finally, the ubiquitin ligase E3 transfers ubiquitin onto a lysine of its substrate (Figure 1A) [6,7]. In humans, there are two E1 enzymes, around 40 different E2 enzymes, which primarily determine the type of ubiquitin chain that is added, as well as over 700 different E3 ubiquitin ligases, which ensure targeting of various substrates and can be separated into two main families: HECT (Homologous to E6-AP Carboxyl Terminal) and RING (Really Interesting New Gene) ubiquitin ligases. DUBs are equally important for maturation, regulation and recycling of ubiquitin [2–5].

A protein can be conjugated to different types of polyubiquitin chains, depending on which of the seven lysine residues of ubiquitin is used to link ubiquitin moieties in the chain. Proteins ubiquitinated by K48-linked chains are mainly destined for proteasomal degradation [8,9]. The 20S proteasome is a barrel-shaped structure composed of four rings: two outer rings composed of seven α-subunits and two inner rings composed of seven β-subunits, which carry the protease activity on the inside of the ring. The 26S proteasome is formed by association of a 20S proteasome with two 19S lids, which ensure specific recognition of ubiquitinated substrates, recycling of ubiquitin through deubiquitination, unfolding of the target protein and translocation through the 20S barrel (Figure 1B) [2,10–14]. While the proteasome represents the main degradation mechanism used in cells, some membrane-associated proteins are degraded by the endo-lysosomal pathway, which can be induced by mono-ubiquitination or K63-linked polyubiquitination. In this pathway, ubiquitinated membrane proteins are endocytosed and are then recognized by the endosomal sorting complexes required for transport (ESCRT), which mediate invagination of the endosomal membrane surrounding the ubiquitinated protein. The core ESCRT machinery consists of the ESCRT-I, ESCRT-II and ESCRT-III complexes, ALIX (Apoptosis-Linked gene 2-Interactiong protein X) and VPS4 (Vacuolar Protein Sorting-associated 4). This results in the formation of small vesicles inside the endosome, thereby generating what is called a multivesicular body (MVB). This MVB can then fuse to the lysosome, where the internal vesicles and their associated proteins are degraded [3,15–17].

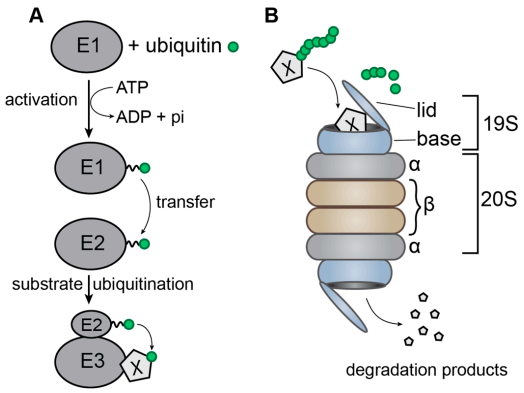


Figure 1. Schematic representation of the ubiquitin-proteasome system. (**A**) Transfer of ubiquitin from the ubiquitin-activating enzyme E1 to the ubiquitin-conjugating enzyme E2 followed by its transfer onto the target protein X by the ubiquitin ligase E3. The broken line symbolizes the thiol-ester bond; (**B**) the 26S proteasome, composed of the 20S barrel and two 19S lids. The ubiquitinated target protein X is recognized by one of the lids and translocated through the barrel where it is degraded by the proteases located on the inside of the β-rings.

The UPS plays a central role in many viral infections (reviewed in [18–20]), with five main modes of action on the viral life cycle:

(1) Some cellular E3 ubiquitin ligases recognize viral proteins and induce their ubiquitination, which can have a positive effect on viral replication. For instance, ubiquitination of the p6 domain of the HIV-1 Gag polyprotein is important for the interaction of p6 with the ESCRT machinery. However, the mono-ubiquitination of lysine residues within the p6 domain (K27 and K33) does not seem to be sufficient to facilitate budding of new virions, the latter being also dependent on the cumulative ubiquitination of NC-p2 (NucleoCapsid-peptide 2) domain [21–24]. Ubiquitination of the HIV-1 accessory protein Tat by cellular E3 ligases stimulates transcription of viral RNA [25,26].

(2) Ubiquitination of viral proteins can also induce their degradation, thereby blocking the viral life cycle. This is a strategy used by certain restriction factors. The polymerase PB1 (Protein Binding 1) of the *Influenza A virus* (IAV) for example is ubiquitinated (K48-linked ubiquitin) by the cellular E3 ubiquitin ligase TRIM32 (TRIpartite Motif-containing protein 32), followed by its degradation by the proteasome [27]. This seems to be a general mechanism as PB1 proteins derived from various IAV serotypes (H1N1 (Hemagglutinin 1 Neuraminidase 1), H3N2, H5N1 or H7N9) associate with TRIM32 in multiple cell types and this suggests that PB1 has not yet adapted to avoid TRIM32 targeting [28]. The *Human herpesvirus type I* (HSV-1) capsid protein Vp5 has also been shown to be degraded by the ubiquitin proteasome system, leaving the viral genome exposed to innate immune sensors [29]. Interestingly, TRIM5α was reported to inhibit HSV-1 and -2 replication at an early stage of the infection cycle [30], suggesting a role for this or related protein in cytosolic sensing of herpesvirus capsids.

(3) Certain viruses have evolved to recruit the cellular E3 ligases to induce the degradation of cellular proteins that might have harmful effects on the viral life cycle. For instance, the protein E6 of *Human papillomavirus* (HPV) recruits the cellular E3 ubiquitin ligase E6-AP to induce ubiquitination and degradation of p53, thereby allowing viral replication [31,32]. The NSP1 (Non-Structural RNA binding protein 1) protein of *Rotaviruses* subverts the Skp1-Cul1-Fbox (SCF) E3 ligase to induce the ubiquitination and degradation of β-TrCP (β-Transducin repeat Containing Protein). β-TrCP is by itself a substrate adaptor of an E3 ligase and its degradation leads to accumulation of the NF-$\varkappa$B inhibitor I$\varkappa$B, resulting in inhibition of the NF-$\varkappa$B induced antiviral responses [33,34]. These mechanisms are important for HIV replication and will be detailed in Section 5.

(4) Other viruses directly encode their own E3 ligases. *Kaposi sarcoma herpesvirus* (KSHV) protein K3 and K5 (RING-CH family of ligases) ubiquitinate MHC-I (Major Histocompatibility Complex I), resulting in its down-regulation from the cell surface through a clathrin-dependent sorting pathway to an endolysosomal compartment [35,36]. This endolysosomal sorting requires K63-linked instead of K48-linked polyubiquitin chains [19]. Another well-known example is the ICP0 protein (Infected Cell Protein 0) of HSV-1, an E3 ubiquitin ligase which induces the degradation of the ND10 (Nuclear Domain 10) nuclear body components PML (Promyelocytic Leukemia Protein) and Sp100 through the UPS, thereby avoiding antiviral sensing [37,38]. ICP0 has also been shown to have a RING-independent E3 ligase activity that polyubiquitinates the E2 enzyme cdc34. ICP0 influences many cellular pathways and is required for the activation of most viral and many cellular genes, for reactivation from latency and suppression of innate immunity [19].

(5) Finally, ubiquitin modifications can be reversed by the isopeptide-bond specific proteolytic activity of DUBs. In addition to cellular DUBs, it has been reported that various virus families code their own DUBs (Coronavirus, Herpesvirus etc.) to evade host antiviral immune response and promote virus replication (for a recent review see [1]). For instance, in the herpesviridae family, a variety of DUBs play an important role in the virus life cycle (e.g., UL36USP (Ubiquitin Ligase 36 Ubiquitin Specific Protease) of HSV-1, tegument protein pUL48 of human cytomegalovirus

(HCMV)). Regarding HIV-1, a recent study reported that several cellular DUBs (USP7 and USP47, Ubiquitin Specific Protease family) play an important role in its replication by regulating Gag processing and thus the infectivity of released virions and simultaneously the entry of Gag into the UPS and MHC-I pathway [39]. Moreover, this study showed that treatment with DUB inhibitors targeting USP47 causes a general Gag processing defect, indicating that USP47 interacts with Gag and prevents its entry into the UPS. Similarly, proteasome inhibitors have been shown to impact HIV-1 replication by reducing the release and maturation of infectious particles [40,41] or by suppressing its transcription [42]. Taken together, these studies suggest a potential antiretroviral activity of DUB and proteasome inhibitors.

The importance of the UPS in antiviral restriction will be discussed here using HIV as an example.

3. The HIV Life Cycle

HIV-1 and 2 are retroviruses of the genus *Lentivirus*. Their genome is composed of two (+) single stranded RNAs encoding the Gag, Pol and Env polyproteins, which correspond to the structural (matrix, capsid, nucleocapsid and p6), enzymatic (protease, reverse-transcriptase and integrase) and envelope (transmembrane and surface) viral proteins. In addition, the genome of these two viruses express two regulatory (Tat and Rev) and four auxiliary (Nef, Vpu/Vpx, Vpr and Vif) proteins, which regulate several steps in the viral life cycle [43,44]. The main difference between HIV-2 and HIV-1 is the lack of the Vpu protein in the former, which is replaced by Vpx [45]. Following viral attachment and entry into the target cell, the dimeric viral genomic RNA is partially uncoated and transported to the cell nucleus. Concomitantly, reverse transcription of the viral genomic RNA takes place to form the pre-proviral DNA, which is then integrated into the cellular genome. The integrated provirus mediates the synthesis of new full-length viral RNA (or unspliced RNA), which will be used as genomic RNA encapsidated into viral particles and as mRNA for structural and enzymatic proteins and mono- and multi-spliced viral mRNAs, which encode the viral envelope and the regulatory and auxiliary proteins in the infected cell. Finally, the components of the viral particle assemble at the plasma membrane, where new viral particles bud, maturate and disseminate to other host cells in the infected organism (Figure 2) [43,44,46–49].

During its life cycle, HIV is subjected to different cellular restriction factors (Figure 2), the first line of defense of cellular immunity. The newly discovered SERINC3 (SERine INCorporator 3) and SERINC5 proteins target the very beginning of the viral life cycle by inhibiting correct fusion of the viral envelope with the plasma membrane, thereby preventing the virus from entering into a new host cell [50,51]. IFITM (InterFeron-Induced TransMembrane) proteins 1, 2 and 3 also target the viral entry into the cell by inhibiting viral fusion with target cells. The exact mechanism of restriction is yet a matter of debate, as well as whether IFITM incorporation in virions or its expression in target cells is responsible for the antiviral effect. IFITM proteins might act on Env to inhibit its functions in viral fusion and it has been shown that some mutations in the Env protein can indeed confer resistance to IFITM restriction [52–57]. Once the virus has entered the cell, TRIM5α (TRIpartite Motif-containing protein 5α) can inhibit the early steps of the viral life cycle in a species-specific manner by accelerating viral uncoating [58–60]. The viral capsid protein also seems to be the target of Myxovirus resistance 2 (Mx2/B), a restriction factor that inhibits nuclear import and subsequent integration of the provirus through an unknown mechanism.

Some mutations in the capsid protein have been shown to confer resistance to Mx2 and particularly some mutations located at the site of interaction with cyclophilin A, an important host factor for HIV-1 infectivity [61–66]. SAMHD1 (Sterile Alpha Motif and Histidine Aspartate domain-containing protein 1) also targets the early phase of viral infection: this deoxynucleotide-triphosphohydrolase inhibits reverse transcription by depleting the pool of cellular dNTPs (deoxy Nucleotide TriPhosphates) [67,68]. During reverse transcription of the viral RNA, the restriction factor APOBEC3G (APOlipoprotein B mRNA Editing enzyme, Catalytic polypeptide-like 3G, or A3G) and other factors from the APOBEC3 family, can induce G to A hypermutations, which prevent production of functional

viral proteins [69–71]. The amount of viral proteins that are produced in an infected cell can be limited by Schlafen11 (SLFN11). Due to the bias of HIV-1 towards A/U rich codons, the virus stimulates production of corresponding tRNAs by the cell to increase viral translation, a mechanism that seems to be partly counteracted by SLFN11, which binds tRNAs in a codon-specific manner [72–74]. The final steps in the viral life cycle can be targeted by Tetherin/BST2 (Bone marrow STromal antigen 2), which inhibits release of new viral particles from the host cell [75–77] and March8 (Membrane-Associated RING-CH 8 protein), which decreases incorporation of envelope proteins into newly produced virions, thereby decreasing their infectivity [78]. Two of these restriction factors, TRIM5α and March8, use the UPS to exert their restricting activity.

HIV is able to counteract restriction factors using its accessory proteins (Figure 2): Nef prevents SERINC5 incorporation into virions by mediating its relocalization to late endosomes through interaction with the clathrin adaptor AP-2 [50,79]. Vif counteracts A3G by inducing its proteasomal degradation as well as by reducing its transcription and translation [69,80–82]. Vpx (and Vpr of certain Simian Immunodeficiency Virus (SIV) strains) counteracts SAMHD1 by inducing its proteasomal degradation [67,83,84]. Vpu (Env for HIV-2 and Nef or Vpu for SIV) counteracts BST2/Tetherin by sequestering it away from sites of viral budding [76,77,85]. Amongst these accessory proteins, Vif, Vpx and Vpu hijack the UPS to exert their counter-defense. In the following section, we will discuss in detail the restriction factors as well as the viral proteins which use the UPS for their respective mechanisms.

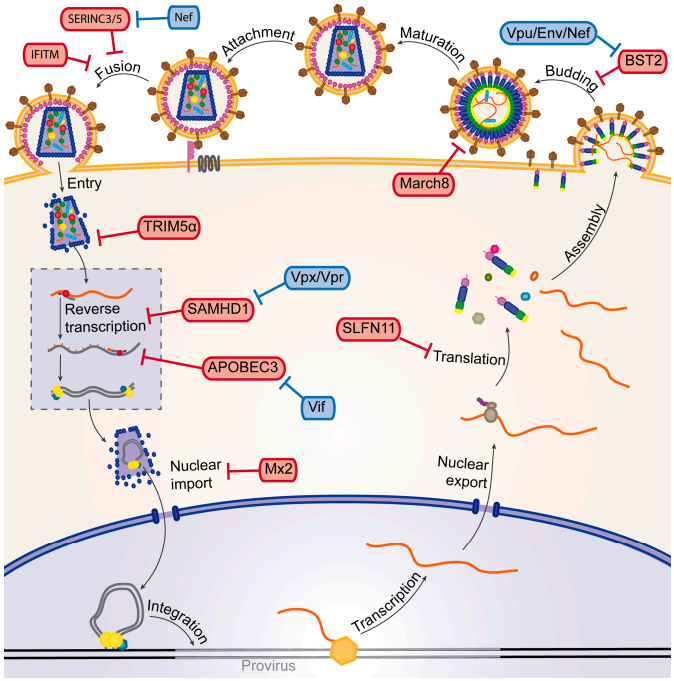


Figure 2. Schematic representation of the HIV-1 life cycle. The main HIV-1 restriction factors and the viral auxiliary proteins that counteract these factors (represented by T bars) are highlighted in red and blue boxes, respectively. See text for a description of the different steps of the life cycle.

4. Cellular Factors Mediating Viral Restriction Using the UPS

4.1. TRIM5α

One example of the cell using the UPS to restrict HIV is TRIM5α, an E3-ubiquitin ligase that interacts with the viral capsid after its entry into the cell. TRIM5α mediates a species-specific block: HIV-1 is restricted by the TRIM5α proteins of old world monkeys like rhesus or cynomolgous

monkeys, while the TRIM5α of human or new world monkeys have no or only a very weak effect on HIV-1 [59,60,86,87]. TRIM5α thereby constitutes one of the factors responsible for the interspecies barrier. The restriction of HIV-1 by TRIM5α is mediated by the interaction of the TRIM5α SPRY (SPIa and RYanodine Receptor) domain (Figure 3A) with the viral capsid in the cytoplasm of newly infected cells [59]. This interaction leads to premature decapsidation of the viral core. Moreover, viral capsid and integrase proteins are degraded (Figure 3C①) and the reverse transcription of the viral genome is inhibited in the presence of a restricting TRIM5α. These effects seem to be mediated by the UPS, since treatment with proteasome inhibitors restores a normal decapsidation rate and reverse transcription. It has also been shown that the proteasome co-localizes with TRIM5α and viral cores in the cytoplasm [88,89]. TRIM5α is also degraded by the proteasome but only in the presence of susceptible viral cores [90], suggesting that TRIM5α recruits the proteasome to the viral cores and induces their degradation. This mechanism seems to be mediated by the E3-ubiquitin ligase activity of TRIM5α, through its RING domain (Figure 3A) [58,91]. Nevertheless, TRIM5α inhibits integration of the proviral DNA independently of the proteasome, suggesting that TRIM5α uses an additional, yet uncharacterized, strategy to block viral infection (Figure 3C②) [92,93]. Finally, the association of TRIM5α with the viral capsid enhances its E3-ubiquitin ligase activity, which, in conjunction with the E2 enzyme UBC13/UEV1A (UBiquitin-Conjugating enzyme 13/Ubiquitin-conjugating Enzyme Variant 1A), leads to the synthesis of free K63-linked ubiquitin chains, thus stimulating TAK1 (Transforming growth factor β-Activated Kinase 1) and finally activating AP1 and NF-κB signaling (Figure 3C③) [94,95]. This indicates that TRIM5α, in addition to its direct antiviral activity, also functions as a sensor that induces a general antiviral state of the cell.

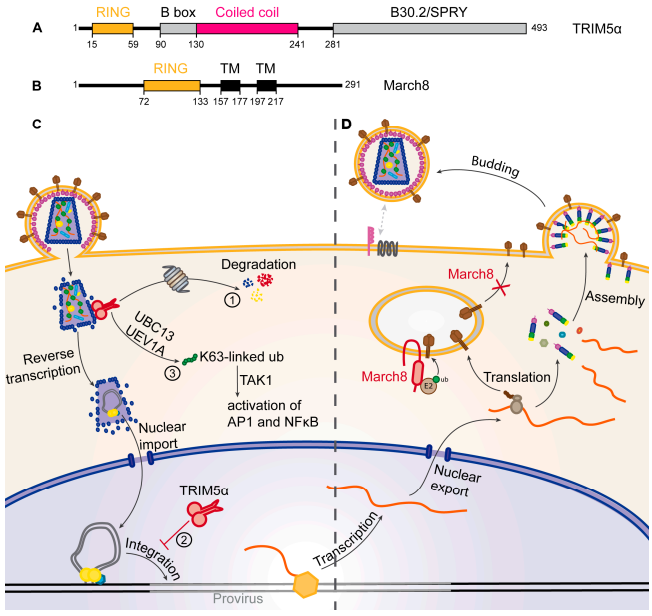


Figure 3. Restriction of HIV by TRIM5α and March8. (**A**,**B**) Schematic representation of the main domains of (**A**) the TRIM5α and (**B**) March8 proteins. Black boxes correspond to transmembrane domains (TM). Amino acid positions of the beginning and end of the domains as well as the total length of the proteins are indicated; (**C**) mechanism of TRIM5α restriction. The dimeric TRIM5α (red) recognizes the viral capsid and ① induces the proteasomal degradation of the capsid (blue), the integrase (yellow) and itself, leading to premature decapsidation of viral RNA. ② TRIM5α also blocks integration of the provirus (red T bar) and ③ induces activation of AP1 and NFκB pathways; (**D**) March8 (red) mediates intracellular retention of envelope proteins (Env, brown), leading to reduced Env incorporation into virions, thereby decreasing infectivity.

4.2. March8

March8 has recently been identified as a restriction factor of HIV-1, expressed by differentiated myeloid cells like monocyte derived macrophages and dendritic cells [78]. March8 significantly reduces infectivity of virions produced from March8-expressing cells by decreasing the number of Env-proteins incorporated into budding virions. March8 is a transmembrane E3-ubiquitin ligase, possessing an N-terminal, cytoplasmic RING domain (Figure 3B), known to downregulate multiple cellular proteins from the plasma membrane by ubiquitination followed by degradation in the endo-lysosomal pathway [96–99]. In the case of HIV-1 restriction, it has been shown that March8 interacts with Env and causes its downregulation from the cell surface. The RING-domain of March8 is necessary for this mechanism, suggesting that ubiquitination plays a role. However, Env does not seem to be degraded in the endo-lysosomal pathway like cellular proteins targeted by March8 but seems rather to be retained in intracellular compartments. March8 thus sequesters Env away from HIV-1 budding sites, thereby reducing Env incorporation into newly formed virions, making them less competent for infection of new target cells (Figure 3D) [78].

5. Counteraction of Restriction Factors by Viral Auxiliary Proteins Using the UPS

5.1. Vif

The family of Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3 (APOBEC3/A3) proteins, is a family of 7 cytosine deaminases (A3A to A3H) which induce transition of cytosine to uracil on single-stranded DNA, with a preferential recognition of CC sequence motifs by A3G and TC motifs by the others [100–102]. A3G (Figure 4A) has been the first member of this family to be identified as a potent antiviral factor. It is incorporated into budding HIV virions and is thereby carried over into the next infected cell [69]. During reverse transcription of the viral genomic RNA, the single stranded negative sense DNA is sensitive to the cytosine-deaminase activity of A3G, leading to C to U transitions [70,71]. These mutations can either be recognized by uracil DNA glycosylases, like the virion-associated UNG2 (Uracyl N-Glycosylase 2), leading to the degradation of the provirus by abasic site endonucleases [103], or they can be conserved in the provirus. Due to the sequence preference of A3G, these mutations very frequently introduce new stop codons in the viral genome, thus leading to the expression of non-functional mutated or/and truncated viral proteins (Figure 4C). HIV-1 counteracts A3G with its Vif protein, which prevents A3G incorporation into virions by inducing its degradation through the proteasome [80]. To do so, Vif recruits an SCF-like E3-ubiquitin ligase, composed of Cullin5, Rbx2, Elongin B and C. In this complex, Vif possesses the role of a substrate adaptor, directly interacting with A3G through its N-terminal domain (Figure 4B), thereby recruiting it for ubiquitination (Figure 4C③) [104].

The recruitment of Cullin5 is mediated by the zinc-binding domain of Vif [105] and Cullin5 in turn recruits the E2-ubiquitin-conjugating enzyme Rbx2. The recruitment of Elongin B and C is mediated by the BC-box domain of Vif (Figure 4B), which can be negatively regulated by phosphorylation. In this complex, not only A3G but also Vif is ubiquitinated, which might contribute to the transport of A3G to the proteasome [106]. The cellular protein HDAC6 (Histone Deacetylase 6) has been shown to play a role in this process, by inducing Vif degradation through autophagosomes as well as by protecting A3G from ubiquitination and degradation [107]. The expression level of Vif is also regulated by Mdm2 (Mouse double minute 2 homolog), an E3-ubiquitin ligase that can induce the ubiquitination of Vif and its proteasomal degradation [108]. CBF-β (Core Binding Factor β), a co-factor of the RUNX transcription factor family, is recruited by Vif and ensures its stability by inhibition of Mdm2 binding [109]. CBF-β is also necessary to allow assembly of the SCF-like E3-ubiquitin ligase mediated by Vif, resulting in the inability of Vif to induce ubiquitination and degradation of A3G in the absence of CBF-β [110,111]. Moreover, by sequestering CBF-β in the E3-ubiquitin ligase complex, Vif indirectly causes a decrease in A3G transcription as the *A3G gene* is regulated by the RUNX transcription factor family, which requires CBF-β as cofactor (Figure 4C①) [81]. Degradation of A3G through the UPS

has been known for a long time as the main mechanism for HIV-1 to counteract cellular restriction; however it has been shown that Vif can also inhibit A3G translation [82,112,113] and this inhibition significantly contributes to the counteraction mechanism (Figure 4C②) [82,112,113]. While A3G is the main member of the A3-family that efficiently restricts HIV, A3D, F and H also showed a restricting activity towards HIV-1 in the absence of Vif, even though to a lesser extent than A3G [114]. Vif is also able to recruit these A3 proteins by different motifs of its N-terminal domain (Figure 4B), thus inducing their degradation by the proteasome similarly to A3G [115–117].

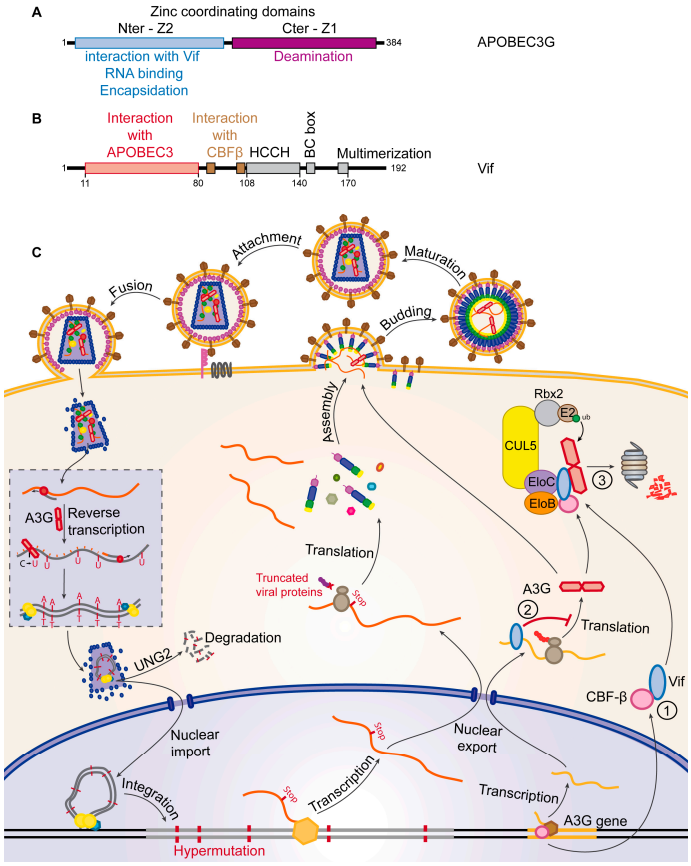


Figure 4. Restriction of HIV by APOBEC3G and counteraction by Vif. (**A**,**B**) Schematic representation of the main domains of (**A**) the APOBEC3G and (**B**) Vif proteins. Amino acid positions of the beginning and end of the domains as well as the total length of the proteins are indicated; (**C**) the mechanism of APOBEC3G restriction and Vif counteraction. APOBEC3G (red) is incorporated into virions and induces hypermutations of the provirus leading either to its degradation or production of truncated viral proteins. Vif (blue) decreases A3G transcription ①, inhibits its translation ② (Red T bar) and induces its degradation by the proteasome ③.

5.2. Vpx

Sterile alpha motif and histidine-aspartate domain-containing protein 1 (SAMHD1, Figure 5A) is a dGTP-regulated deoxynucleoside-triphosphohydrolase that catalyzes the hydrolysis of dNTPs to deoxynucleosides and inorganic triphosphate [118,119]. In non-cycling myeloid cells as well as in resting $CD4^+$ T cells, this restriction factor causes a block in the early steps of the HIV-1 life cycle [67] by depleting the intracellular pool of dNTPs [68], which leads to abortion of the viral genomic RNA reverse transcription and accumulation of defective viral cDNA (Figure 5C) [120]. This block

strongly affects infectivity of HIV-1 in these cell types but has no effect on HIV-2 infectivity [121]. Indeed, HIV-2 possesses the viral protein X (Vpx, Figure 5B) which alleviates the post-entry block mediated by SAMHD1 by inducing its degradation by the proteasome. Vpx has been found to recruit the CUL4A-DDB1-DCAF1(DDB1 and CUL4 Associated Factor 1) E3 ubiquitin ligase through a direct interaction with its substrate recognition protein DCAF1 (DDB1 and CUL4 Associated Factor 1) [122] while also interacting with the C-terminal domain of SAMHD1, thereby loading SAMHD1 onto the E3 complex and inducing its ubiquitination followed by its proteasomal degradation (Figure 5C). The nuclear localization of SAMHD1 is required for its Vpx-induced proteasomal degradation, suggesting the nuclear UPS is important in this mechanism [123,124]. Degradation of SAMHD1 leads to an increase in cellular dNTP levels and the efficiency of proviral DNA synthesis [120]. In this manner, the Vpx protein allows HIV-2 to efficiently infect human dendritic and myeloid cells and it significantly increases the infection by HIV-1 [83]. Vpx therefore seems to be an important protein for viral replication, however it is present exclusively in HIV-2 and some SIV strains.

In these lineages, the *Vpx gene* has evolved from Vpr which is present in all HIV and SIV strains and whose main function is the induction of cell cycle arrest [125–129]. Vpx and Vpr share many similarities, like for example their interaction with the same CUL4A E3 ubiquitin ligase [122,125]. Interestingly, the Vpr protein of some SIV strains has been shown to induce proteasomal degradation of SAMHD1, thereby compensating for the lack of Vpx. Indeed it seems that the ability to degrade SAMHD1 has first been acquired by the Vpr protein in certain lentiviral strains before the evolution of a separate *Vpx gene* which has subsequently conserved the function of SAMHD1 antagonism [84,130]. Nevertheless, many lineages, like HIV-1 for example, lack an anti-SAMHD1 activity. HIV Interestingly, SAMHD1 seems to be regulated in cells by phosphorylation mediated by CDK6-(Cyclin-Dependent Kinase 6) dependent CDK2, which links its activity to cell cycle control. Indeed, SAMHD1 is phosphorylated in cycling cells which blocks its activity as a dNTP hydrolase [131]. This correlates with the permissiveness of cycling cells for HIV-1 infection as opposed to non-cycling cells. Moreover, HIV infection is made possible despite the lack of a viral factor counteracting SAMHD1 by different cellular proteins: CD81 for example has recently been shown to favor HIV-1 infection by interacting with SAMHD1 and stimulation of its proteasome-dependent degradation [132]. Cyclin L2 also induces SAMHD1 proteasomal degradation through interaction with SAMHD1 and DCAF1, a mechanism interestingly similar to the one used by Vpx [133].

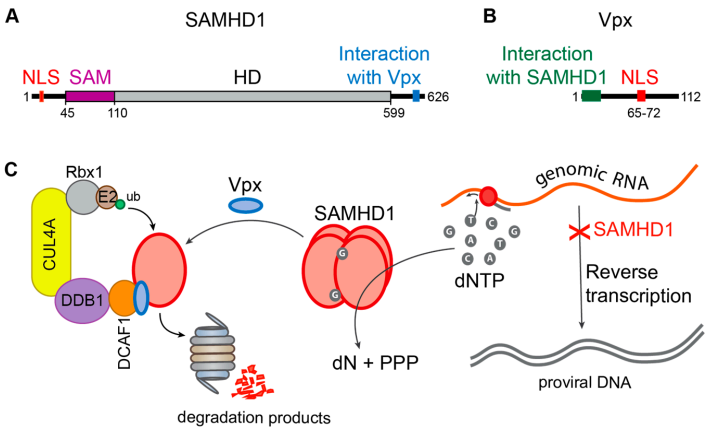


Figure 5. Restriction of HIV by SAMHD1 and counteraction by Vpx. (**A**,**B**) Schematic representation of the main domains of (**A**) SAMHD1 and (**B**) Vpx. The nuclear localization signal (NLS) is indicated in red. Amino acid positions of the beginning and end of the domains as well as the total length of the proteins are indicated; (**C**) the mechanism of SAMHD1 restriction and Vpx counteraction. Tetrameric SAMHD1 (red) hydrolyzes dNTPs, leading to a block of reverse transcription of the viral genome. Vpx (blue) induces SAMHD1 ubiquitination followed by its degradation by the proteasome.

5.3. Vpu

In the absence of Vpu, newly formed virions remain tethered to the plasma membrane of their host cell after budding and are eventually endocytosed and degraded [75]. The cellular restriction factor responsible for the block of virion release is Tetherin/BST-2. BST-2 is found as a disulfide-bond-linked dimer which is anchored into the plasma membrane by two domains: a transmembrane domain close to its N-terminus and an extracellular C-terminal glycosyl-phosphatidylinositol (GPI)-anchor (Figure 6A) [134]. These two domains mediate virion-tethering to the host cell, one remaining in the plasma membrane and the other one being inserted into the viral envelope (Figure 6C). It has been shown that this tethering involves approximately a dozen of BST-2 dimers and that among the two membrane-associated domains, the GPI-anchor is preferentially incorporated into budding virions [135]. The extracellular domain of BST-2 thereby acts like a molecular ruler, maintaining the virus at a constant distance of the plasma membrane, preventing it from disseminating to other target cells [134].

The viral protein Vpu counteracts BST-2 by direct interaction of their transmembrane domains embedded in the plasma membrane [136]. The exact mode of action of Vpu is still a matter of debate, but it seems clear now that Vpu sequesters BST-2 away from virion budding sites, thereby preventing it from incorporation into the viral envelope (Figure 6C①) [77,85,137,138]. Several studies have shown that in the presence of Vpu, newly synthesized BST-2 is sequestered in intracellular compartments, particularly the trans-golgi-network (Figure 6C②). This finally results in the downregulation of surface levels of BST-2, thereby allowing normal rates of virion release in the presence of Vpu [77,85,137,138]. BST-2 is constitutively regulated by ubiquitination and lysosomal degradation mediated by the cellular E3 ubiquitin ligases March8 and NEDD4 (Neural precursor cell Expressed Developmentally Down-regulated protein 4) [139]. It is still a matter of debate however, whether Vpu also uses the endo-lysosomal system for BST-2 counteraction. The E3-ubiquitin ligase adaptor β-TrCP is known to be recruited by the cytoplasmic DSGxxS motif of Vpu (Figure 6B) [140], which might lead to ubiquitination of BST-2 followed by its degradation in the endo-lysosomal system (Figure 6C③) [137,141].

Even though the interaction of Vpu with β-TrCP, as well as the capacity of β-TrCP to recruit an E3-ubiquitin ligase seem to be required for BST-2 counteraction by Vpu [137,142,143], conflicting data have also been reported [144–146]. Certain components of the autophagy pathway, as well as clathrin adaptors AP-1 and 2 and components of the ESCRT system might also be involved in the downregulation of BST-2 by Vpu, which would corroborate transport of BST-2 in the endosomal system [137,147–149]. However, degradation of BST-2 might not be absolutely required for viral counteraction of BST-2, since Vpu is capable of intracellular sequestration of BST-2 independently of its degradation [85,138]. The guanylate binding protein 5 (GBP5) has very recently been discovered as a new restriction factor of HIV-1 infection, that interferes with viral Env proteins, thereby decreasing infectivity of produced virions [150,151]. As Vpu and Env are expressed from the same transcript by leaky scanning, the loss of Vpu expression can in this case lead to an increase of Env expression, as observed in the macrophage tropic AD8 isolate [152], allowing the virus to partly overcome GBP5 restriction. Surprisingly, such Vpu mutants seem to occur frequently despite the presence of BST-2. HIV-2 and SIV are also counteracted by BST-2 proteins expressed by their respective host species in a species-dependent manner, but some of them lack Vpu to counteract this mechanism. It has been shown that the HIV-2 Env protein can enhance virion release in the presence of BST-2 thereby substituting for Vpu [153,154]. Certain SIV strains, like SIVagm, SIVblu and SIVmac also lack the *Vpu gene* and rely on the accessory protein Nef to counteract BST-2. Other SIV strains like SIVmon, SIVmus, SIVgsn and SIVden express Vpu and use it to counteract BST-2. Even though SIVgor and SIVcpz express Vpu, Nef seems to take over the role of BST-2 counteraction. This gives interesting clues about the evolution of HIV and SIV strains [155–157].

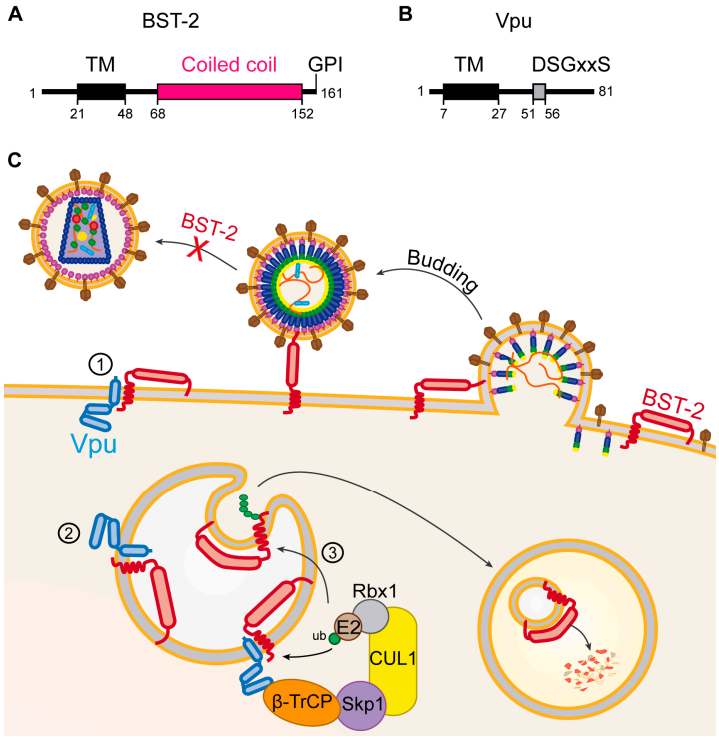


Figure 6. Restriction of HIV by BST-2 and counteraction by Vpu. (**A**,**B**) Schematic representation of the main domains of (**A**) the BST-2 and (**B**) Vpu proteins. Black boxes indicate transmembrane domains (TM). The glycosyl-phosphatidylinositol (GPI) modification at the C-terminal end of BST-2 is indicated. Amino acid positions of the beginning and end of the domains as well as the total length of the proteins are indicated; (**C**) mechanism of BST-2 restriction and Vpu counteraction. BST-2 tethers virions to the plasma membrane, thereby hindering their dissemination. Vpu sequesters BST-2 away from virion budding sites either at the plasma membrane ① or in intracellular compartments ②. Vpu can also induce BST-2 degradation in the endo-lysosomal pathway ③.

6. Other Cellular Proteins Targeted by the Hijacked UPS

The UPS is hijacked by HIV and plays an important role for the viral defense against multiple cellular restriction mechanisms. Apart from restriction factors, several other cellular proteins can also be targeted by HIV through the UPS. The viral auxiliary protein Vpu for example possesses the ability to associate with the CUL1-Skp1 E3 ubiquitin ligase through interaction with its substrate receptor β-TrCP. This association not only seems to play a role in the counteraction of BST-2 but has also been shown to induce degradation of the HIV receptor CD4. Indeed, Vpu induces CD4 ubiquitination followed by its extraction from the Endoplasmic Reticulum (ER) [140,158–162]. The mechanism used by Vpu to induce CD4 depletion involves the cellular ER-associated degradation (ERAD) pathway, which operates as a quality control mechanism to dispose of unwanted ER membrane proteins into the cytosol for subsequent proteasomal degradation. The dislocation of protein from the membrane is achieved by the recruitment of the VCP-UFD1L-NPL4 (Valosin-containing protein-Ubiquitin fusion degradation protein 1-Nuclear protein localization protein 4) complex through recognition by UFD1L of K48-linked poly-ubiquitin chains on the CD4 cytosolic tail. Interestingly, the degradation of CD4 depends also on ubiquitination of serine/threonine residues [140,158–161]. The ATPase activity of VCP then drives dislocation of CD4 from the ER membrane into the cytosol and eventually its degradation in proteasomes. The multiple levels at which Vpu acts to prevent export of CD4 from the ER underscore the importance of ensuring complete depletion of CD4 from the plasma membrane for progression of

the infection [143,160,161,163]. Other targets of Vpu-induced ubiquitination and degradation include the cell surface glycoprotein ICAM-1 and the amino acid transporter SNAT-1, both involved in immune signaling [164,165].

It is well established that the viral auxiliary protein Vpr associates with the CUL4A-RING E3 ligase through interaction with its substrate recognition subunit DCAF1. This complex has been shown to induce ubiquitination followed by proteasomal degradation of the DNA glycosylase UNG2. Thereby Vpr reduces encapsidation of UNG2, ultimately contributing to the protection against the restriction factor A3G. UNG2 recognizes C to U mutations induced by A3G and generates abasic sites, leading to degradation of viral DNA. Indeed a virus lacking Vif can be partially rescued by Vpr-mediated reduction of UNG2 compared to viruses lacking both Vif and Vpr [166–168]. Moreover, it has recently been shown that Vpr can also induce the degradation of A3G itself through the UPS [169]. Vpr seems to also enhance HIV-1 production in macrophages by UPS-mediated degradation of the cellular protein Dicer, which is involved in RNA silencing [170]. The main function of Vpr known to date is the induction of a cell cycle arrest at the G2 phase. The association of Vpr with the CUL4A E3 ubiquitin ligase has been shown to be important for this process, although the exact mechanism is still unknown [125–128]. Cell cycle arrest seems to involve Vpr association with the SLX4-SLX1-MUS81-EME1 complex, leading to SLX4 (Structure-specific endonuclease subunit) activation and ultimately proteasomal degradation of MUS81 (Crossover junction endonuclease) and EME1 (Essential Meiotic Structure-Specific Endonuclease 1) [127,128]. Vpr also induces the degradation of multiple other cellular proteins such as the DNA translocase HLTF (Helicase-Like Transcription Factor) [171], the DNA replication factor MCM10 (Mini Chromosome Maintenance 10) [172], as well as the chromatin associated proteins ZIP (leucine Zipper), sZIP and class I HDACs (Histone Deacetylase 6) [173,174].

7. Conclusions

The UPS plays an important role in viral infections in general and especially in the process of viral restriction and counter-restriction. In this continuous battle between the virus and the cell, the UPS constitutes an efficient tool for both sides. Several HIV auxiliary proteins have evolved the ability to interact with components of the UPS, subverting it for its own means. This allows the targeting of a multitude of different cellular proteins through a single platform. This strategy is not limited to HIV, but is used by a plethora of different viruses to ensure various aspects of their life cycles. Overall, the specific degradation of certain cellular proteins in the UPS allows viruses to generate a favorable environment for their own replication. The almost universal role of the UPS in counteraction of cellular restriction factors by HIV makes the UPS an interesting target for antiviral therapy. One of the main difficulties in therapy-design against HIV is the rapid evolution of the virus, which easily escapes therapeutic molecules by mutation of the targeted viral proteins. Targeting the human UPS represents a promising antiviral strategy because it would allow to avoid the escape through mutations [175,176]. A better knowledge on how the virus hijacks the UPS and which components are involved in viral replication is crucial in this attempt.

Acknowledgments: This work was supported by the CNRS and grants from SIDACTION and the French National Agency for Research on AIDS and Viral Hepatitis (ANRS) to Jean-Christophe Paillart and by a doctoral fellowship from the French Ministry of Higher Education and Research to Tanja Seissler.

Author Contributions: Jean-Christophe Paillart and Tanja Seissler conceived the review topic; Tanja Seissler drafted the manuscript and generated the figures. Jean-Christophe Paillart and Roland Marquet corrected and edited the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bailey-Elkin, B.A.; Knaap, R.C.M.; Kikkert, M.; Mark, B.L. Structure and function of viral deubiquitinating enzymes. *J. Mol. Biol.* **2017**. [CrossRef] [PubMed]

2. Roos-Mattjus, P.; Sistonen, L. The ubiquitin-proteasome pathway. *Ann. Med.* **2004**, *36*, 285–295. [CrossRef] [PubMed]
3. Clague, M.J.; Urbé, S. Ubiquitin: Same molecule, different degradation pathways. *Cell* **2010**, *143*, 682–685. [CrossRef] [PubMed]
4. Davis, M.E.; Gack, M.U. Ubiquitination in the antiviral immune response. *Virology* **2015**, *479–480*, 52–65. [CrossRef] [PubMed]
5. Petroski, M.D.; Deshaies, R.J. Function and regulation of cullin-RING ubiquitin ligases. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 9–20. [CrossRef] [PubMed]
6. Hershko, A.; Heller, H.; Elias, S.; Ciechanover, A. Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. *J. Biol. Chem.* **1983**, *258*, 8206–8214. [PubMed]
7. Scheffner, M.; Nuber, U.; Huibregtse, J.M. Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. *Nature* **1995**, *373*, 81–83. [CrossRef] [PubMed]
8. Gregori, L.; Poosch, M.S.; Cousins, G.; Chau, V. A uniform isopeptide-linked multiubiquitin chain is sufficient to target substrate for degradation in ubiquitin-mediated proteolysis. *J. Biol. Chem.* **1990**, *265*, 8354–8357. [PubMed]
9. Chau, V.; Tobias, J.W.; Bachmair, A.; Marriott, D.; Ecker, D.J.; Gonda, D.K.; Varshavsky, A. A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* **1989**, *243*, 1576–1583. [CrossRef] [PubMed]
10. Collins, G.A.; Goldberg, A.L. The logic of the 26S proteasome. *Cell* **2017**, *169*, 792–806. [CrossRef] [PubMed]
11. Matthews, W.; Driscoll, J.; Tanaka, K.; Ichihara, A.; Goldberg, A.L. Involvement of the proteasome in various degradative processes in mammalian cells. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 2597–2601. [CrossRef] [PubMed]
12. Bashore, C.; Dambacher, C.M.; Goodall, E.A.; Matyskiela, M.E.; Lander, G.C.; Martin, A. Ubp6 deubiquitinase controls conformational dynamics and substrate degradation of the 26S proteasome. *Nat. Struct. Mol. Biol.* **2015**, *22*, 712–719. [CrossRef] [PubMed]
13. Hamazaki, J.; Hirayama, S.; Murata, S. Redundant roles of rpn10 and rpn13 in recognition of ubiquitinated proteins and cellular homeostasis. *PLoS Genet.* **2015**, *11*, e1005401. [CrossRef] [PubMed]
14. Harshbarger, W.; Miller, C.; Diedrich, C.; Sacchettini, J. Crystal structure of the human 20S proteasome in complex with carfilzomib. *Structure* **2015**, *23*, 418–424. [CrossRef] [PubMed]
15. Hurley, J.H.; Emr, S.D. The Escrt complexes: Structure and mechanism of a membrane-trafficking network. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 277–298. [CrossRef] [PubMed]
16. Lauwers, E.; Jacob, C.; André, B. K63-linked ubiquitin chains as a specific signal for protein sorting into the multivesicular body pathway. *J. Cell Biol.* **2009**, *185*, 493–502. [CrossRef] [PubMed]
17. Katzmann, D.J.; Babst, M.; Emr, S.D. Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I. *Cell* **2001**, *106*, 145–155. [CrossRef]
18. Calistri, A.; Munegato, D.; Carli, I.; Parolin, C.; Palù, G. The ubiquitin-conjugating system: Multiple roles in viral replication and infection. *Cells* **2014**, *3*, 386–417. [CrossRef] [PubMed]
19. Randow, F.; Lehner, P.J. Viral avoidance and exploitation of the ubiquitin system. *Nat. Cell Biol.* **2009**, *11*, 527–534. [CrossRef] [PubMed]
20. Biard-Piechaczyk, M.; Borel, S.; Espert, L.; de Bettignies, G.; Coux, O. HIV-1, ubiquitin and ubiquitin-like proteins: The dialectic interactions of a virus with a sophisticated network of post-translational modifications. *Biol. Cell* **2012**, *104*, 165–187. [CrossRef] [PubMed]
21. Ott, D.E.; Coren, L.V.; Copeland, T.D.; Kane, B.P.; Johnson, D.G.; Sowder, R.C.; Yoshinaka, Y.; Oroszlan, S.; Arthur, L.O.; Henderson, L.E. Ubiquitin is covalently attached to the p6Gag proteins of human immunodeficiency virus type 1 and simian immunodeficiency virus and to the p12Gag protein of moloney murine leukemia virus. *J. Virol.* **1998**, *72*, 2962–2968. [PubMed]
22. Gottwein, E.; Jäger, S.; Habermann, A.; Kräusslich, H.-G. Cumulative mutations of ubiquitin acceptor sites in human immunodeficiency virus type 1 gag cause a late budding defect. *J. Virol.* **2006**, *80*, 6267–6275. [CrossRef] [PubMed]
23. Garrus, J.E.; von Schwedler, U.K.; Pornillos, O.W.; Morham, S.G.; Zavitz, K.H.; Wang, H.E.; Wettstein, D.A.; Stray, K.M.; Côté, M.; Rich, R.L.; et al. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* **2001**, *107*, 55–65. [CrossRef]

24. Ott, D.E.; Coren, L.V.; Chertova, E.N.; Gagliardi, T.D.; Schubert, U. Ubiquitination of HIV-1 and MuLV Gag. *Virology* **2000**, *278*, 111–121. [CrossRef] [PubMed]
25. Brès, V.; Kiernan, R.E.; Linares, L.K.; Chable-Bessia, C.; Plechakova, O.; Tréand, C.; Emiliani, S.; Peloponese, J.-M.; Jeang, K.-T.; Coux, O.; et al. A non-proteolytic role for ubiquitin in Tat-mediated transactivation of the HIV-1 promoter. *Nat. Cell Biol.* **2003**, *5*, 754–761. [CrossRef] [PubMed]
26. Faust, T.B.; Li, Y.; Jang, G.M.; Johnson, J.R.; Yang, S.; Weiss, A.; Krogan, N.J.; Frankel, A.D. PJA2 ubiquitinates the HIV-1 Tat protein with atypical chain linkages to activate viral transcription. *Sci. Rep.* **2017**, *7*, 45394. [CrossRef] [PubMed]
27. Fu, B.; Wang, L.; Ding, H.; Schwamborn, J.C.; Li, S.; Dorf, M.E. TRIM32 senses and restricts influenza a virus by ubiquitination of PB1 polymerase. *PLoS Pathog.* **2015**, *11*. [CrossRef] [PubMed]
28. Van Tol, S.; Hage, A.; Giraldo, M.I.; Bharaj, P.; Rajsbaum, R. The TRIMendous role of TRIMs in virus-host interactions. *Vaccines* **2017**, *5*. [CrossRef] [PubMed]
29. Horan, K.A.; Hansen, K.; Jakobsen, M.R.; Holm, C.K.; Søby, S.; Unterholzner, L.; Thompson, M.; West, J.A.; Iversen, M.B.; Rasmussen, S.B.; et al. Proteasomal degradation of herpes simplex virus capsids in macrophages releases DNA to the cytosol for recognition by DNA sensors. *J. Immunol.* **2013**, *190*, 2311–2319. [CrossRef] [PubMed]
30. Reszka, N.; Zhou, C.; Song, B.; Sodroski, J.G.; Knipe, D.M. Simian TRIM5α proteins reduce replication of herpes simplex virus. *Virology* **2010**, *398*, 243–250. [CrossRef] [PubMed]
31. Scheffner, M.; Werness, B.A.; Huibregtse, J.M.; Levine, A.J.; Howley, P.M. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* **1990**, *63*, 1129–1136. [CrossRef]
32. Scheffner, M.; Huibregtse, J.M.; Vierstra, R.D.; Howley, P.M. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell* **1993**, *75*, 495–505. [CrossRef]
33. Graff, J.W.; Ettayebi, K.; Hardy, M.E. Rotavirus NSP1 inhibits NFκB activation by inducing proteasome-dependent degradation of β-TrCP: A novel mechanism of IFN antagonism. *PLoS Pathog.* **2009**, *5*. [CrossRef] [PubMed]
34. Morelli, M.; Dennis, A.F.; Patton, J.T. Putative E3 ubiquitin ligase of human rotavirus inhibits NF-κB activation by using molecular mimicry to Target β-TrCP. *mBio* **2015**, *6*. [CrossRef] [PubMed]
35. Ishido, S.; Wang, C.; Lee, B.S.; Cohen, G.B.; Jung, J.U. Downregulation of major histocompatibility complex class I molecules by Kaposi's sarcoma-associated herpesvirus K3 and K5 proteins. *J. Virol.* **2000**, *74*, 5300–5309. [CrossRef] [PubMed]
36. Lorenzo, M.E.; Jung, J.U.; Ploegh, H.L. Kaposi's sarcoma-associated herpesvirus K3 utilizes the ubiquitin-proteasome system in routing class major histocompatibility complexes to late endocytic compartments. *J. Virol.* **2002**, *76*, 5522–5531. [CrossRef] [PubMed]
37. Chelbi-Alix, M.K.; de Thé, H. Herpes virus induced proteasome-dependent degradation of the nuclear bodies-associated PML and Sp100 proteins. *Oncogene* **1999**, *18*, 935–941. [CrossRef] [PubMed]
38. Everett, R.D.; Rechter, S.; Papior, P.; Tavalai, N.; Stamminger, T.; Orr, A. PML contributes to a cellular mechanism of repression of herpes simplex virus type 1 infection that is inactivated by ICP0. *J. Virol.* **2006**, *80*, 7995–8005. [CrossRef] [PubMed]
39. Setz, C.; Friedrich, M.; Rauch, P.; Fraedrich, K.; Matthaei, A.; Traxdorf, M.; Schubert, U. Inhibitors of deubiquitinating enzymes block HIV-1 replication and augment the presentation of Gag-derived MHC-I epitopes. *Viruses* **2017**, *9*. [CrossRef] [PubMed]
40. Schubert, U.; Ott, D.E.; Chertova, E.N.; Welker, R.; Tessmer, U.; Princiotta, M.F.; Bennink, J.R.; Krausslich, H.G.; Yewdell, J.W. Proteasome inhibition interferes with gag polyprotein processing, release, and maturation of HIV-1 and HIV-2. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 13057–13062. [CrossRef] [PubMed]
41. Ott, D.E.; Coren, L.V.; Sowder, R.C.; Adams, J.; Schubert, U. Retroviruses have differing requirements for proteasome function in the budding process. *J. Virol.* **2003**, *77*, 3384–3393. [CrossRef] [PubMed]
42. Yu, L.; Mohanram, V.; Simonson, O.E.; Smith, C.I.E.; Spetz, A.-L.; Mohamed, A.J. Proteasome inhibitors block HIV-1 replication by affecting both cellular and viral targets. *Biochem. Biophys. Res. Commun.* **2009**, *385*, 100–105. [CrossRef] [PubMed]
43. Frankel, A.D.; Young, J.A. HIV-1: Fifteen proteins and an RNA. *Annu. Rev. Biochem.* **1998**, *67*, 1–25. [CrossRef] [PubMed]
44. Freed, E.O. HIV-1 replication. *Somat. Cell Mol. Genet.* **2001**, *26*, 13–33. [CrossRef] [PubMed]

45. Azevedo-Pereira, J.M.; Santos-Costa, Q. HIV interaction with human host: HIV-2 as a model of a less virulent infection. *AIDS Rev.* **2016**, *18*, 44–53. [PubMed]
46. Jakobsdottir, G.M.; Iliopoulou, M.; Nolan, R.; Alvarez, L.; Compton, A.A.; Padilla-Parra, S. On the whereabouts of HIV-1 cellular entry and its fusion ports. *Trends Mol. Med.* **2017**. [CrossRef] [PubMed]
47. Cimarelli, A.; Darlix, J.-L. HIV-1 reverse transcription. *Methods Mol. Biol.* **2014**, *1087*, 55–70. [CrossRef] [PubMed]
48. Craigie, R.; Bushman, F.D. HIV DNA integration. *Cold Spring Harb. Perspect. Med.* **2012**, *2*, a006890. [CrossRef] [PubMed]
49. Sundquist, W.I.; Kräusslich, H.-G. HIV-1 assembly, budding, and maturation. *Cold Spring Harb. Perspect. Med.* **2012**, *2*, a006924. [CrossRef] [PubMed]
50. Usami, Y.; Wu, Y.; Göttlinger, H.G. SERINC3 and SERINC5 restrict HIV-1 infectivity and are counteracted by Nef. *Nature* **2015**, *526*, 218–223. [CrossRef] [PubMed]
51. Rosa, A.; Chande, A.; Ziglio, S.; de Sanctis, V.; Bertorelli, R.; Goh, S.L.; McCauley, S.M.; Nowosielska, A.; Antonarakis, S.E.; Luban, J.; et al. HIV-1 Nef promotes infection by excluding SERINC5 from virion incorporation. *Nature* **2015**, *526*, 212–217. [CrossRef] [PubMed]
52. Compton, A.A.; Bruel, T.; Porrot, F.; Mallet, A.; Sachse, M.; Euvrard, M.; Liang, C.; Casartelli, N.; Schwartz, O. IFITM proteins incorporated into HIV-1 virions impair viral fusion and spread. *Cell Host Microbe* **2014**, *16*, 736–747. [CrossRef] [PubMed]
53. Lu, J.; Pan, Q.; Rong, L.; Liu, S.-L.; Liang, C. The IFITM proteins inhibit HIV-1 infection. *J. Virol.* **2011**, *85*, 2126–2137. [CrossRef] [PubMed]
54. Wang, Y.; Pan, Q.; Ding, S.; Wang, Z.; Yu, J.; Finzi, A.; Liu, S.-L.; Liang, C. The V3 loop of HIV-1 Env determines viral susceptibility to IFITM3 impairment of viral infectivity. *J. Virol.* **2017**, *91*, e02441-16. [CrossRef] [PubMed]
55. Tartour, K.; Appourchaux, R.; Gaillard, J.; Nguyen, X.-N.; Durand, S.; Turpin, J.; Beaumont, E.; Roch, E.; Berger, G.; Mahieux, R.; et al. IFITM proteins are incorporated onto HIV-1 virion particles and negatively imprint their infectivity. *Retrovirology* **2014**, *11*, 103. [CrossRef] [PubMed]
56. Yu, J.; Li, M.; Wilkins, J.; Ding, S.; Swartz, T.H.; Esposito, A.M.; Zheng, Y.-M.; Freed, E.O.; Liang, C.; Chen, B.K.; et al. IFITM proteins restrict HIV-1 infection by antagonizing the envelope glycoprotein. *Cell Rep.* **2015**, *13*, 145–156. [CrossRef] [PubMed]
57. Foster, T.L.; Wilson, H.; Iyer, S.S.; Coss, K.; Doores, K.; Smith, S.; Kellam, P.; Finzi, A.; Borrow, P.; Hahn, B.H.; et al. Resistance of transmitted founder HIV-1 to IFITM-mediated restriction. *Cell Host Microbe* **2016**, *20*, 429–442. [CrossRef] [PubMed]
58. Roa, A.; Hayashi, F.; Yang, Y.; Lienlaf, M.; Zhou, J.; Shi, J.; Watanabe, S.; Kigawa, T.; Yokoyama, S.; Aiken, C.; et al. RING domain mutations uncouple TRIM5α restriction of HIV-1 from inhibition of reverse transcription and acceleration of uncoating. *J. Virol.* **2012**, *86*, 1717–1727. [CrossRef] [PubMed]
59. Stremlau, M.; Perron, M.; Lee, M.; Li, Y.; Song, B.; Javanbakht, H.; Diaz-Griffero, F.; Anderson, D.J.; Sundquist, W.I.; Sodroski, J. Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5α restriction factor. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5514–5519. [CrossRef] [PubMed]
60. Stremlau, M.; Owens, C.M.; Perron, M.J.; Kiessling, M.; Autissier, P.; Sodroski, J. The cytoplasmic body component TRIM5α restricts HIV-1 infection in old world monkeys. *Nature* **2004**, *427*, 848–853. [CrossRef] [PubMed]
61. Kane, M.; Yadav, S.S.; Bitzegeio, J.; Kutluay, S.B.; Zang, T.; Wilson, S.J.; Schoggins, J.W.; Rice, C.M.; Yamashita, M.; Hatziioannou, T.; et al. MX2 is an interferon-induced inhibitor of HIV-1 infection. *Nature* **2013**, *502*, 563–566. [CrossRef] [PubMed]
62. Goujon, C.; Moncorgé, O.; Bauby, H.; Doyle, T.; Ward, C.C.; Schaller, T.; Hué, S.; Barclay, W.S.; Schulz, R.; Malim, M.H. Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection. *Nature* **2013**, *502*, 559–562. [CrossRef] [PubMed]
63. Schulte, B.; Buffone, C.; Opp, S.; di Nunzio, F.; de Souza Aranha Vieira, D.A.; Brandariz-Nuñez, A.; Diaz-Griffero, F. Restriction of HIV-1 requires the N-terminal region of MxB as a capsid-binding motif but not as a nuclear localization signal. *J. Virol.* **2015**, *89*, 8599–8610. [CrossRef] [PubMed]
64. Liu, Z.; Pan, Q.; Liang, Z.; Qiao, W.; Cen, S.; Liang, C. The highly polymorphic cyclophilin A-binding loop in HIV-1 capsid modulates viral resistance to MxB. *Retrovirology* **2015**, *12*, 1. [CrossRef] [PubMed]

65. Busnadiego, I.; Kane, M.; Rihn, S.J.; Preugschas, H.F.; Hughes, J.; Blanco-Melo, D.; Strouvelle, V.P.; Zang, T.M.; Willett, B.J.; Boutell, C.; et al. Host and viral determinants of Mx2 antiretroviral activity. *J. Virol.* **2014**, *88*, 7738–7752. [CrossRef] [PubMed]
66. Matreyek, K.A.; Wang, W.; Serrao, E.; Singh, P.K.; Levin, H.L.; Engelman, A. Host and viral determinants for MxB restriction of HIV-1 infection. *Retrovirology* **2014**, *11*, 90. [CrossRef] [PubMed]
67. Laguette, N.; Sobhian, B.; Casartelli, N.; Ringeard, M.; Chable-Bessia, C.; Ségéral, E.; Yatim, A.; Emiliani, S.; Schwartz, O.; Benkirane, M. SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* **2011**, *474*, 654–657. [CrossRef] [PubMed]
68. Lahouassa, H.; Daddacha, W.; Hofmann, H.; Ayinde, D.; Logue, E.C.; Dragin, L.; Bloch, N.; Maudet, C.; Bertrand, M.; Gramberg, T.; et al. SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. *Nat. Immunol.* **2012**, *13*, 223–228. [CrossRef] [PubMed]
69. Sheehy, A.M.; Gaddis, N.C.; Choi, J.D.; Malim, M.H. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **2002**, *418*, 646–650. [CrossRef] [PubMed]
70. Zhang, H.; Yang, B.; Pomerantz, R.J.; Zhang, C.; Arunachalam, S.C.; Gao, L. The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* **2003**, *424*, 94–98. [CrossRef] [PubMed]
71. Mangeat, B.; Turelli, P.; Caron, G.; Friedli, M.; Perrin, L.; Trono, D. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* **2003**, *424*, 99–103. [CrossRef] [PubMed]
72. Li, M.; Kao, E.; Gao, X.; Sandig, H.; Limmer, K.; Pavon-Eternod, M.; Jones, T.E.; Landry, S.; Pan, T.; Weitzman, M.D.; et al. Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11. *Nature* **2012**, *491*, 125–128. [CrossRef] [PubMed]
73. Stabell, A.C.; Hawkins, J.; Li, M.; Gao, X.; David, M.; Press, W.H.; Sawyer, S.L. Non-human primate schlafen11 inhibits production of both host and viral proteins. *PLoS Pathog.* **2016**, *12*, e1006066. [CrossRef] [PubMed]
74. Van Weringh, A.; Ragonnet-Cronin, M.; Pranckeviciene, E.; Pavon-Eternod, M.; Kleiman, L.; Xia, X. HIV-1 modulates the tRNA pool to improve translation efficiency. *Mol. Biol. Evol.* **2011**, *28*, 1827–1834. [CrossRef] [PubMed]
75. Neil, S.J.D.; Eastman, S.W.; Jouvenet, N.; Bieniasz, P.D. HIV-1 Vpu promotes release and prevents endocytosis of nascent retrovirus particles from the plasma membrane. *PLoS Pathog.* **2006**, *2*, e39. [CrossRef] [PubMed]
76. Neil, S.J.D.; Zang, T.; Bieniasz, P.D. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* **2008**, *451*, 425–430. [CrossRef] [PubMed]
77. Van Damme, N.; Goff, D.; Katsura, C.; Jorgenson, R.L.; Mitchell, R.; Johnson, M.C.; Stephens, E.B.; Guatelli, J. The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. *Cell Host Microbe* **2008**, *3*, 245–252. [CrossRef] [PubMed]
78. Tada, T.; Zhang, Y.; Koyama, T.; Tobiume, M.; Tsunetsugu-Yokota, Y.; Yamaoka, S.; Fujita, H.; Tokunaga, K. MARCH8 inhibits HIV-1 infection by reducing virion incorporation of envelope glycoproteins. *Nat. Med.* **2015**, *21*, 1502–1507. [CrossRef] [PubMed]
79. Trautz, B.; Pierini, V.; Wombacher, R.; Stolp, B.; Chase, A.J.; Pizzato, M.; Fackler, O.T. The antagonism of HIV-1 Nef to SERINC5 particle infectivity restriction involves the counteraction of virion-associated pools of the restriction factor. *J. Virol.* **2016**, *90*, 10915–10927. [CrossRef] [PubMed]
80. Sheehy, A.M.; Gaddis, N.C.; Malim, M.H. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nat. Med.* **2003**, *9*, 1404–1407. [CrossRef] [PubMed]
81. Anderson, B.D.; Harris, R.S. Transcriptional regulation of APOBEC3 antiviral immunity through the CBF-β/RUNX axis. *Sci. Adv.* **2015**, *1*, e1500296. [CrossRef] [PubMed]
82. Stopak, K.; de Noronha, C.; Yonemoto, W.; Greene, W.C. HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. *Mol. Cell* **2003**, *12*, 591–601. [CrossRef]
83. Hrecka, K.; Hao, C.; Gierszewska, M.; Swanson, S.K.; Kesik-Brodacka, M.; Srivastava, S.; Florens, L.; Washburn, M.P.; Skowronski, J. Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. *Nature* **2011**, *474*, 658–661. [CrossRef] [PubMed]
84. Lim, E.S.; Fregoso, O.I.; McCoy, C.O.; Matsen, F.A.; Malik, H.S.; Emerman, M. The ability of primate lentiviruses to degrade the monocyte restriction factor SAMHD1 preceded the birth of the viral accessory protein Vpx. *Cell Host Microbe* **2012**, *11*, 194–204. [CrossRef] [PubMed]

85. McNatt, M.W.; Zang, T.; Bieniasz, P.D. Vpu binds directly to tetherin and displaces it from nascent virions. *PLoS Pathog.* **2013**, *9*, e1003299. [CrossRef] [PubMed]
86. Shibata, R.; Sakai, H.; Kawamura, M.; Tokunaga, K.; Adachi, A. Early replication block of human immunodeficiency virus type 1 in monkey cells. *J. Gen. Virol.* **1995**, *76 Pt 11*, 2723–2730. [CrossRef] [PubMed]
87. Himathongkham, S.; Luciw, P.A. Restriction of HIV-1 (subtype B) replication at the entry step in rhesus macaque cells. *Virology* **1996**, *219*, 485–488. [CrossRef] [PubMed]
88. Danielson, C.M.; Cianci, G.C.; Hope, T.J. Recruitment and dynamics of proteasome association with rhTRIM5α cytoplasmic complexes during HIV-1 infection. *Traffic* **2012**, *13*, 1206–1217. [CrossRef] [PubMed]
89. Lukic, Z.; Hausmann, S.; Sebastian, S.; Rucci, J.; Sastri, J.; Robia, S.L.; Luban, J.; Campbell, E.M. TRIM5α associates with proteasomal subunits in cells while in complex with HIV-1 virions. *Retrovirology* **2011**, *8*, 93. [CrossRef] [PubMed]
90. Rold, C.J.; Aiken, C. Proteasomal degradation of TRIM5α during retrovirus restriction. *PLoS Pathog.* **2008**, *4*, e1000074. [CrossRef] [PubMed]
91. Campbell, E.M.; Weingart, J.; Sette, P.; Opp, S.; Sastri, J.; O'Connor, S.K.; Talley, S.; Diaz-Griffero, F.; Hirsch, V.; Bouamr, F. TRIM5α-mediated ubiquitin chain conjugation is required for inhibition of HIV-1 reverse transcription and capsid destabilization. *J. Virol.* **2015**, *90*, 1849–1857. [CrossRef] [PubMed]
92. Kutluay, S.B.; Perez-Caballero, D.; Bieniasz, P.D. Fates of retroviral core components during unrestricted and TRIM5-restricted infection. *PLoS Pathog.* **2013**, *9*, e1003214. [CrossRef] [PubMed]
93. Anderson, J.L.; Campbell, E.M.; Wu, X.; Vandegraaff, N.; Engelman, A.; Hope, T.J. Proteasome inhibition reveals that a functional preintegration complex intermediate can be generated during restriction by diverse TRIM5 proteins. *J. Virol.* **2006**, *80*, 9754–9760. [CrossRef] [PubMed]
94. Pertel, T.; Hausmann, S.; Morger, D.; Züger, S.; Guerra, J.; Lascano, J.; Reinhard, C.; Santoni, F.A.; Uchil, P.D.; Chatel, L.; et al. TRIM5 is an innate immune sensor for the retrovirus capsid lattice. *Nature* **2011**, *472*, 361–365. [CrossRef] [PubMed]
95. Yudina, Z.; Roa, A.; Johnson, R.; Biris, N.; de Souza Aranha Vieira, D.A.; Tsiperson, V.; Reszka, N.; Taylor, A.B.; Hart, P.J.; Demeler, B.; et al. RING dimerization links higher-order assembly of TRIM5α to synthesis of K63-linked polyubiquitin. *Cell Rep.* **2015**, *12*, 788–797. [CrossRef] [PubMed]
96. Ohmura-Hoshino, M.; Matsuki, Y.; Aoki, M.; Goto, E.; Mito, M.; Uematsu, M.; Kakiuchi, T.; Hotta, H.; Ishido, S. Inhibition of MHC class II expression and immune responses by c-MIR. *J. Immunol.* **2006**, *177*, 341–354. [CrossRef] [PubMed]
97. Van de Kooij, B.; Verbrugge, I.; de Vries, E.; Gijsen, M.; Montserrat, V.; Maas, C.; Neefjes, J.; Borst, J. Ubiquitination by the membrane-associated RING-CH-8 (MARCH-8) ligase controls steady-state cell surface expression of tumor necrosis factor-related apoptosis inducing ligand (TRAIL) receptor 1. *J. Biol. Chem.* **2013**, *288*, 6617–6628. [CrossRef] [PubMed]
98. Chen, R.; Li, M.; Zhang, Y.; Zhou, Q.; Shu, H.-B. The E3 ubiquitin ligase MARCH8 negatively regulates IL-1β-induced NF-κB activation by targeting the IL1RAP coreceptor for ubiquitination and degradation. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 14128–14133. [CrossRef] [PubMed]
99. Eyster, C.A.; Cole, N.B.; Petersen, S.; Viswanathan, K.; Früh, K.; Donaldson, J.G. MARCH ubiquitin ligases alter the itinerary of clathrin-independent cargo from recycling to degradation. *Mol. Biol. Cell* **2011**, *22*, 3218–3230. [CrossRef] [PubMed]
100. Ito, F.; Fu, Y.; Kao, S.-C.A.; Yang, H.; Chen, X.S. Family-wide comparative analysis of cytidine and methylcytidine deamination by eleven human APOBEC proteins. *J. Mol. Biol.* **2017**, *429*, 1787–1799. [CrossRef] [PubMed]
101. Liddament, M.T.; Brown, W.L.; Schumacher, A.J.; Harris, R.S. APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 In Vivo. *Curr. Biol.* **2004**, *14*, 1385–1391. [CrossRef] [PubMed]
102. Harris, R.S.; Bishop, K.N.; Sheehy, A.M.; Craig, H.M.; Petersen-Mahrt, S.K.; Watt, I.N.; Neuberger, M.S.; Malim, M.H. DNA deamination mediates innate immunity to retroviral infection. *Cell* **2003**, *113*, 803–809. [CrossRef]
103. Yang, B.; Chen, K.; Zhang, C.; Huang, S.; Zhang, H. Virion-associated uracil DNA glycosylase-2 and apurinic/apyrimidinic endonuclease are involved in the degradation of APOBEC3G-edited nascent HIV-1 DNA. *J. Biol. Chem.* **2007**, *282*, 11667–11675. [CrossRef] [PubMed]
104. Yu, X.; Yu, Y.; Liu, B.; Luo, K.; Kong, W.; Mao, P.; Yu, X.-F. Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* **2003**, *302*, 1056–1060. [CrossRef] [PubMed]

105. Mehle, A.; Thomas, E.R.; Rajendran, K.S.; Gabuzda, D. A zinc-binding region in Vif binds Cul5 and determines cullin selection. *J. Biol. Chem.* **2006**, *281*, 17259–17265. [CrossRef] [PubMed]
106. Mehle, A.; Goncalves, J.; Santa-Marta, M.; McPike, M.; Gabuzda, D. Phosphorylation of a novel SOCS-box regulates assembly of the HIV-1 Vif-Cul5 complex that promotes APOBEC3G degradation. *Genes Dev.* **2004**, *18*, 2861–2866. [CrossRef] [PubMed]
107. Valera, M.-S.; de Armas-Rillo, L.; Barroso-González, J.; Ziglio, S.; Batisse, J.; Dubois, N.; Marrero-Hernández, S.; Borel, S.; García-Expósito, L.; Biard-Piechaczyk, M.; et al. The HDAC6/APOBEC3G complex regulates HIV-1 infectiveness by inducing Vif autophagic degradation. *Retrovirology* **2015**, *12*. [CrossRef] [PubMed]
108. Izumi, T.; Takaori-Kondo, A.; Shirakawa, K.; Higashitsuji, H.; Itoh, K.; Io, K.; Matsui, M.; Iwai, K.; Kondoh, H.; Sato, T.; et al. MDM2 is a novel E3 ligase for HIV-1 Vif. *Retrovirology* **2009**, *6*, 1. [CrossRef] [PubMed]
109. Matsui, Y.; Shindo, K.; Nagata, K.; Yoshinaga, N.; Shirakawa, K.; Kobayashi, M.; Takaori-Kondo, A. Core binding factor β protects HIV, type 1 accessory protein viral infectivity factor from MDM2-mediated degradation. *J. Biol. Chem.* **2016**, *291*, 24892–24899. [CrossRef] [PubMed]
110. Jäger, S.; Kim, D.Y.; Hultquist, J.F.; Shindo, K.; LaRue, R.S.; Kwon, E.; Li, M.; Anderson, B.D.; Yen, L.; Stanley, D.; et al. Vif hijacks CBF-β to degrade APOBEC3G and promote HIV-1 infection. *Nature* **2011**, *481*, 371–375. [CrossRef] [PubMed]
111. Zhang, W.; Du, J.; Evans, S.L.; Yu, Y.; Yu, X.-F. T-cell differentiation factor CBF-β regulates HIV-1 Vif-mediated evasion of host restriction. *Nature* **2011**, *481*, 376–379. [CrossRef] [PubMed]
112. Mercenne, G.; Bernacchi, S.; Richer, D.; Bec, G.; Henriet, S.; Paillart, J.-C.; Marquet, R. HIV-1 Vif binds to APOBEC3G mRNA and inhibits its translation. *Nucleic Acids Res.* **2010**, *38*, 633–646. [CrossRef] [PubMed]
113. Guerrero, S.; Libre, C.; Batisse, J.; Mercenne, G.; Richer, D.; Laumond, G.; Decoville, T.; Moog, C.; Marquet, R.; Paillart, J.-C. Translational regulation of APOBEC3G mRNA by Vif requires its 5'UTR and contributes to restoring HIV-1 infectivity. *Sci. Rep.* **2016**, *6*, 39507. [CrossRef] [PubMed]
114. Hultquist, J.F.; Lengyel, J.A.; Refsland, E.W.; LaRue, R.S.; Lackey, L.; Brown, W.L.; Harris, R.S. Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1. *J. Virol.* **2011**, *85*, 11220–11234. [CrossRef] [PubMed]
115. Binka, M.; Ooms, M.; Steward, M.; Simon, V. The activity spectrum of Vif from multiple HIV-1 subtypes against APOBEC3G, APOBEC3F, and APOBEC3H. *J. Virol.* **2012**, *86*, 49–59. [CrossRef] [PubMed]
116. Feng, Y.; Baig, T.T.; Love, R.P.; Chelico, L. Suppression of APOBEC3-mediated restriction of HIV-1 by Vif. *Front. Microbiol.* **2014**, *5*, 450. [CrossRef] [PubMed]
117. Liu, B.; Sarkis, P.T.N.; Luo, K.; Yu, Y.; Yu, X.-F. Regulation of Apobec3F and human immunodeficiency virus type 1 Vif by Vif-Cul5-ElonB/C E3 ubiquitin ligase. *J. Virol.* **2005**, *79*, 9579–9587. [CrossRef] [PubMed]
118. Goldstone, D.C.; Ennis-Adeniran, V.; Hedden, J.J.; Groom, H.C.T.; Rice, G.I.; Christodoulou, E.; Walker, P.A.; Kelly, G.; Haire, L.F.; Yap, M.W.; et al. HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. *Nature* **2011**, *480*, 379–382. [CrossRef] [PubMed]
119. Powell, R.D.; Holland, P.J.; Hollis, T.; Perrino, F.W. Aicardi-goutieres syndrome gene and HIV-1 restriction factor SAMHD1 is a dGTP-regulated deoxynucleotide triphosphohydrolase. *J. Biol. Chem.* **2011**, *286*, 43596–43600. [CrossRef] [PubMed]
120. Kim, B.; Nguyen, L.A.; Daddacha, W.; Hollenbaugh, J.A. Tight interplay among SAMHD1 protein level, cellular dNTP levels, and HIV-1 proviral DNA synthesis kinetics in human primary monocyte-derived macrophages. *J. Biol. Chem.* **2012**, *287*, 21570–21574. [CrossRef] [PubMed]
121. Baldauf, H.-M.; Pan, X.; Erikson, E.; Schmidt, S.; Daddacha, W.; Burggraf, M.; Schenkova, K.; Ambiel, I.; Wabnitz, G.; Gramberg, T.; et al. SAMHD1 restricts HIV-1 infection in resting CD4(+) T cells. *Nat. Med.* **2012**, *18*, 1682–1687. [CrossRef] [PubMed]
122. Bergamaschi, A.; Ayinde, D.; David, A.; le Rouzic, E.; Morel, M.; Collin, G.; Descamps, D.; Damond, F.; Brun-Vezinet, F.; Nisole, S.; et al. The human immunodeficiency virus type 2 Vpx protein usurps the CUL4A-DDB1 DCAF1 ubiquitin ligase to overcome a postentry block in macrophage infection. *J. Virol.* **2009**, *83*, 4854–4860. [CrossRef] [PubMed]
123. Brandariz-Nuñez, A.; Valle-Casuso, J.C.; White, T.E.; Laguette, N.; Benkirane, M.; Brojatsch, J.; Diaz-Griffero, F. Role of SAMHD1 nuclear localization in restriction of HIV-1 and SIVmac. *Retrovirology* **2012**, *9*, 49. [CrossRef] [PubMed]

124. Hofmann, H.; Logue, E.C.; Bloch, N.; Daddacha, W.; Polsky, S.B.; Schultz, M.L.; Kim, B.; Landau, N.R. The Vpx lentiviral accessory protein targets SAMHD1 for degradation in the nucleus. *J. Virol.* **2012**, *86*, 12552–12560. [CrossRef] [PubMed]
125. Tan, L.; Ehrlich, E.; Yu, X.-F. DDB1 and Cul4A are required for human immunodeficiency virus type 1 Vpr-induced G2 arrest. *J. Virol.* **2007**, *81*, 10822–10830. [CrossRef] [PubMed]
126. Belzile, J.-P.; Duisit, G.; Rougeau, N.; Mercier, J.; Finzi, A.; Cohen, E.A. HIV-1 Vpr-mediated G2 arrest involves the DDB1-CUL4AVPRBP E3 ubiquitin ligase. *PLoS Pathog.* **2007**, *3*, e85. [CrossRef] [PubMed]
127. Laguette, N.; Brégnard, C.; Hue, P.; Basbous, J.; Yatim, A.; Larroque, M.; Kirchhoff, F.; Constantinou, A.; Sobhian, B.; Benkirane, M. Premature activation of the SLX4 complex by Vpr promotes G2/M arrest and escape from innate immune sensing. *Cell* **2014**, *156*, 134–145. [CrossRef] [PubMed]
128. Zhou, X.; DeLucia, M.; Ahn, J. SLX4-SLX1 protein-independent down-regulation of MUS81-EME1 protein by HIV-1 viral protein R (Vpr). *J. Biol. Chem.* **2016**, *291*, 16936–16947. [CrossRef] [PubMed]
129. Tristem, M.; Purvis, A.; Quicke, D.L.J. Complex evolutionary history of primate lentiviralvprgenes. *Virology* **1998**, *240*, 232–237. [CrossRef] [PubMed]
130. Fregoso, O.I.; Ahn, J.; Wang, C.; Mehrens, J.; Skowronski, J.; Emerman, M. Evolutionary toggling of Vpx/Vpr specificity results in divergent recognition of the restriction factor SAMHD1. *PLoS Pathog.* **2013**, *9*, e1003496. [CrossRef] [PubMed]
131. Pauls, E.; Ruiz, A.; Badia, R.; Permanyer, M.; Gubern, A.; Riveira-Muñoz, E.; Torres-Torronteras, J.; Alvarez, M.; Mothe, B.; Brander, C.; et al. Cell cycle control and HIV-1 susceptibility are linked by CDK6-dependent CDK2 phosphorylation of SAMHD1 in myeloid and lymphoid cells. *J. Immunol.* **2014**, *193*, 1988–1997. [CrossRef] [PubMed]
132. Rocha-Perugini, V.; Suárez, H.; Álvarez, S.; López-Martín, S.; Lenzi, G.M.; Vences-Catalán, F.; Levy, S.; Kim, B.; Muñoz-Fernández, M.A.; Sánchez-Madrid, F.; et al. CD81 association with SAMHD1 enhances HIV-1 reverse transcription by increasing dNTP levels. *Nat. Microbiol.* **2017**. [CrossRef] [PubMed]
133. Kyei, G.B.; Cheng, X.; Ramani, R.; Ratner, L. Cyclin L2 is a critical HIV dependency factor in macrophages that controls SAMHD1 abundance. *Cell Host Microbe* **2015**, *17*, 98–106. [CrossRef] [PubMed]
134. Hinz, A.; Miguet, N.; Natrajan, G.; Usami, Y.; Yamanaka, H.; Renesto, P.; Hartlieb, B.; McCarthy, A.A.; Simorre, J.-P.; Göttlinger, H.; et al. Structural basis of HIV-1 tethering to membranes by the BST-2/tetherin ectodomain. *Cell Host Microbe* **2010**, *7*, 314–323. [CrossRef] [PubMed]
135. Venkatesh, S.; Bieniasz, P.D. Mechanism of HIV-1 virion entrapment by tetherin. *PLoS Pathog.* **2013**, *9*, e1003483. [CrossRef] [PubMed]
136. Skasko, M.; Wang, Y.; Tian, Y.; Tokarev, A.; Munguia, J.; Ruiz, A.; Stephens, E.B.; Opella, S.J.; Guatelli, J. HIV-1 Vpu protein antagonizes innate restriction factor BST-2 via lipid-embedded helix-helix interactions. *J. Biol. Chem.* **2012**, *287*, 58–67. [CrossRef] [PubMed]
137. Mitchell, R.S.; Katsura, C.; Skasko, M.A.; Fitzpatrick, K.; Lau, D.; Ruiz, A.; Stephens, E.B.; Margottin-Goguet, F.; Benarous, R.; Guatelli, J.C. Vpu antagonizes BST-2-mediated restriction of HIV-1 release via β-TrCP and endo-lysosomal trafficking. *PLoS Pathog.* **2009**, *5*, e1000450. [CrossRef] [PubMed]
138. Dubé, M.; Roy, B.B.; Guiot-Guillain, P.; Binette, J.; Mercier, J.; Chiasson, A.; Cohen, E.A. Antagonism of tetherin restriction of HIV-1 release by Vpu involves binding and sequestration of the restriction factor in a perinuclear compartment. *PLoS Pathog.* **2010**, *6*, e1000856. [CrossRef] [PubMed]
139. Roy, N.; Pacini, G.; Berlioz-Torrent, C.; Janvier, K. Characterization of E3 ligases involved in lysosomal sorting of the HIV-1 restriction factor BST2. *J. Cell Sci.* **2017**, *130*, 1596–1611. [CrossRef] [PubMed]
140. Margottin, F.; Bour, S.P.; Durand, H.; Selig, L.; Benichou, S.; Richard, V.; Thomas, D.; Strebel, K.; Benarous, R. A novel human WD protein, h-β TrCp, that interacts with HIV-1 Vpu connects CD4 to the ER degradation pathway through an F-box motif. *Mol. Cell* **1998**, *1*, 565–574. [CrossRef]
141. Iwabu, Y.; Fujita, H.; Kinomoto, M.; Kaneko, K.; Ishizaka, Y.; Tanaka, Y.; Sata, T.; Tokunaga, K. HIV-1 accessory protein Vpu internalizes cell-surface BST-2/tetherin through transmembrane interactions leading to lysosomes. *J. Biol. Chem.* **2009**, *284*, 35060–35072. [CrossRef] [PubMed]
142. Douglas, J.L.; Viswanathan, K.; McCarroll, M.N.; Gustin, J.K.; Früh, K.; Moses, A.V. Vpu directs the degradation of the human immunodeficiency virus restriction factor BST-2/Tetherin via a {β}TrCP-dependent mechanism. *J. Virol.* **2009**, *83*, 7931–7947. [CrossRef] [PubMed]

143. Mangeat, B.; Gers-Huber, G.; Lehmann, M.; Zufferey, M.; Luban, J.; Piguet, V. HIV-1 Vpu neutralizes the antiviral factor Tetherin/BST-2 by binding it and directing its β-TrCP2-dependent degradation. *PLoS Pathog.* **2009**, *5*, e1000574. [CrossRef] [PubMed]
144. Ramirez, P.W.; DePaula-Silva, A.B.; Szaniawski, M.; Barker, E.; Bosque, A.; Planelles, V. HIV-1 Vpu utilizes both cullin-RING ligase (CRL) dependent and independent mechanisms to downmodulate host proteins. *Retrovirology* **2015**, *12*, 65. [CrossRef] [PubMed]
145. Kueck, T.; Foster, T.L.; Weinelt, J.; Sumner, J.C.; Pickering, S.; Neil, S.J.D. Serine phosphorylation of HIV-1 Vpu and its binding to tetherin regulates interaction with clathrin adaptors. *PLoS Pathog.* **2015**, *11*, e1005141. [CrossRef] [PubMed]
146. Tervo, H.-M.; Homann, S.; Ambiel, I.; Fritz, J.V.; Fackler, O.T.; Keppler, O.T. β-TrCP is dispensable for Vpu's ability to overcome the CD317/Tetherin-imposed restriction to HIV-1 release. *Retrovirology* **2011**, *8*, 9. [CrossRef] [PubMed]
147. Madjo, U.; Leymarie, O.; Frémont, S.; Kuster, A.; Nehlich, M.; Gallois-Montbrun, S.; Janvier, K.; Berlioz-Torrent, C. LC3C contributes to Vpu-mediated antagonism of BST2/Tetherin restriction on HIV-1 release through a non-canonical autophagy pathway. *Cell Rep.* **2016**, *17*, 2221–2233. [CrossRef] [PubMed]
148. Pujol, F.M.; Laketa, V.; Schmidt, F.; Mukenhirn, M.; Müller, B.; Boulant, S.; Grimm, D.; Keppler, O.T.; Fackler, O.T. HIV-1 Vpu antagonizes CD317/Tetherin by adaptor protein-1-mediated exclusion from virus assembly sites. *J. Virol.* **2016**, *90*, 6709–6723. [CrossRef] [PubMed]
149. Janvier, K.; Pelchen-Matthews, A.; Renaud, J.-B.; Caillet, M.; Marsh, M.; Berlioz-Torrent, C. The ESCRT-0 component HRS is required for HIV-1 Vpu-mediated BST-2/tetherin down-regulation. *PLoS Pathog.* **2011**, *7*, e1001265. [CrossRef] [PubMed]
150. Krapp, C.; Hotter, D.; Gawanbacht, A.; McLaren, P.J.; Kluge, S.F.; Stürzel, C.M.; Mack, K.; Reith, E.; Engelhart, S.; Ciuffi, A.; et al. Guanylate binding protein (GBP) 5 is an interferon-inducible inhibitor of HIV-1 infectivity. *Cell Host Microbe* **2016**, *19*, 504–514. [CrossRef] [PubMed]
151. McLaren, P.J.; Gawanbacht, A.; Pyndiah, N.; Krapp, C.; Hotter, D.; Kluge, S.F.; Götz, N.; Heilmann, J.; Mack, K.; Sauter, D.; et al. Identification of potential HIV restriction factors by combining evolutionary genomic signatures with functional analyses. *Retrovirology* **2015**, *12*, 41. [CrossRef] [PubMed]
152. Schubert, U.; Bour, S.; Willey, R.L.; Strebel, K. Regulation of virus release by the macrophage-tropic human immunodeficiency virus type 1 AD8 isolate is redundant and can be controlled by either Vpu or Env. *J. Virol.* **1999**, *73*, 887–896. [PubMed]
153. Bour, S.; Schubert, U.; Peden, K.; Strebel, K. The envelope glycoprotein of human immunodeficiency virus type 2 enhances viral particle release: A Vpu-like factor? *J. Virol.* **1996**, *70*, 820–829. [PubMed]
154. Bour, S.; Strebel, K. The human immunodeficiency virus (HIV) type 2 envelope protein is a functional complement to HIV type 1 Vpu that enhances particle release of heterologous retroviruses. *J. Virol.* **1996**, *70*, 8285–8300. [PubMed]
155. Zhang, F.; Wilson, S.J.; Landford, W.C.; Virgen, B.; Gregory, D.; Johnson, M.C.; Munch, J.; Kirchhoff, F.; Bieniasz, P.D.; Hatziioannou, T. Nef proteins from simian immunodeficiency viruses are tetherin antagonists. *Cell Host Microbe* **2009**, *6*, 54–67. [CrossRef] [PubMed]
156. Jia, B.; Serra-Moreno, R.; Neidermyer, W.; Rahmberg, A.; Mackey, J.; Fofana, I.B.; Johnson, W.E.; Westmoreland, S.; Evans, D.T. Species-specific activity of SIV Nef and HIV-1 Vpu in overcoming restriction by tetherin/BST2. *PLoS Pathog.* **2009**, *5*, e1000429. [CrossRef] [PubMed]
157. Sauter, D.; Schindler, M.; Specht, A.; Landford, W.N.; Münch, J.; Kim, K.-A.; Votteler, J.; Schubert, U.; Bibollet-Ruche, F.; Keele, B.F.; et al. The evolution of pandemic and non-pandemic HIV-1 strains has been driven by Tetherin antagonism. *Cell Host Microbe* **2009**, *6*, 409–421. [CrossRef] [PubMed]
158. Willey, R.L.; Maldarelli, F.; Martin, M.A.; Strebel, K. Human immunodeficiency virus type 1 Vpu protein induces rapid degradation of CD4. *J. Virol.* **1992**, *66*, 7193–7200. [PubMed]
159. Schubert, U.; Antón, L.C.; Bacík, I.; Cox, J.H.; Bour, S.; Bennink, J.R.; Orlowski, M.; Strebel, K.; Yewdell, J.W. CD4 glycoprotein degradation induced by human immunodeficiency virus type 1 Vpu protein requires the function of proteasomes and the ubiquitin-conjugating pathway. *J. Virol.* **1998**, *72*, 2280–2288. [PubMed]
160. Binette, J.; Dubé, M.; Mercier, J.; Halawani, D.; Latterich, M.; Cohen, E.A. Requirements for the selective degradation of CD4 receptor molecules by the human immunodeficiency virus type 1 Vpu protein in the endoplasmic reticulum. *Retrovirology* **2007**, *4*, 75. [CrossRef] [PubMed]

161. Magadán, J.G.; Pérez-Victoria, F.J.; Sougrat, R.; Ye, Y.; Strebel, K.; Bonifacino, J.S. Multilayered mechanism of CD4 downregulation by HIV-1 Vpu involving distinct ER retention and ERAD targeting steps. *PLoS Pathog.* **2010**, *6*, e1000869. [CrossRef] [PubMed]
162. Magadán, J.G.; Bonifacino, J.S. Transmembrane domain determinants of CD4 downregulation by HIV-1 Vpu. *J. Virol.* **2012**, *86*, 757–772. [CrossRef] [PubMed]
163. Petris, G.; Casini, A.; Sasset, L.; Cesaratto, F.; Bestagno, M.; Cereseto, A.; Burrone, O.R. CD4 and BST-2/tetherin proteins retro-translocate from endoplasmic reticulum to cytosol as partially folded and multimeric molecules. *J. Biol. Chem.* **2014**, *289*, 1–12. [CrossRef] [PubMed]
164. Sugden, S.M.; Pham, T.N. Q.; Cohen, É.A. HIV-1 Vpu downmodulates ICAM-1 expression, resulting in decreased killing of infected CD4(+) T cells by NK cells. *J. Virol.* **2017**, *91*. [CrossRef] [PubMed]
165. Matheson, N.J.; Sumner, J.; Wals, K.; Rapiteanu, R.; Weekes, M.P.; Vigan, R.; Weinelt, J.; Schindler, M.; Antrobus, R.; Costa, A.S.H.; et al. Cell surface proteomic map of HIV infection reveals antagonism of amino acid metabolism by Vpu and Nef. *Cell Host Microbe* **2015**, *18*, 409–423. [CrossRef] [PubMed]
166. Ahn, J.; Vu, T.; Novince, Z.; Guerrero-Santoro, J.; Rapic-Otrin, V.; Gronenborn, A.M. HIV-1 Vpr loads uracil DNA glycosylase-2 onto DCAF1, a substrate recognition subunit of a cullin 4A-ring E3 ubiquitin ligase for proteasome-dependent degradation. *J. Biol. Chem.* **2010**, *285*, 37333–37341. [CrossRef] [PubMed]
167. Wu, Y.; Zhou, X.; Barnes, C.O.; DeLucia, M.; Cohen, A.E.; Gronenborn, A.M.; Ahn, J.; Calero, G. The DDB1-DCAF1-Vpr-UNG2 crystal structure reveals how HIV-1 Vpr steers human UNG2 toward destruction. *Nat. Struct. Mol. Biol.* **2016**, *23*, 933–940. [CrossRef] [PubMed]
168. Schröfelbauer, B.; Yu, Q.; Zeitlin, S.G.; Landau, N.R. Human immunodeficiency virus type 1 Vpr induces the degradation of the UNG and SMUG uracil-DNA glycosylases. *J. Virol.* **2005**, *79*, 10978–10987. [CrossRef] [PubMed]
169. Zhou, D.; Wang, Y.; Tokunaga, K.; Huang, F.; Sun, B.; Yang, R. The HIV-1 accessory protein Vpr induces the degradation of the anti-HIV-1 agent APOBEC3G through a VprBP-mediated proteasomal pathway. *Virus Res.* **2015**, *195*, 25–34. [CrossRef] [PubMed]
170. Casey Klockow, L.; Sharifi, H.J.; Wen, X.; Flagg, M.; Furuya, A.K.M.; Nekorchuk, M.; de Noronha, C.M.C. The HIV-1 protein Vpr targets the endoribonuclease Dicer for proteasomal degradation to boost macrophage infection. *Virology* **2013**, *444*, 191–202. [CrossRef] [PubMed]
171. Lahouassa, H.; Blondot, M.-L.; Chauveau, L.; Chougui, G.; Morel, M.; Leduc, M.; Guillonneau, F.; Ramirez, B.C.; Schwartz, O.; Margottin-Goguet, F. HIV-1 Vpr degrades the HLTF DNA translocase in T cells and macrophages. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 5311–5316. [CrossRef] [PubMed]
172. Romani, B.; Shaykh Baygloo, N.; Aghasadeghi, M.R.; Allahbakhshi, E. HIV-1 Vpr protein Enhances proteasomal degradation of MCM10 DNA replication factor through the Cul4-DDB1[VprBP] E3 ubiquitin ligase to induce G2/M cell cycle arrest. *J. Biol. Chem.* **2015**, *290*, 17380–17389. [CrossRef] [PubMed]
173. Maudet, C.; Sourisce, A.; Dragin, L.; Lahouassa, H.; Rain, J.-C.; Bouaziz, S.; Ramirez, B.C.; Margottin-Goguet, F. HIV-1 Vpr induces the degradation of ZIP and sZIP, adaptors of the NuRD chromatin remodeling complex, by hijacking DCAF1/VprBP. *PLoS ONE* **2013**, *8*, e77320. [CrossRef] [PubMed]
174. Romani, B.; Baygloo, N.S.; Hamidi-Fard, M.; Aghasadeghi, M.R.; Allahbakhshi, E. HIV-1 Vpr protein induces proteasomal degradation of chromatin-associated class I HDACs to overcome latent infection of macrophages. *J. Biol. Chem.* **2016**, *291*, 2696–2711. [CrossRef] [PubMed]
175. Klinger, P.P.; Schubert, U. The ubiquitin-proteasome system in HIV replication: Potential targets for antiretroviral therapy. *Expert Rev. Anti-Infect. Ther.* **2005**, *3*, 61–79. [CrossRef] [PubMed]
176. Wang, H.; Guo, H.; Su, J.; Rui, Y.; Zheng, W.; Gao, W.; Zhang, W.; Li, Z.; Liu, G.; Markham, R.B.; et al. Inhibition of Vpx-mediated SAMHD1 and Vpr-mediated host helicase transcription factor degradation by selective disruption of viral CRL4 (DCAF1) E3 ubiquitin ligase assembly. *J. Virol.* **2017**, *91*. [CrossRef] [PubMed]

© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

2.3 APOBEC

The Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like (APOBEC) proteins are a family of cytosine deaminases.

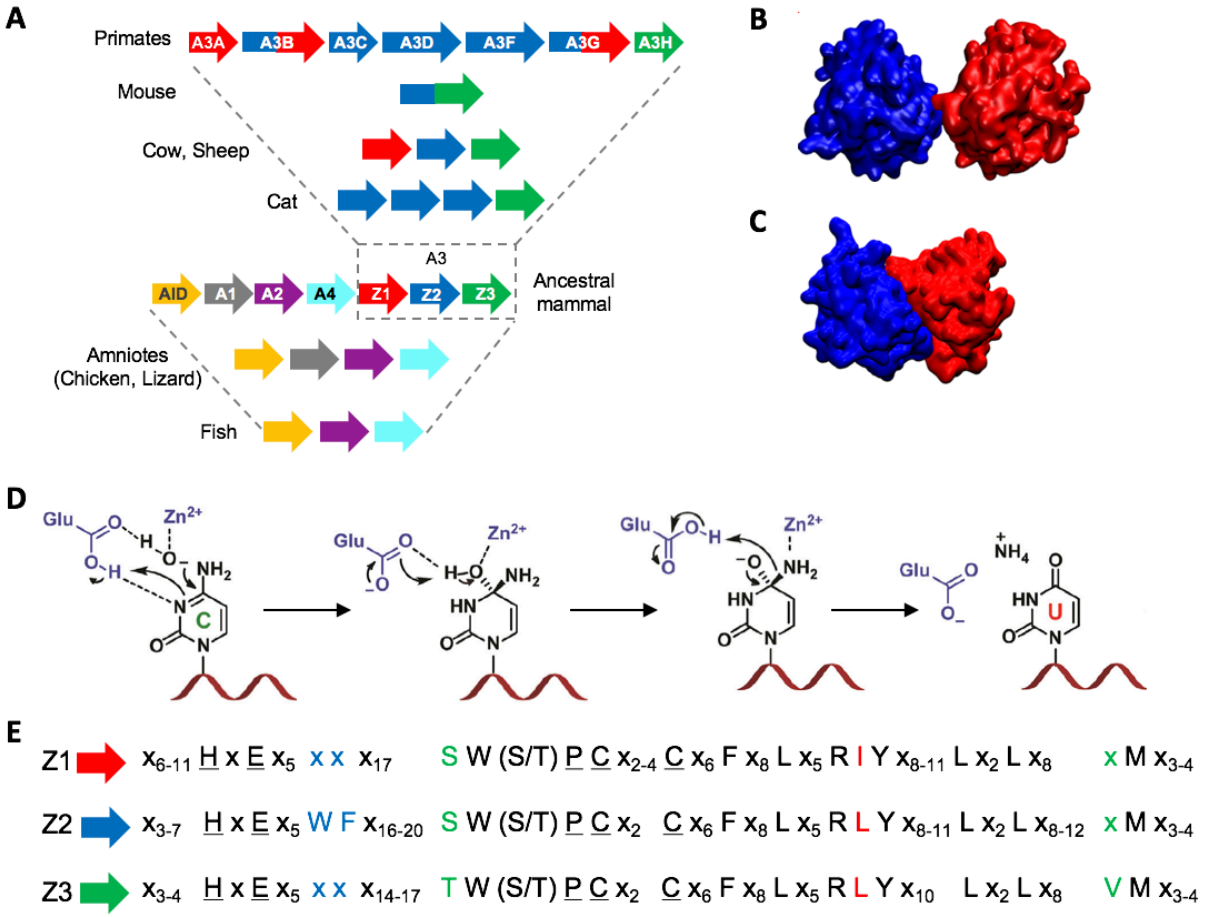


Figure 9: The APOBEC Family. (A) *Evolution of the APOBEC proteins across different species. The APOBEC genes present in different clades and species are shown. For A3 proteins a color code is used with red for Z1 domains, blue for Z2 domains and green for Z3 domains (Adapted from 74, 104).* ***(B)*** *Computational model of the dumbbell and* ***(C)*** *of the globular conformation of A3G (Adapted from 64).* ***(D)*** *Model for the chemical mechanism of cytidine deamination catalyzed by a Zn^{2+} ion and a glutamate (Glu) of the active site (Adapted from 150).* ***(E)*** *Specific sequence signature for Z1, Z2 and Z3 domains. X can be any amino acid. The consensus residues of the zinc coordinating domain are underlined. The amino acids characteristic for each domain are indicated in the corresponding color (Adapted from 178).*

AID is thought to have been the first cytosine deaminase and has likely evolved from the tRNA adenosine deaminases Tad (tRNA-specific adenosine deaminase) and ADAT2 (Adenosine deaminase tRNA specific 2). The subsequently evolved APOBEC proteins are thought to have been generated by duplication of the AID locus. The first APOBEC proteins, APOBEC2 (A2) and A4 have appeared in fish (Fig. 9A). They have been followed by A1, which is present in amniotes and finally A3 which is present in mammals. In placentals the A3 locus has been

further diversified, and duplications have given rise to various numbers of A3 proteins across different species, ranging from one A3 in mice up to seven A3s in primates (Fig. 9A) (82, 112).

The APOBEC proteins have various functions (Table 1). AID is involved in maturation and diversification of antibodies. On the one hand, cytosine deamination by AID promotes somatic hypermutation of antibodies. On the other hand, AID-dependent editing can induce double-stranded DNA breaks which stimulate class switch recombination. A1 interacts with its cofactors A1CF (APOBEC1 complementation factor) and RBM47 (RNA binding motif protein 47) to specifically recognize and edit a precise cytidine in ApoB (Apolipoprotein B) mRNA in the intestine, thereby inserting a premature stop-codon into the coding sequence of ApoB and generating a shorter form of the protein. A2 is expressed in heart and skeletal muscles in mammals and seems to play a role in development and embryogenesis, however no cytosine deaminating activity has yet been described. The A3 proteins play an important role in restriction of transposable elements and viruses. A4 is expressed in testes, however its role is still unknown. A5 also has an unknown function, and is expressed in non-mammalian tetrapods (112).

Each member of the APOBEC family possesses one or two zinc-coordinating domains (Table 1). These domains have a conserved globular structure with a central β-sheet, surrounded by α-helices (188). It comprises the active site, composed of one histidine and two cysteine residues which coordinate one zinc ion, as well as one glutamate residue. This sequence has been conserved as His-X-Glu-X_{23-28}-Pro-Cys-X_{2-4}-Cys (Fig. 9E) (114, 188).

Amongst A3 proteins, the zinc-coordinating domains can be classified in three distinct groups, Z1, Z2 and Z3, defined by specific conserved residues (Fig. 9E). Some A3s possess only one zinc-coordinating domain (A3A, A3C, A3H), while others are composed of two (Table 1). In the case of A3 proteins which possess two domains (A3B, A3D, A3F and A3G), only the C-terminal domain has retained catalytic activity while the N-terminal domain may only serve for substrate binding and multimerization (121, 188). The structure of A3 proteins with two domains has never been solved due to the flexible linker connecting both domains. However, recent computational modelling and microcopy has suggested, that A3G is present 16% of the time in an extended dumbbell structure (Fig. 9B), and 84% in a compact globular structure (Fig. 9C) (72).

In the process of cytidine deamination, the glutamate residue at the active site generates a hydroxide ion which is stabilized by the zinc and attacks the fourth position of the cytosine base, leading to the replacement of an amine (NH2) group by a carbonyl (CO) group, thereby transforming the cytosine into a uracil (Fig. 9D). Most APOBEC proteins specifically recognize

single-stranded DNA as the substrate for cytosine deamination, although some of them can also edit RNA (Table 1). The target cytosine is chosen depending on the sequence context, with most APOBEC proteins having a specific preference for the base in the -1 position (Table 1) (82, 189).

	Function	Chromosome	Number of domains (type)	Substrate	Target site	Target elements	Expression	Subcellular localization
AID	somatic hypermutation and class switch recombination	12	1	ssDNA	RC (R = purine)	immunoglobulin genes	Activated B cells	n/C
A1	ApoB mRNA editing	12	1	RNA	TC (specific 11 nucleotide sequence)	ApoB mRNA	Gastrointestinal tract	n/C
A2	development and embryogenesis	6	1	?			Heart, skeletal muscle, TNF-alpha activated liver cells	N/C
A3		22						
A3A	restriction of transposable elements and viruses		1 (Z1)	ssDNA, RNA	TC	HIV-1, SIV, HTLV-1, AAV, HPV, transposons	monocytes, macrophages, non-progenitor cells	N/C
A3B			2 (Z2-Z1)	ssDNA	TC	HIV-1, SIV, HTLV-1, MLV, RSV, transposons	PKC-induced cancer cells, IFN-alpha activated liver cells	N
A3C			1 (Z2)	ssDNA	TC	HIV-1, SIV, PFV, MLV, HBV, HSV-1, EBV transposons	immune centers, peripheral blood cells	N/C
A3D			2 (Z2-Z2	ssDNA	TC	HIV-1, SIV, transposons	immune centers, peripheral blood cells	C
A3F			2 (Z2-Z2)	ssDNA	TC	HIV-1, SIV, PFV, MLV, XMRV, PFV, MLV, RSV, MPMV, transposons	immune centers, peripheral blood cells, IFN-alpha activated liver cells	C
A3G			2 (Z2-Z1)	ssDNA, RNA	CC	HIV-1, SIV, MLV, XMRV, PFV, MMTV, HTLV1, EIAV, RSV, MPMV, HBV, transposons	immune centers, peripheral blood cells, IFN-alpha activated liver cells	C
A3H			1 (Z3)	ssDNA	TC	HIV-1, HTLV1, HPV, HBV, transposons	immune centers, peripheral blood cells	N/C
A4	role unknown	1	1	?				

***Table 1: Overview of the main characteristics of APOBEC proteins**. Each protein of the human APOBEC family is indicated with its main function, the chromosome on which the corresponding gene can be found, the number of protein domains and their classification (Z1, Z2 and Z3 type domain), the preferred substrate as well as the preferred sequence for cytosine deamination, the targeted elements, the expression pattern and the subcellular localization of the protein (n/C: mainly in the cytoplasm, N/C: equally in the cytoplasm and the nucleus, N: in the nucleus, C: in the cytoplasm). AAV: Adeno-associated virus; EBV: Epstein-Barr virus; EIAV: Equine infectious anemia virus; HBV: Hepatitis B virus; HPV: Human papillomavirus; HSV: Herpex simplex virus; HTLV: Human T-cell leukemia virus; MLV: Murine leukemia virus; MMTV: Mouse mammary tumor virus; MPMV: Mason Pfizer monkey virus; PFV: Prototype foamy virus; RSV: Rous sarcoma virus; XMRV: Xenotropic Moloney murine leukemia virus-related virus (Adapted from 8, 106, 178, 191).*

A3 proteins represent at the same time protection and threat for genomic integrity. While repressing mutational elements like transposons and viruses, they also have a mutational potential themselves. This is why they have to be tightly controlled, for example by subcellular localization (A3A deaminates its substrate in the nucleus, but its activity is retained in the

cytoplasm when not needed) or controlled expression (A3B is usually poorly expressed, its overexpression has been shown to contribute to certain cancers) (Table 1) (159, 188).

A3 proteins are important restriction factors, which are active against different viruses (Table 1). A3G has first been discovered as an HIV restriction factor in 2002 (206). It has since been shown that A3G is incorporated into budding virions in an infected cell by interaction with the HIV-1 RNA, cellular RNAs which are co-encapsidated as well as the viral protein Gag through its NC domain (24, 189). During the reverse transcription step in the newly infected cell, A3G can then restrict the virus, using different mechanisms. The predominant mechanism of A3G restriction is hypermutation. During reverse transcription, the viral genome is present as a single-stranded DNA, which is the substrate for A3G cytosine deamination. Cytosine deamination in the negative-sense ssDNA intermediate leads to G-to-A mutations in the positive strand (82, 235). There is a declining number of mutations from 5' to 3' of the viral ssDNA, which correlates with the time that the ssDNA remains available to A3G cytosine deamination. It has been shown that A3G can induce mutation of around 10% of the guanines of the viral genome to adenines in a single cycle and thereby repress viral infectivity by several logs (82, 159). On the one hand, mutation of important regulatory elements, like for example the TAR in the 5'-UTR of the HIV genome, can significantly reduce efficiency of the subsequent viral life cycle. On the other hand, A3G cytosine deamination very often leads to introduction of premature stop codons in viral coding sequences, because of the nature of A3G preferred target sites (CC), which often coincides with tryptophan codons (UGG). The appearance of premature stop codons leads to production of truncated viral proteins and results in the abortion of the viral life cycle (152, 158). There is some controversy surrounding the question, whether A3G-induced C to U mutations in viral DNA can also induce degradation of the viral genome. Indeed, some studies suggest that the uracils in viral DNA are detected by UNG2, which leads to excision of the base and subsequent cleavage and degradation of the viral genome by apurinic/apyrimidinic endonucleases. However, contrary results have also been reported (158, 170).

A3G has also been shown to have restrictive activity independently of its catalytic activity, although to a much lesser extent. Indeed, A3G has been shown to directly inhibit reverse transcription and integration of the viral genome. This is thought to be influenced on the one hand by direct interaction of A3G with the viral RT and on the other hand A3G can bind to the HIV gRNA and form oligomers, which might constitute a roadblock for RT enzymes (24, 152, 158, 170).

Amongst A3 family members, A3G represents the strongest restriction for HIV-1. However, A3F and A3H also restrict HIV, even if it is to a lesser extent. Finally, A3D is the weakest inhibitor of HIV infection (188).

2.4 Vif

As described above, the HIV-1 viral infectivity factor (Vif) downregulates A3G levels in the infected cell and thereby counteracts A3G-mediated restriction of HIV-1.
Vif downregulates A3G using three different mechanisms: (i) Vif recruits A3G to an E3-ubiquitin ligase complex and induces its ubiquitination followed by its degradation by the 26S proteasome (Fig. 10A, B) (207, 233); (ii) By sequestering the RUNX (Runt-related transcription factor) transcription cofactor CBF-β (Core-binding factor β), Vif indirectly reduces transcription of A3G (5, 111); (iii) Vif inhibits A3G translation (77, 147, 219). While the ubiquitination and degradation of A3G by Vif have been studied extensively (reviewed above in 188), little is known about A3G translational inhibition by Vif.

It has been shown that Vif binds to the A3G mRNA and this binding seems to be cooperative. Filter binding assays as well as fluorescence spectroscopy have shown strongest Vif binding to the 3'-UTR, closely followed by binding to the 5'-UTR and significantly weaker Vif binding to the A3G CDS. The secondary structure of A3G mRNA 5'- and 3'-UTRs has been elucidated by chemical probing which showed that both are organized in three stem-loop structures (Fig. 10C). Vif binding sites have been localized to stem loops (SL) 1 and SL3 in the 5'-UTR as well as all three SLs in the 3'-UTR (Fig. 10C). However, only the 5'-UTR seems to be necessary and sufficient to allow A3G translational inhibition by Vif (147).
The importance of the 5'-UTR of A3G has further been confirmed in non-infected and infected cells as well as with different isolates of Vif. Moreover, it has been shown that only SL2 and SL3 in the 5'-UTR of A3G mRNA are necessary for A3G translational inhibition by Vif. Furthermore, it has been shown that A3G degradation and translational inhibition are two independent pathways, because a Vif mutant deficient for A3G degradation is still able to significantly decrease A3G levels in the cell through translational inhibition (77). In summary, A3G translational inhibition is an important redundant pathway used by Vif to counteract A3G. This pathway is in itself sufficient to significantly decrease A3G levels in the cell and to rescue HIV-1 infectivity compared to a Vif-deficient context (77).
Current therapeutic strategies aim to inhibit Vif-dependent degradation of A3G in order to restore A3G-mediated viral restriction. This is achieved for example by inhibition of A3G-Vif

interaction (N4.1 (168) and IMB-26/35 (38)) or inhibition of Vif interaction with EloC (RN18 analogs (238) and VEC-5 (241)). These strategies lead to an increase in cellular A3G levels due to inhibition of A3G degradation and thereby mediate a decrease in viral infectivity (24, 129). However, A3G translational inhibition by Vif also has to be taken into account in these strategies, as it can in itself significantly repress A3G restriction. Indeed, when A3G derepression is incomplete, low levels of A3G-induced mutagenesis can have a beneficial effect on viral evolution. Therefore, it is important to further study A3G translational inhibition by Vif in order to better understand the underlying mechanism and eventually develop means to block this pathway.

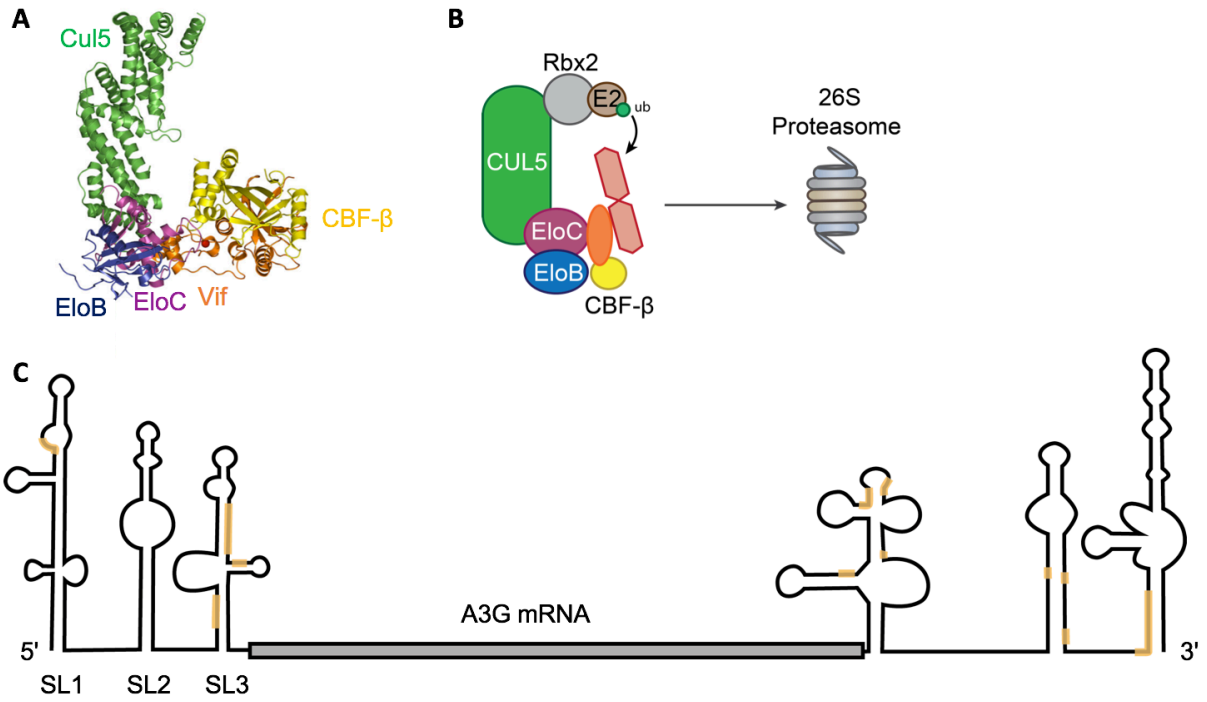


Figure 10: A3G counteraction by Vif. (A) *Crystal structure of Vif in complex with the SCF-like E3 ubiquitin ligase (Adapted from 72, 102); PDB id: 4N9F.* ***(B)*** *Vif-mediated recruitment of A3G to the E3-ubiquitin ligase leads to its ubiquitination and degradation in the proteasome. Vif is represented in orange and A3G in red, other proteins are annotated.* ***(C)*** *Secondary structure of the A3G mRNA UTRs and binding sites of Vif (indicated in orange). The A3G CDS is indicated as a grey box (Adapted from 138).*

3. Translation

3.1 The canonical mechanism of translation in eukaryotes

The translation in eukaryotes is a highly regulated mechanism that takes place in four distinct steps: (i) initiation, (ii) elongation, (iii) termination and (iv) recycling.
For canonical translation, the mRNA is required to possess a cap structure at its 5'-extremity, formed by a m^7G(5')ppp(5')N, as well as a poly-A-tail at its 3'-extremity which is bound by a series of poly-A binding proteins (PABPs) (118).

(i) Translation initiation
The translation competent mRNA is bound to a number of proteins. The cap structure of the mRNA is bound by a complex of initiation factors, called eIF4F. eIF4F is composed of the cap-binding factor eIF4E, the RNA-helicase eIF4A, and eIF4G, a scaffold protein which links eIF4E, eIF4A as well as the PABP at the 3'-end of the mRNA. The small subunit of the ribosome (40S) is bound to the translation initiation factors eIF1, eIF1A, eIF3 and eIF5 (69, 98, 195).
The first step is the recruitment of a ternary complex (TC) by the 40S ribosomal subunit in order to form the 43S preinitiation complex (PIC; Fig. 11A①). The TC is composed of the translation initiation factor eIF2 bound to GTP and to the initiating methionyl-tRNA (Met-$tRNA_i^{Met}$). The PIC is then recruited onto the mRNA 5'-extremity through direct interaction between eIF3 and eIF4G, which leads to the formation of a 48S complex (Fig. 11A②). In this complex, the small ribosomal subunit, which is maintained in an open latch conformation by eIF1 and eIF1A, is positioned onto the mRNA. This allows entry of the mRNA into the mRNA channel of the small ribosomal subunit, which can then begin scanning from 5' to 3' in search for the initiation codon, defined by the sequence AUG. The scanning mechanism is supported by the RNA helicases eIF4A and DHX29 (DExH-Box helicase 29), which unwind double-stranded RNA regions to allow the progression of the PIC. The start codon is recognized by base pairing with the anticodon sequence of the Met-$tRNA_i^{Met}$ (Fig. 11A③). Correct recognition of the start codon is ensured by eIF1 as well as stabilizing interactions with the ribosomal RNA of the small subunit (12, 89, 148, 208).
Structural rearrangements upon start codon recognition lead to eIF5-stimulated hydrolysis of eIF2-bound GTP followed by detachment of most initiation factors (Fig. 11A④). eIF5B can then join the complex and stimulate recruitment of the large ribosomal subunit (60S) in order to form

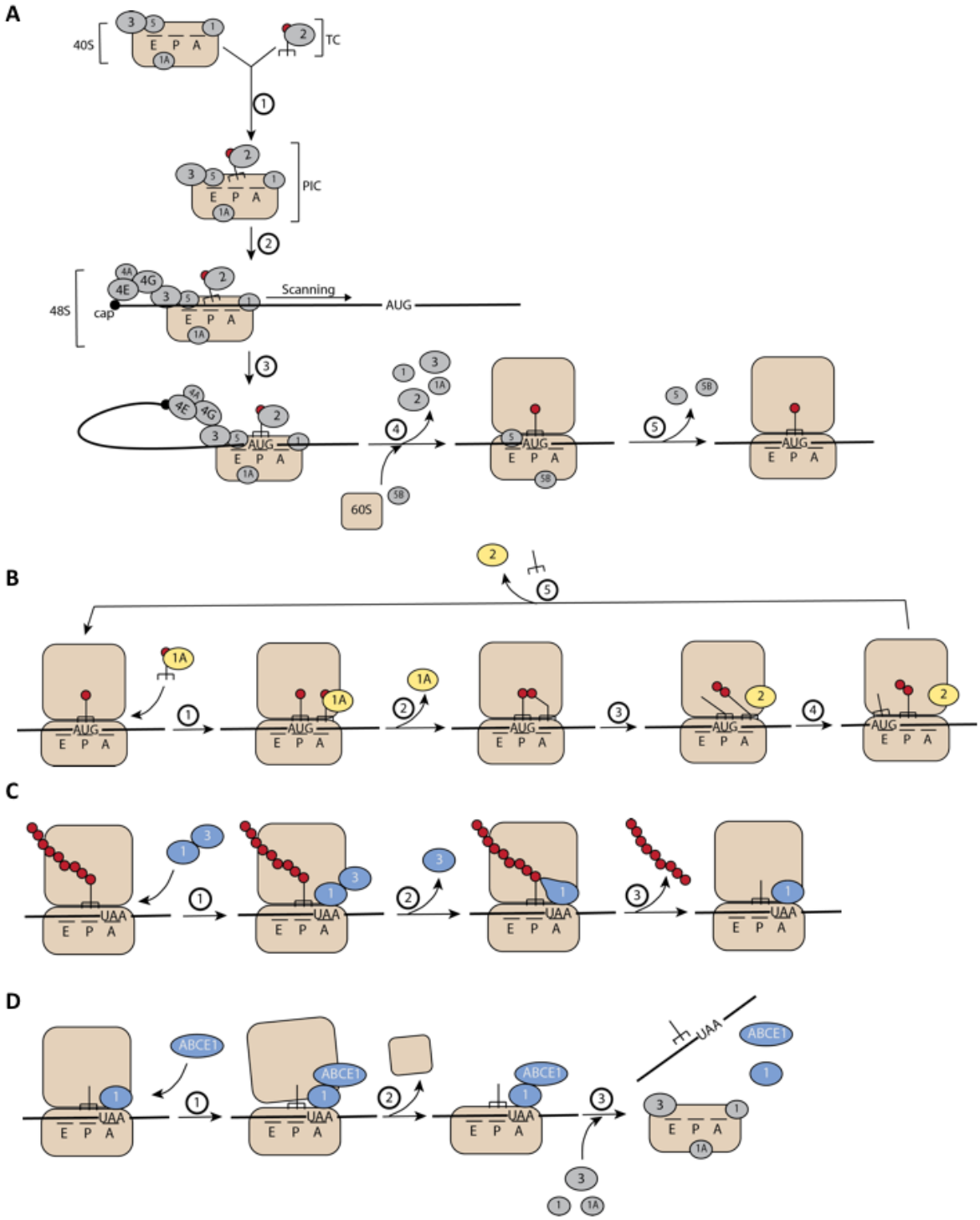


Figure 11: Mechanism of eukaryotic translation. (A) *Translation initiation. eIFs are depicted in grey. 1- Assembly of the TC and the 40S to form the PIC. 2- Recruitment of the PIC onto the mRNA by eIF4G-eIF3 interaction. 3- Scanning and recognition of the start codon (ie. AUG). 4- Hydrolysis of eIF2-bound GTP and dissociation of initiation factors. Recruitment of eIF5B and the 60S subunit. 5- Hydrolysis of eIF5B-bound GTP and detachment of eIF5 and eIF5B.* ***(B)*** *Elongation of translation. eEFs are depicted in yellow. 1- Recruitment of the tRNA that hybridizes with the codon in the A-site. 2- Hydrolysis of eEF1A-bound GTP and detachment. Placement of the new aminoacyl in the PTC. 3- Peptide bond formation, placement of the tRNAs in a hybrid state and recruitment of eEF2. 4- Translocation mediated by eEF2-bound GTP hydrolysis. 5- Detachment of eEF2 and the deacylated tRNA to allow beginning of a new cycle.* ***(C)*** *Termination of translation. eRFs are depicted in blue. 1- Detection of a stop codon (ie. UAA) in the A-site by eRF1 and binding of eRF3. 2- Hydrolysis of eRF3-bound GTP, detachment of eRF3 and structural rearrangement of eRF1. 3- eRF1-stimulated hydrolysis of the peptide-tRNA bond and detachment.* ***(D)*** *Recycling of the translating ribosome. 1- Recruitment of ABCE1. 2- Hydrolysis of ABCE1-bound ATP and detachment of the 60S subunit. 3- Recruitment of new eIFs and release.*

the translation competent ribosome (80S). eIF5B then detaches through GTP-hydrolysis (Fig. 11A⑤) (78, 89, 98).

(ii) Elongation

The ribosome is composed of three sites: (a) the P-site, which is initially bound by the Met-$tRNA_i^{Met}$ and will later accommodate the tRNA carrying the growing peptide-chain; (b) the A-site, where new aminoacyl-tRNAs are recruited through codon-anticodon interaction; (c) the E-site, which accommodates exiting, deacylated tRNAs (12, 126).

Elongation starts with the arrival of a new aminoacyl-tRNA (aa-tRNA) in the A site of the ribosome (Fig. 11B①). The aa-tRNA is bound by the elongation factor eEF1A which is associated with GTP. The codon-anticodon base pairing of the new aa-tRNA in the A-site of the ribosome induces conformational changes which lead to hydrolysis of eEF1A-bound GTP, followed by departure of eEF1A (Fig. 11B②). The aminoacyl-moiety of the aa-tRNA is then positioned at the peptidyl transferase centre (PTC) of the ribosome. Here, formation of a peptide bond between the peptide bound to the tRNA in the P-site and the new aminoacid in the A-site is catalyzed. This induces the rotation of the two tRNAs into a hybrid state (Fig. 11B③). The deacylated tRNA that was in the P-site now adopts a hybrid P/E position, while the new peptidyl-tRNA finds itself in a hybrid A/P position. eEF2-GTP can then bind to the A-site where hydrolysis of its GTP to GDP catalyzes the full translocation of the two tRNAs into the E- and P-sites respectively (Fig. 11B④). eEF2 is then released from the ribosome and a new cycle of elongation can begin (Fig. 11B⑤) (54, 55, 126, 198).

(iii) Termination

The presence of one of the three possible stop codons UAA, UGA and UAG in the A-site of the ribosome initiates translation termination. The release factor eRF1 recognizes the stop codon and binds in the A-site of the ribosome. It then recruits eRF3-GTP (Fig. 11C①). eRF1 stimulates hydrolysis of eRF3-bound GTP, which leads to detachment of eRF3 but also induces conformational changes in eRF1 (Fig. 11C②). eRF1 extends into the PTC, where it stimulates cleavage of the peptide chain from the tRNA (Fig. 11C③) (55, 84, 99, 198).

(iv) Recycling

After cleavage and departure of the newly translated peptide from the tRNA associated with the P-site of the ribosome, eRF1 recruits ABCE1-ATP (ATP-binding cassette sub-family E member 1; Fig. 11D①). ATP hydrolysis by ABCE1 induces detachment of the 60S ribosomal subunit (Fig. 11D②). Subsequent binding of new translation initiation factors (eIF1, eIF1A and

eIF3) to the remaining 40S ribosomal subunit stimulates detachment of the deacylated tRNA from the P-site as well as detachment of the mRNA (Fig. 11D③) (78, 84, 98).
The circularization of the mRNA, thanks to the interaction of eIF4G with the cap-bound eIF4E and the poly-A-bound PABP, is thought to enhance translation efficiency by allowing for a ribosome to directly resume translation after termination of a previous round of translation (208).

3.2 Regulation of translation initiation

Initiation of translation is a highly regulated process that depends on many different factors. There are several mechanisms that fine-tune initiation of translation both under physiological conditions as well as under stress or in case of an infection.

3.2.1 RNA binding proteins

mRNA forms interactions with a large variety of different proteins in the cell. These proteins, called RNA binding proteins (RBPs), are often characterized by the presence of a typical RNA-interacting domain like for example RNA recognition motifs (RRM), K-homology domains (KH), Zinc fingers, double-stranded RNA-binding domains (dsRBD), DEAD box helicase domains or Pumilio domains. RNA binding proteins regulate a large number of processes, which influence mRNA biogenesis, subcellular localization, degradation but also translation (21, 135, 151, 176).
One of the primary translation initiation-regulating RBPs are translation initiation factors (eIFs) themselves. Almost all eIFs can be phosphorylated, which is thought to regulate their activity. One well studied example is eIF2A, which is used to transport the Met-tRNA$_i^{Met}$ to the PIC. eIF2A-bound GTP is hydrolyzed to GDP in the process and has to be replaced by another GTP molecule, before eIF2A can initiate another round of translation. This exchange is very inefficient on its own and needs to be stimulated by eIF2B. However, under certain conditions, like for example stress, eIF2A is phosphorylated by eIF2α-kinase and eIF2A-bound GDP cannot be recycled anymore. This leads to a decrease of the available pool of eIF2A-GTP in the cell and resulting reduction of translation initiation. Another example is eIF4E which can be sequestered by eIF4E-BP (eIF4E-binding protein). Only when eIF4E-BP is phosphorylated, eIF4E is released and becomes available to bind to the cap and to eIF4G (89, 98).

There are many other proteins, apart from eIFs, that can interact with mRNAs and modulate translation initiation efficiency. Certain RBPs can for example inhibit correct recruitment of the cap-binding complex and therefore abolish translation initiation. LARP1 (La-related protein 1) for example is an RBP that binds to TOP motifs (Terminal oligopyrimidine tract), characterized by an invariable C followed by 4-15 pyrimidines at the 5'-extremity of mRNAs, and inhibits eIF4E interaction with the cap (83). FMRP (Fragile X mental retardation protein) binds to select mRNAs like for example map1B (microtubule associated protein 1B) and recruits CYFIP1 (Cytoplasmic FMR1 interacting protein 1) which inhibits eIF4G binding by interaction with eIF4E (39).

Translational initiation can also be regulated by RBPs at the PIC-recruitment step. Indeed, IRP (iron regulatory protein) for example binds to mRNA sequences in the 5'-UTR of its target mRNA and inhibits recruitment of the 43S PIC (153). The GAIT (IFN-γ-activated inhibitor of translation) complex, which is composed of 4 subunits (EPRS (Glutamyl-prolyl-tRNA synthetase), NSAP1 (NS1 associated protein 1), L13a (60S ribosomal protein) and GAPDH (Glyceraldehyde 3-phosphate dehydrogenase)) and assembles in response to IFNγ, binds to targeted mRNAs (often mRNAs coding for inflammatory factors) and inhibits 43S PIC recruitment through interaction of L13a with eIF4G. This allows regulation of translation of specific transcripts in response to IFNγ (10).

Other RBPs inhibit translation initiation at later steps, like for example hnRNPs K and E1 which bind to the 3'-UTR of lox (lysyl oxidase) mRNA and inhibit recruitment of the 60S ribosomal subunit, thereby preventing formation of the functional ribosome (160).

There are many more examples of RBPs regulating translation in eukaryotes, which suggests that RBPs play a central role in translational control mechanisms (176). However, the complete interactome of a particular mRNA has seldom be characterized, and many things are yet to be discovered in this domain.

3.2.2 Non canonical translation initiation

While the above described canonical mechanism of translation is the preferential pathway for translation of eukaryotic mRNAs, other mechanisms exist in the cell.

The canonical start codon for translation initiation is AUG, however initiation at other codons, mainly CUG and GUG has also been demonstrated (97, 215).

Canonical translation initiation is dependent on the 7-methylguanosine cap at the 5’-end of mRNAs, however, cap-independent initiation mechanisms also exist. One of these mechanisms is mediated by an internal ribosome entry site (IRES). In general, an IRES is a

highly structured RNA sequence, which can often be found in viral mRNAs but is also used in some cellular mRNAs. IRESes can be classified into 4 different groups with increasing structural complexity and decreasing dependency of eIFs. IRESes of class I for example directly interact with eIF4G and thereby recruit the PIC while bypassing the requirement for a cap and eIF4E. While the initiation mechanism of class I IRESes largely resembles canonical translation, class II and III IRESes necessitate less and less elements of the canonical translation initiation machinery and finally, class IV IRESes completely bypass most of the canonical mechanism. The class IV IRES directly interacts with elongation competent ribosomes and positions the P-site onto an RNA structure, which mimics a codon-anticodon interaction, thereby completely bypassing all eIFs (27, 123, 132).
While IRESes represent an important intrinsic part of an mRNA in order to allow its non-canonical translation, post-transcriptional modifications can be a more subtle and flexible way of translation regulation. Methylation of adenosines on position 6 (m6A) is the most prevalent post-transcriptional modification in mRNAs. While most m6A modifications are found in the 3'-UTR and regulate mRNA stability, m6A modifications in the 5'-UTR can allow cap-independent translation initiation by direct recruitment of eIF3 (40, 149, 150).

3.2.3 Upstream open reading frames

During the scanning of the 5'-UTR by the PIC, the AUG of the mRNA's main coding sequence (mAUG) is often not the first start codon encountered. Indeed, it has been shown that 49% of all cellular mRNAs contain small ORFs upstream of or sometimes overlapping with the main coding sequence (Fig. 12A). These upstream ORFs (uORF) generally have an inhibitory effect on the translation of the main coding sequence (mORF) and sometimes allow fine-tuning of translational regulation of the mORF (14, 34, 41, 234).
The translation of the mORF in presence of one or more uORFs can take place by different mechanisms:

(a) Leaky scanning
The mechanism of leaky scanning is defined by the scanning PIC which fails to recognize a start codon and instead continues scanning until the next possible AUG (Fig. 12B). The recognition of a start codon relies on its base pairing with the Met-tRNA$_i^{Met}$, but also depends on other sequence motifs surrounding the AUG. A favourable sequence context, also called Kozak context, is represented by an adenine in position -3 and a purine in position +4 of the start codon. The probability of leaky scanning decreases with increasing fidelity to this

favourable sequence context. Thereby, a uORF with a strong Kozak context will be preferentially recognized by the scanning PIC and leads to a strong inhibition of mORF translation, while a uORF with a weak Kozak context will often be bypassed by the PIC and can only exert weak translational repression (14, 48, 115).

The msl-2 (male-specific lethal 2) mRNA is an example for leaky scanning. This mRNA contains one uORF in its 5'-UTR but the mORF can be translated through leaky scanning. The leaky scanning mechanism can also be influenced by RBPs. In the case of the msl-2 mRNA for example, the SXL (sex-lethal) protein binds to a region downstream of the uORF, thereby stimulating translation initiation at the uORF while decreasing leaky scanning and translation of the mORF (22, 70, 143).

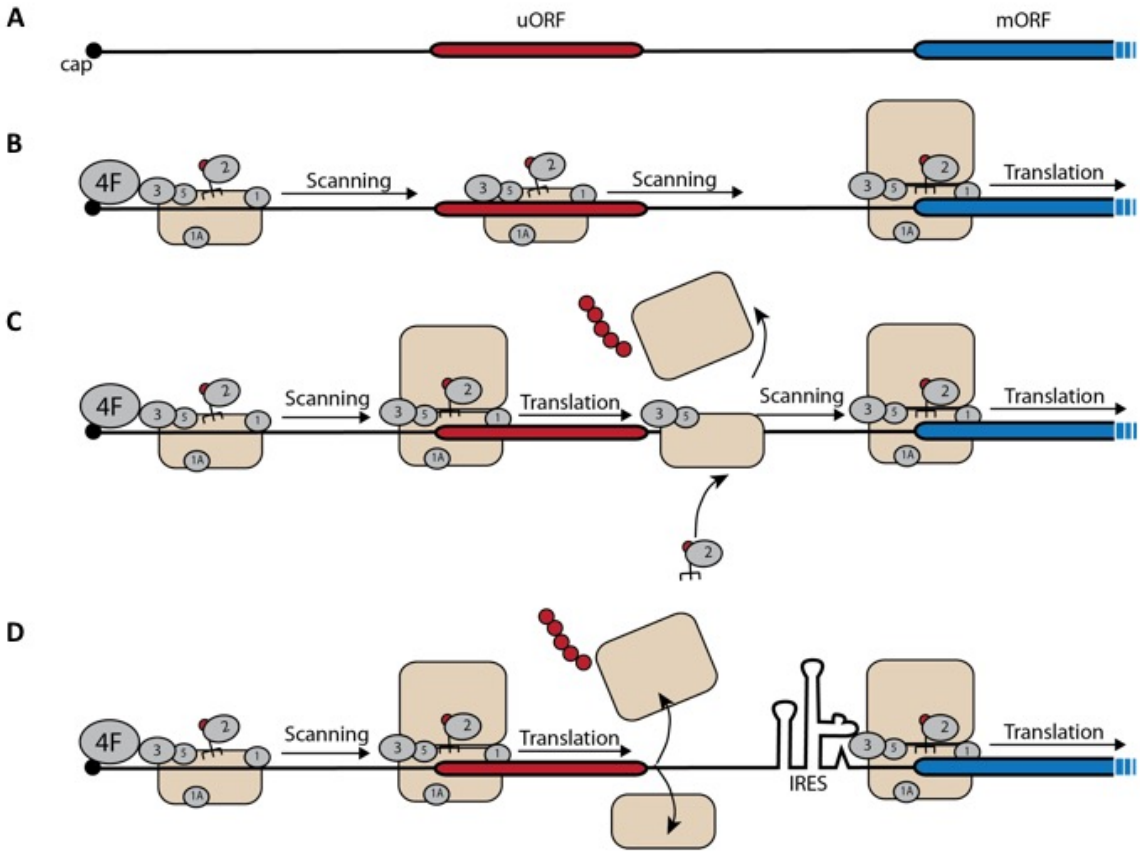


Figure 12: Mechanisms of translation of uORF-containing transcripts. (A) *Representation of the 5'-UTR of an mRNA that contains a uORF (red) upstream of the mORF (blue).* ***(B)*** *Leaky scanning: The 48S PIC bypasses the uAUG and instead continues scanning until recognition of the mAUG.* ***(C)*** *Reinitiation: The ribosome first recognizes the uAUG and translates the uORF. After translation termination, the 60S disassembles, but certain initiation factors have remained bound to the 40S, which resumes scanning, recruits a new TC and can reinitiate translation at the mAUG.* ***(D)*** *While the uORF is translated in a cap-dependent manner, an IRES allows direct recruitment of the ribosome onto the mORF.*

(b) Re-initiation

Re-initiation is a process by which the ribosome can resume scanning after translation of a uORF in order to re-initiate translation at the mAUG (Fig. 12C). One prerequisite for this

mechanism is a relatively short uORF with few rare codons. This ensures that the uORF is quickly translated and allows the ribosome to remain associated with some eIFs, notably eIF3. After translation termination of the uORF, the remaining eIFs can stimulate the 40S complex to not undergo recycling, but to resume scanning instead. During intercistronic scanning (between the uORF and the mORF), the 40S complex has to reacquire the remaining eIFs and most importantly the TC. The longer the intercistronic region, the bigger the probability that the ribosome reacquires a new TC and is able to initiate translation at the mAUG. However, when the intercistronic region is too long, the scanning 40S might drop off before being able to re-initiate translation (14, 34, 78, 116, 117).

Re-initiation can allow specific regulation of mORF translation as it is the case for example for the transcription factor ATF4 (Activating transcription factor 4). The ATF4 mRNA contains two uORFs, one short and the second longer and overlapping with the mAUG. Under normal conditions, the first uORF is translated and re-initiation occurs, which allows translation of the second uORF, thereby bypassing the mAUG. However, under stress conditions, the availability of the TC decreases and its reacquisition after translation of the first uORF takes longer. Under these conditions, the scanning re-initiation complex bypasses the second uORF and instead re-initiates at the mAUG (14, 48, 215).

Re-initiation can also be regulated by RBPs, like for example by the DENR-MCT1 (Density regulated re-initiation and release factor; Monocarboxylate transporter 1) complex which stimulates re-initiation by binding to the ribosome (127, 193).

(c) IRES

One possibility for a mORF to be translated despite the presence of a uORF is through an IRES, which allows positioning of the PIC downstream of the uORF (Fig. 12D). This strategy is used for example by the CAT-1 (Cationic aminoacid transporter 1) mRNA which contains a uORF as well as an IRES. However, cap-dependent translation of the uORF is still necessary, because it induces structural changes of the mRNA which leads to active conformation of the IRES (215).

(d) uORF-encoded peptides

While most uORF-encoded peptides do not seem to play a role in translation regulation, some examples have been described, where the uORF-encoded peptide itself inhibits translation of the main ORF, for example by interaction with the peptide channel of the ribosome and induction of ribosome stalling on the uORF sequence. This is for example the case for the uORF of CPA1 (carboxypeptidase A1) mRNA (101, 234).

(e) uORF-induced non-sense mediated decay

The presence of premature stop codons on an mRNA usually leads to its non-sense mediated decay (NMD). The NMD pathway protects the cell from overexpression of truncated proteins from mutated transcripts. This mechanism is triggered by the presence of a stop codon upstream of an exon junction complex (EJC). The EJC recruits the NMD factors UPF2 and 3 (regulator of nonsense transcripts 2 and 3), while eRF1 and 3 on the terminating ribosome recruit UPF1 and SMG1 (Serine/threonine-protein kinase SMG1) to form the SURF (SMG1-UPF1-eRFs) complex. Interaction between SURF and UPF2 and 3 activates SMG1, which phosphorylates UPF1. Phosphorylated UPF1 can then recruit the RNA degradation machinery. While some uORF-containing transcripts seem to be targets of NMD, most of them are protected from this mechanism. This is thought to be due to the proximity of the poly-A tail to the 5'-UTR due to circularization of the mRNA. This leads to interaction of PABP with eRF3 and inhibits recruitment of UPF1. Therefore, most uORFs mediate their translation inhibition of the mORF without decreasing the mRNA level (14, 101, 120).

(f) uORF variability and disease

While uORFs regulate translation of their associated mORFs, the presence of uORFs on the transcripts itself can also be regulated. The use of alternative start sites during transcription, as well as alternative splicing can lead to the differential inclusion or exclusion of uORFs into a transcript. Moreover, some single nucleotide polymorphisms (SNP) can lead to appearance or disappearance of uORFs on certain transcripts. This variability of uORFs in the 5'-UTR of transcripts significantly influences expression levels of associated mORF-encoded proteins. In some cases, appearance or disappearance of uORFs in transcripts can cause diseases like it is the case in hereditary pancreatitis, where a SNP in the *SPINK1* (Serine peptidase inhibitor, Kazal type 1) gene leads to the appearance of a second uORF in the mRNA resulting in strong inhibition of SPINK1 protein production (34).

3.2.4 RNA granules

When their translation is inhibited, mRNAs are usually degraded or stored away. P-bodies are cytoplasmic RNPs which store mRNAs whose translation is repressed. P-bodies are mainly composed of proteins of the degradation machinery, including decapping enzymes (DCP1 and 2), deadenylases (CCR4, Carbon catabolite repressor protein 4) and exonucleases (XRN1, 5'-3' exoribonuclease 1). Under normal circumstances, translationnally inactive transcripts

aggregate in these granules while awaiting either degradation or release to resume translation (60, 105, 164, 169).

When a stress-induced translational block occurs however, stalled 48S complexes are stored in stress granules. In contrast to P-bodies, stress granules contain the components of the translation initiation machinery, including eIFs and the 40S ribosomal subunit, as well as translation inhibitors (TIA-1 (T-cell-restricted intracellular antigen 1), TIAR (TIA-1-related protein), FMRP, G3BP (Ras-GAP SH3 domain binding protein)) and immune sensors (PKR (Protein kinase R), RIG-I). Stress granules allow storage of inactive mRNAs until the translational block is lifted and their translation can resume. Stress granules are often induced in response to elevated eIF2 phosphorylation for example (6, 33, 139, 169).

Interestingly it has been described that A3G and A3F localise to P-bodies and stress-granules (68, 228).

3.3 Translational regulation of HIV-1

HIV-1 translation only partially uses the canonical mechanism described above. Cap-dependent translation of the HIV transcripts is complicated by the highly structured 5'-UTR, containing stable stem-loops like for example TAR, which can inhibit the recruitment and hinder scanning of the ribosome. Therefore, HIV recruits cellular factors like La and the helicases DDX9 (DEAD-box protein 9) and DDX3 which unwind stable structures in the 5'-UTR and stimulate cap-dependent translation of HIV transcripts. Moreover, HIV infection generates a disfavourable cellular environment for canonical translation. This is due to viral targeting of translation initiation factors like eIF4G and the eIF3 subunit d, which are cleaved by the viral protease in late time points of the infection, as well as eIF4E, which is sequestered by hypophosphorylated eIF4E-BP in the context of Vpr-induced G2/M cell cycle arrest. It has been shown however that instead of canonical translation initiation factors, HIV transcripts are preferentially associated with the cap binding complex (CBC), which is composed of CBP80 and CBP20. This complex is deposited on cellular mRNAs in the nucleus and stimulates their splicing and nuclear export. It also induces a pioneering round of translation for each mRNA before being replaced by eIF4F. HIV might have hijacked this complex to bypass canonical eIFs for translation. In addition to the cap-dependent translation, HIV transcripts also contain two IRESes, one in the 5'-UTR which allows translation of $p55^{Gag}$, as well as a second one slightly downstream which allows translation of a shortened $p40^{Gag}$, which is expressed at lower levels (29, 30, 37, 87, 157, 202).

m6A modifications of HIV transcripts have been shown to have a stimulating role on viral gene expression. On the one hand, m6A in the RRE stimulates Rev binding and enhances nuclear export of HIV RNAs. On the other hand, other m6A modifications can be found all along the viral gRNA and seem to have a positive effect on translation (106, 107, 125, 166, 179, 223).
Some viral proteins have to be translated through a particular mechanism. One of them is the polyprotein Pol, which can only be translated as a fusion with Gag through a -1 frame shift of the ribosome on a slippery heptanucleotide sequence (UUUUUUA) at the 3'-end of the Gag coding sequence (31).
Vpu and Env are encoded by the same transcript with the Vpu coding sequence overlapping the Env start codon. It has been shown that a small uORF composed of only a start and a stop codon, which overlaps with the Vpu start codon is necessary to allow Env translation (119).

Objectives

II Objectives of this thesis

APOBEC3G is a cellular restriction factor which induces abortion of the HIV replication cycle through hypermutation of the viral genome. This restriction is prevented by the HIV Vif protein which counteracts A3G using three different mechanisms: (i) A3G ubiquitination and degradation (ii) A3G transcriptional inhibition and (iii) A3G translational inhibition. A3G translational inhibition is responsible for around 50 % of the decrease in A3G levels in response to Vif and is sufficient in itself to markedly reduce the capacity of A3G to restrict HIV infection (77). My laboratory has studied A3G translational inhibition by Vif for several years and has shown the importance of this mode of A3G counteraction. Previous studies in the lab have also identified the 5'-UTR of A3G mRNA as an essential element in A3G translational inhibition by Vif (77). Moreover, Vif has been shown to bind to the A3G 5'-UTR (147). Despite its importance, the mechanism of A3G translational inhibition by Vif is still largely unknown. The overall aim of this thesis was therefore to contribute to a better understanding of this mode of action by Vif.

To this end I have followed different objectives:

(i) The first objective was to evaluate the importance of a short uORF in the 5'-UTR of the A3G mRNA in A3G translation and regulation by Vif. The aim was to better understand which characteristics of the uORF are important for A3G translation and Vif-mediated regulation. We also wanted to compare A3G translational inhibition to other A3 proteins. Finally, we aimed to generate a T-cell line where the A3G uORF is deleted in order to characterize its importance in viral infection.

(ii) The second objective was to identify cellular factors that might play a role in Vif-mediated translational inhibition of A3G. Indeed, several examples are known of RBPs that regulate translation, notably in the presence of a uORF. Our hypothesis was that inhibitory cellular proteins might for instance be recruited on the 5'-UTR of A3G mRNA by Vif or that Vif might interact with stimulatory protein complexes and cause their drop-off from A3G mRNA. To test this hypothesis, I have first characterized the proteic interactome of the 5'-UTR of A3G mRNA. By comparison of proteins in the presence and absence of Vif and the uORF, the objective

was to identify cellular factors whose presence on A3G mRNA depends on one or the other. Subsequently I have developed a protocol with the aim of studying the full-length A3G mRNA.

(iii) Various proteins have been identified that specifically interact with A3G mRNA and are modulated by Vif. The final objective of my thesis was to evaluate their direct interaction with Vif as well as their effect on A3G translation and Vif-mediated translation inhibition.

Importance of a conserved uORF in the 5'-UTR of A3G and A3F mRNA

III. Importance of a conserved uORF in the 5'-UTR of A3G and A3F mRNA

1. A conserved uORF in APOBEC3G mRNA impacts its own translation and is essential to its translational inhibition by the HIV-1 Vif protein

The SL2 and 3 in the 5'-UTR of A3G mRNA have previously been shown to be essential for Vif-mediated translational inhibition (77). A short upstream ORF (uORF) spanning these two SLs has then been discovered in the lab. This has opened a number of possibilities for A3G translational regulation. In the following study, I have contributed to the better understanding of the mode of translation of A3G in the presence of the uORF. Moreover, we have studied the importance of different characteristics of the uORF on Vif-mediated translational inhibition of A3G. Comparison of A3G with other A3 proteins has shown that translational inhibition is unique to A3G and A3F which contain the uORF in their 5'-UTR. Finally, the localization of A3G mRNA in the cell has been studied and showed a Vif-mediated and uORF-dependent relocalization into stress granules. All of this has contributed to characterize the role of the uORF in the 5'-UTR of A3G and to determine its importance in translation and regulation by Vif. The results are described in the following article, which will soon be submitted for publication.

A conserved uORF in APOBEC3G mRNA impacts its own translation and is essential to its translational inhibition by the HIV-1 Vif protein

Camille Libre[1,3,#], Tanja Seissler[1,#], Santiago Guerrero[1,4], Julien Batisse[1,5], Benjamin Stupfler[1], Orian Gilmer[1], Melanie Weber[1], Andrea Cimarelli[2], Lucie Etienne[2], Roland Marquet[1] & Jean-Christophe Paillart[1,*]

[1]Université de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, UPR 9002, F-67000 Strasbourg, France
[2]CIRI-International Center for Infectiology Research, INSERM U1111, Université Claude Bernard Lyon 1, CNRS, UMR5308, Ecole Normale Supérieure de Lyon, Université Lyon, F-69000 Lyon, France

*To whom correspondence should be addressed: Tel: (+33) (0)3 88 41 70 35; Fax: (+33) (0)3 88 60 22 18
E-mail: jc.paillart@ibmc-cnrs.unistra.fr

[#] These authors must be considered as co-first authors

Present address:
[3]Cancer Research Center of Lyon, Clinical & Experimental models of lymphomagenesis, INSERM U1052 CNRS 5286, Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux. Université Claude Bernard Lyon 1, F-69000 Lyon, France
[4]Centro de Investigación Genética y Genómica, Facultad de Ciencias de la Salud Eugenio Espejo, Universidad UTE, 170129 Quito, Ecuador.
[5]IGBMC, Université de Strasbourg, CNRS UMR 7104, INSERM U1258, F-67404 Illkirch-Graffenstaden, France

ABSTRACT

The HIV-1 viral infectivity factor (Vif) is essential for viral fitness and pathogenicity. Vif targets cellular cytosine deaminases, APOBEC3G (A3G), A3F, A3D and A3H which inhibit HIV-1 replication by inducing extensive mutation (G to A) in the (-) strand DNA during reverse transcription and by reducing the efficiency of reverse transcription. Vif counteracts A3G (1) by recruiting an E3-ubiquitin complex and inducing its degradation through the proteasome pathway, (2) by reducing A3G transcription, and (3) by inhibiting its translation at the mRNA level, thus preventing its incorporation into viral particles. Recently, we showed that two distal stem-loop (SL2-SL3) structures within the 5' untranslated region (5'-UTR) of A3G mRNA are important for the Vif-mediated translational inhibition of A3G. Here, we uncovered the importance of a short and conserved uORF (upstream Open Reading Frame) within SL2-SL3. Through an extensive mutagenesis study of the A3G 5'-UTR and uORF, combined to an analysis of their translational level after HEK293T transfection, we defined that A3G mRNA was translated through an original dual leaky-scanning and re-initiation mechanism. Interestingly, complete disruption of the uORF abrogated the Vif-mediated downregulation of A3G translation. Other A3 proteins were not regulated at the translational level by Vif. Furthermore, in stress conditions, we showed that the targeting of A3G mRNA into stress granules was dependent not only on Vif, but also on the uORF, thus participating to an eventual suppression of A3G expression. Taken together, we show for the first time that A3G translation is regulated by a small uORF conserved in the human population within A3G 5'-UTR and that Vif uses this specific motif to repress the translation and to target it to stress granules.

INTRODUCTION

The Human Immunodeficiency virus Vif (Viral infectivity factor) protein is essential for infectious particles production in targeted cells (1). Indeed, early studies demonstrated that Vif was necessary for virus replication in primary lymphoid and myeloid cells (also called non-permissive cells) but dispensable in a subset of immortalized T cell lines (permissive cells) (2–4). This original characteristic is due to the expression of a dominant inhibitor of HIV-1 replication in non-permissive cells (5, 6), lately identified as APOBEC3G (Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3) or A3G (7). A3G belongs to a larger family of cytidine deaminases (A3A to A3H) that interfere with the reverse transcription by inducing mutations during the synthesis of the viral single-stranded DNA thus leading to cytidine to uridine transitions and the production of non-infectious viral particles (8). Amongst these deaminases, A3D, A3F, A3G and A3H have been shown to efficiently block HIV-1 replication after viral entry (9–14). HIV-1 expresses the auxiliary Vif protein to counteract the highly potent, intrinsic antiviral activities of A3 proteins, and A3F and A3G in particular which are the most active against HIV-1 (9, 14, 15). In the absence of Vif, A3G and A3F are packaged into viral particles (16) to induce viral DNA hypermutations at the next replication cycle, which in turn results in non-functional virions. Furthermore, there is also evidence that A3G and A3F can inhibit the reverse transcription and the integration steps through a deamination-independent mechanism (17–21). In HIV-1, three distinct and mutually reinforcing mechanisms are employed by Vif to reduce A3G expression and counteract its antiviral activity (22). First, Vif recruits a Cullin-RING E3 ligase 5 (CRL5) complex (composed of the Cullin 5, Elongin B/C, and Rbx2 proteins) to A3 proteins to promote their polyubiquitination and subsequent proteasomal degradation (23). This pathway has been well characterized and is the principal mechanism of A3G and A3F inhibition by Vif (24–26). Secondly, the interaction of the transcriptional cofactor CBFβ (Core Binding Factor β) with Vif-CRL5 affects its association with the RUNX family of transcription factors, leading to the downregulation of RUNX-dependent genes of which A3G belongs (27, 28). And third, Vif counteracts A3G expression by reducing its translation (29, 30). Indeed, it has been shown that two stem-loop structures (SL2-SL3) within the 5'-untranslated region (UTR) of A3G mRNA are essential for the translational inhibition of A3G by Vif (31). Importantly, both proteasomal degradation and translational inhibition of A3G by Vif participate to reduce the intracellular level of A3G and inhibit its packaging into viral

particles, demonstrating that HIV-1 has evolved redundant mechanisms to specifically inhibit the potent antiviral activity of A3G.

Regulation of translation represents a critical layer of gene expression control in essentially all cells. It allows rapid and localized changes in the expression of proteins in response to extra- and intracellular stimuli, thus being crucial for a large number of important cellular processes. Translational control can occur on a global basis by modifications of the basic translation machinery, or selectively target defined subsets of mRNAs. The latter commonly involves the sequence-specific recognition of target mRNAs by trans-acting factors such as miRNA complexes or RBP (RNA-binding protein) (32, 33). To better understand the translational regulation of A3G by Vif, and the role of the two stem-loop structures within the 5'-UTR of A3G mRNA, we searched for *cis*-acting regulatory elements within this region. We uncovered the importance of a short and conserved uORF (upstream Open Reading Frame) in the distal part (SL2-SL3) of the 5'-UTR of A3G mRNA sequences. Given that uORFs usually correlate with reduced protein expression of the downstream ORF, we studied the impact of this uORF on A3G translation and on its Vif-mediated translational repression. Through an extensive mutagenesis study of the A3G 5'-UTR and uORF, combined to an analysis of their translational level after HEK293T transfection, we defined that A3G mRNA was translated through an original dual leaky-scanning and re-initiation mechanism. Interestingly, complete disruption of the uORF abrogated the Vif-mediated downregulation of A3G translation. Other A3 proteins did not seem to be regulated at the translational level by Vif. Furthermore, in stress conditions, we showed that the targeting of A3G mRNA into stress granules was dependent not only on Vif, but also on the presence of the uORF, thus participating to an eventual suppression of A3G expression. Taken together, we show for the first time that A3G translation is regulated by a small uORF embedded in its 5'-UTR and that Vif uses this specific motif to repress the translation and target it to stress granules.

MATERIAL AND METHODS

Plasmids

Plasmid pCMV-hA3G has been previously described (Mercenne et al, 2010). The pCMV-hA3F plasmid was obtained after mRNA isolation from H9 cells and RT-PCR analysis. Amplified PCR products containing the A3F mRNA sequence were digested by EcoRI and XbaI and ligated into pCMV6-XL4 previously digested with the same

restriction enzymes. Mutated plasmids were generated by Quick-Change Site-directed Mutagenesis (table 1) (Agilent Technology) based on the secondary structure model of the 5'UTR of A3G mRNA (Mercenne et al, 2010) and verified by DNA sequencing (Eurofins, Germany). The A3 uORF2 mutants were constructed by adding a G after the uAUG (translation initiation codon of the uORF) and deleting a G before the uUGA (termination codon of the uORF) in order to place the uUGA in frame with the uAUG. Vif was expressed from pcDNA-hVif expression vector encoding codon-optimized NL4.3 Vif (Nguyen et al 2004).

RACE-PCR

Rapid Amplification of cDNA-ends by PCR (RACE-PCR) was performed following the instructions of the supplier in the 5'/3' RACE Kit, 2nd Generation (Roche). For the 5'-RACE-PCR, 0.2-0.5 µg of human spleen total RNA (Life Technologies) served as matrix to synthesize the cDNA corresponding to the 5'-end of A3 RNAs by using the Transcriptor Reverse Transcriptase and a specific primer 1 (SP1) according to manufacturer recommendations. The cDNAs were produced for 1 h at 55°C and the reaction was stopped by heating the mixture at 85°C for 5 min. After a purification step, a poly-A tail was added to cDNAs which were then amplified by a second PCR. This PCR used a dT-Anchor Primer and a second SP2 (0.25 µM each) to amplify 5 µl of polyadenylated cDNAs in a 50 µL mix containing 1 U of Phusion Polymerase, 0.2 mM dNTPs and 1.5 mM $MgCl_2$. The PCR protocol was the following: 3 min at 98°C and 10 cycles of 15 s at 98°C, 30 s at the optimal melting temperature and 1 min at 72°C. The elongation step was then increased by 20 s until it reached 2 min and 23 cycles were performed. A final elongation step of 7 min ends the amplification. A nested PCR was performed with 1 µl of amplicons from the last PCR, in an identical reactional mixture, except for the SP3 used.

For the 3'-RACE-PCR, as mRNAs for the total RNA extract were already polyadenylated, the purification and poly-A tailing steps were not necessary. Using 1 µg of total RNA and a dT-Anchor primer, the cDNAs were synthetized according to the manufacturer protocol and used for a PCR amplification with a PCR Anchor Primer and SP5 primer. Both PCRs were performed as described above. The nested PCR, missing from the initial protocol, was added with a SP6 primer and the same PCR Anchor Primer used in the last PCR to obtain enough material for the bacterial transformation.

Cell culture

HEK 293T cells were maintained in Dulbecco's modified Eagle's medium (DMEM, Life Technologies) supplemented with 10% fetal bovine serum (PAA) and 100 U/ml penicillin/streptomycin (Life Technologies) at 37°C and 5% CO_2 atmosphere. Transfections of HEK 293T cells were carried out using the X-tremeGene 9 DNA Transfection Reagent (SIGMA) as recommended by the manufacturer. Briefly, 500,000 cells/well were seeded at 70% confluence in a 6-well plate and co-transfected with 100 ng of pCMV-hA3G constructs and 1 ug of pcDNA-hVif. Cells were also exposed to the chemical proteasome inhibitor ALLN (25 µM) for 14 h.

For FISH and immunofluorescence (IF) experiments, cells were plated on a glass coverslip in 6 wells-microplate (500,000 cells/well for HEK 293T cells and 350,000 cells/well for HeLa cells). Cells were cultured in DMEM. Transfections of HEK 293T or Hela cells were carried out using the X-tremeGENE 9 DNA Transfection Reagent (Roche) as described above. The exception was the co-transfection of 0.5 µg of stress granules (GFP-PABP or GFP-TIA1) or P-Bodies (GFP-DCP1 or GFP-AGO2) markers when required. Twenty-four hours post transfection, stress induction was performed either by treating the cells with 500 µM sodium arsenic (AsNa, Sigma-Aldrich) or by incubating them at 44°C (heat shock) for 30 min before for further analysis.

Immunoblotting

Twenty-four hours post-transfection, cells were washed in PBS 1X (140 mM NaCl, 8 mM NaH_2PO_4, 2 mM Na_2HPO_4) and lysed for 10 min at 4°C in RIPA 1X (PBS 1X, 1% NP40, 0.5% sodium deoxycholate, 0.05% SDS) supplemented with protease inhibitors (complete EDTA Free cocktail, Roche). After 1 h centrifugation at 14,000 g, cell lysates were adjusted to equivalent protein concentration (Bradford assay, Bio-Rad), fractionated on Criterion TGX 4-15% gels (Bio-Rad) and transferred onto a 22 µm PVDF membranes using the Trans-Blot Turbo™ Transfer System (Bio-Rad). Blots were probed with appropriate primary antibodies. Polyclonal anti-A3G (#9968), anti-A3F (#11226) and monoclonal anti-Vif (#319) antibodies were obtained through the NIH AIDS Research and Reference Reagent Program. Monoclonal anti-β-actin antibody was purchased from SIGMA (#A5316). The PVDF membranes were incubated with horseradish peroxidase-conjugated secondary antibodies (Bio-Rad), and the proteins were visualized by enhanced chemiluminescence (ECL) using the ECL Prime Western blotting detection reagent (GE Healthcare) and the ChemiDoc™

Touch Imaging System (Bio-Rad). Bands were quantified using Image J. Student's T-test was used to determine statistical significance.

Real-time qPCR

Twenty-four hours post-transfection, total RNA was isolated from HEK293T cells using TRI-Reagent (SIGMA). After RNase-free DNase treatment (Roche), total RNA was isolated by phenol/chloroform extraction followed by ethanol precipitation. Total RNA (1 ug) was reverse-transcribed using the iScript^{TM} Reverse Transcription Supermix (Bio-Rad) as recommended by the manufacturer. Subsequent qPCR analysis was performed using the SYBR Green qPCR Master Mix (ThermoFisher) and was monitored on a CFX Real Time System (Bio-Rad). Gene-specific primers were: A3G forward primer, 5'-GGATCCACCCACATTCACTT-3', and reverse primer, 5'-ATGCGCTCCACCTCATAAC-3'; A3F forward primer, 5'-CCCGATGGAGGCAA TGTATC-3', and reverse primer, 5'-GAGATAGGTGAGTGGTGCTTTAC-3'; β-actin forward primer, 5'-GGACTTCGAGCAAGAGATGG-3', and reverse primer, 5'-AGCACTGTGTTGGCGTACAG-3'. The A3G and A3F mRNA levels were normalized to those of actin mRNA and relative quantification was determined using the standard curve-based method.

FISH and immunofluorescence (IF) assays

The FISH probe was obtained as follow: fragment from nucleotide position 100 to 406 of A3G mRNA was cloned between EcoRI and Xba1 sites into a pcDNA vector. After linearization by XbaI, T7 *in vitro* transcription was performed in presence of conjugated DIG-11-UTP (Roche) following the manufacturer instruction (1 mM of each dNTP except dUTP at 0.65 mM and 0.35 mM of the labeled DIG-11-UTP). After DNAse I treatment, A3G specific probe was purified by phenol-chloroform extraction and ethanol precipitation. Pellet were resuspended in 50 µl milliQ water giving a 100X concentrated A3G mRNA probe. Aliquots of 1 µl were stored at -80°C.

For FISH and IF assays, cells were fixed with 4% (wt/vol) paraformaldehyde/PBS for 20 min at RT. Fixation was stopped in 100 mM glycine for 10 min at room temperature. Cells were then permeabilized with 0.2% (wt/vol) triton X-100/PBS solution at room temperature (RT) for 5 min. For FISH assays, after 2 washes in PBS, coverslip was treated for 15 min at room temperature with DNAse I solution (Roche - 25U/coverslip). Coverslips were washed with 1 ml of PBS. One µl of specific A3G mRNA probe was

diluted 100x in milliQ water (concentration around 5 ng/µl). Coverslips were loaded with 50 µl of pre-warmed Hybridization Solution (formamide 50%; tRNA 0.1 µg/µl; SSPE 2X; Denharts solution 5X; RNAse OUT 0.1 U/µl and 25 ng of specific probe) for 16 h at 42°C, in humid atmosphere. Coverslips were then washed with 50 µl of pre-warmed buffer containing 50% formamide and SSPE2X for 15 min at 42°C, and then twice with 50 µl SSPE2X for 5 min at 42°C. Finally, for both FISH and IF assays, after a wash in PBS, coverslips were blocked with 3% (wt/vol) bovine serum albumin (BSA) in PBS for 1 h at room temperature. Primary antibodies were diluted in 3% (wt/vol) BSA/PBS and incubated 3 h at 37°C, followed by incubation of secondary antibodies at room temperature for 2 h in the dark. After a brief wash in PBS, coverslips were mounted in one drop of SlowFade gold antifade reagent (Thermo Fisher) with (IF) or without (FISH) DAPI (Thermo Fisher).

The following primary antibodies were used: sheep polyclonal anti-DIG (Sigma Aldrich); goat and rabbit polyclonal anti-TIA-1; goat and rabbit polyclonal anti-DCP-1 were purchased from Santa Cruz Biotechnology. Rabbit polyclonal anti-PABP (ab21060) and rabbit monoclonal anti-phospho-S51-EIF2alpha (ab32157) was purchased from Abcam. Rabbit polyclonal anti-hA3G-C17 (#10082) and mouse monoclonal anti-HIV-1 Vif (#319) antibodies were obtained through the NIH AIDS Research and Reference Reagent Program. All were used at a $1/500^{e}$ dilution. The following dye-conjugated secondary antibodies were used: Alexa Fluor 647 anti-rabbit, Alexa Fluor 647 anti-goat, Alexa Fluor 488 anti-mouse, Alexa Fluor 546 anti-sheep, Alexa Fluor 405 anti-mouse, Alexa Fluor 594 anti-goat (Life Technologies). All were used at a $1/200^{e}$ dilution.

Microscopy and image analysis

Images were acquired on a Zeiss LSM780 spectral microscope running Zen software, with a 63x,1.4NA Plan-Apochromat oil objective at the IBMP microscopy and cellular imaging platform (Strasbourg, France). Excitation and emission settings were spectrally selected among the 4 laser wavelengths available in the microscope (405; 488; 561; and 633 nm) according to the secondary antibody used. Image processing (contrast, brightness and merges) were performed with ImageJ 1.43m software (ref: Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/, 1997-2012.). An additional macro allowing the multichannel profile plot were design by Jerome Mutterer from the Microscopy

platform. Percentage of cells showing a co-localization between FISH signal (A3G mRNA) and stress or P-Body markers were counted on subsets of 100 cells (n=3).

APOBEC3 genotype data

The NCBI dbSNP database (https://www.ncbi.nlm.nih.gov/snp, July 31st 2019) as well as the UCSC Genome Browser database (https://genome.ucsc.edu/, July 31st 2019) were mined for polymorphic variants in human APOBEC3G and APOBEC3F mRNA 5'UTR regions.

RESULTS

***In silico* analysis of A3G mRNA 5'UTR: identification of a conserved uORF**

To further understand the mechanism by which Vif inhibits A3G mRNA in a 5'UTR-dependent manner (30, 31), we searched for *cis*-acting elements within the 5'-UTR that could contribute to this repression. Herein, using a computational platform capable of identifying RNA regulatory elements (RegRNA 2.0) (34), we identified a terminal oligopyrimidine (TOP) element and an uORF within the 5'-UTR of A3G mRNA sequence (Figure 1). Interestingly, TOP elements are thought to serve as *cis*-regulatory sequences that inhibit the binding of proteins implicated in the translational regulation (35). Indirectly, we showed that this TOP element is not required in either A3G translation or Vif-mediated A3G translation inhibition (31). Indeed, RNA mutant A3G SL2-SL3 (deletion of SL1, which contains the TOP element) showed similar expression profile and Vif-mediated inhibition as wild-type (wt) A3G RNA after transfection of HEK 2913T cells (see figure 5 in (31)). uORFs are also *cis*-regulatory elements that negatively regulate translation of downstream ORF (36). Concerning A3G mRNAs, our analysis showed that this element is encoded within the SL2 and SL3 domains of the 5'-UTR (figure 1). Of note, we previously showed that these two SLs are required for Vif-mediated A3G translation inhibition (31), suggesting this uORF could account for the translational inhibition previously observed. The uORF is positioned between nucleotides 177 (initiation codon uAUG) and 246 (termination codon uUGA) within the 5'-UTR, forty-nine nucleotides upstream the main AUG (mAUG) codon of A3G (nucleotide 298), and encodes a putative peptide of 23 amino acid (figure 1) without any particular motif according to the PROSITE database (https://prosite.expasy.org). We also observed that the Kozak context around the

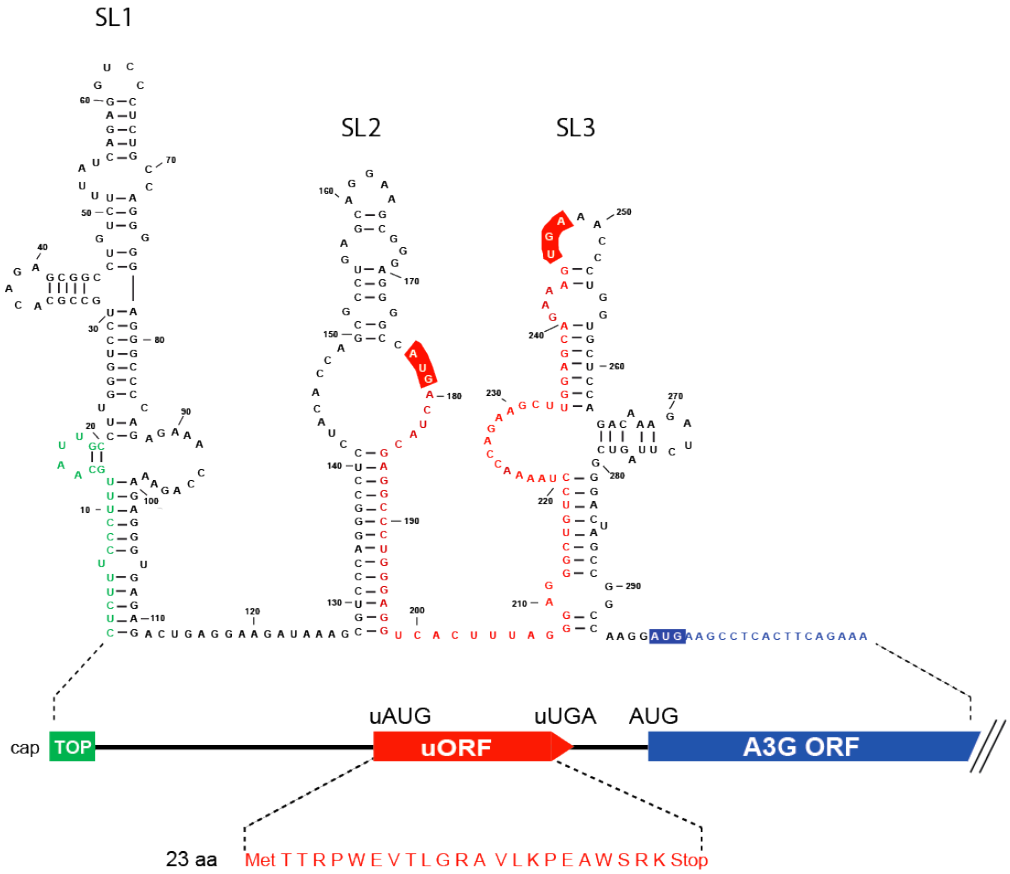


Figure 1: Schematic representation of the 5'-UTR of A3G mRNA. The secondary structure of the 5'-UTRs of A3G mRNA is indicated, as well as the TOP element (green) and of the uORF (red). The putative peptide expressed from the uORF is also indicated.

upstream and major ORFs is fairly strong (GGGGCCAUGA and CCAAGGAUGA for the uAUG and mAUG, respectively; the initiation codon of translation is underlined), suggesting the uAUG is functional. This uORF is also present in A3F mRNA with 94% and 96% identity at the nucleotide and amino acid level, respectively. To determine if these uORFs are conserved in the human population, we mined the NCBI dbSNP (single nucleotide polymorphism database). In the A3G uORF (+/- 10 nts) region, which included the Kozak context, we did not find any variant with a Minor Allele Frequency (MAF) above 0.005 (Table S1). In the A3F uORF region, there was a single common SNP (MAF>0.01, rs35898507) that would induce an amino acid change from G to V (G being the ancestral allele present at >96% in the population) (Table S1). We did not find any variant with a MAF>0.0005 that would impact the ORF of the A3G and A3F uORFs (i.e. that would impact the start codon, the stop codon, or would induce a frameshift). Lastly, the distance of the uORF to the mAUG of both genes was also

conserved. This shows that A3G and A3F uORFs are conserved in the human population.

The uORF represses the translation of A3G mRNA

Since uORFs reduce protein expression by approximatively 30-80% (36), we asked whether this motif was also responsible for the regulation of A3G mRNA expression. First, we inactivated the uORF of this mRNA by either deleting the whole uORF (A3G ΔuORF), or by substituting the uAUG (A3G suAUG) (figure 2A). The expression of these mutants was examined after transfection of HEK293T cells and immunoblotting against A3G protein (figure 2B). As expected, uORF inactivation significantly increased A3G protein expression (figure 2B, compare lanes 2-3 to lane 1), suggesting the uORF intrinsically repressed the translation of A3G mRNA.

Next, we wondered whether this translational regulation could be extended to other A3 proteins. Thus, we performed 5'- and 3'-RACE (5'-rapid amplification of cDNA ends analysis) from human spleen total RNA with A3B, A3C, A3D and A3H specific primers (table 2) in order to identify 5'- and 3'-UTRs (A3A vector was a generous gift from Dr. Vincent Caval, Pasteur Institute, Paris; A3F and A3G were already available in the lab). After cloning and sequencing of the various transformants, corresponding full-length mRNA sequences were synthesized (Proteogenix, France) and consequently cloned into pCMV vectors (figure 3A). The 5'-UTR sequences comprised 44, 106, 407 and 134 nucleotides for A3B, A3C, A3D and A3H, respectively, while their 3'-UTR sequences contained 572, 431, 934 and 372 nucleotides, respectively (figure 3A). Sequences obtained for the 5'-UTR of A3C and A3D correspond to the one published in GenBank™ (accession number NM_014508.2 for A3C, and NM_152426.3 for A3D). The 5'-UTR sequence of A3B is shorter (63 nucleotides) than the reference (NM_004900.4) while the one from A3H is longer (7 nucleotides) than the reference (NM_001166003.1). All A3 clones were transfected into HEK293T cells in presence or absence of Vif and with or without ALLN (proteasome inhibitor) in order to discriminate the translational inhibition from the proteasomal degradation (31). Our results showed that all A3 proteins are well

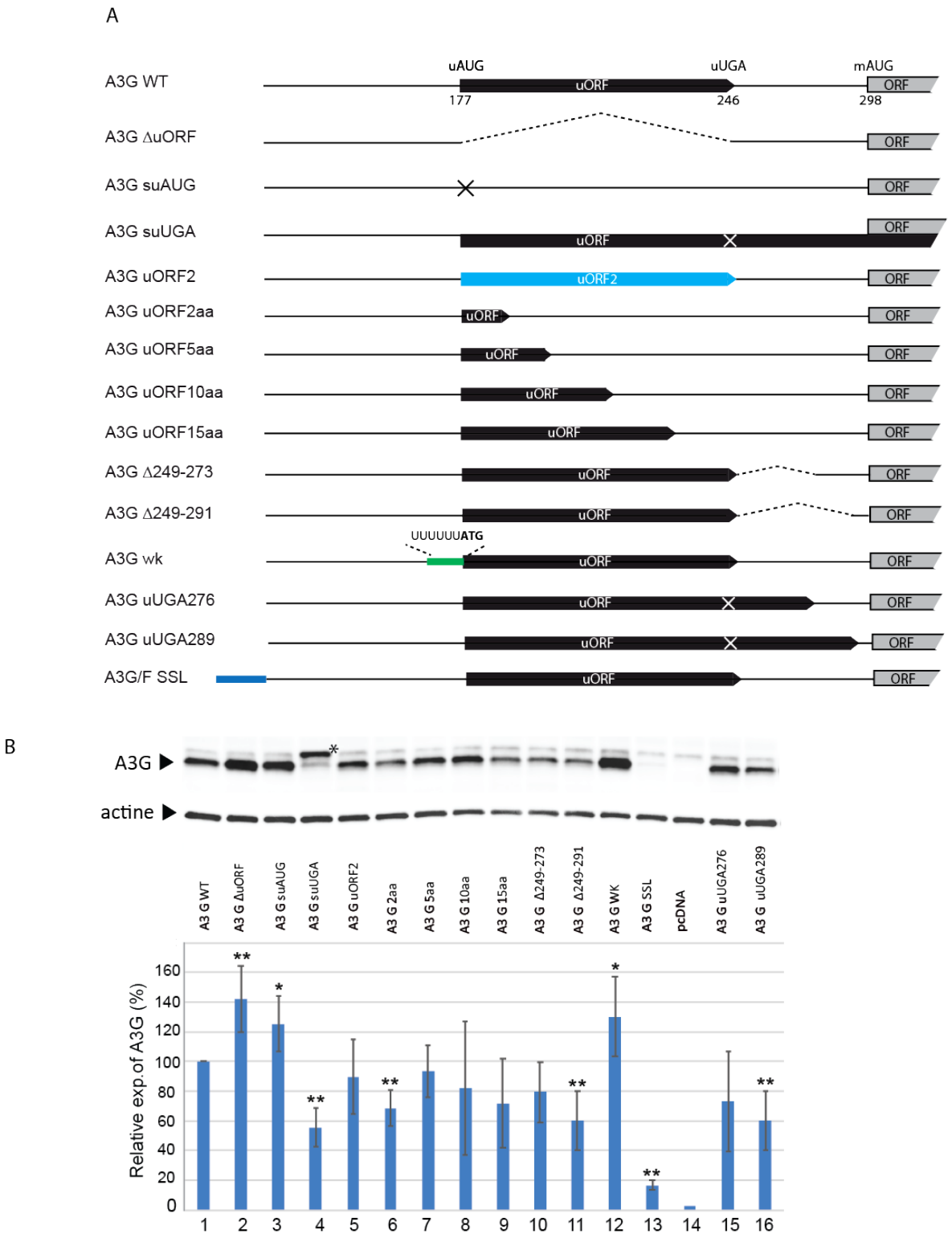


Figure 2: Importance of the uORF for A3G mRNA expression. A) Schematic representation of the different A3G 5'-UTR constructs used in this study. Wild-type A3G mRNA and uORF mutants are represented. B) HEK293T cells were transfected with wild-type or mutated plasmids. Proteins were separated by SDS/PAGE and analyzed by immunoblotting. Bands were quantified using Image J and relative expression of A3G is represented in a histogram. Standard deviations are representative for at least three independent experiments. P-values are indicated as follows: *<0,05; **<0,01.

expressed from full-length mRNA constructs (figure 3B, lanes 1) and are degraded by Vif as expected (figure 3, lanes 2), with the exception of A3A and A3B. However, in presence of ALLN, we only observed a significant decrease of A3G and A3F expression when Vif was present (figure 3, compare lanes 3 & 4), suggesting that these two A3 proteins are the only ones to be regulated at the translational level by Vif. Moreover, sequence analysis of the 5'-UTR of A3A, A3B, A3C, A3D, and A3H did not reveal the presence of any uORFs or other regulatory motifs.

Mechanism of A3G mRNA translation

The results presented above show that the uORF negatively regulates the translation of A3G mRNA. According to the literature (37), uORFs can regulate the translation of the main ORF in different manners, in *cis* or in *trans* through the peptide synthesized, and mechanisms such as leaky-scanning, direct translation through an IRES or re-initiation can be envisaged. In order to analyze several aspects of the translation process, we constructed a series of A3G mRNAs with mutations in the uORF sequence or its surrounding nucleotides and analyzed their expression after transfection of HEK293T cells (figure 2).

First, we define the importance of the uORF peptide as a *trans*-acting element capable of regulating the main ORF translation. To achieve this, we changed the uORF amino acid sequence (see material & Methods) and studied its effect on A3G expression (mutant A3G uORF2). We observed no effect of the putative peptide in A3G expression (figure 2B, lane 5), consistent with the fact that only in rare cases peptide expressed from an uORF has regulatory effects (38, 39). This result also suggests that the functional importance of the uORF is dependent on features that drive uORF translation rather than the specific peptide produced.

Next, we examined if translation initiation at the uAUG is required for the regulatory mechanism. The frequency of ribosomal recognition of a translation initiation codon is determined by its sequence context (40), and positioning of a translation initiation codon within a "poor" sequence context will result in inefficient ribosomal recognition and bypassing (leaky scanning). We thus replaced the uAUG Kozak consensus sequence (GGGGCCAUGA) by a weak non favorable context (UUUUUUAUGA, mutant A3G wk) (figure 2A). As expected, more ribosomes fail to recognize the uAUG, and we observed a significant increase of protein expression

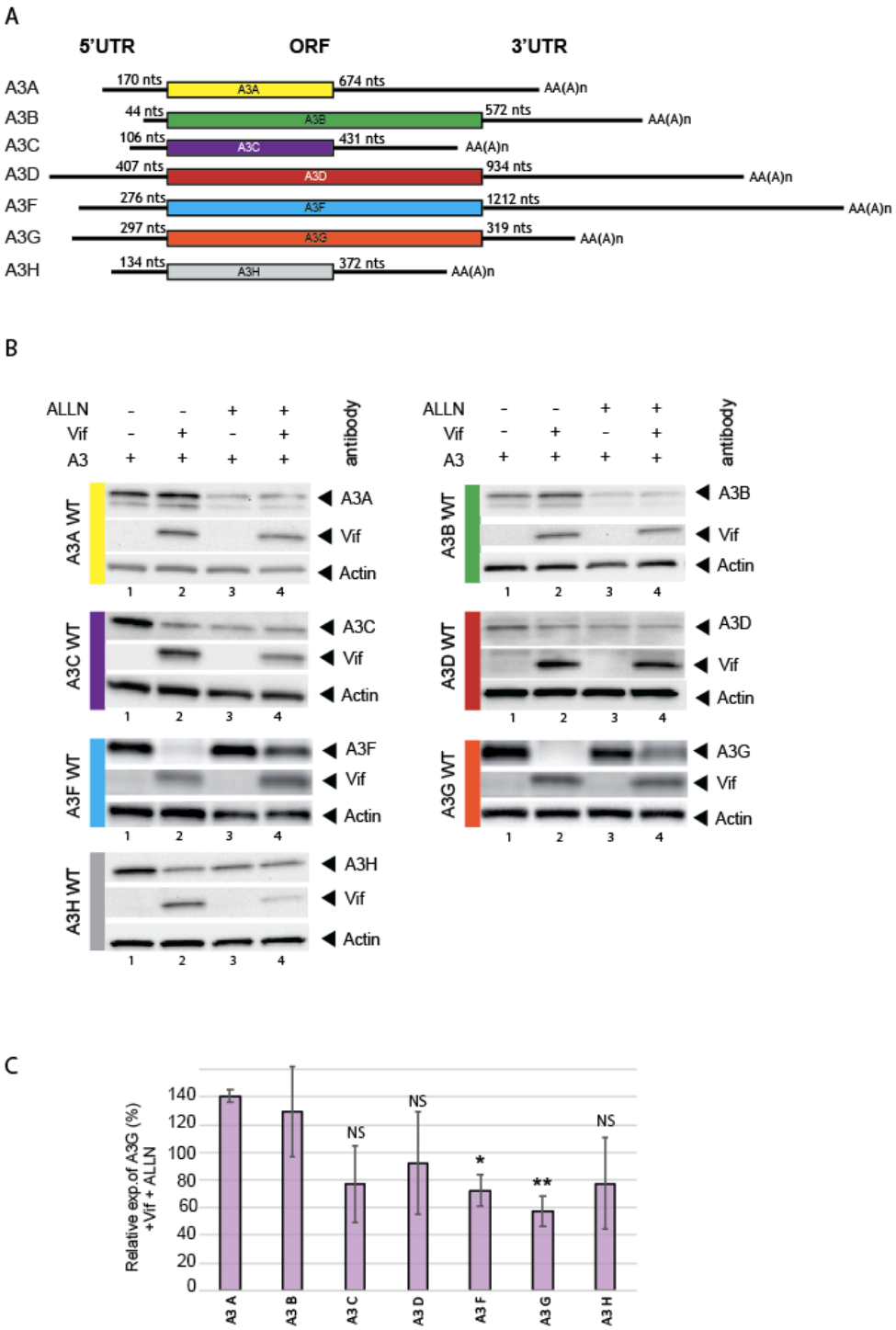


Figure 3: Vif does not inhibit the translation of all APOBEC3 mRNAs. A) Schematic representation of A3 mRNAs identified by RACE-PCR. The lengths of the various 5'- and 3'-UTRs are indicated. B) HEK293T cells were transfected with plasmids expressing the different A3 proteins in the presence or absence of Vif and in the presence or absence of a proteasome inhibitor (ALLN). Proteins were separated by SDS/PAGE and analyzed by immunoblotting. C) Bands were quantified using Image J and relative expression of A3 proteins is represented. Standard deviations are representative for at least three independent experiments. P-values are indicated as follows: *<0,05; **<0,01.

(figure 2B, lanes 12), suggesting that the natural uAUG Kozak context is an essential element for the translational repression.

In a natural context of A3G mRNA, the uORF encodes a peptide of 23 amino acids. Next, we created mutant RNA where the translation termination codon of the uORF (uUGA) is inactivated and the uAUG placed in frame with the downstream major ORF (mAUG). Thus, this construct allows to directly monitor upstream translation initiation, prevent uORF termination and, hence, re-initiation, and downstream translation becomes solely dependent on leaky scanning. Interestingly, two protein bands can now be observed. Indeed, in addition to the standard A3G protein encoded by the main ORF (figure 2B, lanes 4, lower band - 384 amino acids), a N-terminally elongated version (424 amino acids) derived from initiation at the uAUG can be detected (figure 2B, lanes 4, asterisk), demonstrating that the initiation codon of the uORF (uAUG) is functional and in a functional Kozak context (see above). Moreover, we observed a diminution of the expression of the main A3G protein for suUGA mutant (about 50%) (figure 2, lanes 4). Whereas re-initiation is impossible in this case due to the mutated uORF stop-codon, these results suggest that A3G may partially be translated by re-initiation. However, the fact that wild-type A3G protein is still detected in this mutant suggests that the translation of the main ORF could be accomplished by leaky-scanning or through a direct entry of the ribosome (IRES - Internal Ribosome Entry Site) at the mAUG.

To test the IRES hypothesis, we inserted a stable stem-loop at the 5'-end of A3G mRNA (figure 2A, mutant A3G-SSL) so that a 40S ribosomal subunit cannot bind and ribosome scanning is inhibited (41). Then, if an IRES is present within the 5'-UTR of A3G mRNA, ribosomes should load directly on this site and initiate translation. As seen in figure 2B (lane 13), the expression of this mutant was very low, ruling out the hypothesis of an IRES-dependent A3G translation.

To go further concerning the possibility of a re-initiation mechanism, we constructed different mutants based on previous reports. Indeed, it is well known that re-initiation is more efficient when uORF sequences are small (42) whereas leaky-scanning is not dependent on the length of the uORF (43). Then, reducing the uORF length should enhance A3G expression if a re-initiation mechanism is involved. We thus tested mutants of A3G uORF where the putative 23 amino acids peptide was reduced to 2, 5, 10 or 15 amino acids (figure 2A). The results showed that these mutants did not significantly impact A3G protein expression whereas A3G 2aa mutant decreased it by

30 % (figure 2B, lanes 6-9), suggesting that the length of the uORF somehow regulates the translation of A3G mRNA. To validate this hypothesis and because re-initiation is also dependent on the distance between the stop-codon uUGA and the main AUG of A3G (43), we then reduced the distance between the two ORFs by either deleting half (A3G Δ249-273) or the entire (A3G Δ249-291) inter-ORF sequence, or by inserting a stop codon at positions 276 (A3G uUGA276) or 289 (A3G uUGA289) (figure 2A). In both cases, inter-ORF sequence is reduced to 25 and 9 nucleotides. If A3G translation initiation is accomplished by a re-initiation mechanism, reducing this distance should decrease protein expression. Interestingly, A3G protein expression was only significantly decreased when the inter-ORF region was reduced to its minimum, i.e. for mutants A3G Δ249-291 and A3G uUGA289 (figure 2B, lanes 11-12 and 15-16), suggesting that A3G may also be translated through a re-initiation mechanism.

To exclude effects of the uORF and mutations on mRNA stability or differential transcription rates, we tested whether the various constructions affect mRNA levels by RTqPCR. Of note, none of the mutations impacted on RNA stability (figure S1), suggesting that regulation occurs at the translation level.

Initiation at the uORF is required in Vif-mediated A3G translation inhibition

To further characterize the molecular mechanism of Vif-mediated A3G translation inhibition, we determined the uORF effect on this translational repression. We therefore transfected the different A3G mutants in HEK293T in presence or absence of Vif and with or without ALLN (proteasome inhibitor) in order to discriminate the translational inhibition from the proteasomal degradation (31). First, when we co-transfected wt A3G construct with Vif in presence of ALLN, we observed a typical 40% reduction in A3G synthesis due to A3G translational repression by Vif (figure 4, blue bar), as previously observed (31) (see also figure S2 for western blots). In a second step, we analyzed similarly the effect of Vif (+/- ALLN) on the different A3G mRNA constructs. The results showed that the ΔuORF mutant present the same expression as the wild-type A3G protein in presence of Vif (between 20-30%) (figures 4, mutants ΔuORF and figure S2) due to both proteasomal degradation and translational inhibition. However, in presence of ALLN, we did not observe any significant decrease of A3G expression when Vif was present (figure 4, blue bars, and figure S2), suggesting that the uORF is required for the Vif-mediated translational inhibition. To validate this hypothesis, we transfected A3G suAUG construct containing a single

substitution at the upstream initiation codon. As expected, translational inhibition was not observed with this mutant (figure 4, blue bar).

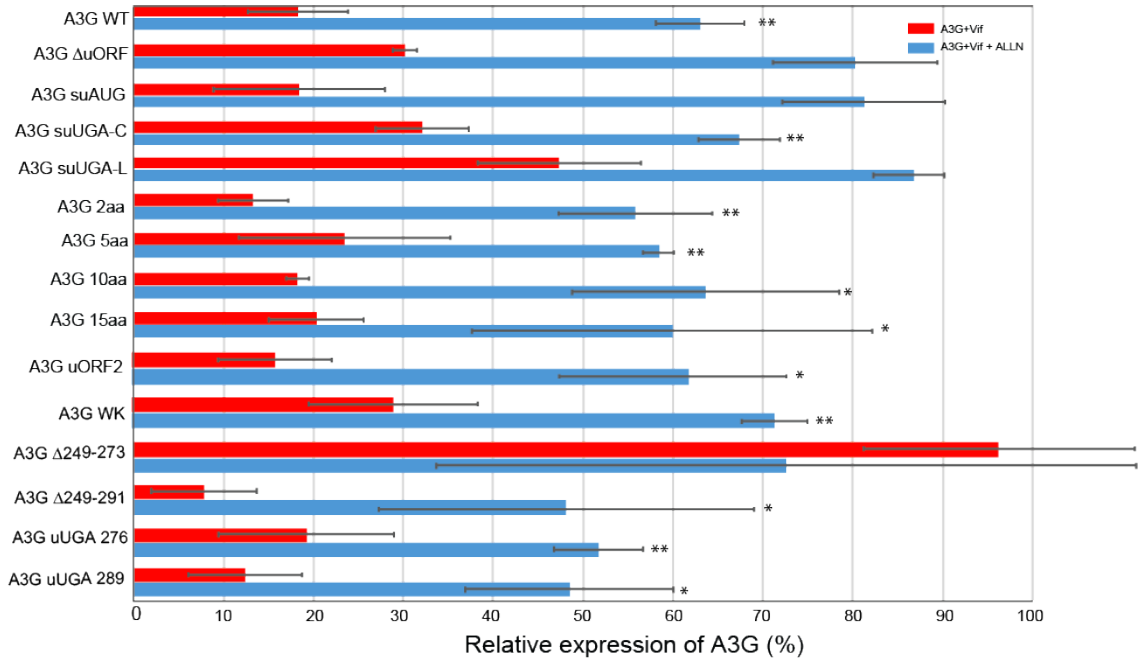


Figure 4: Effect of the uORF on Vif-mediated A3G translation inhibition. HEK293T cells were transfected with plasmids expressing wild-type and mutated A3G constructs in the presence or absence of Vif and in the presence or absence of a proteasome inhibitor (ALLN). Proteins were separated by SDS/PAGE and analyzed by immunoblotting (see supporting figure 2). Bands were quantified using Image J and relative expression of A3G proteins is represented. Histograms represent the effects of Vif on A3G: degradation and translation (red bars) and translation only (in presence of ALLN, blue bars). Standard deviations are representative for at least three independent experiments. P-values are indicated as follows: *<0,05; **<0,01.

Next, we asked whether the peptide sequence (A3G uORF2), and its size (A3G 2aa, 5aa, 10aa, 15aa uUGA276 (32 aa), uUGA289 (36aa)) were important for the translational inhibition mechanism. These different constructs were transfected into HEK293T cells as above and analyze by western blot (figure S2). Interestingly, we showed that these mutants present an expression profile similar to wild-type A3G in presence of Vif and ALLN (figure 4), with a reduction of protein expression of about 30-40% (figure 4), suggesting that the amino acid sequence, and its length are not important for the translational mechanism, ruling out a possible role of this peptide, in conjunction with Vif, to inhibit A3G translation.

Finally, we wanted to know whether the distance between the two ORFs are important for the translational inhibition of A3G by Vif. To test this effect, we co-

transfected mutants A3G uUGA276, uUGA289, Δ249-273, and Δ249-291 with Vif and ALLN as described above. As observed in figure 4, the four A3G mutants showed a regulation of their translation by Vif similar the one observed for wild-type A3G (figure 4), suggesting that the inter-ORF region does not play a role in the Vif-mediated translational repression of A3G.

Taken together, these results indicate that translation initiation at the uORF is essential for the Vif-mediated A3G translational inhibition.

Relocation of A3G mRNA to stress granules is dependent on the uORF and Vif

Mechanisms of translational control dictate which mRNA transcripts gain access to ribosomes, and this process is highly regulated by the interplay of RNA binding proteins (RBPs) and RNA granules, such as processing bodies (P-bodies) or stress granules (SG). SGs are transient foci enriched in translation initiation factors and 40S ribosomal subunits, whereas P-bodies are enriched in RNA-decay machinery. Thus, SGs and P-bodies can be considered as extensions of the messenger ribonucleoprotein (mRNP) translational control cycle, i.e. as compartments where translationally silenced mRNPs are stored (44, 45). Then, we asked whether Vif, through the uORF, could relocate A3G mRNA into these storage compartments and participate to the downregulation of A3G translation. To test this hypothesis, we analyzed the wild-type A3G mRNA and two mutant constructs: A3G suAUG and A3G Δ5'UTR. These two mutants (a single substitution of the uAUG and a complete deletion of the 5'-UTR) have been chosen because their translation is not down-regulated by Vif (this study and (31)). We began by a careful examination of the intracellular localization of A3G mRNAs and proteins by FISH and immunofluorescence (IF) analysis, respectively, after transfection of HEK293T cells (figure 5). As expected, full-length wild-type A3G mRNA and protein were detected in the cytoplasm (figure 5, lane 4). Similarly, mRNAs and proteins expressed from mutants A3G Δ5'-UTR and suAUG were also present in the cytoplasm (figure 5, lanes 5 & 6), suggesting that mutations do not impact on mRNA localization.

In a following step, we analyzed the co-localization of A3G mRNAs and proteins with SG and P-bodies markers in presence or absence of Vif. We co-transfected HEK293T cells with A3G constructs +/- Vif, and then briefly exposed them to sodium arsenite (ARS) or high temperature (44°C) to induce a stress condition. We then

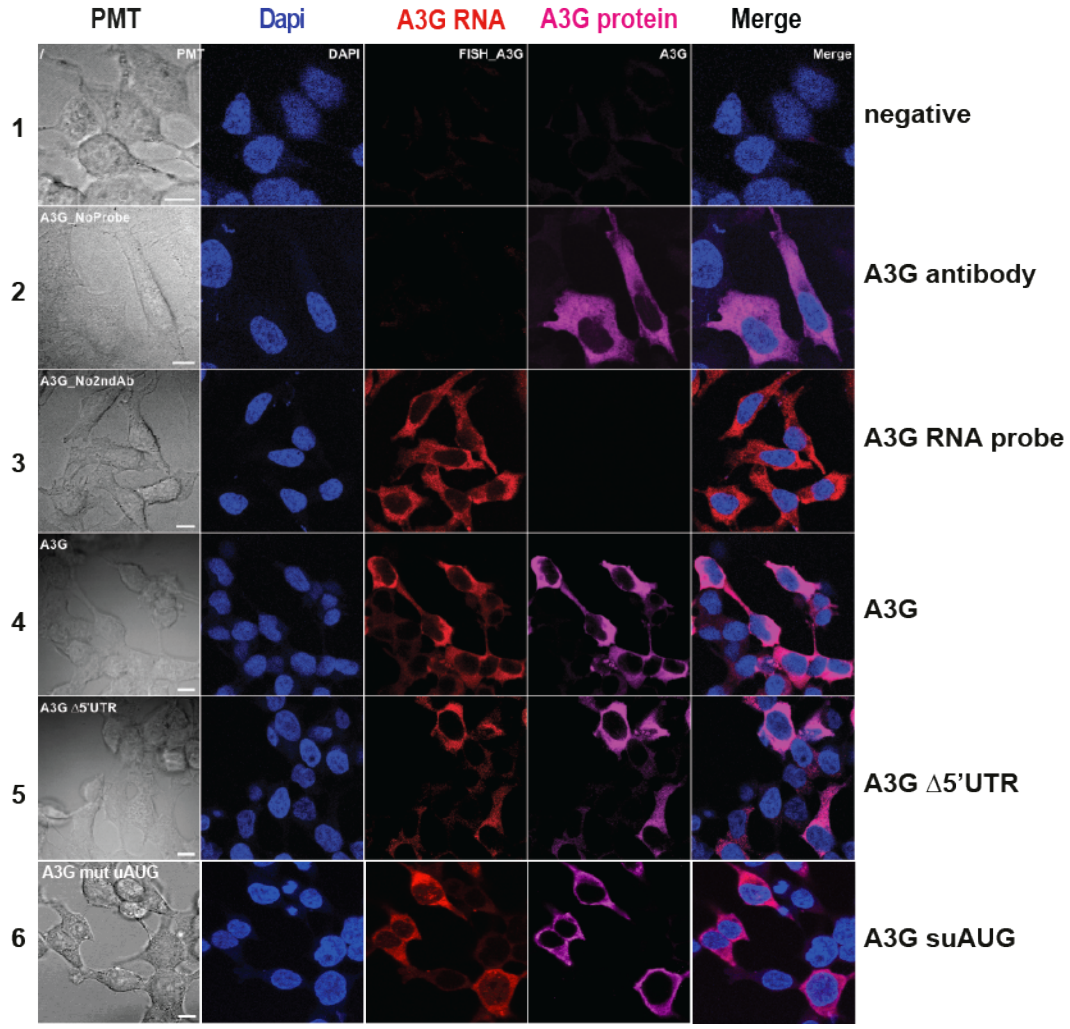


Figure 5: Analyses of A3G mRNAs and protein by FISH and immunofluorescence. HEK293T cells were transfected with plasmids expressing A3G wild-type as well as Δ5'-UTR and suAUG mRNAs. Cells were fixed and probed with anti-DIG (A3G mRNAs) and anti-A3G (A3G protein) antibodies. Cells were stained with Dapi to visualize nuclei and the images were merged digitally. Controls and A3G mRNAs (lines 1 to 7) are indicated on the right of the panels.

performed FISH analysis (A3G mRNAs) and immunofluorescence staining using antibodies directed against Vif, A3G, AGO2 and DCP1 (P-body markers) and PABP1 and TIA-1 (SG markers) proteins. First, concerning wild-type A3G mRNA, we observed that, as expected, PABP1 was localized into SGs regardless of stress conditions (figure 6A and figure S3). Interestingly, under stress conditions, while we observed a clear co-localization of A3G and PABP1 proteins into punctate granules in presence or absence of Vif (figure 6A, column 7), co-localization of A3G mRNA and PABP1 was only observed in presence of Vif (figure 6A, column 6). We obtained between 50-70% and 45-55% co-localization for PABP1/A3G mRNA (figure 6D, blue

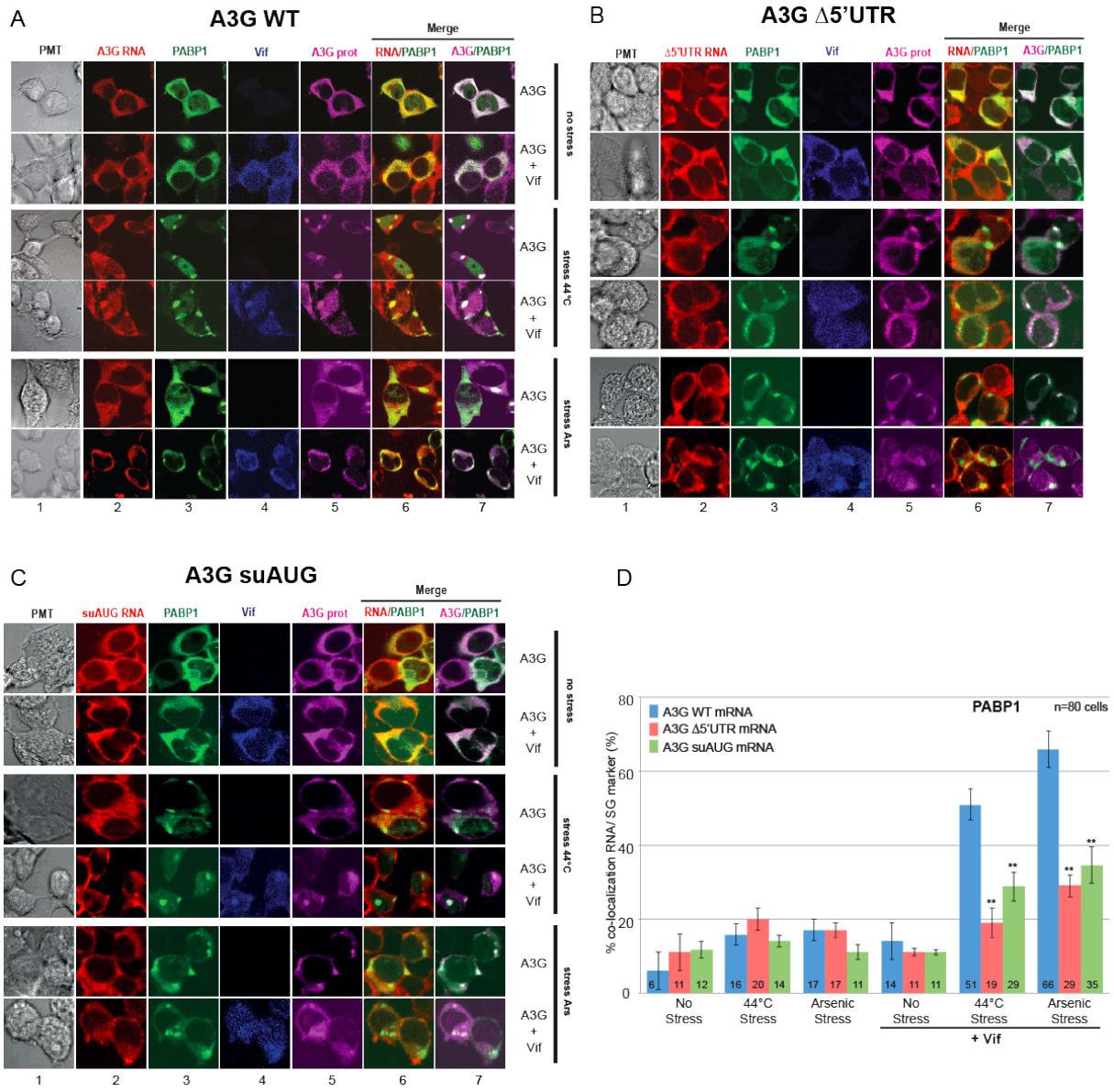


Figure 6: Importance of the uORF for the relocation of A3G mRNA into stress granule by Vif. HEK293T cells were transfected with plasmids expressing A3G wild-type (panel A) as well as Δ5'-UTR (panel B) and suAUG (panel C) mRNAs, in absence or presence of Vif. Cells were cultured in various conditions: (i) untreated (no stress) or stressed by (ii) incubation at 44°C or (iii) with arsenite sodium (ARS). Cells were fixed and probed with anti-DIG (A3G mRNAs), anti-A3G (A3G protein), anti-Vif and anti-PABP1 antibodies. Cells were stained with Dapi to visualize nuclei and the images were merged digitally. D) Histograms represent the percentage of co-localization of A3G mRNAs with PABP1. Standard deviations are representative for at least three independent experiments. P-values are indicated as follows: **<0,01.

bars) and TIA-1/A3G mRNA (figure S4), respectively. Meanwhile, we also performed FISH and IF analysis for A3G Δ5'-UTR (figure 6B) and suAUG (figure 6C) mRNA

mutants. Interestingly, whereas a co-localization of A3G and PABP1 proteins was still detected for these two constructs (figure 6B & C, column 7), we observed a significant decrease (around 30%) of the presence of A3G mRNA Δ5'-UTR and suAUG into SGs in the presence of Vif (figure 6D, red and green bars), suggesting that relocation of A3G mRNA into SGs depends not only of Vif but also of the uORF.

Furthermore, we performed similar experiments with P-bodies markers (AGO2 and DCP1) (figure 7). Under physiological conditions, while A3G and AGO2 proteins co-localized (figure 7A, column 7), A3G mRNAs (wild-type or mutants) are rarely observed co-localizing with P-bodies markers (less than 20% co-localization) regardless of the presence of Vif (figure 7B).

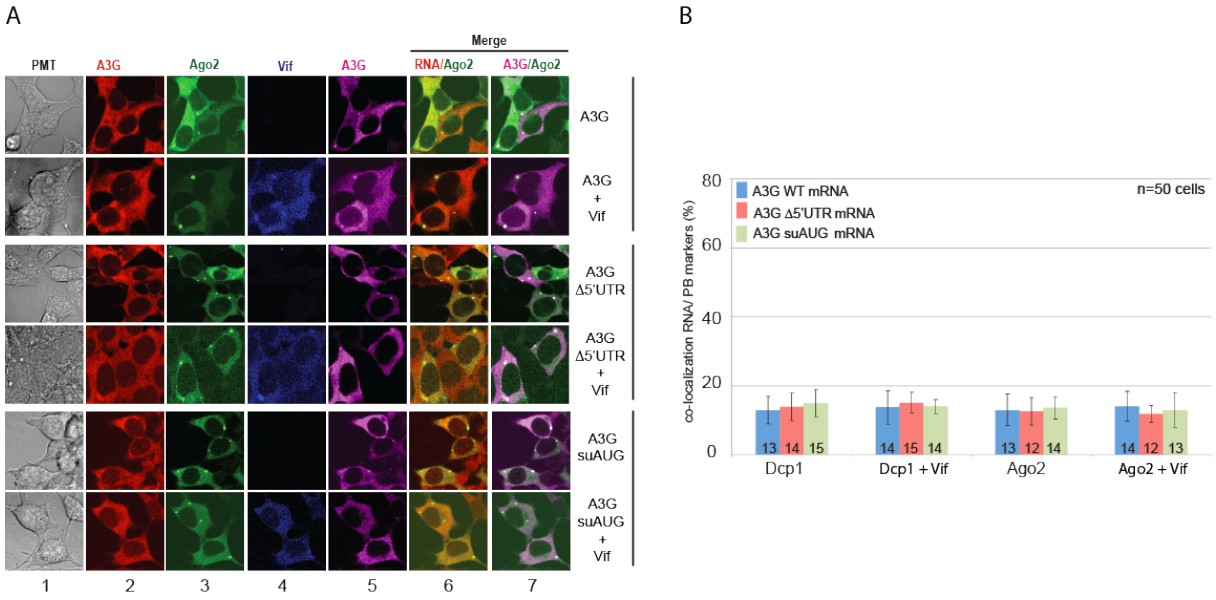


Figure 7: The uORF does not contribute to A3G mRNA relocation into P-bodies by Vif. A) HEK293T cells were transfected with plasmids expressing A3G wild-type as well as Δ5'-UTR and suAUG mRNAs, in absence or presence of Vif. Cells were fixed and probed with anti-DIG (A3G mRNAs), anti-A3G (A3G protein), anti-Vif and anti-Ago2 antibodies. Cells were stained with Dapi to visualize nuclei and the images were merged digitally. B) Histograms represent the percentage of co-localization of A3G mRNAs with Ago2 or Dcp1. Standard deviations are representative for at least three independent experiments.

DISCUSSION

The role of the Vif protein is of major importance for an efficient viral infection in non-permissive cells by antagonizing the antiviral activity of A3G proteins. While mechanisms leading to A3G degradation through the recruitment of an E3 ubiquitin ligase complex by Vif are rather well understood, little is known concerning its

translational regulation by Vif. Recently, we showed that the 5'-UTR of A3G mRNA, and specifically the SL2-SL3 domain, was required for the Vif-induced translation inhibition of A3G (31). In the present study, we reported that an uORF embedded within these structures plays a crucial role, not only for the repression of A3G translation by Vif, but also for its own translation.

To understand how Vif inhibits A3G translation, we first determined the importance of the uORF in A3G translation. uORFs are common features found in many eukaryotic mRNAs (36, 46–48). They are widely recognized as *cis*-regulatory elements than can affect mRNA translation from the main physiological ORF, thus fine-tuning the level of protein expression. In humans and rodents approximatively half of the transcripts contain uAUGs, and their presence generally correlates with reduced protein expression (46, 47). Here, by disrupting the uORF (deletion of the uORF or substitution of the uAUG), we showed that translation of A3G was significantly increased (figure 2), suggesting that the uORF was indeed a repressive element, as previously observed for several other genes like the tyrosine kinases HCK, LCK, ZAP70, YES1 or the oncogenes MDM2 and CDK4 (49). uORFs can regulate translation by multiple mechanisms which may depends on several variables such as 1) the distance between the 5'cap and the uORF; 2) the context of the uORF initiation codon; 3) the length of the uORF; 4) the position of the uORF stop codon; and 5) the length of the intercistronic sequence (50, 51). Using mutated A3G mRNA constructs to analyze the impact of the uORF on these various mechanisms, we first showed that the distance between the 5'cap and the uORF does not impact on A3G translation since deletion of SL1, thus reducing this distance from 176 to 50 nucleotides, gave the same level of A3G expression (mutant A3G SL2-SL3 in (31)). Second, the context around the uAUG seems to be favorable to initiate translation of the putative 23 amino acids peptide as the weakening of its Kozak sequence enhanced A3G expression (figure 2). Likewise, substitution of the uORF stop codon (uUGA) in order to put the uORF in frame with the major ORF of A3G, gave rise to the expression of two A3G proteins: a large one, initiated at the uAUG, and the standard one initiated at the main AUG (figure 2). Taken together, these results suggest the uAUG is efficiently recognized by the scanning ribosome (even though the natural Kozak context is not optimal) but can also be leaky-scanned by the ribosome which will then initiate at the main AUG to express A3G. Third, by using A3G mRNA constructs that reduced or increased the length of the uORF (insertion of stop codon at different position in the uORF and in the inter-ORF

sequence, figure 2) and potentially expressed peptides of 2, 5, 10, 15, 32 and 36 amino acids, we showed that peptide length does not contribute to A3G translation, with the exception of the putative expression of a 2 and 36 amino acids peptide that significantly decreased A3G expression (figure 2B). This is not completely surprising as long uORFs and short distance between uORF and the main ORF (which is the case for mutant A3G suUGA289) have been shown to increase repressiveness of the uORF (52–54). Besides, we showed that the nature of the 23 amino acids peptide expressed from the uORF (mutant uORF2) has no function in the translation of A3G. Unfortunately, we failed to identify the uORF-expressed peptide in transfected cell by mass spectrometry (data not shown). Finally, by reducing the length between the uORF stop codon and the main AUG after deletion (A3G ∆249-273 and 249-291) or stop codon insertion (A3G uUGA276 and uUGA289), we observed a significant decrease of A3G expression (40 %) only for the two mutants that reduced the most this distance (to 9 nucleotides, figure 2), suggesting that A3G may also be translated through a re-initiation mechanism. In this case, and as previously observed, the distance would be too short for the scanning 40S to reacquire a new eIF2•GTP•Met-tRNA^{iMet} complex and initiate translation at the main AUG (55–57). Finally, we showed that the highly structured 5'-UTR of A3G mRNA (30) did not drive any translation through a potential IRES, as the insertion of a stable stem-loop at the 5'-end of the mRNA thus avoiding the loading of the 43S subunit, almost completely inhibited A3G translation (figure 2). Moreover, a genome-wide search yielding a large number of mammalian cellular IRESs did not identify the 5'-UTR of A3G as potential IRES (Weingarten-Gabbay, Science 2016). Altogether, these results suggest that A3G is translated through a dual leaky-scanning and re-initiation mechanism. Further studies will be needed to evaluate: (i) the re-initiation efficiency; accurate determination will be difficult as it is hard to assess how much is really true re-initiation or a combination of leaky scanning and re-initiation; (ii) how the ribosome is stalled, through a specific RNA structure or the termination context (58). Studies in mammalian cells and yeast showed that uORF-bearing mRNAs are susceptible to be targeted by NMD which is attributed to the termination events occurring at uORF stop codons (59, 60). This does not seem to be the case for A3G as we did not observe any reduction of wt or mutated mRNA levels.

Beyond effects of the uORF on A3G translation, we also observed a link between uORF presence (functionality) and the Vif-mediated translational inhibition of A3G.

Indeed, under conditions where the proteasomal pathway was inhibited, we showed that Vif was able to significantly reduce the translation of A3G expressed from mainly all mutated mRNA constructs, with the exception of mutants ΔuORF and suAUG, suggesting that initiation at the upstream AUG is important for the repression induced by Vif. Besides, we also noticed that the putative peptide expressed from the uORF, whatever its size, is not implicated in the repression excluding a *trans*-peptide effect with Vif, likewise the distance between the two ORFs. Therefore, according to these observations, it seems that Vif induces A3G translational repression during the initiation steps at the uORF. Whereas additional experiments will be needed to clearly understand this mechanism, one can imagine that Vif interacts with components of the eukaryotic translation initiation machinery to reduce the translational rate of A3G or may recruit cellular factors involved in the negative regulation of the translation. Such studies are ongoing in the laboratory.

Finally, we observed a strong correlation between the presence/functionality of the uORF on A3G mRNA and its co-localization into SGs in presence of Vif in stress conditions. Indeed, when the uORF is deleted or the uAUG not functional, the presence of mutated A3G mRNAs was significantly decreased into SGs when Vif was co-expressed (figure 6). However, A3G protein expressed from these constructs can be found into P-bodies and SGs, as previously observed (61–64). The presence of wt A3G mRNA into SGs in presence of Vif is perhaps not so surprising, since SGs play an important role in the regulation of gene expression at the translational level in response to a variety of external stimuli (65). Therefore, these results validate the notion the uORF acts a negative regulator of A3G expression by helping its relocation into storage compartments in presence of Vif.

In sum, we have identified a uORF within the 5'-UTR of A3G mRNA that is conserved in the human population and drives not only its own translation but is also required by the HIV-1 Vif protein to repress its translation. While the relocation into SGs of A3G mRNA by Vif through the uORF could explain in part the down-regulation of A3G expression, much work will be required to fully understand this mechanism. Deciphering the mechanisms of the Vif-mediated translational inhibition of A3G mRNAs will be important to find new molecular inhibitors able to counteract Vif activity and reduce viral infectivity.

ACKNOWLEDGEMENT

The following reagents were obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: A3G polyclonal antibody (#9968) from Dr. Warner Greene and Vif monoclonal antibody (#319) from Dr. Michael H. Malim. This work was supported by a grant from the French National Agency for Research on AIDS and Viral Hepatitis (ANRS) and SIDACTION to JCP, and by post-doctoral (JB) and doctoral fellowships from ANRS (SG, CL) and the French Ministry of Research and Higher Education (TS).

III. Importance of a conserved uORF in the 5'-UTR of A3G and A3F mRNA

Table 1: Description of the primers used in this study (A3G and A3F constructs)

Mutants	Primers	Sequences (5' to 3')
A3G ΔuORF	pS-ΔuORF	GAAGCGGGAGGGGCCAACCCTGGTGCTCCA
	pAS-ΔuORF	TGGAGCACCAGGGTTGGCCCCTCCCGCTTC
A3G suAUG	pS-suAUG	GAAGGGGGAGGGGCCAAGACTACGAGGCCCTGG
	pAS-suAUG	CCAGGGCCTCGTAGTCTTGGCCCCTCCCCCTTC
A3G suUGA	pS-suUGA	GCCTGGAGCAGAAAGGAAACCCTGGTGCTCCA
	pAS-suUGA	TGGAGCACCAGGGTTTCCTTTCTGCTCCAGGC
A3G 2aa	pS-A3G2aa	GCCATGACTACGTGATGATGGGAGGTCACT
	pAS- A3G2aa	AGTGACCTCCCATCATCACGTAGTCATGGC
A3G 5aa	pS-A3G5aa	ACGAGGCCCTGGTGATGAACTTTAGGGAGG
	pAS-A3G5aa	CCTCCCTAAAGTTCATCACCAGGGCCTCGT
A3G 10aa	pS-A3G10aa	GTCACTTTAGGGTGATGAGTCCTAAAACCA
	pAS-A3G10aa	TGGTTTTAGGACTCATCACCCTAAAGTGAC
A3G 15aa	pS-A3G15aa	GCTGTCCTAAAATGATGAGCTTGGAGCAGA
	pAS-A3G15aa	TCTGCTCCAAGCTCATCATTTTAGGACAGC
A3G Δ249-273	pS-Δ25	TGGAGCAGAAAGTGATTAGTCGGGACTAGC
	pAS-Δ25	GCTAGTCCCGACTAATCACTTTCTGCTCCA
A3G Δ249- 291	pS-Δ50	TGGAGCAGAAAGTGACCAAGGATGAAGCCT
	pAS-Δ50	AGGCTTCATCCTTGGTCACTTTCTGCTCCA
A3G WK	pS-WK	GAAGCGGGAAAAAAAATGGCTACGAGGCCCT
	pAS-WK	AGGGCCTCGTAGCCATAAAAAATCCCGCTTC
A3G uORF2	pS-uORF2	GGGAGGGGCCATGGACTACGAGGCCCTGG
	pAS-uORF2	CCAGGGCCTCGTAGTCCATGGCCCCTCCC
	pS-uORF2	CTTGGAGCAGAAATGAAACCCTGGTGCTCC
	pAS-uORF2	GGAGCACCAGGGTTTCATTTCTGCTCCAAG
A3G uUGA276	pS-25	CTCCAGACAAAGATCTGATTAGTCGGGACTAGC
	pAS-25	GCTAGTCCCGACTAATCAGATCTTTGTCTGGAG
A3G uUGA289	pS-uUGA289	TTAGTCGGGACTAGCTGACGGCCAAGGATGAAG
	pAS-uUGA289	CTTCATCCTTGGCCGTCAGCTAGTCCCGACTAA
A3G WT+50	pS-TB	GTACAGCTGTCATTTGCACCACTAATGCTATAAGGATGAAGC CTCACTTCA
	pAS-TB	TACCAGCTGTCTCTGCACCACTTGGGTGCTATGGCCGGCTA GTCCCGACTAA

III. Importance of a conserved uORF in the 5'-UTR of A3G and A3F mRNA

Mutants	Primers	Sequences (5’ to 3’)
A3F ΔuORF	pS-ΔuORF	GAAGGGGGAGGGGCCAACCCTGGTGCTCCA
	pAS-ΔuORF	TGGAGCACCAGGGTTGGCCCCTCCCCCTTC
A3F suAUG	pS-suAUG	GAAGGGGGAGGGGCCAAGACTACGAGGCCCTGG
	pAS-suAUG	CCAGGGCCTCGTAGTCTTGGCCCCTCCCCCTTC
A3F suUGA	pS-suUGA	GCCTGGAGCAGAAAGGAAACCCTGGTGCTCCA
	pAS-suUGA	TGGAGCACCAGGGTTTCCTTTCTGCTCCAGGC
A3F 2aa	pS-A3F2aa	GGGGCCATGACTACGTGATGATGGGAGGTCACTTTA
	pAS- A3F2aa	TAAAGTGACCTCCCATCATCACGTAGTCATGGCCCC
A3F 5aa	pS-A3F5aa	ACTACGAGGCCCTGGTGATGAACTTTAGGGAGGGCT
	pAS-5aa	AGCCCTCCCTAAAGTTCATCACCAGGGCCTCGTAGT
A3F 10aa	pS-10aa	GAGGTCACTTTAGGGTGATGAGTCCTGAAACCTGGA
	pAS-10aa	TCCAGGTTTCAGGACTCATCACCCTAAAGTGACCTC
A3F 15aa	pS-15aa	AGGGCTGTCCTGAAATGATGAGCCTGGAGCAGAAAG
	pAS-15aa	CTTTCTGCTCCAGGCTCATCATTTCAGGACACCCT
A3F Δ249-273	pS-Δ25	TGGAGCAGAAAGTGATTAGTCGGGACTAGC
	pAS-Δ25	GCTAGTCCCGACTAATCACTTTCTGCTCCA
A3F Δ249- 291	pS-Δ50	TGGAGCAGAAAGTGACCAAGGATGAAGCCT
	pAS-Δ50	AGGCTTCATCCTTGGTCACTTTCTGCTCCA
A3F uORF2	pS-uORF2	GGGAGGGGCCATGGACTACGAGGCCCTGG
	pAS-uORF2	CCAGGGCCTCGTAGTCCATGGCCCCTCCC
	pS-uORF2	AGCCTGGAGCAGAAATGAAACCCTGGTGCT
	pAS-uORF2	AGCACCAGGGTTTCATTTCTGCTCCAGGCT

Table 2: Description of the primers used for the RACE PCRs

Target Gene	Primer Name	Primer Sequence	Tm
APOBEC3B	A3B SP1	GCA CAG CCC CAG GAG AAG CA	62.7°C
	A3B SP2	GAC CCT GTA GAT CTG GGC CG	59.6°C
	A3B SP3	GGC GCT CCA CCT CAT AGC AC	60.7°C
	A3B SP5	CGG CCC AGA TCT ACA GGG TC	59.6°C
	A3B SP6	ACC AGC AAA GCA ATG TGC TC	56.6°C
APOBEC3C	A3C SP1	GAG ACT CTC CCG TAG CCT TC	56.5°C
	A3C SP2	CAT GAT CTC CAC AGC GAC CC	57.9°C
	A3C SP3	AGA GGC GGG CGG TGA AGA TG	62.3°C
	A3C SP5	GGG TCG CTG TGG AGA TCA TG	57.9°C
	A3C SP6	ATC CAT CCA CCC CCA CAG AC	59.2°C
APOBEC3D	A3DE SP1	CAT TGG GGT GCT CAG CCA AG	59.1°C
	A3DE SP2	AGG TGA TCT GGA AGC GCC TG	59.7°C
	A3DE SP3	CAC ATT TCT GCG TGG TTC TC	54.2°C
	A3DE SP5	TGC AGC CTG AGT CAG GAA GG	59.5°C
	A3DE SP6	TAG AGT GCA ATG GCT GGA TC	55.6°C
APOBEC3H	A3H SP1	AGC GGT TTC TCG TGG TCC AC	60°C
	A3H SP2	TCC ACA CAG AAG CCG CAG CC	63°C
	A3H SP3	GTC AAC CAG CTC CCA GGC AC	61°C
	A3H SP5	GGC TGC GGC TTC TGT GTG GA	63°C
	A3H SP6	GGT CCC GGT GGA GGT CAT GG	62.5°C
	PCR Anchor Primer	GAC CAC GCG TAT CGA TGT CGA C	59.8°C
	dT Anchor Primer	GAC CAC GCG TAT CGA TGT CGA CTT TTT TTT TTT TTT TTV	

REFERENCES

1. Strebel K, Daugherty D, Clouse K, Cohen D, Folks T, Martin MA. 2003. The HIV A (sor) gene product is essential for virus infectivity. Nature.
2. Gabuzda DH, Li H, Lawrence K, Vasir BS, Crawford K, Langhoff E. 1994. Essential role of vif in establishing productive HIV-1 infection in peripheral blood T lymphocytes and monocyte/macrophages. J Acquir Immune Defic Syndr.
3. Sakai H, Shibata R, Sakuragi J, Sakuragi S, Kawamura M, Adachi A. 1993. Cell-dependent requirement of human immunodeficiency virus type 1 Vif protein for maturation of virus particles. J Virol.
4. von Schwedler U, Song J, Aiken C, Trono D, May M, Mangeat B, Alobwede I, Trono D, Vlahov D, Donfield S, Goedert JJ, Phair J, Buchbinder S, O'Brien SJ, Telenti A, Winkler CA. 1993. Vif is crucial for human immunodeficiency virus type 1 proviral DNA synthesis in infected cells. J Virol.
5. Madani N, Kabat D. 1998. An Endogenous Inhibitor of Human Immunodeficiency Virus in Human Lymphocytes Is Overcome by the Viral Vif Protein. J Virol 72:10251–10255.
6. Simon JHM, Gaddis NC, Fouchier RAM, Malim MH. 1998. Evidence for a newly discovered cellular anti-HIV-1 phenotype. Nat Med.
7. Sheehy AM, Gaddis NC, Choi JD, Malim MH. 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature.
8. Harris RS, Bishop KN, Sheehy AM, Craig HM, Petersen-Mahrt SK, Watt IN, Neuberger MS, Malim MH. 2003. DNA deamination mediates innate immunity to retroviral infection. Cell.
9. Malim MH. 2009. APOBEC proteins and intrinsic resistance to HIV-1 infectionPhilosophical Transactions of the Royal Society B: Biological Sciences.
10. Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D. 2003. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature.
11. Mbisa JL, Barr R, Thomas JA, Vandegraaff N, Dorweiler IJ, Svarovskaia ES, Brown WL, Mansky LM, Gorelick RJ, Harris RS, Engelman A, Pathak VK. 2007. Human Immunodeficiency Virus Type 1 cDNAs Produced in the Presence of APOBEC3G Exhibit Defects in Plus-Strand DNA Transfer and Integration. J Virol.
12. Ooms M, Brayton B, Letko M, Maio SM, Pilcher CD, Hecht FM, Barbour JD, Simon V. 2013. HIV-1 Vif adaptation to human APOBEC3H haplotypes. Cell Host Microbe.
13. Sato K, Izumi T, Misawa N, Kobayashi T, Yamashita Y, Ohmichi M, Ito M, Takaori-Kondo A, Koyanagi Y. 2010. Remarkable Lethal G-to-A Mutations in vif-Proficient HIV-1 Provirus by Individual APOBEC3 Proteins in Humanized Mice. J Virol.
14. Refsland EW, Hultquist JF, Luengas EM, Ikeda T, Shaban NM, Law EK, Brown WL, Reilly C, Emerman M, Harris RS. 2014. Natural Polymorphisms in Human APOBEC3H and HIV-1 Vif Combine in Primary T Lymphocytes to Affect Viral G-to-A Mutation Levels and Infectivity. PLoS Genet.
15. Sato K, Takeuchi JS, Misawa N, Izumi T, Kobayashi T, Kimura Y, Iwami S, Takaori-Kondo A, Hu WS, Aihara K, Ito M, An DS, Pathak VK, Koyanagi Y. 2014. APOBEC3D and APOBEC3F Potently Promote HIV-1 Diversification and Evolution in Humanized Mouse Model. PLoS Pathog.
16. Zennou V, Perez-caballero D, Bieniasz PD. 2004. APOBEC3G Incorporation into Human Immunodeficiency Virus Type 1 Particles APOBEC3G Incorporation into Human Immunodeficiency Virus Type 1 Particles. J Virol.
17. Bishop KN, Verma M, Kim EY, Wolinsky SM, Malim MH. 2008. APOBEC3G inhibits elongation of HIV-1 reverse transcripts. PLoS Pathog.
18. Iwatani Y, Chan DSB, Wang F, Maynard KS, Sugiura W, Gronenborn AM, Rouzina I, Williams MC, Musier-Forsyth K, Levin JG. 2007. Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G. Nucleic Acids Res.
19. Mbisa JL, Bu W, Pathak VK. 2010. APOBEC3F and APOBEC3G Inhibit HIV-1 DNA Integration by Different Mechanisms. J Virol.
20. Gillick K, Pollpeter D, Phalora P, Kim E-Y, Wolinsky SM, Malim MH. 2013.

Suppression of HIV-1 Infection by APOBEC3 Proteins in Primary Human CD4 + T Cells Is Associated with Inhibition of Processive Reverse Transcription as Well as Excessive Cytidine Deamination . J Virol.
21. Pollpeter D, Parsons M, Sobala AE, Coxhead S, Lang RD, Bruns AM, Papaioannou S, McDonnell JM, Apolonia L, Chowdhury JA, Horvath CM, Malim MH. 2018. Deep sequencing of HIV-1 reverse transcripts reveals the multifaceted antiviral functions of APOBEC3G. Nat Microbiol.
22. Seissler T, Marquet R, Paillart JC. 2017. Hijacking of the ubiquitin/proteasome pathway by the hiv auxiliary proteins. Viruses.
23. Yu X, Yu Y, Liu B, Luo K, Kong W, Mao P, Yu XF. 2003. Induction of APOBEC3G Ubiquitination and Degradation by an HIV-1 Vif-Cul5-SCF Complex. Science (80-).
24. Stanley BJ, Ehrlich ES, Short L, Yu Y, Xiao Z, Yu X-FF, Xiong Y. 2008. Structural Insight into the HIV Vif SOCS Box and Its Role in Human E3 Ubiquitin Ligase Assembly. J Virol.
25. Bergeron JRC, Huthoff H, Veselkov DA, Beavil RL, Simpson PJ, Matthews SJ, Malim MH, Sanderson MR. 2010. The SOCS-Box of HIV-1 vif interacts with elonginBC by induced-folding to recruit its Cul5-containing ubiquitin ligase complex. PLoS Pathog.
26. Guo Y, Dong L, Qiu X, Wang Y, Zhang B, Liu H, Yu Y, Zang Y, Yang M, Huang Z. 2014. Structural basis for hijacking CBF-β and CUL5 E3 ligase complex by HIV-1 Vif. Nature.
27. Kim DY, Kwon E, Hartley PD, Crosby DC, Mann S, Krogan NJ, Gross JD. 2013. CBFβ Stabilizes HIV Vif to Counteract APOBEC3 at the Expense of RUNX1 Target Gene Expression. Mol Cell.
28. Anderson BD, Harris RS. 2015. Transcriptional regulation of APOBEC3 antiviral immunity through the CBF-b/RUNX axis. Sci Adv.
29. Stopak K, De Noronha C, Yonemoto W, Greene WC. 2003. HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. Mol Cell.
30. Mercenne G, Bernacchi S, Richer D, Bec G, Henriet S, Paillart JC, Marquet R. 2009. HIV-1 Vif binds to APOBEC3G mRNA and inhibits its translation. Nucleic Acids Res.
31. Guerrero S, Libre C, Batisse J, Mercenne G, Richer D, Laumond G, Decoville T, Moog C, Marquet R, Paillart JC. 2016. Translational regulation of APOBEC3G mRNA by Vif requires its 5′UTR and contributes to restoring HIV-1 infectivity. Sci Rep.
32. Jackson RJ, Hellen CUT, Pestova T V. 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol.
33. Sonenberg N, Hinnebusch AG. 2009. Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. Cell.
34. Chang TH, Huang HY, Hsu JB, Weng SL, Horng JT HH. 2013. RegRNA 2.0. BMC Bioinformatics.
35. Yamashita R, Suzuki Y, Takeuchi N, Wakaguri H, Ueda T, Sugano S, Nakai K. 2008. Comprehensive detection of human terminal oligo-pyrimidine (TOP) genes and analysis of their characteristics. Nucleic Acids Res.
36. Calvo SE, Pagliarini DJ, Mootha VK. 2009. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proc Natl Acad Sci.
37. Barbosa C, Peixeiro I, Romão L. 2013. Gene Expression Regulation by Upstream Open Reading Frames and Human Disease. PLoS Genet.
38. Law GL, Raney A, Heusner C, Morris DR. 2001. Polyamine Regulation of Ribosome Pausing at the Upstream Open Reading Frame of S-Adenosylmethionine Decarboxylase. J Biol Chem.
39. Fang P, Spevak CC, Wu C, Sachs MS. 2004. A nascent polypeptide domain that can regulate translation elongation. Proc Natl Acad Sci.
40. Kozak M. 1987. At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. J Mol Biol.
41. Kozak M. 2015. Circumstances and mechanisms of inhibition of translation by

secondary structure in eucaryotic mRNAs. Mol Cell Biol.
42. Hinnebusch AG. 2005. Translational regulation of GCN4 and the general amino acid control of yeast. Annu Rev Microbiol.
43. Kozak M. 2005. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. Gene.
44. Guzikowski AR, Chen YS, Zid BM. 2019. Stress-induced mRNP granules: Form and function of processing bodies and stress granules. Wiley Interdiscip Rev RNA.
45. White JP, Lloyd RE. 2012. Regulation of stress granules in virus systems. Trends Microbiol.
46. Johnstone TG, Bazzini AA, Giraldez AJ. 2016. Upstream ORFs are prevalent translational repressors in vertebrates. EMBO J.
47. Chew GL, Pauli A, Schier AF. 2016. Conservation of uORF repressiveness and sequence features in mouse, human and zebrafish. Nat Commun.
48. Ruiz-Orera J, Albà MM. 2019. Translation of Small Open Reading Frames: Roles in Regulation and Evolutionary Innovation. Trends Genet.
49. Wethmar K, Schulz J, Muro EM, Talyan S, Andrade-Navarro MA, Leutz A. 2016. Comprehensive translational control of tyrosine kinase expression by upstream open reading frames. Oncogene.
50. Morris DR, Geballe AP. 2000. Upstream Open Reading Frames as Regulators of mRNA Translation MINIREVIEW Upstream Open Reading Frames as Regulators of mRNA Translation. Mol Cell Biol.
51. Young SK, Wek RC. 2016. Upstream open reading frames differentially regulate genespecific translation in the integrated stress response. J Biol Chem.
52. Andreev DE, O'connor PB, Fahey C, Kenny EM, Terenin IM, Dmitriev SE, Cormican P, Morris DW, Shatsky IN, Baranov P V. 2015. Translation of 5' leaders is pervasive in genes resistant to eIF2 repression. Elife.
53. Terenin IM, Akulich KA, Andreev DE, Polyanskaya SA, Shatsky IN, Dmitriev SE. 2015. Sliding of a 43S ribosomal complex from the recognized AUG codon triggered by a delay in eIF2-bound GTP hydrolysis. Nucleic Acids Res.
54. Kozak M. 2002. Pushing the limits of the scanning mechanism for initiation of translation. Gene.
55. Vattem KM, Wek RC. 2004. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. Proc Natl Acad Sci.
56. Harding HP, Novoa I, Zhang Y, Zeng H, Wek R, Schapira M, Ron D. 2000. Regulated translation initiation controls stress-induced gene expression in mammalian cells. Mol Cell.
57. Lu PD, Harding HP, Ron D. 2004. Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. J Cell Biol.
58. Wethmar K. 2014. The regulatory potential of upstream open reading frames in eukaryotic gene expression. Wiley Interdiscip Rev RNA.
59. He F, Li X, Spatrick P, Casillo R, Dong S, Jacobson A. 2003. Genome-Wide Analysis of mRNAs Regulated by the Nonsense-Mediated and 5′ to 3′ mRNA Decay Pathways in Yeast. Mol Cell.
60. Mendell JT, Sharifi NA, Meyers JL, Martinez-Murillo F, Dietz HC. 2004. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat Genet.
61. Gallois-Montbrun S, Kramer B, Swanson CM, Byers H, Lynham S, Ward M, Malim MH. 2007. Antiviral Protein APOBEC3G Localizes to Ribonucleoprotein Complexes Found in P Bodies and Stress Granules. J Virol.
62. Marin M, Golem S, Rose KM, Kozak SL, Kabat D. 2008. Human Immunodeficiency Virus Type 1 Vif Functionally Interacts with Diverse APOBEC3 Cytidine Deaminases and Moves with Them between Cytoplasmic Sites of mRNA Metabolism. J Virol.
63. Wichroski MJ, Robb GB, Rana TM. 2006. Human retroviral host restriction factors APOBEC3G and APOBEC3F localize to mRNA processing bodies. PLoS Pathog.
64. Kozak SL, Marin M, Rose KM, Bystrom C, Kabat D. 2006. The anti-HIV-1 editing

enzyme APOBEC3G binds HIV-1 RNA and messenger RNAs that shuttle between polysomes and stress granules. J Biol Chem.

65. Kedersha NL, Anderson P. 2002. Stress granules: sites of mRNA triage that regulate mRNA stability and translatability. Biochem Soc Trans.

Supporting figures

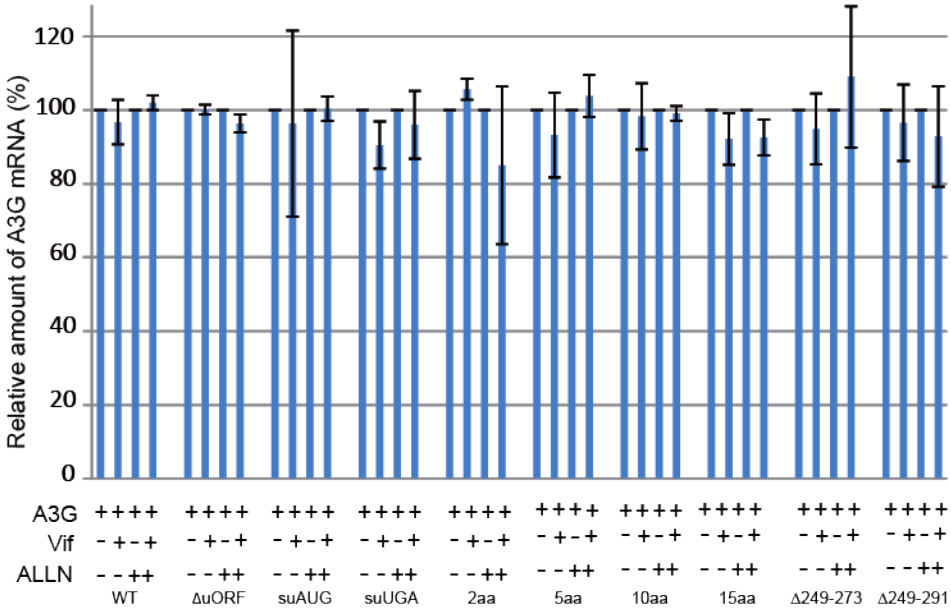


Figure S1: APOBEC3G mRNA expression level in HEK293T transfected cells. Total RNA was extracted from transfected HEK293T cells and A3G qPCR were performed to study the expression of wild-type and mutated A3G constructs. Standard deviations are representative of at least three independent experiments. P-values are indicated as follows: *<0,05; **<0,01; NS: non-significant.

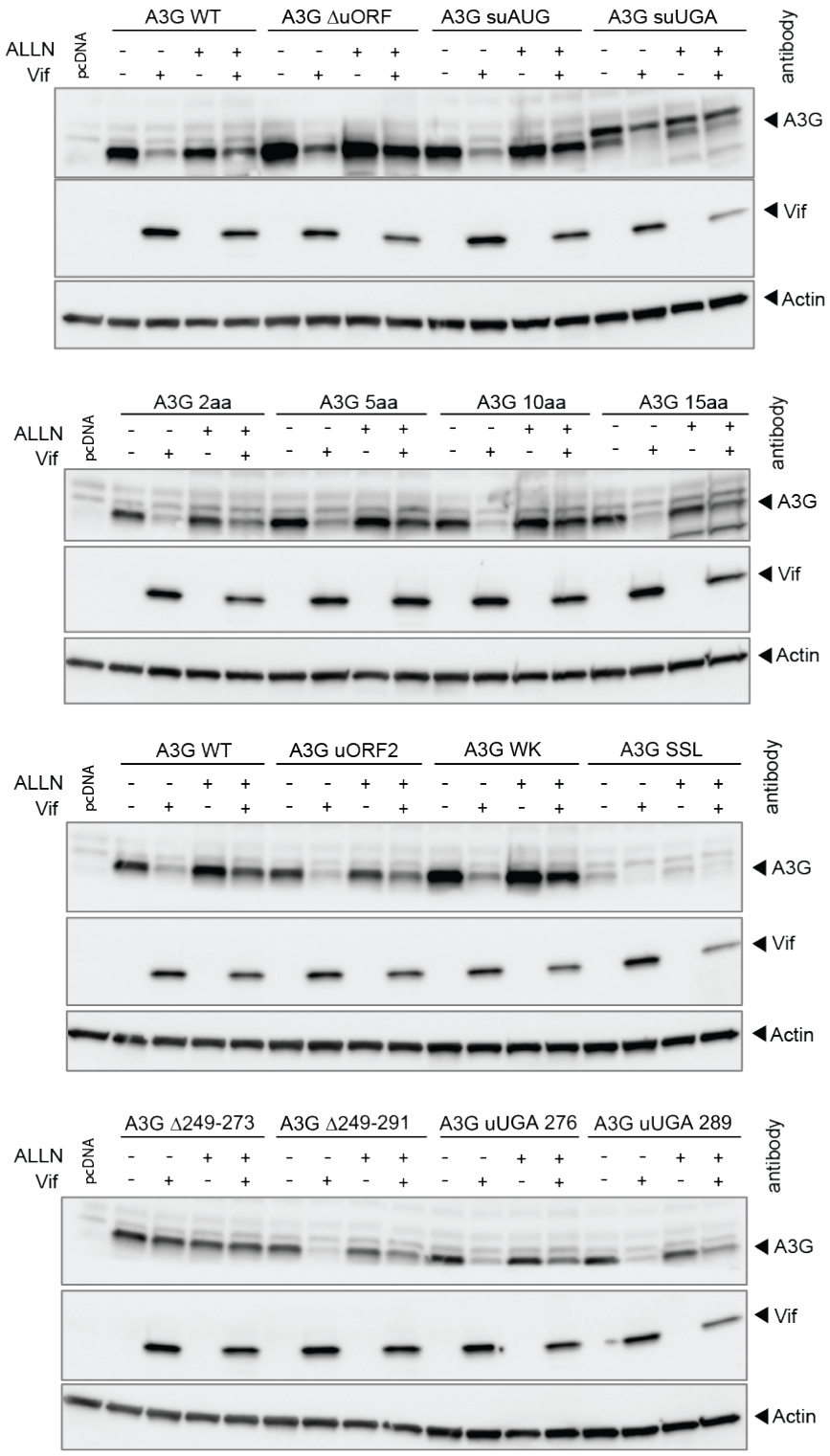


Figure S2: Effect of Vif on the translation of the different A3G mRNA constructs. HEK293T cells were transfected with plasmids expressing wild-type and mutated A3G constructs in the presence or absence of Vif and in the presence or absence of a proteasome inhibitor (ALLN). Proteins were separated by SDS/PAGE and analyzed by immunoblotting. Bands were quantified using Image J and relative expression of A3G proteins are represented in histograms (see figure 5).

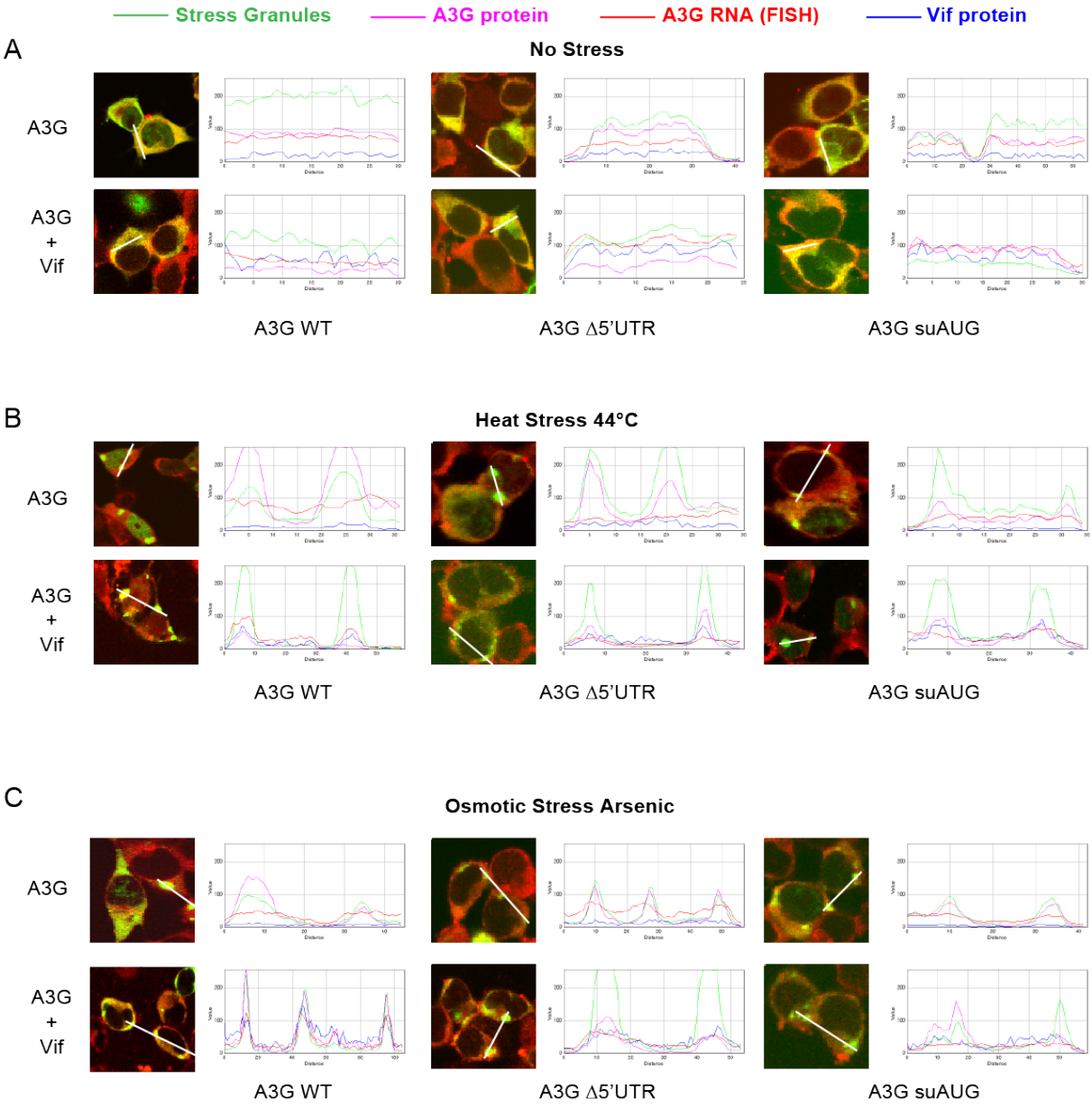


Figure S3: **Intensity plots (FISH)** obtained from the cytoplasmic signals of PAPB1 (green lines), A3G (pink lines), and A3G mRNA (red lines) in the absence or presence of Vif in physiological (A), or in stress conditions: 44°C (B) or sodium arsenite (C).

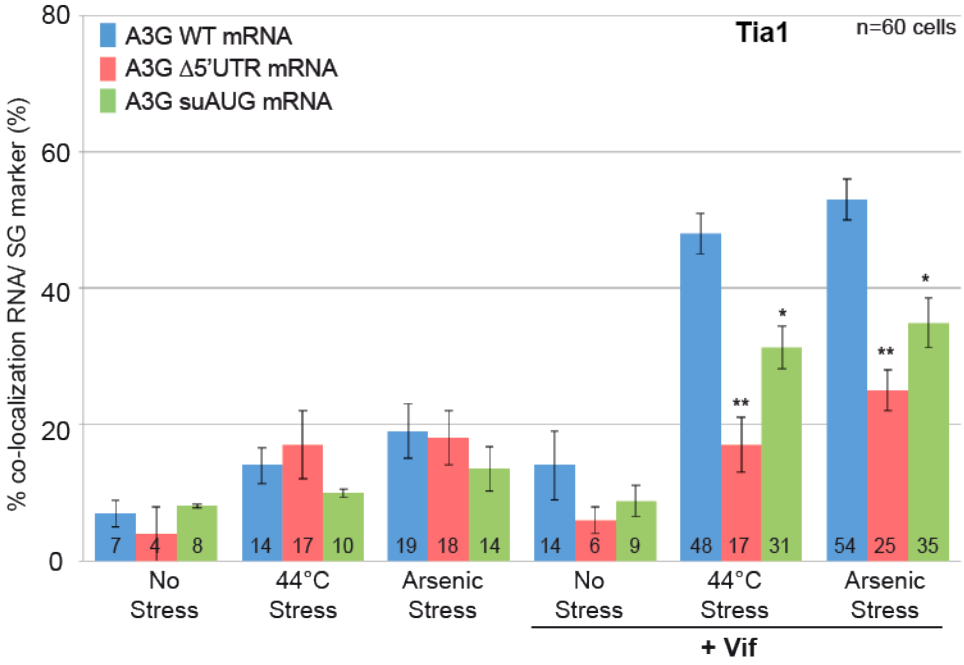


Figure S4: Importance of the uORF for the relocation of A3G mRNA into stress granule by Vif. HEK293T cells were transfected with plasmids expressing A3G wild-type as well as Δ5'-UTR and suAUG mRNAs, in absence (left) or presence (right) of Vif. Cells were cultured in various conditions (see figure 7). Histograms represent the percentage of co-localization of A3G mRNAs with TIA1. Standard deviations are representative for at least three independent experiments. P-values are indicated as follows: *<0,05; **<0,01.

Table S1. Variants identified in the human APOBEC3G and APOBEC3F uORF (+/- 10nt) regions

In bold, the common SNP (MAF>1%); In italics, variants that would impact uORF (i.e. at the uAUG position or inducing frameshift)

Data from the NCBI dbSNP database ; Homo sapiens:GRCh38.p12 (GCF_000001405.38)Chr 22 (NC_000022.11)

Variant ID	Location	Variant type	Molecular consequences	1000G MAF	ExAC MAF	More Info
A - APOBEC3G region						
rs1411087966	39,077,233	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs1352285213	39,077,236	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs1045489276	*39,077,241*	*single nucleotide variant*	*nc transcript variant, 5 prime UTR variant, intron variant*			*G=0.00020 (25/125568, TOPMED) ; G=0.0004 (13/30904, GnomAD)*
rs906717989	39,077,248	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs191099504	39,077,251	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant	A = 0.000199681		
rs1404085347	39,077,253	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs1313614097	39,077,254	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs865892647	39,077,258	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs1174076649	39,077,262	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs1479564191	39,077,263	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs994717582	39,077,269	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs1183197095	39,077,270	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs1235861207	39,077,272	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs1025905410	39,077,275	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs369893025	39,077,285	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant	G = 0.000399361		
rs1347789559	39,077,291	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs1209430250	39,077,293	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs1568997572	39,077,296	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs1254594028	39,077,297	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs1323315229	39,077,298	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs1288537138	39,077,301	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs576774669	39,077,302	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant	A = 0.000199681		
rs1251049749	39,077,315	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs1357485481	39,077,319	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
rs1212080289	39,077,320	single nucleotide variant	nc transcript variant, 5 prime UTR variant, intron variant			
B- APOBEC3F region						
rs1419286464	39,040,830	single nucleotide variant	5 prime UTR variant, intron variant			
rs1386275494	*39,040,840*	*single nucleotide variant*	*5 prime UTR variant, intron variant*			*G=0.00001 (1/125568, TOPMED)*
rs1165840091	*39,040,843*	*deletion*	*5 prime UTR variant, intron variant*			*delA=0.00001 (1/125568, TOPMED)*
rs1443373746	39,040,846	single nucleotide variant	5 prime UTR variant, intron variant			
rs550109499	39,040,847	single nucleotide variant	5 prime UTR variant, intron variant	G = 0.000199681		
rs1340838824	39,040,854	single nucleotide variant	5 prime UTR variant, intron variant			
rs998327947	39,040,857	single nucleotide variant	5 prime UTR variant, intron variant			
rs1243866239	39,040,866	single nucleotide variant	nc transcript variant, 5 prime UTR variant			
rs568479877	39,040,869	single nucleotide variant	nc transcript variant, 5 prime UTR variant	C = 0.000798722		
rs535530104	39,040,870	single nucleotide variant	nc transcript variant, 5 prime UTR variant	C = 0.000599042		
rs35898507	**39,040,871**	**single nucleotide variant**	**nc transcript variant, 5 prime UTR variant**	**T = 0.0371406**		
rs566315206	39,040,873	single nucleotide variant	nc transcript variant, 5 prime UTR variant	T = 0.000399361		
rs1219970355	39,040,877	single nucleotide variant	nc transcript variant, 5 prime UTR variant			
rs1321107161	39,040,882	single nucleotide variant	nc transcript variant, 5 prime UTR variant			
rs1216187563	39,040,884	single nucleotide variant	nc transcript variant, 5 prime UTR variant			
rs34383565	*39,040,885 - 39,040,888*	*indel*	*nc transcript variant, 5 prime UTR variant*			*delA=0.00006 (8/125568, TOPMED)*
rs533876307	39,040,889	single nucleotide variant	nc transcript variant, 5 prime UTR variant	A = 0.000199681		
rs1268443526	39,040,895	single nucleotide variant	nc transcript variant, 5 prime UTR variant			
rs1399683719	39,040,901	single nucleotide variant	nc transcript variant, 5 prime UTR variant			
rs371758756	39,040,915	single nucleotide variant	nc transcript variant, 5 prime UTR variant		A = 3.65979e-05	
rs545407115	39,040,916	single nucleotide variant	nc transcript variant, 5 prime UTR variant			
rs756220246	39,040,921	single nucleotide variant	nc transcript variant, 5 prime UTR variant		A = 3.64724e-05	
rs764161274	39,040,923	single nucleotide variant	nc transcript variant, 5 prime UTR variant		C = 0.0	

2. Role of the A3G uORF in viral infection

2.1 Introduction

The uORF in the 5'-UTR of A3G mRNA is necessary for A3G translational inhibition by Vif. This mechanism is responsible for around 50 % of the decrease in A3G levels in cells in the presence of Vif and therefore represents a crucial pathway in addition to A3G ubiquitination and degradation in the proteasome. It has been shown that when the part of the 5'-UTR that contains the uORF is altered, more A3G is encapsidated into newly formed virions and this increase in packaged A3G is sufficient to significantly decrease virus infectivity (77). Since this study has been performed on a single viral cycle in human embryo kidney (HEK293T) cells, we were interested in studying the importance of the uORF in $CD4^+$ target cells and in the context of a productive viral infection. H9 cells are a human $CD4^+$ T-cell line that intrinsically expresses A3G. Therefore, this cell line is qualified as non-permissive, as virions produced by these cells in the absence of Vif are non-infectious. In the presence of Vif, A3G restriction is counteracted and produced virions are infectious (67, 206). In order to study the effect of the A3G uORF on viral infection, our objective was to generate a cell line where the uORF is deleted using the CRISPR/Cas9 technology.

2.2 Material and methods

2.2.1 Generation of a ΔuORF cell line using nanoblades

Cells were cultivated in RPMI medium (Gibco) supplemented with 10 % fetal calf serum (Pan Biotech), 100 U/ml penicillin and 100 µg/ml streptomycin (Gibco) at 37 °C and 5 % CO_2. After seeding of $5*10^5$ cells in a 12 well plate, cells were treated with nanoblades in the presence of 4 µg/µl of polybrene (134). The Nanoblades used for this experiment were loaded with two different sgRNAs (Table 2) directed against regions at either side of the A3G uORF and were provided by Dr. Emiliano Ricci. After 6 h of treatment with the nanoblades, cells were resuspended in fresh medium, diluted and seeded in a 96 well plate in order to have only one cell every 4 wells. Cells were left to grow for 1-2 months.

Name	Sequence	Target	Targeted position	Hybridization temperature
sgRNA1	GCG CCU GAG CAG GAA GCG GG	A3G	150 to 169	
sgRNA2	GAU CUU UGU CUG GAG CAC CA	A3G	254 to 273	
A3G-DNA-s	GAT GGC TAC CTT GCA GGC TG	A3G	-611 to -592	63 °C
A3G-DNA-a	CGG CGG AGA CTG TAG TGA GC	A3G	1719 to 1738	
A3G-DNA-nest-s	CAG CCT GTC TGT CTT GAT GGT GG	A3G	-125 to -102	64 °C
A3G-DNA-nest-a	GTT CAA TTC TCC CAG TTC CAG AGT TCA	A3G	631 to 657	
A3G-RNA-s	TGC AAT TGC CTT GGG TCC TG	A3G	12 to 31	55 °C
A3G-RNA-a	TGA ATG TGG GTG GAT CCA TC	A3G	885 to 904	
A3G-RNA-nest-s	GGA AGA TAA AGC GTC CCA GGG CCT	A3G	117 to 140	60 °C
A3G-RNA-nest-a	GTG GCC ATA TCC CTT GTA CAC TTT GTG C	A3G	587 to 614	
ESPs	TTC TCC AGA ATC AGG AAA AC	A3G	1429-1449	61 °C
ESPa	GTG TCT GTG ATC AGC TGG AG	A3G	1569-1551	
Gs	AAC CTG CCA AGT ACG ATG ACA TC	GAPDH	828-850	55 °C
Ga	GTA GCC CAG GAT GCC CTT GA	GAPDH	885-904	
HIVs	TTG AAA GCG AAA GTA AAG CCA G	HIV gRNA	199 to 220	60 °C
HIVa	CTT AAT ACC GAC GCT CTC GCA C	HIV gRNA	339 to 360	

Table 2: List of oligonucleotides. *The name, sequence and target for each oligo is indicated. sgRNA1 and 2 are RNAs, the rest are DNAs. The position targeted by each oligo is indicated relative to the transcription start site of the target mRNA. For each couple of primers used for PCR, their orientation (s- sense, a- antisense) and the optimal hybridization temperature is indicated.*

2.2.2 Screening of the ΔuORF deletion by RT-PCR on A3G mRNA and PCR on genomic DNA

After amplification of the obtained clones, RNA was isolated from cells using TriReagent (Molecular Research Center) following manufacturer's instructions. For the reverse transcription, about 1 µg of RNA was adjusted to a final volume of 16 µl. Four µl of iScript Reverse Transcription Supermix for RT-qPCR (BioRad) were added and the mixture was incubated at 25 °C for 5 min, 42 °C for 30 min and 85 °C for 5 min. Genomic DNA was extracted from cells using the PureLink Genomic DNA Kit (Thermo Fisher) following manufacturer's instructions. An aliquot of 100-200 ng of genomic DNA or 5 µl of the RT-reaction was mixed with 0.2 µM of sense and antisense primers (A3G-DNA-s and A3G-DNA-a for PCR on genomic DNA and A3G-RNA-s and A3G-RNA-a for PCR on RT reactions; Table 2), 1x Taq buffer (75 mM Tris-HCl (pH 8.8), 20 mM ammonium sulfate, 0.01 % Tween20, 2 mM $MgCl_2$), 0.25 mM of each dNTP and 1 U of Taq polymerase (homemade). The total volume was adjusted to 50 µl and the mixture was incubated at 95 °C for 3 min, then 35 cycles of 30 s at 95 °C, 30 s at the hybridization temperature (Table 2) and 1 min at 72 °C were performed. This was followed by 5 min at 72 °C. To increase signal and specificity, 1 µl of the PCR reaction was used for a

second PCR with nested primers (A3G-DNA-nest-s and A3G-DNA-nest-a or A3G-RNA-nest-s and A3G-RNA-nest-a; Table 2) following the same protocol as before. Ten µl of the PCR product were run on a 1.5 % agarose (TBE 0.5x) gel at 130 V for around 4 h. Major bands were cut out and purified from gel slices using the Nucleospin Gel and PCR Clean-Up (Macherey Nagel) following manufacturer's instruction and then cloned into a pJET vector using the CloneJET PCR Cloning Kit (Thermo Scientific) following manufacturer's instructions. The obtained clones were sequenced by Eurofins.

2.2.3 Infection of cells

H9 cells were cultured in RPMI medium (Gibco) supplemented with 10 % fetal calf serum at 37 °C and 5 % CO_2. For infection, $1*10^7$ cells were incubated with HIV-1 LAI C27.4 infectious particles at a multiplicy of infection (MOI) of 0.05 for 30 min at 37°C. They were then cultivated in a total volume of 20 ml of RPMI medium for 48 h. Cells (i) and supernatants (ii) were recovered by centrifugation.

(i) Cell pellets were washed twice with PBS and were then split in two aliquots. One aliquot was mixed with 200 µl of RIPA 1x (1x PBS, 1 % NP40, 0.5 % Na-DOC, 0.05 % SDS, 1x Halt Protease Inhibitor Cocktail (ThermoScientific)) and incubated at 4 °C for 10 min. Cells were centrifuged at 18,188 x g for 1 h at 4 °C, then the supernatant was recovered and stored at -20 °C. The second aliquot of cells was mixed with 1 ml of TriReagent (Molecular Research Company), incubated for 10 min at room temperature and then stored at -20 °C.

(ii) The supernatant of infected cell cultures was split into 3 aliquots. The first aliquot was used to quantify the amount of p24 using the Innotest HIV antigen mAb (Fujirebio) ELISA p24 kit. The second aliquot allowed measurement of the infectivity of the produced viruses using titration on TZM-bl cells. For this, viruses were diluted sequentially and mixed with $1*10^4$ cells in DMEM medium supplemented with 10 % of fetal calf serum and 18.75 µg/ml DEAE dextran (Sigma). The mixture was incubated for 36 h and the luminescence was then detected by incubation for 10 min with Bright-Glo (Promega). The third aliquot was mixed with 1 ml of TriReagent (Molecular Research Company), incubated for 10 min at room temperature and then stored at -20 °C.

These steps were performed by G. Laumond (Institute of Virology, Research Team of Dr. Christiane Moog, INSERM UMR 1109, Strasbourg, France) in a P3 laboratory.

2.2.4 Western blotting

The protein concentration of the cellular lysates was determined by a Bradford assay. In a 96 well plate, 1 µl of each lysate was diluted to a final volume of 200 µl. In parallel, a concentration range of 0 to 70 µg/ml of BSA was also established at a final volume of 200 µl. Fifty µl of Bradford Assay Dye (BioRad) were added and the absorbance was measured at 595 nm. An equivalent of 80 µg of protein of cellular lysate were migrated on a 4-15 % Criterion TGX Precast Midi Protein Gel (BioRad) in 1x TGS (25 mM Tris, 200 mM Glycine, 1 % SDS) for 35 min at 200 V. Then proteins were transferred onto a PVDF membrane using a Midi PVDF Transfer Pack (BioRad) at 25 V and 2.5 A for 10 min in a Trans-Blot Turbo (BioRad). The membrane was then incubated in WB blocking solution (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 % Triton x100, 5 % milk (Regilait)) for 1 h. Primary and secondary antibodies were added as described in table 3. After the primary antibody, the membrane was washed twice with TNT (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 % Triton x100). After the secondary antibody, the membrane was washed twice with TNT and once with TN (50 mM Tris-HCl (pH 7.5), 150 mM NaCl). One ml of Peroxide Solution and 1 ml Luminol Enhancer Solution (Amersham ECL Prime Western Blotting Detection Reagent, GE Healthcare) were mixed and added onto the membrane. Chemiluminescent signal was revealed using a Chemidoc (BioRad). For detection of multiple proteins on the same membrane, membranes were stripped for 25 min at room temperature under rotation with Antibody Stripping Buffer (Geba), then washed thoroughly with demineralized water. The detected bands were quantified using integrated density measurement with the ImageJ software (https://imagej.nih.gov/ij/) (196). A rectangular selection of the same size has been drawn around the bands of the expected size. An empty lane or a region directly below or above the bands has been used to quantify the background signal.

Name	Reference	Company	Dilution	Incubation
HIV-1 positive patient serum		Provided by Prof. J. Mak	1:10,000	ON
Polyclonal Rabbit @A3G	9968	NIH AIDS Reagent Program	1:10,000	ON
Monoclonal Mouse @Vif	319	NIH AIDS Reagent Program	1:10,000	ON
HRP - Monoclonal Mouse @GAPDH	MCA4739P	BioRad	1:10,000	1-2 h
HRP - Goat @Rabbit-IgG	1706515	BioRad	1:10,000	1-2 h
HRP - Goat @Mouse-IgG	1706516	BioRad	1:10,000	1-2 h

Table 3: List of antibodies. *Antibody names, references and company are indicated. @GAPDH antibody as well as secondary antibodies are conjugated to horseradish peroxidase (HRP, indicated in the name). Primary antibodies are diluted in WB blocking solution and secondary antibodies in TNT at the indicated dilutions and incubated for the indicated time at 4 °C. The following reagents were obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: Anti-Human APOBEC3G Polyclonal from Dr. Warner C. Greene (219); HIV-1 Vif Monoclonal Antibody (#319) from Dr. Michael H. Malim (209, 210). The HIV-1 positive patient serum was provided by Prof. Johnson Mak.*

2.2.5 RT-qPCR

RNA was isolated from cells and from viruses using TriReagent (Molecular Research Center) following manufacturer's instructions. Total RNA was treated with DNase I (Roche) in 1x DNase buffer (Roche) for 1 h at 37 °C and was then purified using phenol-chloroform extraction followed by ethanol precipitation. For reverse transcription, 1 µl of RNA was adjusted to a final volume of 16 µl. Four µl of iScript Reverse Transcription Supermix for RT-qPCR (BioRad) were added and the mixture was incubated at 25 °C for 5 min, 42 °C for 30 min and 85 °C for 5 min. For quantification of A3G mRNA, 5 µl of the RT-reaction was mixed with 0.2 µM of ESP sense and antisense primers (Table 2), 1x Taq buffer, 0.25 mM of each dNTP, 2.5 U of Taq polymerase (homemade) and 1.25 µl of EvaGreen Dye (Biotium). The mixture was incubated at 95 °C for 3 min, then 40 cycles of 30 s at 95 °C, 30 s at 61 °C and 1 min at 72 °C were performed and the amount of DNA was monitored at each cycle using Evagreen fluorescence measurement. For RT-qPCR of GAPDH mRNA or HIV-1 gRNA, 5 µl of the RT reaction were mixed with 10 µM of sense and antisense primers (Table 2) and 7.5 µl of Maxima SYBR qPCR Mastermix (ThermoFisher). The final volume was adjusted to 15 µl. The mix was incubated at 95 °C for 5 min, then 40 cycles of 95 °C for 15 s, 55 °C for 13 s and 72 °C for 30 s were performed with acquisition of the SYBR green fluorescent signal at every cycle. A titration curve was generated using between 10^9 and 10^3 copies of pCMV-A3G, pCMV-GAPDH or pNL4.3ΔEnv plasmids. The fluorescent signal was analysed using the CFX Maestro software (BioRad) and the baseline threshold was fixed at 200 relative fluorescence units.

2.3 Results

2.3.1 Deletion of the A3G uORF in an H9 T-cell line

To delete the uORF in the A3G gene in H9 cells, we used a recently developed Nanoblade technology. Nanoblades are virus-like particles (VLPs) which contain the Cas9 protein in association with a given sgRNA. The VLPs allow efficient delivery of Cas9 into the target cells, where genome editing can take place (134). Here, we used two sgRNAs which target regions 7 nts upstream and 6 nts downstream of the uORF, respectively (Table 2). The used sgRNAs were as far as possible specific to the A3G uORF, however, many clones with partially mutated A3F genes were also identified due to evolutionary conservation. Screening of a total of 337 clones has allowed us to identify a cell line with a heterozygous genotype which possesses one WT A3G allele and one allele where the uORF of the A3G gene is deleted (Fig. 13A). Despite the heterozygous genotype, at the mRNA level we have identified almost exclusively

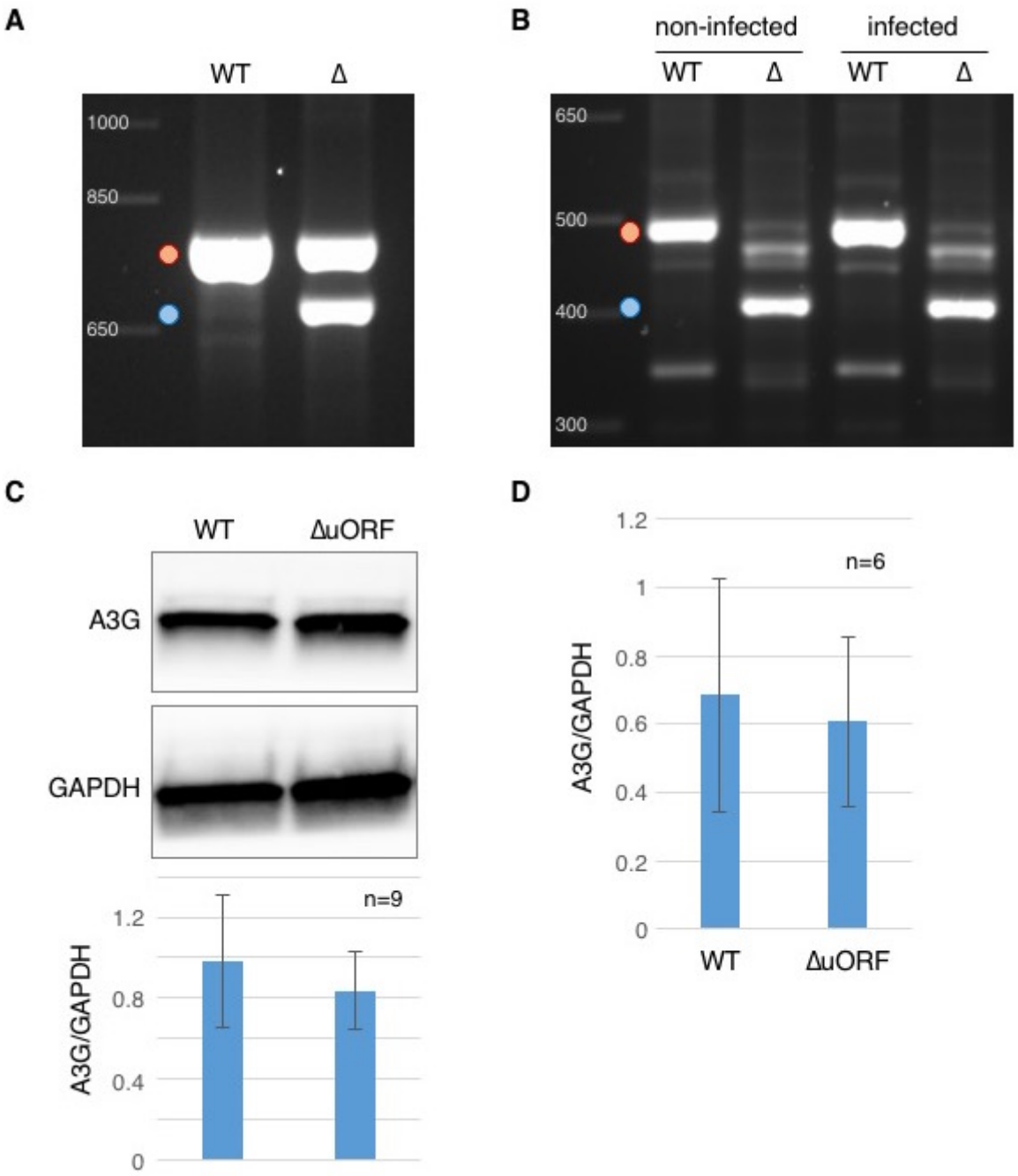


Figure 13: Identification and characterization of a ΔuORF H9 cell line. (A) *Genomic DNA and* ***(B)*** *total RNA of non-infected H9 WT and ΔuORF cells or cells infected for 48 h by HIV-1 LAI at a MOI of 0.05 have been purified and amplified using two rounds of PCR with nested primers. PCR products have been separated on a 1.5 % agarose (TBE 0.5 x) gel. The red dot indicates the WT allele of A3G and the blue dot indicates a shorter allele where the uORF of A3G is deleted. Sizes are indicated on the left.* ***(C)*** *Detection of the A3G protein in H9 WT and ΔuORF cells by western blot. GAPDH is shown as a loading control. Bands obtained in 9 different experiments have been quantified using ImageJ and the mean A3G/GAPDH ratio is shown in the histogram.* ***(D)*** *Ratio of the A3G mRNA relative to the housekeeping gene GAPDH quantified by RT-qPCR in H9 WT and ΔuORF cells in 6 biological replicates.*

the ΔuORF mRNA while the WT mRNA seems to be only marginally expressed (Fig. 13B). This expression pattern does not change over time and even under stress conditions, like for example infection of the cells with HIV-1 (Fig. 13B). The A3F gene of this clone was verified

and did not show any mutations (data not shown). As a homozygous clone with specific deletion of the A3G uORF without simultaneous mutation of the A3F gene could not be obtained, the present clone was selected for further experiments. Expression of A3G protein was approximately the same for the WT and ΔuORF cell lines (Fig. 13C). At the mRNA level, A3G mRNA seemed to be slightly less expressed in ΔuORF cells than in WT cells, but this difference did not seem to be significant (Fig. 13D).

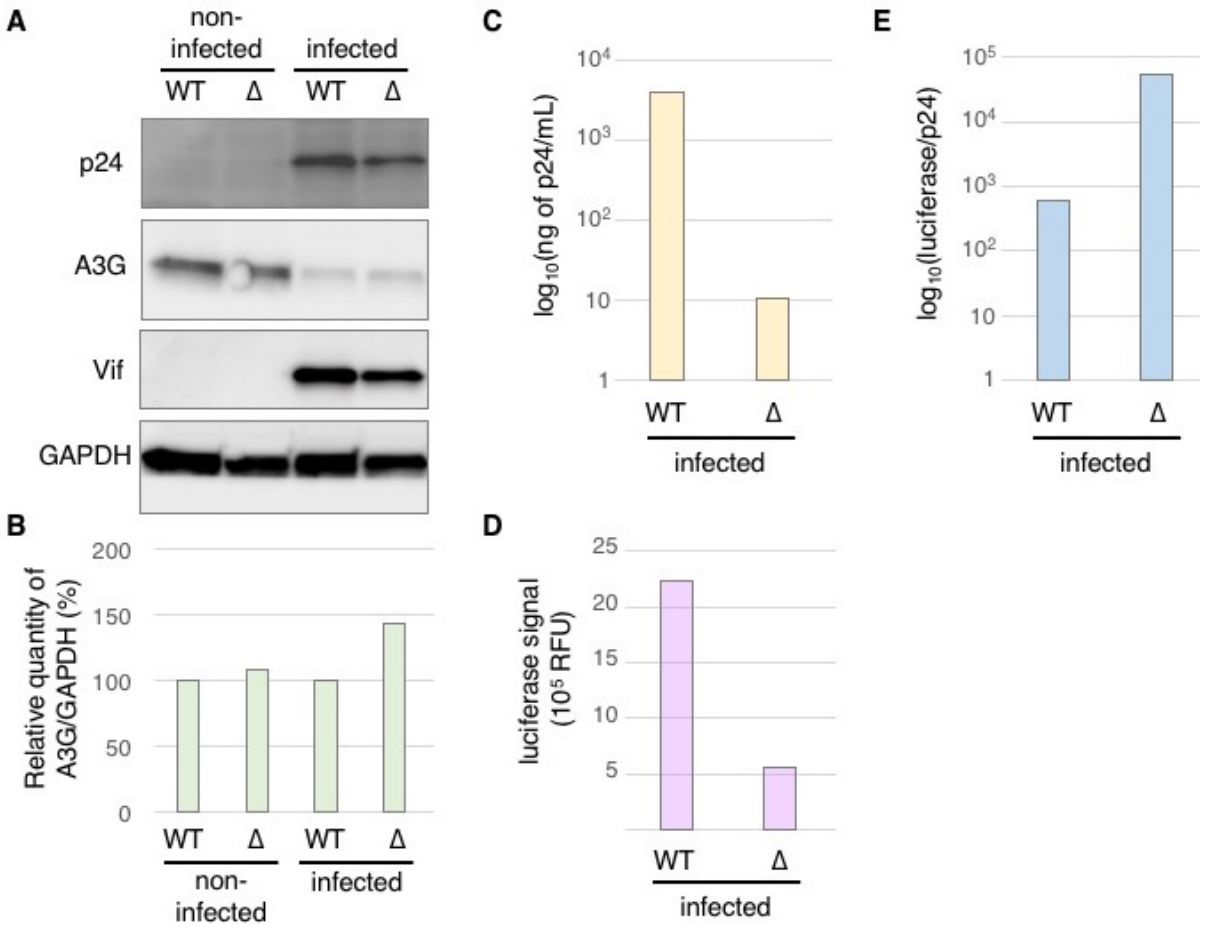


Figure 14: Effect of the A3G uORF on HIV-1 infection. *H9 WT and ΔuORF cells were infected or not with HIV-1 LAI for 48 h at a MOI of 0.05. Cells and supernatants were then recovered and analyzed.* ***(A)*** *Expression of p24, A3G and Vif in infected and non-infected cells was analyzed by western blot. GAPDH was used as a loading control.* ***(B)*** *Bands obtained for A3G were quantified using ImageJ and normalized to GAPDH levels. Expression of A3G in the WT cells was set to 100 %.* ***(C)*** *The p24 protein was quantified in the supernatant of infected cells by ELISA.* ***(D)*** *Supernatants of infected cells were analyzed by titration on TZM-bl cells and viral infection was measured by the produced luciferase signal.* ***(E)*** *Infectivity of virions produced in WT and ΔuORF cells obtained by normalization of the luciferase signal with the amount of p24.*

2.3.2 Effect of the uORF on viral production and infectivity

The ΔuORF as well as the parental WT cell lines were infected with the LAI isolate of HIV-1. Upon infection, we could observe a decrease in A3G protein levels in WT cells and a little less in ΔuORF cells (Fig. 14A, B). Production of the viral proteins Vif and p24 seemed to be lower

in ΔuORF cells compared to WT cells (Fig. 14A). Quantification of p24 in the supernatant of infected cells by ELISA showed a striking decrease of several logs in ΔuORF cells (Fig. 14C), which suggests that these cells produce much less virions. When the produced virions are titrated on the TZM-bl reporter cell line, a significantly lower signal could also be obtained in ΔuORF cells compared to WT cells (Fig. 14D), which indicates that less cells were infected by supernatants of ΔuORF cells. Surprisingly, when looking at the infectivity, obtained by normalization of the luciferase signal by the amount of p24, there seems to be an increase in ΔuORF cells.

2.4 Discussion

Using nanoblades, we have successfully generated an H9 T-cell line where the uORF of the A3G gene is deleted. Even though this cell line is heterozygous, it seems to almost exclusively express ΔuORF A3G mRNA. The first surprising result was that upon deletion of the uORF, the A3G protein is not overexpressed compared to the WT cells. Indeed, our previous results with plasmid constructs showed that the uORF negatively regulates A3G translation (Libre, Seissler et al., in preparation). However, there seems to be no difference in A3G protein levels in the WT and ΔuORF H9 cell lines. This might be due to other regulatory mechanisms in the cell which regulate a homeostatic protein production, but could not be explained by the mRNA level, which did not seem to be significantly different in the two cell types. As expected, our preliminary results suggest that A3G protein levels were more strongly decreased in the WT than in ΔuORF cells upon infection, even though this effect was very moderate and still has to be confirmed. This might be due to translational inhibition of A3G which is only possible in WT and not ΔuORF cells. While Vif can induce ubiquitination and degradation of A3G in both cell lines, A3G translation can be inhibited only in WT cells which would explain increased A3G protein levels in ΔuORF cells. In order to confirm this, it would be interesting to repeat the experiment in the presence of a proteasome inhibitor, which would limit Vif-mediated counteraction of A3G to translational inhibiton.

While the effect on A3G protein levels in cells is rather mild, viral production seemed to be strongly reduced in ΔuORF cells compared to the WT. This reduction could be observed by p24 quantification as well as by infection of a reporter cell line. Indeed, viruses produced from WT cells are largely protected from the deleterious effects of A3G thanks to efficient downregulation of A3G through Vif and can therefore be amplified in subsequent infectious cycles. Viruses produced from ΔuORF cells might have packaged more A3G due to a less efficient counteraction of A3G in producer cells. Therefore, many of these viruses might be

hypermutated in subsequent cycles and be unable to produce new virions. This could explain an overall decreased viral production over the course of multiple cycles in ΔuORF cells.
While these results are preliminary and will have to be confirmed, they seem to suggest that the uORF has an effect on viral infection. However, as the effect on cellular A3G protein levels is rather modest, it will have to be confirmed that the observed effect on viral production is not due to off-target modifications by the CRISPR machinery. Moreover, encapsidated A3G protein levels will have to be evaluated in order to confirm whether virion-associated A3G levels increase as expected in virions produced from ΔuORF cells. It would also be interesting to analyze proviral DNA from WT and ΔuORF cells to identify A3G-induced hypermutation. Surprisingly, infectivity of virions produced in ΔuORF cells seemed increased. While the observed effect on the A3G protein as well as on the luciferase signal is rather modest and consistent, the decrease on p24 seems massive in comparison. This large discrepancy has probably hampered normalization of the luciferase signal with the p24 and the obtained result for infectivity is not really meaningful. There could have been a problem with the ELISA and the experiment will have to be repeated in order to confirm the effect.
Moreover, in order to improve this study, it would be interesting to purify intact virions through a sucrose cushion prior to quantification of p24 and western blot analysis of encapsidated A3G protein.

Identification of proteins interacting with the 5'-UTR of A3G mRNA

IV. Identification of proteins interacting with the 5'-UTR of A3G mRNA

1. Introduction

In order to better understand the mechanism used by Vif to inhibit the translation of A3G we were interested in identifying potential cellular partners of Vif. Indeed, it has been shown that Vif binds to the 5'-UTR of A3G mRNA. There, it might recruit cellular factors that have an inhibitory effect on translation or interact with A3G mRNA-associated protein complexes and stimulate the drop off of cellular factors that have a stimulatory effect on A3G translation. In order to identify these proteins which might play a role in the translational inhibition mediated by Vif, a recently developed protocol for RNA-protein interaction detection (RaPID) has been used (172).

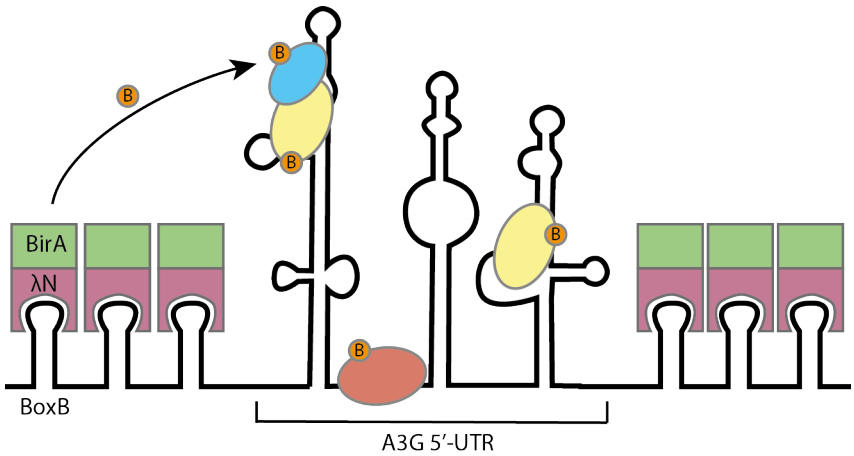


Figure 15: Detection of proteins interacting with the 5'-UTR of A3G mRNA. *The biotin ligase BirA is fused to the N protein of bacteriophage λ. The 5'-UTR of A3G mRNA is surrounded on both sides by 3 BoxB stem-loops. λN binds to BoxB and thereby brings BirA into proximity with the 5'-UTR of A3G mRNA. BirA can then catalyze biotinylation of proteins that are associated with the 5'-UTR of A3G mRNA.*

This protocol is based on a promiscuous biotin ligase called BirA that has been isolated from E. coli. BirA catalyzes biotinylation of proteins within a range of 10 nm and has first been used in fusion with different proteins of interest in order to identify interacting proteins (109, 178, 183). For the study of RNA-associated proteins, BirA has been fused to the protein N of

bacteriophage λ. λN specifically binds to a stem-loop, named BoxB, which is present in the phage genome and allows transcription antitermination (180). For RaPID, the RNA of interest, in this case the 5'-UTR of A3G mRNA, is surrounded on both extremities with 3 BoxB stem loops (Fig. 15). The expression of both BirA-λN and the chimeric RNA in cells allows binding of BirA-λN to the RNA of interest through the λN-BoxB interaction and is followed by biotinylation of proteins bound to the RNA of interest. These proteins can then be isolated using streptavidin-bound magnetic beads and identified by mass spectrometry (172).

2. Material and methods

2.1 Plasmids

BASU and BoxB plasmids were obtained from Dr. Karim Majzoub (Department of Neurosurgery, Stanford University School of Medicine, Stanford, California, USA). The sequence of the 5'-UTR of A3G-WT, A3G-ΔuORF and the scrambled sequence (Appendix 1) were cloned into the BoxB plasmid using the Esp3I cloning site.

The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: HIV-1 NL4-3 Vif expression vector (pcDNA-hVif) from Dr. Stephan Bour and Dr. Klaus Strebel. For co-immunoprecipitation, a FLAG-tag was added at the C-terminus of hVif, resulting in the pcDNA-Vif-FLAG construct. The empty pcDNA vector was used as a negative control (pcDNA-Ø). The pCMV-A3G plasmid contains the full-length A3G cDNA (both UTRs and the CDS) under the control of the CMV promoter.

2.2 Preparation of cells

HEK293T cells at confluence were washed once with DMEM medium (Gibco) supplemented with 10 % fetal calf serum (Pan Biotech), 100 U/ml penicillin and 100 µg/ml streptomycin (Gibco) and once with EDTA/0.25 % trypsin (Gibco). Then they were detached from the culture support using trypsin and resuspended in DMEM. Cells were counted using 0.4 % trypan blue (Sigma) in a *Luna* cell counter (logos) and dispatched into 10 cm dishes at $2*10^6$ cells per dish in a final volume of 5 ml DMEM. Cells were incubated at 37 °C and 5 % CO_2.

Cells were transfected with expression plasmids around 16 h after their seeding. Two µg of pcDNA-Vif or pcDNA-Ø, 12 µg of the Box plasmid (BoxB-A3Gwt, BoxB-A3GΔuORF or BoxB-Scr) and 2 µg of BASU plasmids were used per dish. The plasmid preparation is gently mixed with 45 µl of XtremeGene9 (Roche) and 155 µl of Opti-MEM (Gibco) and incubated at room temperature for 10-15 min, then pipetted onto the cells.

Thirty-one h after transfection, old medium was discarded and fresh DMEM medium supplemented with 200 µM biotin (Sigma) was added to cells. Eighteen h after addition of biotin, the medium was discarded and cells were detached from the dishes in 1 mL of dPBS (Gibco) using cell scrapers. They were then centrifuged at 288 xg for 5 min and washed once with cold dPBS. Cell pellets were kept at -80 °C.

2.3 Pull-down of biotinylated proteins

Cells from one dish were resuspended in 1333 µl of lysis buffer B (50 mM Tris (pH 7.5), 500 mM NaCl, 0.2 % SDS, 1 mM DTT) at room temperature, then 2 % Triton x100 were added and thoroughly mixed. 1467 µl of wash buffer B4 (50 mM Tris (pH 7.5)) were added and the mixture was sonicated twice for 10 s with a 10 s pause in between. The extract was centrifuged at 15,000 xg at 4 °C for 10 min and the protein concentration of the supernatant was determined by a Bradford assay. In a 96 well plate, 1 µl of each lysate was diluted to a final volume of 200 µl. In parallel, a concentration range of 0 to 70 µg/ml of BSA was also established at a final volume of 200 µl. Fifty µl of Bradford Assay Dye (BioRad) were added and the absorbance was measured at 595 nm. A volume corresponding to an equal amount of proteins for each sample of the same replicate (1 mg for replicate #1, 1.5 mg for replicates #2 and #3) was diluted at 1/2 with wash buffer B4 and used for pull-down.

For pull-down, 300 µg of Dynabeads MyOne Streptavidin C1 (ThermoFisher) per sample were washed twice with wash buffer B4, resuspended with the cellular lysate and incubated over night at 4 °C under rotation. Beads were washed twice with wash buffer B1 (2 % SDS), once with wash buffer B2 (50 mM HEPES, 500 mM NaCl, 0.1 % Na-DOC, 1 % Triton x100, 1 µM EDTA), once with wash buffer B3 (10 mM Tris (pH 7.5), 250 µM LiCl, 0.5 % Na-DOC, 0.5 % NP40, 1 µM EDTA) and once with wash buffer B4. Beads were then resuspended in 100 µl of Laemmli and prepared for mass spectrometry.

2.4 Identification of proteins by mass spectrometry

The beads were heated in Laemmli at 95 °C for 10 min and were then discarded. Each sample was precipitated with 0.1 M ammonium acetate in 100% methanol, and proteins were resuspended in 50 mM ammonium bicarbonate. After a reduction-alkylation step (dithiothreitol 5 mM – iodoacetamide 10 mM), proteins were digested overnight with 1:25 (w/w) sequencing-grade porcine trypsin (Promega). One fifth of the peptide mixture was analysed by nanoLC-MS/MS in an Easy-nanoLC-1000 system coupled to a Q-Exactive Plus mass spectrometer (ThermoFisher). Each sample was separated with an analytical C18 column (75 µm ID × 25 cm nanoViper, 3 µm Acclaim PepMap; ThermoFisher) with a 160 min 300 nL/min gradient of

acetonitrile. The obtained data was searched against the Swissprot database with human taxonomy using the Mascot algorithm (version 2.5, Matrix Science). Mascot files were then imported into Proline v1.4 package (http://proline.profiproteomics.fr/) for post-processing. Proteins were validated with 1 % false discovery rate (FDR) and the total number of MS/MS fragmentation spectra (Spectral Count, SpC) was used to quantify each protein in the different samples.

These steps have been performed by the team of the mass spectrometry platform (Plateforme de proteomique Esplanade).

2.5 Bioinformatic analysis of identified proteins

Mass spectrometry data obtained for each sample, including the proteins identified and their associated spectral counts, were stored in a local MongoDB database and several pairwise comparisons were then performed through a Shiny Application (developed by Béatrice Chane-Woon-Ming, UPR9002, CNRS) built upon the msmsEDA (76) and msmsTests (75) R/Bioconductor packages. The latter were respectively used to conduct exploratory data analyses of LC-MS/MS data by spectral counts and differential expression tests using a negative-binomial regression model. The p-values were adjusted with FDR control by the Benjamini-Hochberg method and 3 parameters were used (adjusted p-value < 0.1 or 0.05, a minimum of 2 SpC in the most abundant condition, and a minimum fold change of 2) to define differentially expressed proteins.

Total proteins, as well as proteins significantly enriched in samples compared to the corresponding scramble controls were analysed using functional annotation clustering with medium to high classification stringency against the background of the total human proteome on the Database for Annotation, Visualization and Integrated Discovery (DAVID; https://david.ncifcrf.gov) (94).

2.6 Co-immunoprecipitation (without crosslinking) of Vif with its potential cellular partners

HEK293T cells were seeded in 10 cm dishes at 2.5 million cells per dish. After 16-18 h of seeding, cells were transfected with 0.4 µg of pCMV-A3G and 4 µg of pcDNA- Ø or pcDNA-Vif-FLAG. For the transfection, the plasmid preparation is gently mixed with 12 µl of XtremeGene9 (Roche) and 188 µl of Opti-MEM (Gibco), incubated at room temperature for 10-15 min, then pipetted onto the cells. Ten h after transfection, cells were treated with 25 µM ALLN. Twenty-four h after transfection, cells were washed once with dPBS (Gibco), then 250

µl of RIPA 1x (1x PBS, 1 % NP40, 0.5 % Na-DOC, 0.05 % SDS, 1x Halt Protease Inhibitor Cocktail (ThermoScientific)) were added and cells were incubated at 4 °C for 10 min. Cells were centrifuged at 18,188 xg for 1 h at 4 °C, then the supernatant was recovered.
900 µg protein G Dynabeads (Invitrogen) per sample were washed 3 times with RIPA 1x, then resuspended in 30 µl RIPA 1x. Two µg of Anti-FLAG M2 Mouse monoclonal antibody (F1804, Sigma) were added and incubated for 2 h at room temperature under rotation. Then beads were washed with RIPA 1x, resuspended in the cellular lysate (equivalent of 1.5 culture dishes per sample) and incubated at 4 °C for 2 h under rotation. Beads were washed 5 times with RIPA 1x and bound proteins were eluted in WB buffer (1x NuPAGE LDS Sample Buffer (Invitrogen), 1x NuPAGE Sample Reducing Agent (Invitrogen)) at 70 °C for 10 min.

Name	Reference	Company	Dilution	Incubation
Polyclonal Rabbit @PDIP3	NBP1-92392	Novus Biologicals	1:200	ON
Polyclonal Rabbit @MDC1	NB100-395	Novus Biologicals	1:1000	ON
Polyclonal Rabbit @RPL37A	NBP1-55204	Novus Biologicals	1:500	ON
Polyclonal Rabbit @A3G	9968	NIH AIDS Reagent Program	1:10,000	ON
Monoclonal Mouse @Vif	319	NIH AIDS Reagent Program	1:10,000	ON
HRP - Goat @Rabbit-IgG	1706515	BioRad	1:10,000	1-2 h
HRP - Goat @Mouse-IgG	1706516	BioRad	1:10,000	1-2 h

Table 4: List of antibodies. *Antibody names, references and company are indicated. Secondary antibodies are conjugated to horseradish peroxidase (HRP, indicated in the name). Primary antibodies are diluted in WB blocking solution and secondary antibodies in TNT at the indicated dilutions and incubated for the indicated time at 4°C. The following reagents were obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: Anti-Human APOBEC3G Polyclonal from Dr. Warner C. Greene (219); HIV-1 Vif Monoclonal Antibody (#319) from Dr. Michael H. Malim (209, 210).*

2.7 Co-immunoprecipitation (with crosslinking) of Vif with its potential cellular partners

HEK293T cells were seeded into 10 cm dishes at 2.5 million cells per dish in a final volume of 5 ml DMEM medium (Gibco), supplemented with 10% fetal calf serum (Pan Biotech), 100 U/ml penicillin and 100 µg/ml streptomycin (Gibco). Cells were incubated at 37 °C and 5 % CO_2. Cells were transfected with expression plasmids around 16-18 h after their seeding with 0.4 µg of pCMV-A3G and 4 µg of pcDNA-Vif-FLAG per dish. The total amount of plasmid is adjusted to 4.4 µg using the pcDNA-Ø. The plasmid preparation is gently mixed with 12 µl of XtremeGene9 (Roche) and 188 µl of Opti-MEM (Gibco) and incubated at room temperature for 10-15 min, then pipetted onto the cells. Ten h after transfection, cells were treated with 25 µM ALLN. 24 h after transfection, cells were washed twice with cold PBSCM buffer (dPBS (Gibco) supplemented with 0.1 mM $CaCl_2$ and 1 mM $MgCl_2$), and then incubated for 6 h at 4

°C in PBSCM supplemented with 1 mM dithiobis-succinimidyl-propionate (DSP, ThermoFisher). PBSCM-DSP was discarded and cells were incubated for 15 min at 4 °C in PBSCM supplemented with 20 mM Tris (pH 7.5). Cells were washed twice with cold PBSCM buffer and then incubated for 30 min at 4 °C in lysis buffer (50 mM Tris (pH 7.5), 300 mM NaCl, 5 mM EDTA, 1 % Triton x100, 1 x HALT Protease Inhibitor Cocktail (ThermoFisher)). Cells were spun down at 17,000 xg for 15 min at 4°C and the supernatant was recovered. Sixty µl of Pierce Anti-DYKDDDDK Magnetic Agarose Beads (Invitrogen) were used per sample. Beads were washed 3 times with PBS, resuspended with the cell lysate and incubated over night at 4 °C under rotation. Beads were washed 5 times with wash buffer (50 mM Tris (pH 7.5), 300 mM NaCl, 5 mM EDTA, 0.1 % Triton x100) and once with PBS, then incubated at 37 °C for 15 min in 2x LDS Sample Buffer (NuPage) supplemented with 50 mM DTT. The supernatant was recovered and analysed by western blotting.

2.8 Western blotting

Western blotting was performed as described in III-2.2.4. For the input, cellular lysate equivalent to 20-30 µg of proteins was loaded. Primary and secondary antibodies were added as described in table 4.

3. Results

The RaPID approach has allowed us to identify between 581 and 916 proteins in each sample. Amongst the total proteins identified in all samples, a large proportion has known functions in mRNA metabolism (Fig. 16A). This indicates that BirA seems to indeed biotinylate mainly proteins associated with RNA, thereby partially validating the functionality of the approach, as in the case of random biotinylation in the cell, no particular functional clusters would have been enriched above the basic expression level of the human proteome. Many of the identified proteins are implicated in translation, splicing and export of mRNAs while others possess known RNA binding domains (Fig. 16A). Surprisingly, proteins implicated in cell-cell junctions constitute the most enriched functional cluster (Fig. 16A).

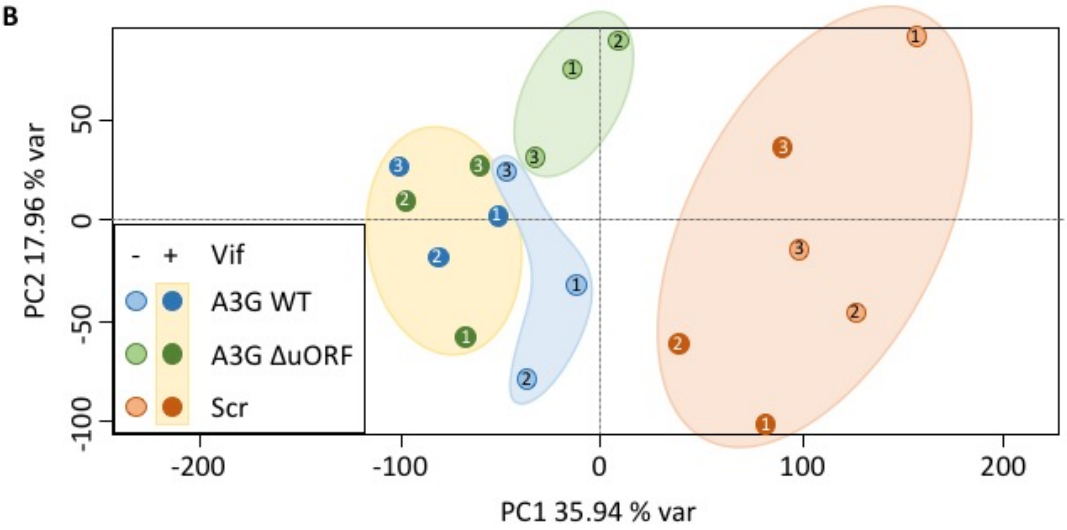
The identified proteins across all samples have first been subjected to a principal component analysis (PCA), a tool for the high-level visualization of sample similarities. The PCA identified two principal components (PCs) that respectively captured ~36% and ~18% of the total variability in the data (Fig. 16B). The experimental groups tend to separate across the principal components, while biological replicates do not, indicating stronger experimental effects than

replication effects or noise, thus validating the experimental design. Reassuringly, there does not appear to be any overall outlier samples, which may have indicated experimental issues. It can be noted that PC1 clearly separates controls performed with the scrambled RNA

A

Functional cluster	Enrichment score
cell-cell junction	89.69
splicing	46.87
translation, ribosome	43.50
RNA binding	38.41
Helicase	17.24
mRNA export	17.17

B



C

adjusted p-value	Number of proteins enriched on A3G mRNA: WT -	WT +	ΔuORF -	ΔuORF +	Vif
0.05	8	18	3	14	
0.1	15	30	5	9	

D

Functional cluster	Enrichment score
Translation, ribosome	10.47
Splicing	5.11
RNA or nucleotide binding	3.57

Figure 16: Identification of proteins that interact with the 5'-UTR of A3G mRNA WT or ΔuORF. *Proteins have been isolated by RaPID and identified by MS/MS.* ***(A)*** *DAVID functional annotation clustering with medium classification stringency of total proteins identified in all samples. The 6 most enriched clusters are shown.* ***(B)*** *Principal component (PC) analysis of all samples. The colour code is represented in the bottom left corner and numbers indicate the replicate.* ***(C)*** *Number of proteins that are significantly enriched in indicated samples compared to the corresponding scramble control with an adjusted p-value < 0.1 or < 0.05 and a minimum fold change of 2. Analysis has been performed with a differential expression test using a negative-binomial regression model.* ***(D)*** *DAVID functional annotation clustering with high classification stringency of proteins significantly enriched compared to the scramble controls. The 3 most enriched clusters are shown.*

sequence from the other samples (Fig. 16B). This suggests that the majority of proteins retrieved with the A3G 5'-UTR are not random RNA-binding proteins, as those that could bind the scrambled RNA control, but were specifically associated with our RNA of interest. In the

absence of Vif, samples obtained for the WT A3G 5'-UTR (Fig. 16B, blue field) are separated from ΔuORF samples (Fig. 16B, green field) by PC2, which suggests that the composition of these samples is dependent on the presence of the uORF. WT replicate #3 seems to be an outlier because it clusters closer to the ΔuORF samples than the other WT replicates. This sample will therefore be excluded for differential expression tests comparing A3G WT with ΔuORF or A3G WT in the presence and absence of Vif.

In the presence of Vif, WT and ΔuORF samples mix with no clearly distinctive clustering depending on the uORF. Samples in the presence of Vif (Fig. 16B, yellow field) seem to cluster separately from samples in the absence of Vif (Fig. 16B, green and blue fields), which indicates that the composition of samples depends on Vif.

Despite the large number of total identified proteins and a strongly marked separation of samples from the negative controls observed in PCA, statistical analysis of proteins that are enriched in each sample compared to the scramble control has allowed identification of only a relatively small number of significant proteins. Indeed, between 5 and 30 proteins have been identified to be significantly enriched on each 5'-UTR with an adjusted p-value below 0.1 (Fig. 16C). Functional annotation clustering of these proteins showed that proteins implicated in translation and splicing are amongst the most enriched functional clusters (Fig. 16D). The cluster of cell-cell junction proteins, which was the most enriched cluster amongst the total proteins (Fig. 16A) has not been identified amongst proteins specifically enriched in our samples (Fig. 16D), which indicated that this cluster contains unspecific proteins.

Comparison of proteins present on the A3G 5'-UTR in the presence and absence of the uORF allowed identification of 11 proteins which are significantly enriched in the presence, and 15 proteins significantly enriched in the absence of the uORF (Fig. 17A). Amongst these proteins, some have known functions in regulation of translation (EIF3A, EIF3C, IF4G1, RL24, RL27, PDIP3), mRNA stability and degradation (ATX2L, VIME), nuclear mRNA export (THOC2), splicing (NUP98, PRP31) or are known to bind RNAs (HNRPK, SND1). A brief description of each of these proteins is given in table 5.

Comparison of proteins associated with the WT 5'-UTR of A3G in the presence and absence of Vif allows identification of potential cellular partners of Vif. Proteins which are upregulated on the A3G 5'-UTR by Vif are potentially recruited and proteins which are downregulated in the presence of Vif might be dropped off. We have identified 14 proteins that are significantly downregulated from the A3G 5'-UTR by Vif (Fig. 17B). Importantly, 4 of these proteins (RL27,

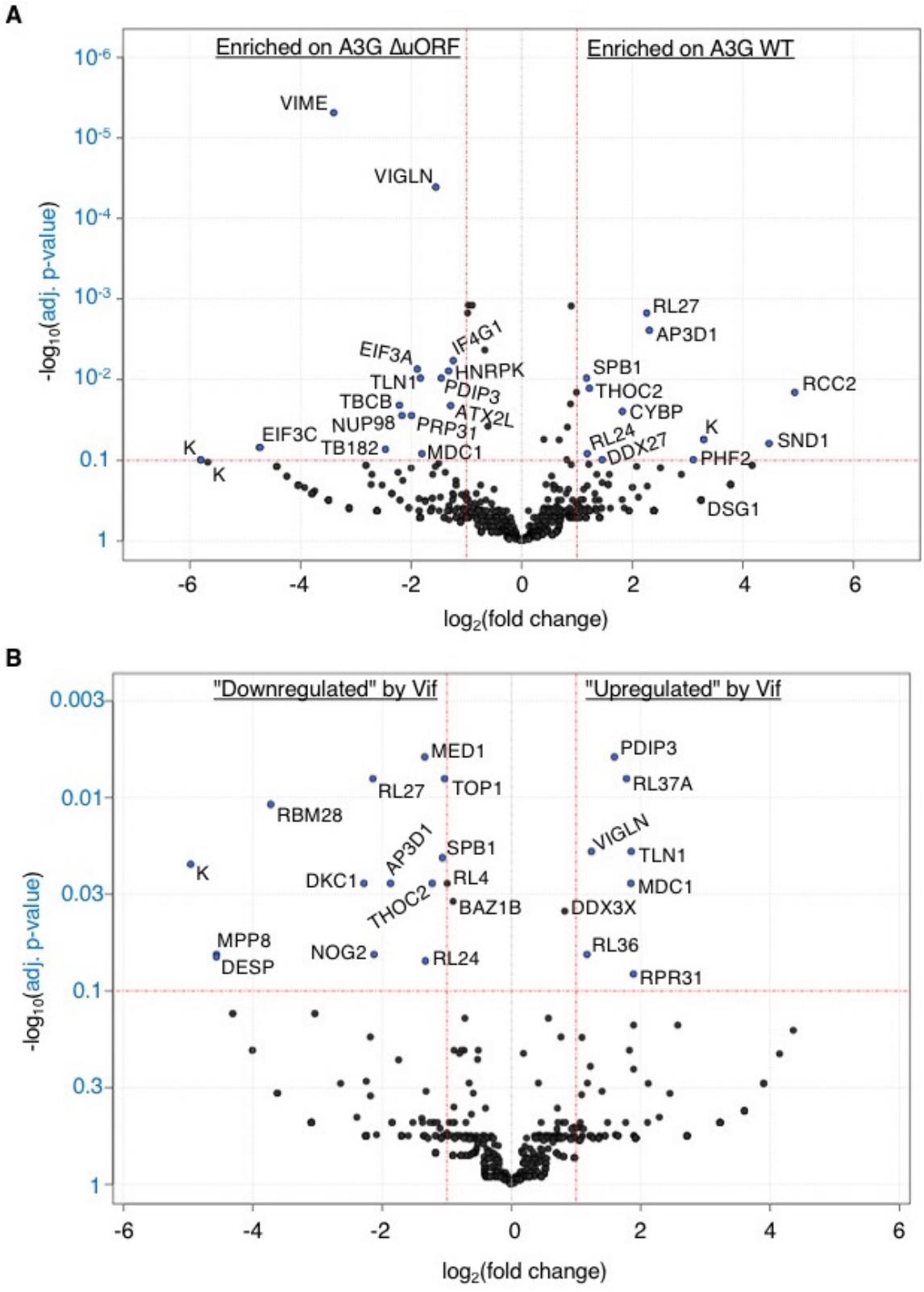


Figure 17: Influence of Vif and the uORF on A3G 5'-UTR-bound proteins. *Comparison of proteins bound to the 5'-UTR of A3G mRNA in two different conditions was realized using a negative-binomial regression model. An adjusted p-value < 0.1 and a minimal fold change of 2 were used as cut-off and are indicated as red lines on the y-axis and the x-axis respectively. Protein names are indicated next to the corresponding spot, keratins are marked as "K".* ***(A)*** *Comparison of proteins bound to the 5'-UTR of WT and ΔuORF A3G mRNA in the absence of Vif. Proteins enriched on WT A3G are indicated on the right and proteins enriched on ΔuORF A3G on the left.* ***(B)*** *Comparison of proteins bound to the 5'-UTR of WT A3G mRNA in the presence and absence of Vif. Proteins enriched in the presence of Vif are indicated on the right and proteins enriched in the absence of Vif on the left.*

RL24, AP3D1 and THOC2) were also specifically associated with WT A3G in comparison to the ΔuORF (Fig. 17A). Seven proteins seem to be significantly recruited by Vif. Interestingly, amongst these proteins two (PDIP3 and MDC1) were also identified to be preferentially associated with the ΔuORF A3G 5'-UTR (Fig. 17B). These proteins, as well as RPL37A, which was the most statistically significant after PDIP3, were selected to test whether Vif directly interacts with them or not. Immuno-precipitation of a C-terminal FLAG-tagged Vif from crosslinked or non-crosslinked HEK293T cell lysates was performed with an @-FLAG antibody and associated proteins were analysed by western blot. As expected, A3G co-immunoprecipitated with Vif in both conditions, however, none of the tested proteins seemed to directly interact with Vif (Fig. 18).

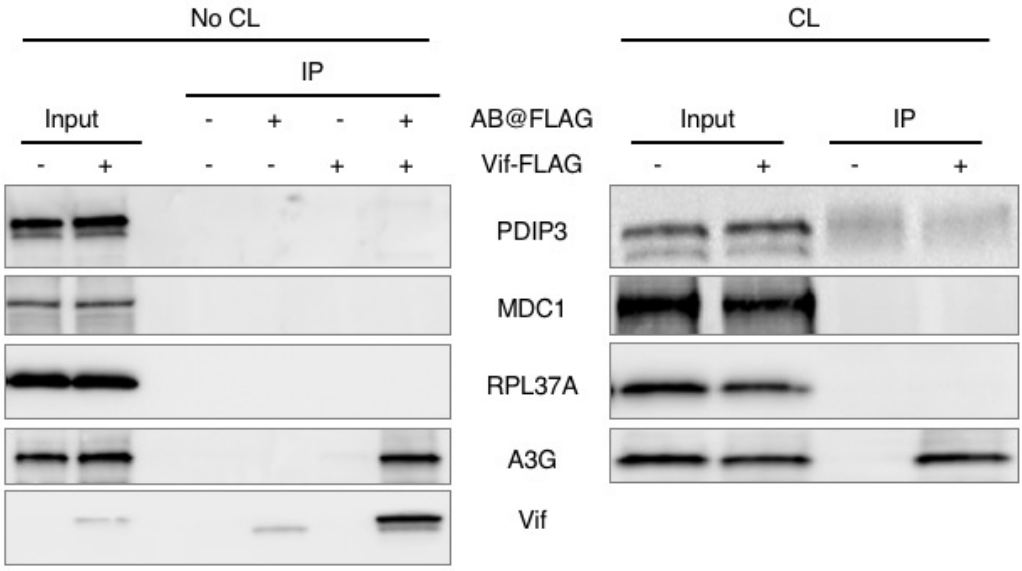


Figure 18: Co-immunoprecipitation of Vif with its potential cellular partners. *Immunoprecipitation of Vif-FLAG from HEK293T cell lysates has been performed on non-crosslinked (No CL) or cross-linked (CL) samples. Non-crosslinked samples were immunoprecipitated in the presence or absence of an anti-Flag antibody and Vif-Flag. Crosslinked samples were immunoprecipitated on anti-flag magnetic beads in the presence or absence of Vif-Flag. Approximately 5 % of the samples was charged as "Input". Proteins were revealed by western blotting.*

3. Discussion

The RaPID protocol has allowed us to identify between 3 and 30 proteins which significantly associate with the 5'-UTR of A3G mRNA in different conditions. Amongst the identified proteins, many are known to typically interact with mRNAs. Samples of the different conditions mostly cluster together which indicates that their composition depends on the presence of Vif and the uORF. While this validates the strength of the approach and shows that RaPID is able to identify proteins specifically associated with our RNA of interest, the number of significantly

enriched proteins is rather low. One possible explanation is that many proteins that are associated with the RNA of interest might be too far away from the BirA enzyme to allow biotinylation and are therefore lost during the pull-down. Indeed, the RaPID protocol has been published using rather small RNA sequences of around 38 nts (172), while RNAs studied in our case have a total length of 304 (WT) and 232 (ΔuORF) nts. Given that BirA can only biotinylate proteins at a maximum distance of 10 nm, only a portion of the proteins associated with the 5'-UTR of A3G might have been reached by BirA. Interestingly, when Vif is added, the samples don't cluster anymore depending on the presence of the uORF. This is unexpected because as shown previously (Libre, Seissler et al., in preparation), the uORF is necessary to allow Vif-mediated translational control of A3G.

When comparing the proteins associated with the A3G WT and ΔuORF 5'-UTRs, several proteins are significantly enriched on either the one or the other. It can be noted that several translation initiation factors (EIF3C, EIF3A and IF4G1) are enriched on the ΔuORF RNA. EIF3A is particularly interesting because it is the subunit of the eIF3 complex that regulates binding to the mRNA and can stimulate, as well as repress, translation (122). PDIP3, known to be a translation stimulator (130), is also enriched on the ΔuORF 5'-UTR. This might be due to a more active translation of the A3G mORF in the absence of the uORF, while translation is repressed in its presence.

Two ribosomal proteins (RL24 and RL27) are enriched on WT RNA. This is particularly interesting because ribosomes with a different composition in ribosomal proteins exist and these variants are known to have different characteristics (79). For example, it has been shown in plants that ribosomes containing the RL24 protein more efficiently re-initiate translation in the presence of a uORF as well as in polycistronic mRNAs (155, 163). The RL24-containing ribosome might specifically associate with the 5'-UTR of WT A3G mRNA to allow translation of the A3G mORF by re-initiation after translation of the uORF.

When comparing proteins associated with the WT A3G 5'-UTR in the presence and absence of Vif, several proteins seem to be dropped off from or recruited on the A3G 5'-UTR in the presence of Vif. Amongst these, we found 3 ribosomal proteins (RL4, RL24 and RL27) that are dropped off, while 2 others (RL37A and RL36) are recruited. This might be an example for a Vif-induced switch in ribosome composition with an effect on translation. Interestingly, RL24 and RL27 have been identified to be specifically enriched on the A3G 5'-UTR in the presence of the uORF, but to be downregulated in the presence of Vif. RL24 for example might be necessary for translation of the A3G mORF by stimulating re-initiation after uORF translation. Vif downregulates the association of this protein with the A3G 5'-UTR which in consequence might lead to a decrease in re-initiation and results in decreased translation of the A3G mORF.

UniprotKB ID	UniProtKB Accession	Protein Name	Function
AP3D1_HUMAN	O14617	AP-3 complex subunit delta-1	Role in trafficking of vesicles from the Golgi.
ATX2L_HUMAN	Q8WWM7	Ataxin-2-like protein	Regulation of stress granule and P-body formation.
BAZ1B_HUMAN		Tyrosine-protein kinase BAZ1B	Role in chromatin remodelling, transcription regulation and DNA damage response.
CYBP_HUMAN	Q9HB71	Calcyclin-binding protein	Potential component of ubiquitin E3 complexes.
DDX27_HUMAN	Q96GQ7	Probable ATP-dependent RNA helicase DDX27	ATP-dependent RNA helicase involved in rRNA biogenesis.
DDX3X_HUMAN	O00571	ATP-dependent RNA helicase DDX3X	Stimulates translation of mRNAs with structured 5'-UTRs, can compete with eIF4E and inhibit translation, involved in P-body formation.
DESP_HUMAN	P15924	Desmoplakin	Component of desmosomes.
DKC1_HUMAN	O60832	H/ACA ribonucleoprotein complex subunit DKC1	snoRNP complex that catalyzes pseudouridylation of rRNA.
DSG1_HUMAN	Q02413	Desmoglein-1	Component of intercellular desmosome junctions.
EIF3A_HUMAN	Q14152	Eukaryotic translation initiation factor 3 subunit A	Component of the eIF-3 complex, can bind to RNA (especially stem-loops) and activate or repress translation.
EIF3C_HUMAN	Q99613	Eukaryotic translation initiation factor 3 subunit C	Component of the eIF-3 complex.
HNRPK_HUMAN	P61978	Heterogeneous nuclear ribonucleoprotein K	Binds to poly-C RNA sequences.
IF4G1_HUMAN	Q04637	Eukaryotic translation initiation factor 4 gamma 1	Component of the eIF4F complex.
MDC1_HUMAN	Q14676	Mediator of DNA damage checkpoint protein 1	Role in cell cycle arrest in response to DNA damage.
MED1_HUMAN	Q15648	Mediator of RNA polymerase II transcription subunit 1	Activation of transcription of RNA polymerase II-dependent genes.
MPP8_HUMAN	Q99549	M-phase phosphoprotein 8	Mediates recruitment of the HUSH complex to methylates histones. The HUSH complex is also involved in the silencing of unintegrated retroviral DNA.
NOG2_HUMAN	Q13823	Nucleolar GTP-binding protein 2	Required for nuclear export and maturation of pre-60S ribosomal subunits.
NUP98_HUMAN	P52948	Nuclear pore complex protein Nup98	Role in transcription and alternative splicing activation.
PDIP3_HUMAN	Q9BY77	Polymerase delta-interacting protein 3	Stimulates translation by recruitment of ribosomal protein S6 kinase beta-1 I/RPS6KB1 to newly synthesized mRNA.
PHF2_HUMAN	O75151	Lysine-specific demethylase PHF2	Histone lysine demethylase activated through phosphorylation by PKA.
PRP31_HUMAN	Q8WWY3	U4/U6 small nuclear ribonucleoprotein Prp31	Component of the spliceosome.
RBM28_HUMAN	Q9NW13	RNA-binding protein 28	Component of the spliceosome.
RCC2_HUMAN	Q9P258	Protein RCC2	Multifunctional protein, regulates GTPases such as RAC1 and RALA.
RL24_HUMAN	P83731	60S ribosomal protein L24	Component of the ribosome.
RL27_HUMAN	P61353	60S ribosomal protein L27	Component of the large ribosomal subunit.
RL36_HUMAN	Q9Y3U8	60S ribosomal protein L36	Component of the ribosome.
RL37A_HUMAN	P61513	60S ribosomal protein L37a	Component of the ribosome.
RL4_HUMAN	P36578	60S ribosomal protein L4	Component of the ribosome.
SND1_HUMAN	Q7KZF4	Staphylococcal nuclease domain-containing protein 1	Endonuclease that mediates miRNA decay, transcriptional coactivator for STAT5.
SPB1_HUMAN	Q8IY81	pre-rRNA 2'-O-ribose RNA methyltransferase	Catalyzes 2'-O-methylation of mRNAs. The HIV-1 gRNA recruits this enzyme to escape the innate immune system.
TB182_HUMAN	Q9C0C2	182 kDa tankyrase-1-binding protein	Unknown
TBCB_HUMAN	Q99426	Tubulin-folding cofactor B	Binds to alpha-tubulin folding intermediates.
THOC2_HUMAN	Q8NI27	THO complex subunit 2	Role in nuclear export of mRNAs.
TLN1_HUMAN	Q9Y490	Talin-1	Connection of the cytoskeleton to the plasma membrane.
TOP1_HUMAN	P11387	DNA topoisomerase 1	Releases the supercoiling of DNA introduced by replication of transcription by cleaving and rejoining one strand.
VIGLN_HUMAN	Q00341	Vigilin	Role in cell sterol metabolism. Component of stress granules in yeast.
VIME_HUMAN	P08670	Vimentin	Component of intermediate filaments, involved in the stabilization of CO1A1 and CO1A2 mRNAs.

Table 5: List of all mentioned proteins. *The UniProtKB ID, accession, full name and a short summary of the function are indicated for each protein (https://www.uniprot.org) (242).*

This would also be consistent with the fact that Vif can only inhibit translation of A3G in the presence of the uORF. In the absence of the uORF, RL24 would not be necessary and

therefore Vif could not affect A3G translation. Interestingly, two of the proteins (PDIP3 and MDC1) specifically associated with the ΔuORF 5'-UTR in the absence of Vif, were also found enriched on the WT 5'-UTR in the presence of Vif. PDIP3 is particularly interesting because it is known to regulate translation. It is difficult to imagine however, how the recruitment of these proteins on A3G 5'-UTR might help Vif in inhibiting A3G translation, given that PDIP3 is associated with translation stimulation (130). Moreover, a direct interaction between Vif and PDIP3, which would have explained PDIP3 recruitment to the A3G 5'-UTR, could not be detected by co-immunoprecipitation. In order to confirm the role of the proteins identified in this study in translational regulation of A3G, it would be interesting to decrease their expression using RNA silencing, or to overexpress them in cells and see whether this has an impact on A3G expression.

Overall, this study has generated some interesting results, however, the small number of significantly enriched proteins on each mRNA has been an important downside for data analysis. Indeed, most of the proteins that seemed interesting in the comparison between the WT and ΔuORF as well as the comparison with and without Vif are not significantly enriched in each of these conditions separately in comparison to the corresponding scramble controls, which hampers the overall significance of these proteins. Moreover, it is probable that recruitment of the ribosome or even translation initiation is not actually possible on the chimeric BoxB-RNA used in this experiment due to Box-B stem loops which are rather tightly associated with the λN proteins. This, as well as the maximum biotinylation distance of BirA discussed above might have contributed to the low number of proteins identified in this study.

In conclusion, the technique has been useful to get a first view on the different types of proteins that might be associated with A3G mRNA and it promotes the hypothesis, that the interactome of this mRNA is different depending on the uORF and Vif. However, in order to have a more complete view of the A3G mRNA interactome, it would be better to study the native full-length mRNA by RNA pull-down from cellular lysates. This would not only allow to study the entire A3G mRNA in a more physiological context, but also bypass the experimental limitations encountered with the RaPID technique.

Development of a protocol to identify proteins associated with the full-length A3G mRNA

V. Development of a protocol to identify proteins associated with the full-length A3G mRNA

1. Study of RNA-protein complexes: state of the art

RNA-protein interactions are essential in cells and regulate many important mechanisms in the life of an mRNA including transcription, nuclear export, cellular localization, translation and stability. A multitude of protocols have been developed in recent years in order to fully apprehend the RNA-protein interactome in cells with the aim to decipher diverse cellular processes (16, 46, 141, 173, 176).

On the one hand, protein-centric analyses aim at finding all different RNAs bound by a given RNA-binding protein. Most of these studies are based on cross-linking of RNA-protein complexes in cells followed by immunoprecipitation of the protein of interest and sequencing of associated RNAs. Cross-linking can be performed using different strategies: (i) irradiation of cells with UV-light at a wavelength of 254 nm allows cross-linking of pyrimidines with directly interacting proteins, mainly through their aromatic residues (CLIP); (ii) cross-linking efficiency can be increased by the incorporation of a photo-activable residue into RNA which will react directly with proteins after irradiation at a wavelength of 365 nm (PAR-CLIP); (iii) cross-linking with formaldehyde is also possible, but this does not only cross-link RNA-protein but also protein-protein interactions and therefore allows stabilization of larger RNP complexes (16, 46, 141, 176).

On the other hand, RNA-centric approaches try to identify all different proteins interacting with a given RNA. While highly specific antibodies are readily available for CLIP and PAR-CLIP, RNA-centric approaches are more complicated due to the difficulty of specifically isolating one RNA of interest.

Multiple different protocols exist for capture of RNA-binding proteins on *in vitro*-transcribed RNA. This RNA is often biotinylated, either at its 5' or 3' extremity (15, 237), or by incorporation of biotinylated UTP or CTP (13, 26, 185, 218), which allows binding of the RNA to streptavidin coated beads. The RNA can also contain an aptamer, allowing specific binding of an adaptor

protein bound to beads. These protocols exploit for example the specific interaction between the S1 aptamer with streptavidin or the MS2 stem loop with its coat protein (124). Once the RNA is bound to the beads, those are incubated with cellular lysate in order to capture interacting proteins, which can then be analysed either by western blotting or mass spectrometry (173, 218, 237).

While these protocols are rather convenient because they allow rapid testing of different RNA constructs, the disadvantage is that RNA-protein interactions are formed *in vitro* in a non-physiological context, which might favour non-specific interactions. Therefore, different protocols have been developed, which capture RNP complexes directly formed *in cellula*. For example, RNAs containing an aptamer can be expressed in cells after transfection of expression plasmids. The RNA-protein interactions are formed in cells under physiological conditions followed by purification of the RNP of interest with the corresponding adaptor protein for the used aptamer (92, 124). The adaptor protein can also be fused to a biotinylase and be co-expressed in cells with the targeted transcript. This allows biotinylation and isolation of proteins that bind on the RNA of interest in proximity of the aptamer (172). However, these approaches necessitate ectopic introduction and expression of the aptamer-containing transcript which could increase non-physiologic interactions or hinder binding of certain factors. The capture of intrinsically expressed transcripts is also possible using complementary oligonucleotides (1, 16, 36, 141, 211). These oligonucleotides can have different lengths, ranging from 10 to 120 nucleotides and are generally biotinylated to allow capture on streptavidin beads. A single oligo, complementary to the RNA of interest can be used, but most of the time, multiple oligos are tiled across the entire transcript to maximize pull-down efficiency (1, 16, 36, 141, 211).

Vif interacts with the A3G mRNA and might induce drop off or recruitment of cellular proteins. Previous results have shown that the proteic interactome of the 5'-UTR of A3G mRNA changes depending on the presence of Vif and the uORF. In order to characterize the interactome of the entire A3G mRNA in more physiological conditions we aimed at developing a protocol to pull-down full length A3G mRNA from cells and to study the associated protein complexes.

2. Material and methods

2.1 Pull-down of A3G mRNA using complementary, biotinylated oligonucleotides

2.1.1 Design of oligos for pull-down

The structure of the A3G coding region was predicted using the mfold website with default settings (http://unafold.rna.albany.edu//?q=mfold/RNA-Folding-Form) (240). The structure with a calculated ΔG of -353.3 has been selected and 10 oligonucleotides have been designed to hybridize with entirely or partially single stranded regions, covering the entire sequence of the A3G coding region (Table 6 B-1 to 10; Appendix 2).

Name	Sequence (5'-3')	Modification	Region of A3G mRNA
B-1	tgg gtc tat tat aaa agt tg	5' biotin	354-373
B-2	ttc gga ata cac ctg gcc tc	5' biotin	461-480
B-3	gta cca ggt gac ctc ata ct	5' biotin	551-570
B-4	ggc aac gaa gat ggt cag gg	5' biotin	641-660
B-5	aaa ttc gtc ata att cat ga	5' biotin	748-767
B-6	ttg tta aag ttg aaa gtg aa	5' biotin	901-920
B-7	cct ggt tgc ata gaa agc cc	5' biotin	1014-1033
B-8	ctc acg tgt ttg ttt ttt ga	5' biotin	1195-1214
B-9	att att gaa att ttg gcc cc	5' biotin	1291-1309
B-10	atg gcc cgc agc ctc cca ct	5' biotin	1511-1530
D-1	cct ggt tgc ata gaa agc cc	5' desthiobiotin	1014-1033
D-2	ctc acg tgt ttg ttt ttt ga	5' desthiobiotin	1195-1214
D-3	att att gaa att ttg gcc cc	5' desthiobiotin	1291-1309
D-4	atg gcc cgc agc ctc cca ct	5' desthiobiotin	1511-1530
cap1	caa ttg aga aca gtg ctg aaa ttc gtc ata att cat ga	3' biotin	749-778
scr1	ggt aag ctc cta aat aaa tag agt ctc gat act tgt aa	3' biotin	
rel1	tca tga att atg acg aat ttc agc act gtt ctc aat tg	/	
cap2	caa ttg aca cag gct cac gtg ttt gtt ttt tga aat ga	3' biotin	1190-1219
scr2	tga gca tat aat agt gcg tgg tct cat aca tca gtt tt	3' biotin	
rel2	tca ttt caa aaa aca aac acg tga gcc tgt gtc aat tg	/	

Table 6: Oligonucleotides used for A3G mRNA pull-down. *For each oligo, a name, the sequence of the oligo from 5' to 3', the type of modification if there is one and the position on A3G mRNA to which the oligo hybridizes are indicated.*

2.1.2 Preparation of cells

HEK293T cells at confluence were washed once with DMEM medium (Gibco), supplemented with 10% fetal calf serum (Pan Biotech), 100 U/ml penicillin and 100 µg/ml streptomycin (Gibco), and once with EDTA/0.25 % trypsin (Gibco). Then they were detached from the culture support using trypsin and resuspended in DMEM. Cells were counted using 0.4 % trypan blue (Sigma) in a *Luna* cell counter (logos) and dispatched into 10 cm dishes at $2.5*10^6$ cells per dish in a final volume of 10 ml DMEM. Cells were incubated at 37 °C and 5 % CO_2.
Cells were transfected with expression plasmids around 16-18 h after their seeding. 0.1 µg of pCMV-A3G and 1 µg of pcDNA-Vif were used per dish. The total amount of plasmid is adjusted to 2 µg using the pcDNA-Ø. The plasmid preparation is gently mixed with 6 µl of XtremeGene9 (Roche) and 94 µl of Opti-MEM (Gibco) and incubated at room temperature for 10-15 min, then pipetted onto the cells.
Cells were detached from the dishes 24 h after transfection using trypsin. Cells were resuspended in DMEM and counted. They were then centrifuged at 288 xg for 5 min and washed twice with cold dPBS (Gibco). Cell pellets were kept on ice.

2.1.3 Cell lysis in a syringe

The cells were resuspended in lysis buffer S (20 mM HEPES (pH 7.5); 100 mM KCl; 10 mM $MgCl_2$; 1 mM DTT; 0.5 % NP40; 0.1 u/ml RNAsin (Promega); 5 mM EDTA; 1x Halt protease inhibitor cocktail (ThermoScientific)) at 100 µl per 10^7 cells. Cells were aspirated ten times through a 27G needle, and the resulting total cell lysate was centrifuged at 153 xg for 10 min at 4 °C to eliminate cell debris.
The protein concentration of the supernatant was determined by a Bradford assay. In a 96 well plate, 1 µl of each lysate was diluted to a final volume of 200 µl. In parallel, a concentration range of 0 to 70 µg/ml of BSA was also established at a final volume of 200 µl. Fifty µl of Bradford Assay Dye (BioRad) were added and the absorbance was measured at 595 nm. The lysate was then aliquoted at 500 µg of protein per tube, 15 mM of DTT were added and the final volume was adjusted to 200 µl with 100-KCl buffer (20 mM HEPES (pH 7.5); 100 mM KCl; 10 mM $MgCl_2$; 0.01 % NP40).

2.1.4 Pull-down

Fifty µg of MagsiSTA-600 streptavidin-coated magnetic beads (MagnaMedics) were incubated with 500 pmol of an equimolar mixture of ten biotinylated oligonucleotides complementary to A3G mRNA. Alternatively, equivalent amounts of other beads in the Magsi-STA trial kit (MagnaMedics) were used. After 1 h in a thermomix at 25 °C and 650 rpm shaking, beads

were collected on a magnetic stand, washed once in 100-KCl buffer and then incubated for 15 min at 37 °C in 100 µl saturation buffer (0.2 µg/µl total yeast tRNA; 0.5 µg/µl heparin; 0.64 U/µl RNasin; 20 mM HEPES (pH 7.5); 100 mM KCl; 10 mM $MgCl_2$; 0.01 % NP40). One aliquot of cell lysate was added and incubated in a thermomix for 1 h at 37 °C and 650 rpm, then the tubes were kept on ice for at least 15 min. Beads were washed twice with 100-KCl buffer and resuspended in 30 µl H_20.

To test for unspecific binding, 0.3 U of DNase (Roche) and 1x DNase buffer (Roche) have been added to beads after pull-down and incubated at 37 °C for 1 h. The mix was then subjected to phenol-chloroform extraction followed by ethanol precipitation.

For more stringency, a third washing step with a 500-KCl buffer (20 mM HEPES (pH 7.5); 500 mM KCl; 10 mM $MgCl_2$; 0.01 % NP40) or the 1000-KCl buffer (20 mM HEPES (pH 7.5); 1 M KCl; 10 mM $MgCl_2$; 0.01 % NP40) was added after incubation of the beads with the lysate.

For preclearing, the lysate was incubated twice with 150 µg of beads at 37 °C for 30 min and the beads were eliminated. The pre-cleared lysate was then used for pull-down as described above.

2.1.5 Analysis of oligo retention

Fifteen µl of the final bead suspension was mixed with 15 µl of formamide blue (95 % formamide, 0.025 % xylene cyanol, 0.025 % bromophenol blue) and incubated at 90 °C for 2 min. The mixture was loaded onto a 0.8 % agarose gel (TBE 0.5X) and run at 100 V for 30 min and revealed using ethidium bromide staining.

2.1.6 RT-PCR

Fifteen µl of the final bead suspension were used for RT-PCR.

RT-BioRad: the final volume of beads was adjusted with H_2O to 16 µl. Four µl of iScript Reverse Transcription Supermix for RT-qPCR (BioRad) or the corresponding No-RT control were added and the mixture was incubated at 25 °C for 5 min, 42 °C for 30 min and 85 °C for 5 min.

RT-AMV: 1 µl of primer Spe (Table 7) was added to the beads and the mixture was heated at 90 °C for 2 min and then kept on ice for at least 2 min. RT Buffer 1x (Life Sciences), 0.25 mM of each dNTP and 4 U of RT-AMV (Life Sciences) were added and the total volume was adjusted to 20 µl. The mixture was incubated at 42 °C for 20 min, 50 °C for 30 min and 60 °C for 10 min.

After RT, 500 µg of RNase A (Roche) were added and incubated for 15 min at 60 °C in order to digest template RNA.

PCR: An aliquot of the RT-reaction was mixed with 0.2 µM of sense and antisense primers (Table 7), 1x Taq buffer (75 mM Tris (pH 8.8), 20 mM ammonium sulfate, 0.01 % Tween20, 2 mM $MgCl_2$), 0.25 mM of each dNTP, 2.5 U of Taq polymerase (homemade). The total volume was adjusted to 50 µl and the mixture was incubated at 95 °C for 3 min, then 35 cycles of 30 s at 95 °C, 30 s at the hybridization temperature (Table 7) and 1 min at 72 °C were performed. This was followed by 5 min at 72 °C.

PCR products were migrated on a 0.8% agarose gel (TBE 0.5x) at 100 V for 45 min and revealed by ethidium bromide staining.

Name	Sequence (5'-3')	Target	Position on target	Size of amplicon (bp)	Hybridization temperature (°C)
Spe	aag atg cac agg ctc acg tg	A3G	1207-1226		
Es	ttg caa ttg cct tgg gtc ct	A3G	11-30	375	65
Ea	gac gag aaa gga tgg gtc ta	A3G	366-385		
Cs	gga tcc acc cac att cac tt	A3G	888-907	89	66
Ca	atg cgc tcc acc tca taa c	A3G	959-977		
ESPs	ttc tcc aga atc agg aaa ac	A3G	1429-1449	122	61
ESPa	gtg tct gtg atc agc tgg ag	A3G	1569-1551		
Gs	aac ctg cca agt acg atg aca tc	GAPDH	828-850	76	55
Ga	gta gcc cag gat gcc ctt ga	GAPDH	885-904		

Table 7: Primers used for PCR and qPCR. *For each primer a name, the sequence of the primer from 5' to 3', the target mRNA and the position of the target to which the oligo hybridizes are indicated. Spe is used as an RT-primer, the others go by pairs (s-sense and a-antisense) and are used for PCR amplification of an amplicon with the indicated size (bp=base pairs).*

2.1.7 Quantification of bands on agarose gel

An image of the gel was analysed with ImageJ (https://imagej.nih.gov/ij/) (196). A rectangular selection of the same size has been drawn around the bands of the expected size. An empty lane or a region directly below or above the bands has been used to quantify the background signal. Integrated density was measured for each rectangle.

2.1.8 *In vitro* transcription of A3G mRNA

First, 5 µg of the pCMV-T7-A3G has been linearized using 10 U XbaI in 1x Tango buffer for 2 h at 37 °C. The digested plasmid has then been purified using phenol-chloroform extraction followed by ethanol precipitation and linearization has been verified on an agarose gel. *In vitro* transcription has been performed using 1 µg of linearized plasmid and the Megascript T7 kit (Invitrogen) according to manufacturer's instructions at 37 °C for 3 h. The reaction mix was

then treated with Turbo DNase for 15 min at 37 °C and the transcripts were purified using phenol-chloroform extraction followed by ethanol precipitation.

2.1.9 Preparation of A3G mRNA-associated proteins for mass spectrometry

Five 10 cm dishes of cells were prepared for each pull-down sample as described above. Cells were treated with 25 µM of ALLN (Calbiochem) 10 h after their transfection and with 0.1 mg/ml cycloheximide 30 min before harvesting. Cycloheximide allows stabilization of translational complexes on the mRNA and has been added at a concentration of 0.1 mg/ml to all buffers of the pull-down protocol.

Cells were detached from the culture dishes in cold PBS using cell scrapers, spun down at 288 xg for 5 min and washed once in cold PBS. The volume of the cell pellet was estimated by eye and cells were resuspended in 3 times their volume of hypotonic lysis buffer (10 mM HEPES (pH 7.5), 1.5 mM $MgCl_2$, 10 mM KCl, protease inhibitor mix, RNasin). Cells were left to swell on ice for 10-15 min and were then broken with 15-20 strokes with a tight piston in a dounce homogenizer. NP40 0.4 % was added and gently mixed. Cell debris were spun down at 153 xg for 10 min at 4 °C and the supernatant was recovered.

The pull-down was performed as described above with 10 µg of beads, 500 pmol of the 10 biotinylated oligonucleotides, the complete amount of the five 10 cm dishes of cellular lysate, 3 washes with 100-KCl buffer for each sample. At the end, beads were resuspended in 30 µl of H_2O and 5 µl were used to check the presence of A3G mRNA by RT-PCR.

These steps were performed in a biological triplicate.

2.1.10 Mass spectrometry

Twenty-five µl of beads suspension was precipitated with 0.1 M ammonium acetate in 100 % methanol and were then resuspended in 50 mM ammonium bicarbonate. After a reduction-alkylation step (dithiothreitol 5 mM – iodoacetamide 10 mM), proteins were digested overnight with 1:25 (w/w) sequencing-grade porcine trypsin (Promega). Beads were then discarded and one fifth of the peptide mixture was analysed by nanoLC-MS/MS in an Easy-nanoLC-1000 system coupled to a Q-Exactive Plus mass spectrometer (ThermoFisher). Each sample was separated with an analytical C18 column (75 µm ID × 25 cm nanoViper, 3 µm Acclaim PepMap; ThermoFisher) with a 160 min 300 nL/min gradient of acetonitrile. The obtained data was searched against the Swissprot database with human taxonomy using the Mascot algorithm (version 2.5, Matrix Science). Mascot files were then imported into Proline v1.4 package (http://proline.profiproteomics.fr/) for post-processing. Proteins were validated with 1 % FDR

and the total number of MS/MS fragmentation spectra (Spectral Count) was used to quantify each protein in the different samples.
These steps have been performed by the team of the mass spectrometry platform (Plateforme de proteomique Esplanade).
Bioinformatic analyses were performed as described in IV-2.5.

2.2 Pull-down and specific elution of A3G mRNA using complementary, desthiobiotinylated oligonucleotides

Cellular lysates were prepared as described in 2.1.2 and 2.1.3.
Fifty µg of MagsiSTA-600 streptavidin-coated magnetic beads (MagnaMedics) were incubated with 500 pmol of an equimolar mixture of 4 desthiobiotinylated oligonucleotides complementary to A3G mRNA (Table 6). After 1 h at 4 °C, the beads were collected on a magnetic stand and washed once in 100-KCl buffer. One aliquot of cell lysate was added and incubated in a thermomix for 1 h at 37 °C and 650 rpm. The beads were washed once with 100-KCl buffer and once with 500-KCl buffer. The beads were then resuspended in 30 µl H_20 supplemented with 250-1000 pmol of free biotin and incubated for 1 h at 25 °C and 650 rpm. The eluate was recovered, and the beads were resuspended in 30 µl H_2O.
An aliquot of 15 µl of each fraction (eluate and beads) was used for recovery of oligos by incubation with 15 µl of formamide blue at 90 °C for 2 min. The remaining 15 µl of each fraction was used for detection of A3G mRNA by RT-PCR as described in 2.1.6. Quantification has been performed as described in 2.1.7.

2.3 Pull-down of A3G mRNA using complementary, biotinylated oligonucleotides and elution by RNase H

2.3.1 Digestion of *in vitro* transcribed A3G mRNA by RNase H followed by migration on a denaturing gel

Five pmol of *in vitro* transcribed A3G mRNA (prepared as in 2.1.8) have been incubated with 500 pmol of oligo B7 or an equimolar mix of nine oligos (B1 to B9, Table 6) at 90 °C for 2 min. After 2 min on ice, RNase H buffer (20 mM HEPES (pH 7.5); 50 mM KCl; 5 mM $MgCl_2$; 1 mM DTT ; 50 µg/µl BSA; ThermoScientific) and 0.5 µl RNase H (ThermoScientific) have been added in a final volume of 100 µl and incubated at 37 °C for 20 min. The RNA was then purified using a phenol-chloroform extraction followed by ethanol precipitation and resuspended in 15

µl urea blue before being run of an 8 % PAGE-8 M urea gel (TBE 1x) at 200 V for 2 h. RNA bands on the gel have been revealed by incubation in stains-all (Sigma).

2.3.2 Pull-down followed by elution with RNase H

Cellular lysates have been prepared by Dounce homogenization in a hypotonic lysis buffer as described in 2.1.9. Ten µg of MagsiSTA-600 beads have been incubated with 100 pmol of one or nine oligos at 25 °C for 1 h with shaking. Beads were washed once with 100-KCl buffer and resuspended in 20 mM HEPES (pH 7.5); 100 mM KCl; 10 mM $MgCl_2$; 0.01 % NP40 ; 15 mM DTT supplemented either with 10 pmol of *in vitro* transcribed A3G mRNA, 200 µg of total cellular RNA (extraction with TriReagent (Molecular Research center) following manufacturer's instructions) or a total cellular lysate equivalent to five 10 cm dishes of cells in a final volume of 200 µl. The mixture was incubated for 2 h at 37 °C. Beads were then washed twice with 100-KCl buffer, once with 500-KCl buffer and then resuspended in RNase H buffer and transferred into a new tube. Fresh RNase H buffer supplemented with 0.25 µl RNase H (ThermoScientific) was added. After incubation at 37 °C for 30 or 60 min, RNase H was inactivated at 90 °C for 2 min. The supernatant was recovered, and the beads were resuspended in 16 µl H_2O.

2.3.3 RT-qPCR

RT-BioRad: the final volume of the eluate or the beads was adjusted with H_2O to 16 µl. Four µl of iScript Reverse Transcription Supermix for RT-qPCR (BioRad) or of the corresponding No-RT control were added and the mixture was incubated at 25 °C for 5 min, 42 °C for 30 min and 85 °C for 5 min.

qPCR: the RT-reaction was mixed with 0.2 µM of ESP sense and antisense primers (Table 7), 1x Taq buffer, 0.25 mM of each dNTP, 2.5 U of Taq polymerase (homemade) and 1.25 µl of EvaGreen Dye (Biotium). The mixture was incubated at 95 °C for 3 min, then 40 cycles of 30 s at 95 °C, 30 s at 61 °C and 1 min at 72 °C were performed and the amount of DNA was monitored at each cycle using Evagreen fluorescence measurement. In parallel, a titration curve was generated using between 10^9 and 10^3 copies of the pCMV-A3G.

2.4 Pull-down of A3G mRNA using a complementary, biotinylated oligo and elution by a competitor oligo

Cells were prepared as described in 2.1.2. The following protocol was adapted from (113). The cell pellet was resuspended in lysis buffer K (469 mM LiCl ; 62.5 mM Tris HCl (pH 7.5); 1.25 % LiDS ; 1.25 % Triton X100 ; 12.5 mM RVC ; 12.5 mM DTT ; 125 U/mL RNasin ; 1.25 x Halt Protease Inhibitor Cocktail (ThermoScientific)) to a final concentration of $5*10^6$ cells/mL and kept on ice for 10 min with intermittent vortexing. Then, cells were sonicated on ice during 4 s with a 4 s break for 8 cycles and centrifuged at 1,000 xg for 5 min. The supernatant was recovered and 0.25 times its volume of H_2O was added. Cell lysate equivalent to around five 10 cm cell culture dishes (approximately $1\text{-}2*10^7$ cells) were used per sample. Alternatively, 1 fmol of an *in vitro* transcribed RNA (see 2.1.8) or total cellular RNAs from $1\text{-}2*10^7$ cells (isolated using TriReagent (Molecular Research Center) following manufacturer's instructions) were diluted in a final volume of 1 ml at 375 mM LiCl; 50 mM Tris HCl (pH 7.5); 1 % LiDS; 1 % Triton X100; 10 mM RVC; 10 mM DTT; 100 U/mL RNasin; 1 x Halt Protease Inhibitor Cocktail (ThermoScientific).

Then, 10-1,000 fmol of capture or scramble oligo (for sequences see table 6) were added and the mixture was incubated for 3 h at 37 °C. SpeedBeads Magnetic streptavidin coated particles (GE Healthcare) were washed with bead washing buffer (375 mM LiCl; 50 mM Tris HCl (pH 7.5); 1 % LiDS; 1 % Triton X100) and 6 to 60 µg of beads were added to the lysate and incubated 1 h at 37 °C. Beads were then washed for 15 min at 37 °C with washing buffer K (100 mM LiCl; 50 mM Tris-HCl (pH 7.5); 0.2 % LiDS; 0.2 % Triton X100), resuspended in release buffer K (375 mM LiCl; 50 mM Tris-HCl (pH 7.5); 0.1 % LiDS) and transferred into a new tube. The release oligo was added at 10^3 times the amount of capture oligo in a final volume of 16 µl and incubated at room temperature for 30 min. The supernatant was recovered and used for RT-qPCR as described in 2.3.3.

For RT-qPCR of GAPDH mRNA, samples were mixed with 10 µM of sense and antisense primers (Table 7) and 7.5 µl of Maxima SYBR qPCR Mastermix (ThermoFisher). The final volume was adjusted to 15 µl. The mix was incubated at 95 °C for 5 min, then 40 cycles of 95 °C for 15 s, 55 °C for 13 s and 72 °C for 30 s were performed with acquisition of the SYBR green fluorescent signal at every cycle.

2.5 Pull-down of cellular proteins on an in vitro transcribed, biotinylated A3G mRNA

2.5.1 *In vitro* transcription of biotinylated, capped, poly-adenylated A3G mRNA

First, 5 µg of the pCMV-T7-A3G has been linearized using 10 U XbaI in 1x Tango buffer for 2 h at 37 °C. The digested plasmid has then been purified using phenol-chloroform extraction followed by ethanol precipitation and linearization has been verified on agarose gel.

For transcription of a capped, poly-adenylated and biotinylated transcript, 1 µg of linearized plasmid was used with the HiScribe T7 ARCA mRNA kit with tailing (NEB) according to manufacturer's instructions. The transcription mix was supplemented with 1.25 mM Bio-16-UTP (ThermoFisher) and incubated at 37 °C overnight. The plasmid was digested with 4 U DNase I (NEB) at 37 °C for 15 min followed by a tailing reaction at 37 °C for 30 min. The transcripts were precipitated by 2.5 M LiCl (NEB) according to manufacturer's instructions.

2.5.2 Preparation of cellular lysates using a nitrogen cell bomb

HEK293T cells at confluence were washed once with DMEM medium (Gibco), supplemented with 10 % fetal calf serum (Pan Biotech), 100 U/ml penicillin and 100 µg/ml streptomycin (Gibco), and once with EDTA/0.25 % trypsin (Gibco). Then they were detached from the culture support using trypsin and resuspended in DMEM. Living cells were counted using 0.4 % trypan blue (Sigma) in a *Luna* cell counter (logos) and dispatched into 15 cm dishes at $5.8*10^6$ cells per dish in a final volume of 20 ml DMEM. Cells were incubated at 37 °C and 5 % CO_2.

Cells were transfected with expression plasmids around 16-18 h after their seeding. One µg of pCMV-A3G-ΔUTR and 8 µg of pcDNA-Vif were used per dish. The total amount of plasmid was adjusted to 9 µg using the pcDNA-Ø. The plasmid preparation was gently mixed with 27 µl of XtremeGene9 (Roche) and 373 µl of Opti-MEM (Gibco) and incubated at room temperature for 10-15 min, then pipetted onto the cells.

Cells were detached in cold dPBS (Gibco) using cell scrapers 24 h after transfection. They were centrifuged at 300 xg for 5 min. The cell pellet was weighed and then resuspended in lysis buffer P (20 mM HEPES, 100 mM KAc, 2 mM MgAc, 1 mM DTT, 100 U/mL RNasin; 1 x Halt Protease Inhibitor Cocktail (ThermoScientific)) at 2.5 ml per g of cells. Cells were then introduced into a nitrogen cell bomb (Parr) and incubated at 500 psi for 30 min at 4 °C. The resulting cell lysate was spun down in two subsequent cycles at 1000 xg for 5 min at 4 °C and the supernatant was recovered and stored at -80 °C.

2.5.3 Pull-down of cellular proteins on biotinylated A3G mRNA

SpeedBeads Magnetic streptavidin coated particles (GE Healthcare) were washed with 100-KCl buffer and 40 µg of beads were incubated with 1 pmol of biotinylated A3G mRNA in 100-KCl buffer for 1 h at room temperature under rotation. Beads were then washed with 100-KCl buffer, resuspended in cellular lysate (per sample, an equivalent of 1.5 cell dishes, diluted at 1/2 in lysis buffer P was used) and incubated for 3 h at 4 °C under rotation. Beads were washed twice with 100-KCl buffer, once with ND buffer (20 mM HEPES, 100 mM KCl, 10 mM $MgCl_2$) and transferred into a new tube. Beads were then resuspended in 15 µl H_20 and 2 µl were used for RT-qPCR as described in 2.3.3.

3. Results

3.1 Pull-down of A3G mRNA using complementary, biotinylated oligonucleotides

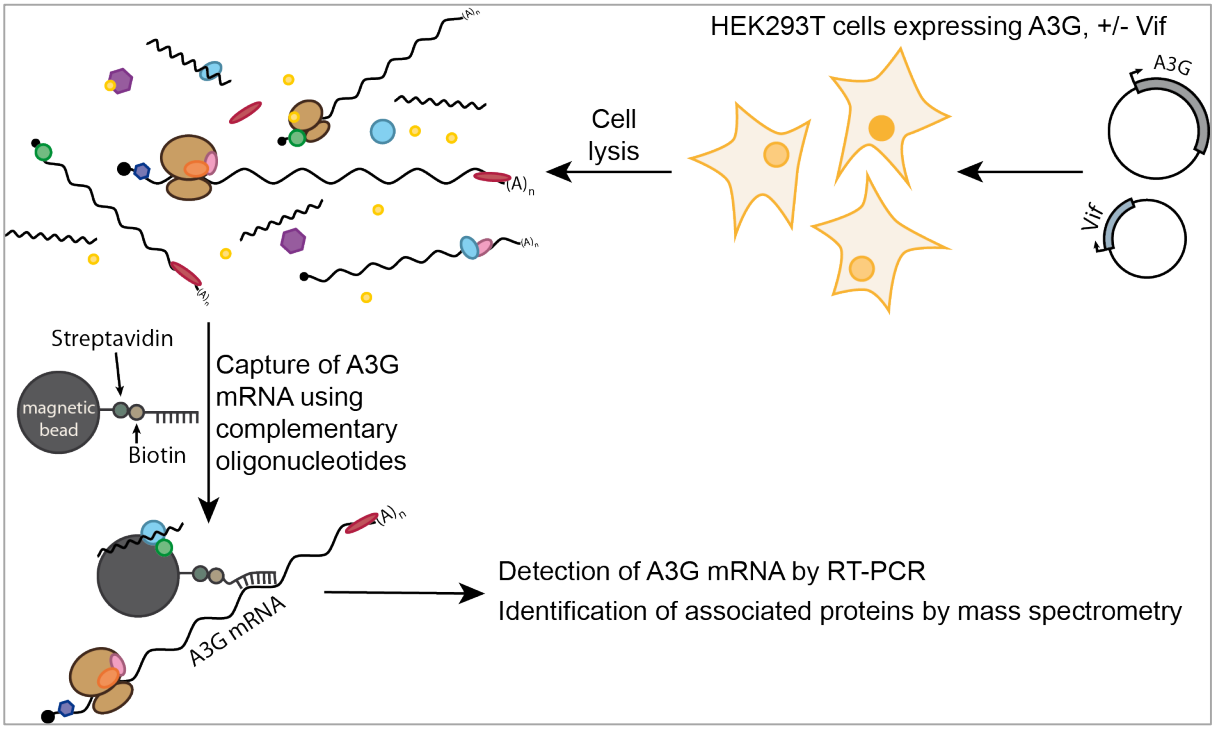


Figure 19: Experimental setup for pull-down of A3G mRNA using complementary, biotinylated oligonucleotides. *A3G mRNA was expressed in HEK293T cells and then pulled down from the total cellular lysate using magnetic beads coupled to complementary, biotinylated oligonucleotides, followed by identification of A3G mRNA-associated proteins by mass spectrometry.*

3.1.1 Experimental setup

The secondary structure of the A3G mRNA has been predicted using the mfold website (240). The secondary structure of the 5'- and 3'-UTRs corresponded globally to their experimentally

validated structures (147). This allowed us to choose ten mainly or partially single-stranded regions in the coding sequence of A3G mRNA for the design of complementary oligonucleotides that are regularly spaced across the entire coding sequence of A3G (Appendix 2). A search has been conducted on the NCBI database against the entire human genome in order to exclude any non-specific interactions with other mRNAs. An equimolar mix of the ten oligonucleotides, biotinylated at their 5'-end, were bound to streptavidin-covered magnetic beads (Fig. 19). In parallel, a cellular lysate was produced from HEK293T cells, which expressed A3G mRNA. This lysate was then incubated with the oligo-bound beads in order to allow specific capture of A3G mRNA by hybridization with the complementary oligos (Fig. 19). Beads without oligos as well as non-transfected cells were used as a negative control. Following several washing steps, the aim of this protocol was to specifically isolate A3G mRNA as well as the associated proteins.

3.1.2 Binding of the biotinylated oligonucleotides to different types of beads

First of all, the correct binding of the biotinylated oligonucleotides on different types of streptavidin-coated magnetic beads has been tested in the absence of cellular lysate. For this, different beads, mainly divergent in their size, surface coating and number of bound molecules (Fig. 20A), have been compared. The quantity of beads used in these experiments has been adjusted in order to obtain the same theoretical binding capacity of 250 pmol of biotin. Thus, 500 pmol of oligonucleotides have been incubated with the beads as input in order to saturate the beads and maximize the binding. At the end of the protocol, the beads were recovered and oligos were detached by heating in formamide buffer. Migration of the recovered oligos on an agarose gel showed that MagsiSTA-3TL beads are the only ones that were able to bind 50 % of the input oligos, which corresponds to their calculated maximum binding capacity, while the other beads show less binding, ranging from 40 % to 17 % of the input (Fig. 20B).

3.1.3 Retention of the A3G mRNA on magnetic beads in presence and absence of complementary oligonucleotides

Next, pull-down of A3G mRNA from a cellular lysate was performed using the same types of beads as before. The quantity of A3G mRNA retained on the beads at the end of the protocol was estimated by RT-PCR. While MagsiSTA-3TL beads were able to bind the largest quantity of biotinylated oligonucleotides (Fig. 20B), they seemed not to pull-down A3G mRNA at all (Fig. 20C). Globally, beads with a carboxyl surface coating showed better performance than those with tosyl surface coating. Actually, it were MagsiSTA-600 beads that retained A3G most efficiently, closely followed by MagsiSTA-600BI, -3L and -1 beads (Fig 20C).

A

Name	600	600BI	1L	1	1TL	1TS	3L	3TL
Size	600 nm		1 µm				3 µm	
Surface	carboxyl				tosyl		carboxyl	tosyl
Binding	5000	6800	2000	5000	2000	5000	1200	900

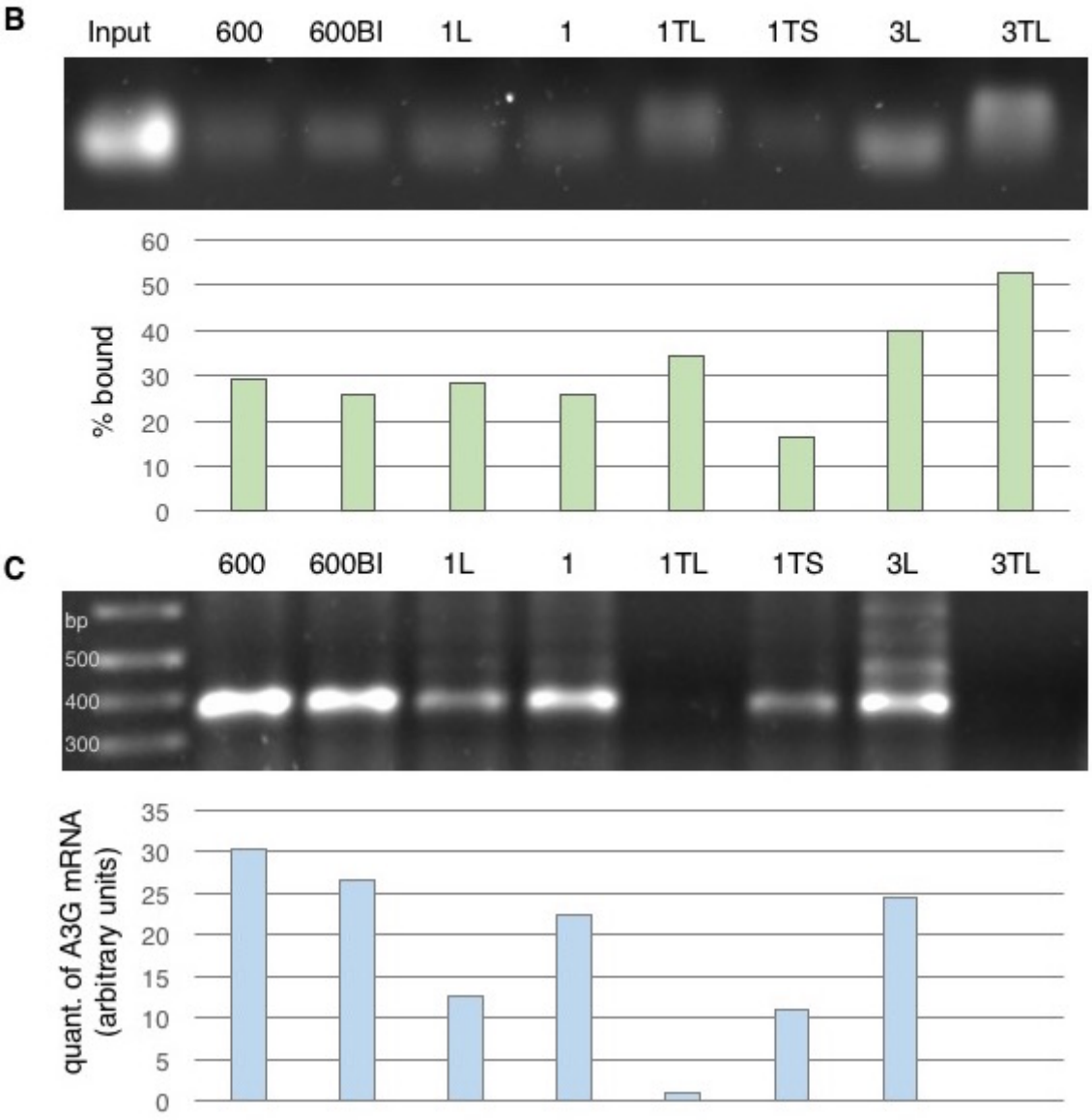


Figure 20: Binding of biotinylated oligonucleotides on different types of beads. (A) *Overview over the different types of beads with their size, type of surface coating and biotin binding capacity (indicated in pmol of biotin per mg of beads).* ***(B)*** 500 pmol of biotinylated oligonucleotides have been incubated with different types of streptavidin-covered magnetic beads. The quantity of beads used has been adjusted to obtain 250 pmol of binding capacity for each type. The p*ull*-down protocol has been performed in the absence of cellular lysate. At the end of the protocol, oligos were eluted using formamide blue and run on a 0.8 % agarose (TBE 0.5x) gel. 500 pmol of the oligo mixture were run as an input control. Each band was quantified using ImageJ and the percentage of the input was calculated. **(C)** The same protocol has been repeated in the presence of cellular lysate and at the end, A3G mRNA was amplified using RT-PCR (RT-AMV with primer Spe, PCR with primers E). PCR products were run on a 0.8 % agarose (TBE 0.5x) gel and bands were quantified using ImageJ.

These four types of beads were then tested for their A3G mRNA binding specificity. For this, A3G mRNA retention was analysed on oligo-bound beads as well as on a beads-only control without oligos. None of the beads showed a very good specificity, with a lot of A3G mRNA bound to beads in the absence of complementary oligos. The ratio of A3G mRNA bound in the presence compared to the absence of oligos ranged from 1.47 to 0.91 (Fig. 21A).

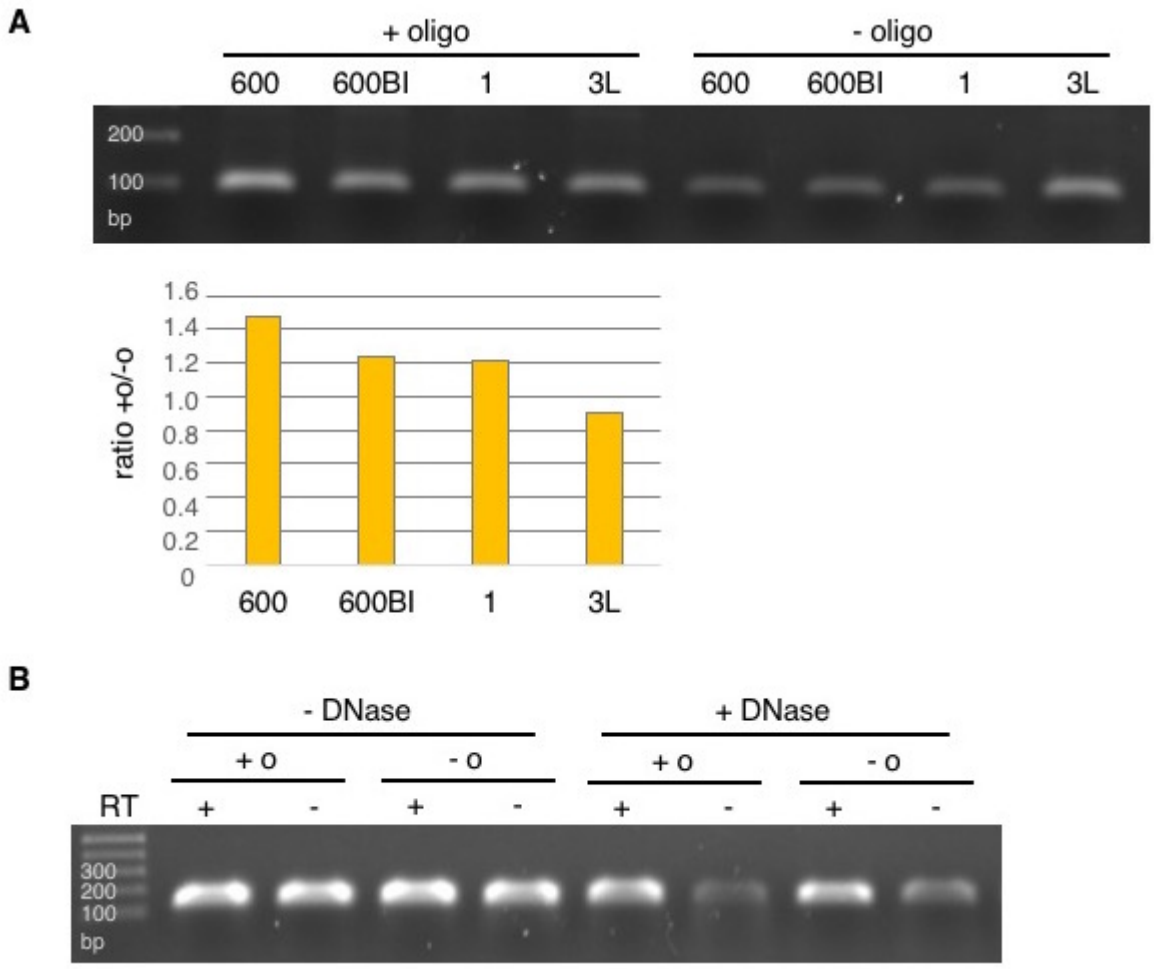


Figure 21: Specificity of retention of A3G mRNA on oligo-bound magnetic beads compared to the beads-only negative control. ***(A)*** *The pull-down protocol was performed using 10 µg of beads coupled to the capture oligos (+oligo) or not (-oligo). One µg of an in vitro transcribed A3G mRNA was used for capture instead of the cellular lysate. At the end, A3G mRNA was amplified using RT-PCR (RT-AMV primer Spe, PCR primers C). The PCR products were run on a 0.8 % agarose (TBE 0.5x) gel and bands were quantified using ImageJ. The ratio between the band obtained in presence of capture oligos compared to the beads-only control was calculated.* ***(B)*** *The pull-down protocol was performed using 50 µg of Magsi-600 beads in the presence (+ o) or absence (- o) of capture oligos. An additional washing step with 500-KCl buffer was added. At the end of the protocol, beads were treated with DNase (+ DNase) or not (-DNase). The A3G mRNA was amplified using RT-PCR (RT-BioRad or no-RT control, PCR primers ESP).*

As the MagsiSTA-600 beads seemed to be the best overall in oligo-binding, A3G mRNA binding and specificity, they were selected for subsequent experiments. In order to understand what was causing the unspecific binding to beads in the absence of oligos, a no-RT control was performed, which revealed indeed a contamination by DNA (Fig. 21B lanes 2 and 4).

However, while a DNAse treatment of the samples resulted in a rather clean no-RT control (Fig. 21B lanes 6 and 8), it did not decrease the signal in the absence of oligos (Fig. 21B lane 7). This indicated that it was RNA and not DNA which bound directly and unspecifically to the beads in the absence of oligos. The fragment amplified by RT-PCR in samples with and without specific capture oligos has been purified and cloned into a pJET vector. Sequencing confirmed that in both cases, A3G mRNA was detected on the beads and not another contaminant (data not shown).

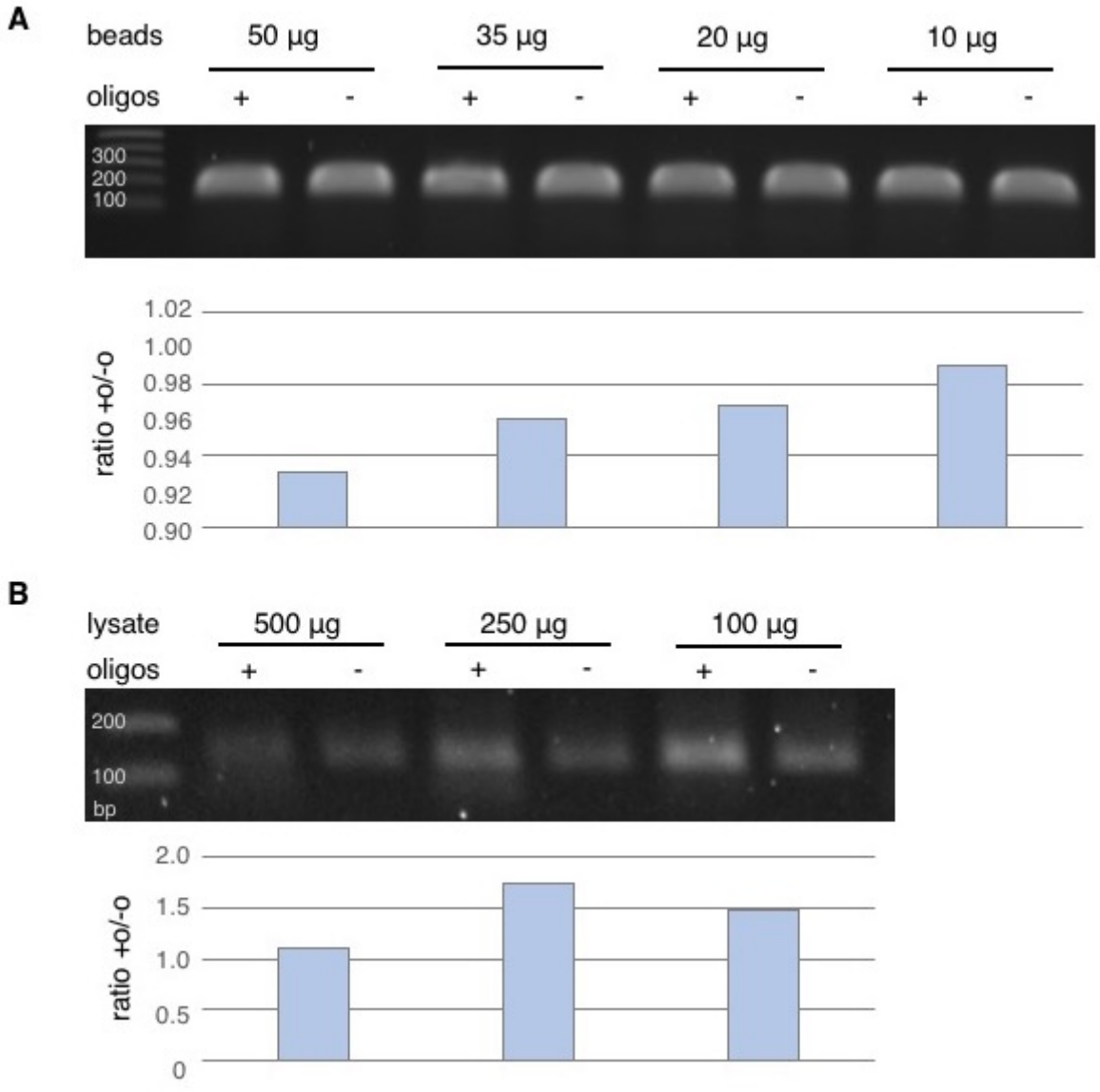


Figure 22: Pull-down of A3G mRNA from different quantities of cellular lysate with different amounts of magnetic beads. (A) *A3G mRNA was pulled down from 500 µg of cellular lysate using decreasing amounts of beads as indicated in the presence or absence of biotinylated oligos. Beads were washed once with 100-KCl and once with 500-KCl buffer. The A3G mRNA was amplified using RT-PCR (RT-BioRad, PCR primers ESP), migrated on a 0.8 % agarose (TBE 0.5 x) gel and bands were quantified using ImageJ.* ***(B)*** *Pull-down of A3G mRNA was performed on 50 µg of beads from decreasing amounts of cellular lysate as indicated in the presence or absence of biotinylated oligos. Beads were washed twice with 100-KCl and once with 500-KCl buffer using an Extractman (Gilson). The A3G mRNA was amplified using RT-PCR (RT-BioRad, PCR primers ESP), migrated on a 0.8 % agarose (TBE 0.5 x) gel and bands were quantified using ImageJ.*

3.1.4 Tests to eliminate aspecifically bound RNAs from the magnetic beads

The aforementioned experiment suggested that A3G mRNA was to a certain extent able to bind to the magnetic beads independently of the complementary oligonucleotides. In order to improve the specificity of the pull-down, several different approaches were tested.

In a first experiment, the quantity of beads used for pull-down was decreased, from the 50 µg used previously down to 10 µg, the hypothesis being that with a decreased amount of beads, the surface available for non-specific binding would decrease too. Indeed, the ratio between A3G mRNA bound to beads in the presence compared to the absence of biotinylated oligos increased slightly with a decreasing amount of beads; however, this was not sufficient to solve the problem as even with 10 µg of beads, A3G mRNA bound to the oligos was hardly higher than in the negative control without oligos (Fig. 22A).

With the aim of optimizing the lysate to beads ratio, decreased amounts of lysate were also tested, the hypothesis being that at 500 µg of lysate, the beads might be over-saturated. A slightly increasing specificity could be noted with a decreasing amount of lysate, with a maximum at 250 µg of proteins (Fig. 22B).

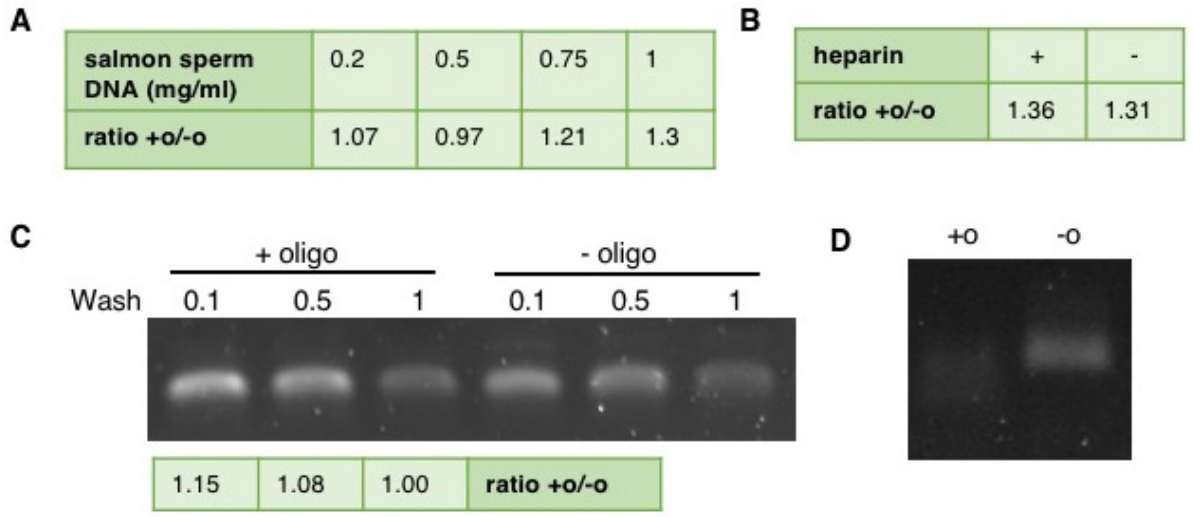


Figure 23: Tests to improve specificity of A3G mRNA pull-down by saturation, increasing washing stringency or preclearing. (A) *Different concentrations of salmon sperm DNA were added to the saturation buffer. After pull-down, A3G mRNA was amplified using RT-PCR (RT-BioRad, PCR primers ESP) and migrated on a 0.8 % agarose (TBE 0.5x) gel. Bands were quantified using ImageJ and ratios in the presence (+o) and absence (-o) of capture oligos were calculated.* ***(B)*** *The saturation buffer was prepared with or without heparin. After pull-down, A3G mRNA was amplified using RT-PCR (RT-BioRad, PCR primers E) and migrated on a 0.8% agarose (TBE 0.5x) gel. Bands were quantified using ImageJ and ratios were calculated.* ***(C)*** *Beads were washed twice with 100-KCl buffer (0.1) and a third washing step with 500-KCl or 1000-KCl buffer (0.5 and 1, respectively) was added. After pull-down, A3G mRNA was amplified using RT-PCR (RT-AMV with primer Spe, PCR primers C) and migrated on a 0.8 % agarose (TBE 0.5x) gel. Bands were quantified using ImageJ and ratios were calculated.* ***(D)*** *Cellular lysates were incubated twice with 150 µg of MagsiSta-600 beads for 30 min at 37 °C. Beads were eliminated before using the lysate for pull-down. The A3G mRNA was amplified using RT-PCR (RT-BioRad, PCR primers ESP) and migrated on a 0.8 % agarose (TBE 0.5x) gel.*

Next, different agents were tested in order to saturate the surface of magnetic beads, the objective being to mask potential non-specific binding sites of A3G mRNA on the beads (185). However, no significant change in the amount of A3G mRNA compared to the beads-only control could be noted, neither after saturation with salmon sperm DNA (Fig. 23A), nor with heparin (Fig. 23B).

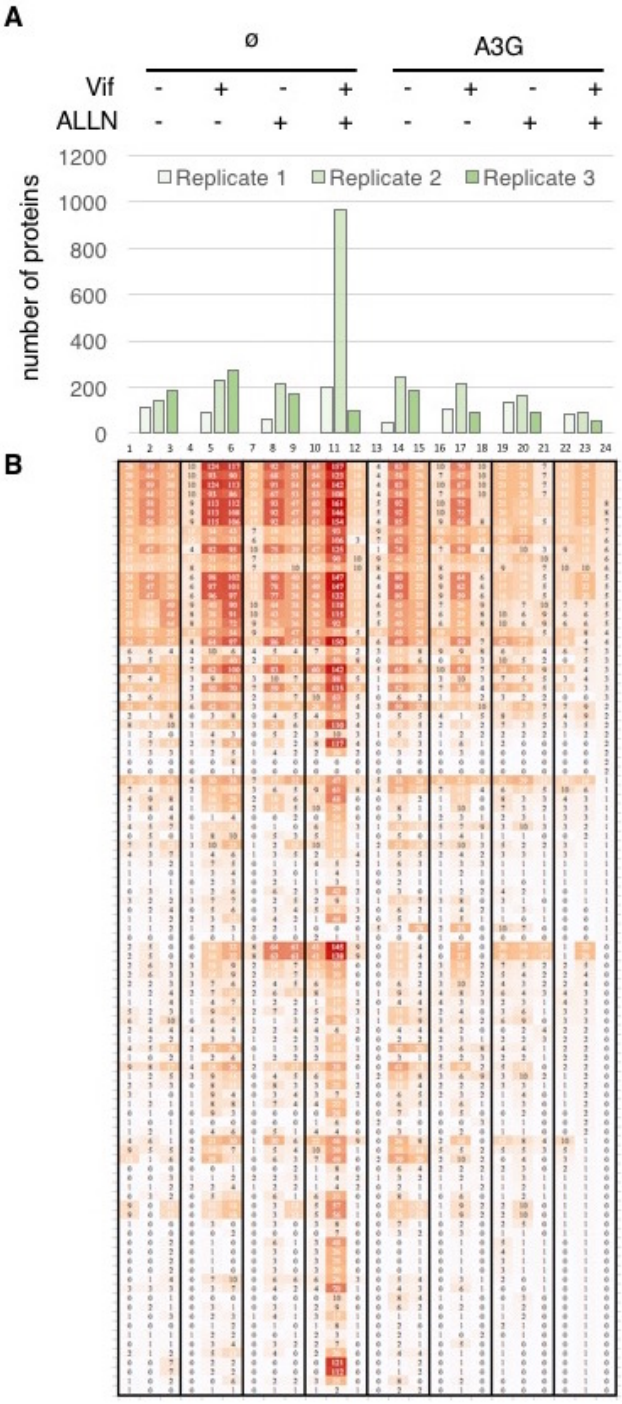
Next, supplementary washing steps of the beads after incubation with the cellular lysate were added, using washing buffers with increasing salt concentration. These washing steps were more stringent than the initial washing, and it was hypothesized that this might help to detach all mRNAs non-specifically bound to the beads while only A3G mRNA specifically bound by oligos would be able to subsist (218). However, more stringent washing led to a comparable loss of bead-bound A3G mRNA, both in the presence and in the absence of complementary oligos (Fig. 23C) and therefore didn't help to improve specificity of the pull-down.

Finally, lysates were incubated two times with uncoupled beads prior to the pull-down. This process, known as preclearing, is thought to clean the lysate of all molecules able to bind directly to the beads (185). However, preclearing only led to a loss of total material and did not help in decreasing the signal obtained in the absence of oligos (Fig. 23D).

Even though some of these conditions seemed to slightly increase the ratio of A3G mRNA recovered in the presence compared to the absence of complementary oligos, the overall variability of the pull-down is still too high. The standard conditions reproduced in every test show a large variability, with ratios ranging from 1.47 (Fig. 21A) to 0.94 (Fig. 22A). Therefore, even though some particular conditions seemed to show a slight increase, with a maximum ratio of 1.74 obtained by decreasing the quantity of lysate (Fig. 22B), this increase can't be considered significant.

3.1.5 Identification of proteins associated with A3G mRNA by mass spectrometry

In order to identify proteins that specifically interact with A3G mRNA, the pull-down was performed with 10 complementary, biotinylated oligos on cellular lysates transfected with pCMV-A3G following the protocol described above. As shown above, a condition without the complementary oligos could not be used as a negative control because a significant amount of A3G mRNA binds directly to the beads. Therefore, cells transfected with an empty vector have been used as a negative control because they don't express A3G mRNA. While this control will allow to identify background signal from proteins that non-specifically bind to the magnetic beads, A3G mRNA as well as its associated proteins are expected to be enriched solely on samples generated from A3G-expressing cells.



C

Number of proteins enriched on A3G mRNA

		-	+	-	+	
		-	-	+	+	ALLN
p-value	0.05	0	5	10	0	
	0.1	0	5	13	0	

D

Total proteins	
Functional cluster	Enrichment score
Ribosome, translation	54.6
Splicing	39.89
Cell-cell adhesion	35.79
Nucleotide binding	26.56
RNA recognition domain	18.96
Helicase	14.84
mRNA export	14.48
Mitochondia	8.64
Proteins enriched on A3G mRNA	
Name	Function
MYH9	Myosin heavy chain
LC7L2	RNA binding
RSRC2	Poly-A binding
NKAP	Transcriptional repressor
5 histones	
8 keratins	

Figure 24: Proteins pulled-down in the presence or absence of A3G mRNA. *After pull-down of A3G mRNA from cellular lysates, associated proteins were analyzed by mass spectrometry. Cells were transfected to express A3G and/or Vif and were treated or not with the proteasome inhibitor ALLN. Cellular lysates not expressing A3G (ø) were used as a negative control. All samples were done as a biological triplicate.* ***(A)*** *Total number of identified proteins in each sample.* ***(B)*** *Heatmap of the number of spectra for the 100 most abundant proteins. The darker the color, the more spectra were detected. Each row represents a protein.* ***(C)*** *Number of proteins that are significantly enriched in indicated samples compared to the corresponding negative control with an adjusted p-value < 0.1 or < 0.05 and a minimum fold change of 2. Analysis has been performed with a differential expression test using a negative-binomial regression model.* ***(D)*** *Functional clustering of total identified proteins as well as proteins specifically and significantly associated with the presence of A3G mRNA using DAVID.*

The total number of proteins identified by nanoLC-MS/MS was rather low overall with less than 250 proteins in most samples (Fig. 24A). There was also a considerable variability across different samples with a number of identified proteins ranging from 50 to 965 (Fig. 24A). The functional clustering of the total identified proteins revealed that known functions associated with mRNA metabolism (i.e. translation, splicing and nuclear export) or other known nucleic acid-associated domains (nucleotide binding domains, RNA recognition motifs and helicase domains) are amongst the most enriched functions (Fig. 24D).

A heatmap of the number of spectra identified for the 100 most abundant proteins also shows that there is a high variability even amongst different replicates of the same condition (Fig. 24B, compare lane 4 to 5 and 6 or 17 to 16 and 18 for example). Only few proteins are identified with a large number of spectra while most are detected with a low number of spectra in only a few samples (Fig. 24B). The heatmap also shows the tendency that a little less proteins are identified in the presence of A3G mRNA than in its absence (Fig. 24B, compare lanes 13-24 to 1-12).

Each condition was then compared in the presence of A3G mRNA with the negative control and proteins significantly enriched were determined using a differential expression test. This analysis allowed identification of very few proteins specifically enriched in the presence of A3G mRNA only in 2 of the 4 conditions tested (Fig. 24C), most of these being histones or keratins (Fig. 24D). Comparison of samples in the presence and absence of Vif showed no significantly downregulated or upregulated proteins (data not shown).

Overall, this analysis shows that the quality of generated samples is probably insufficient to reliably identify A3G mRNA-associated protein complexes. The pull-down protocol used in this experiment is probably not sufficient to gather enough A3G mRNA in order to stand out from the considerable background. Therefore, it was necessary to further improve the pull-down protocol before continuing analysis by mass spectrometry.

3.2 Pull-down and specific elution of A3G mRNA using complementary, desthiobiotinylated oligonucleotides

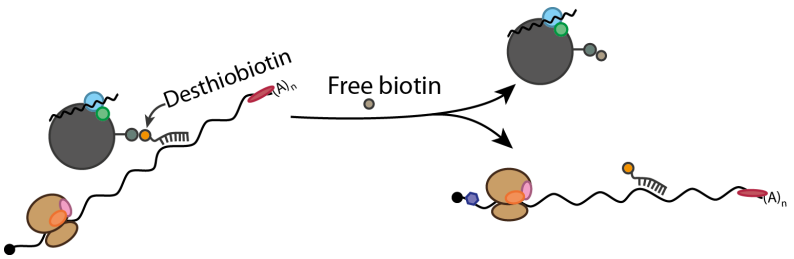


Figure 25: Elution of A3G mRNA captured by desthiobiotinylated oligos. *Free biotin competes with desthiobiotinylated oligos on strepavidin beads and thereby displaces the oligo-mRNA complex, while unspecifically bound RNAs and proteins remain on the beads.*

3.2.1 Experimental setup

Previous tests have shown that A3G mRNA seems to be retained on magnetic beads not only in the presence of complementary oligos, but also in their absence due to direct binding to the beads themselves. In order to specifically enrich samples in A3G mRNA only in the presence of complementary oligos, a specific elution step was added to the protocol. Therefore, a set of 4 oligos was used. These oligos were complementary to A3G mRNA and desthiobiotinylated at their 5'-end. Desthiobiotin binds to streptavidin but with a lesser affinity than biotin. In consequence, free biotin added to the sample at the end of the pull-down will compete with desthiobiotinylated oligos on the beads and displace the oligo-RNA complex (90). This allows to specifically elute A3G mRNA into the supernatant, while unspecifically bound RNAs and proteins remain bound to the beads and can be eliminated (Fig. 25).

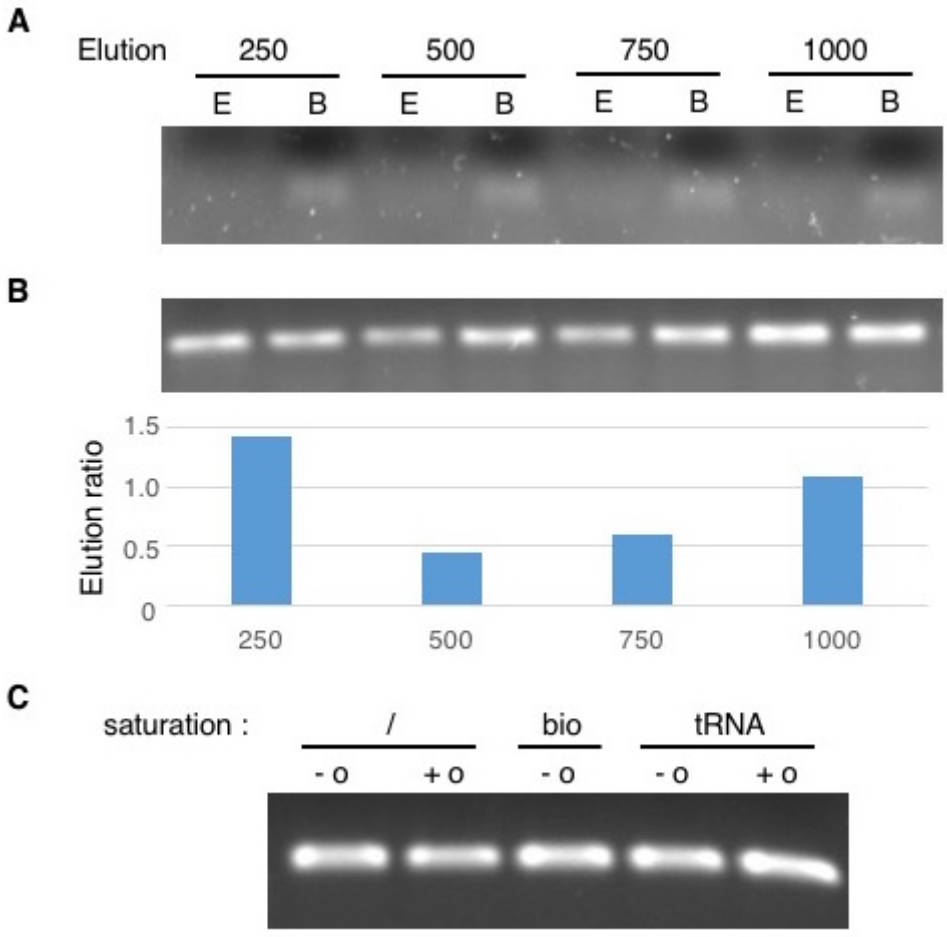


Figure 26: Pull-down of A3G mRNA using complementary, desthiobiotinylated oligos. *The pull-down was performed and A3G mRNA was eluted using between 250 and 1000 pmol of free biotin. The eluate (E) as well as the beads (B) were recovered and* ***(A)*** *half was incubated with formamide blue to reveal bound oligos, while* ***(B)*** *A3G mRNA was amplified from the other half by RT-PCR (RT-BioRad, PCR primers ESP). The obtained fractions were migrated on a 0.8% agarose (TBE 0.5x) gel, bands were quantified using ImageJ and elution ratios were calculated.* ***(C)*** *Pull-down in the presence or absence (/) of a saturation buffer containing total yeast tRNAs (tRNA) or biotin (bio) was performed with desthiobiotinylated oligos (+o) or without (-o). A3G mRNA in eluates was amplified by RT-PCR (RT-BioRad, PCR primers ESP) and migrated on a 0.8% agarose (TBE 0.5x) gel.*

3.2.2 Tests

First, different elution conditions were tested in order to obtain optimal elution of A3G mRNA from the beads at the end of the pull-down. Beads with a theoretic binding capacity of 250 pmol of biotin were used and elution of desthiobiotinylated oligos was performed with increasing amounts of free biotin, ranging from 250 pmol to 1 nmol. The elution of desthiobiotinylated oligos from the beads seemed very inefficient even at high quantities of free biotin and most of the oligos appeared to remain bound to the beads (Fig. 26A). Nevertheless, some of the A3G mRNA was still eluted from the beads with the highest elution efficiency with 250 pmol of free biotin (Fig. 26B).

However, elution of A3G mRNA was not dependent on the presence of desthiobiotinylated oligos; indeed, after elution, an equivalent amount of A3G mRNA was found in the supernatant in the presence and in the absence of oligos (Fig. 26C, lanes 4 and 5). Beads were saturated either with total yeast tRNAs or with free biotin, but this did not decrease the amount of A3G mRNA retrieved in the beads-only control (Fig. 26C, lanes 1-3).

3.3 Pull-down of A3G mRNA using complementary, biotinylated oligonucleotides and elution by RNAse H

3.3.1 Experimental setup

Another strategy to specifically elute A3G mRNA from complementary oligos while unspecifically bound RNAs remain on the beads is digestion by RNase H (211). RNase H only digests RNA hybridized to DNA and therefore allows release from the beads of A3G mRNA fragments that were hybridized to complementary oligos (Fig. 27). In order to be able to detect A3G mRNA fragments in the supernatant by RT-qPCR, one of the ten previously used biotinylated oligos (Table 6, B-10) had to be excluded from the pull-down, as it would have led to digestion of the RNA region used as a template for PCR amplification.

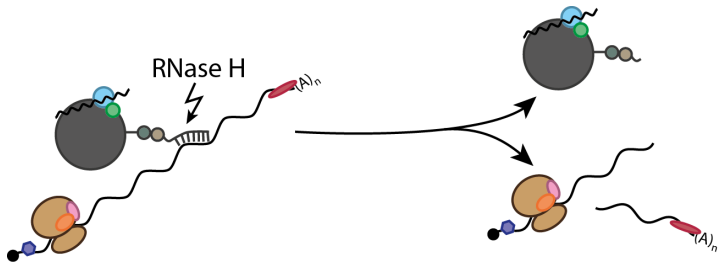


Figure 27: Elution of A3G mRNA hybridized to biotinylated capture oligos by RNase H treatment. *RNase H digests RNA hybridized to DNA which allows specific elution of A3G mRNA bound to its specific capture oligos.*

3.3.2 Tests

The feasibility of this approach has first been tested using an *in vitro*-transcribed A3G mRNA. The RNA has been incubated for 20 min without oligos, as well as with one (B7; Table 6) or the total set of nine oligos (B1-9; Table 6) in the presence of RNase and analysed on a denaturing PAGE. RNA is indeed digested for 20 min by RNase H in the presence of at least one oligo (Fig. 28A). However, it is difficult to evaluate the size of the generated RNA fragments and it seems like some full-length RNA subsists in all conditions. This might indicate that the digestion is not complete and longer incubation with RNase H might be necessary.

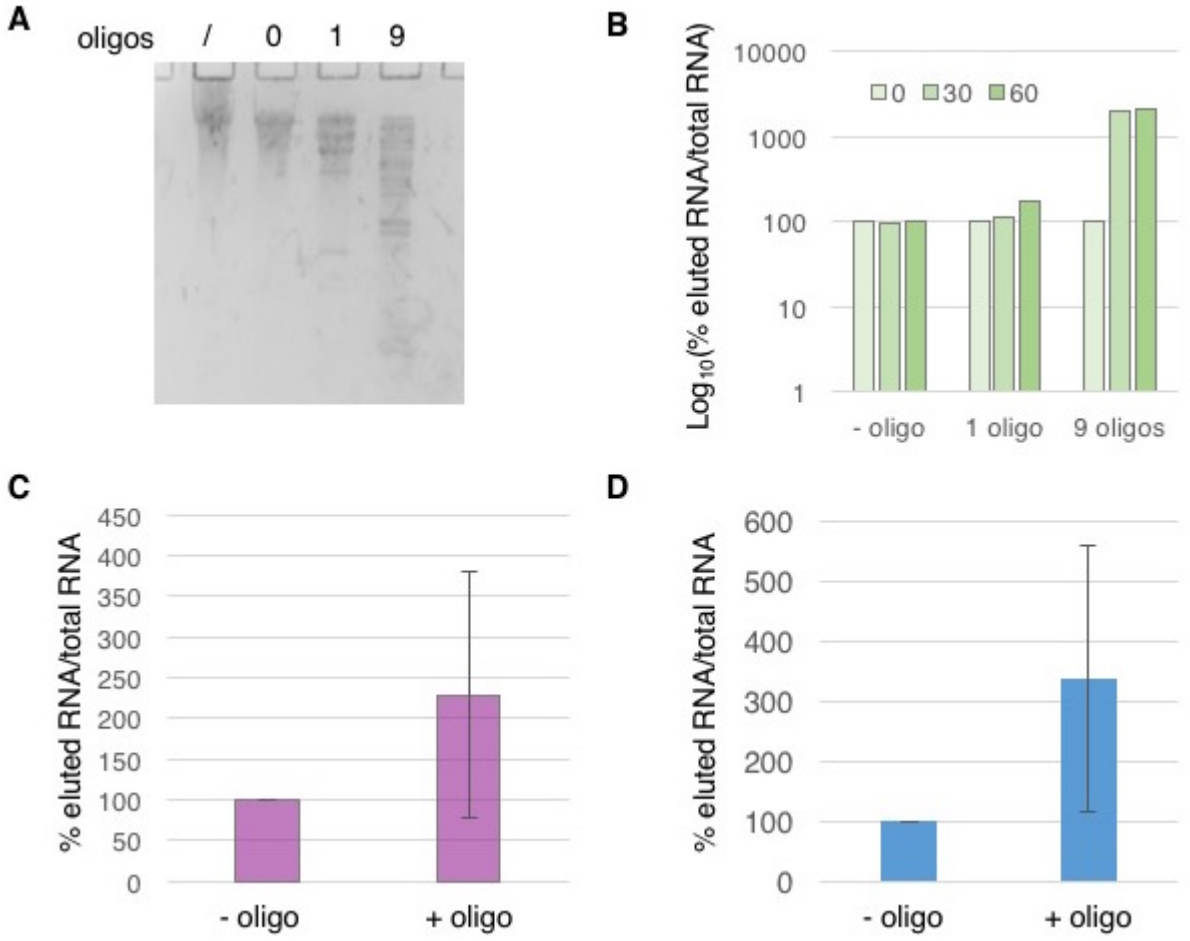


Figure 28: Pull-down of A3G mRNA followed by elution by RNase H digestion. (A) *Five pmol of in vitro transcribed A3G mRNA was digested by RNase H in the presence or absence of one (B-7, Table 1) or nine (B1-9, Table 1) oligos for 20 min at 37°C and migrated on a 8% PAGE/8M urea gel with a control RNA that has not been incubated with RNase H (/)* ***(B)*** *Pull-down of in vitro transcribed A3G mRNA has been performed in the presence or absence of 1-9 complementary oligos followed by elution with RNase H for 0 to 60 min (color gradient as indicated).* ***(C)*** *Pull-down of A3G mRNA from total cellular RNAs followed by elution by RNase H digestion. n=2.* ***(D)*** *Pull-down of A3G mRNA from a total cellular lysate (RNAs + proteins) followed by elution by RNase H digestion. n=4.* ***(B-D)*** *Eluted A3G mRNA fragments as well as A3G mRNA remaining bound to the beads has been quantified by RT-qPCR and ratios have been calculated. The ratio of eluted RNA in the absence of oligos has been set to 100 %.*

The *in vitro*-transcribed A3G mRNA has then been used for pull-down in the presence of one or nine oligos followed by RNase digestion for 30 to 60 min. The amount of eluted A3G mRNA has been quantified by RT-qPCR and normalized to the total amount of RNA (eluted RNA + RNA retained on the beads). After 30 min of incubation with RNase H, a significant amount of A3G mRNA is eluted from the beads only in the presence of nine oligos (Fig. 28B). A small amount is eluted in the presence of one oligo while the amount of eluted RNA remained constant independently of the RNase H incubation time in the absence of oligos (Fig. 28B).
The protocol has then been tested with total cellular RNAs extracted from cells. While the percentage of eluted RNA seems to clearly increase in the presence of oligos, a great variability could also be noted across replicates (Fig. 28C). Overall, the specific elution in the presence of oligos is not reproducible.
Finally, pull-down of A3G mRNA from total cell extracts was tested. The same tendency of an increased RNA elution in the presence of oligos could be noted but was not reproducible across replicates (Fig. 28D).

3.4 Pull-down of A3G mRNA using a complementary, biotinylated oligo and elution by a competitor oligo

3.4.1 Experimental strategy

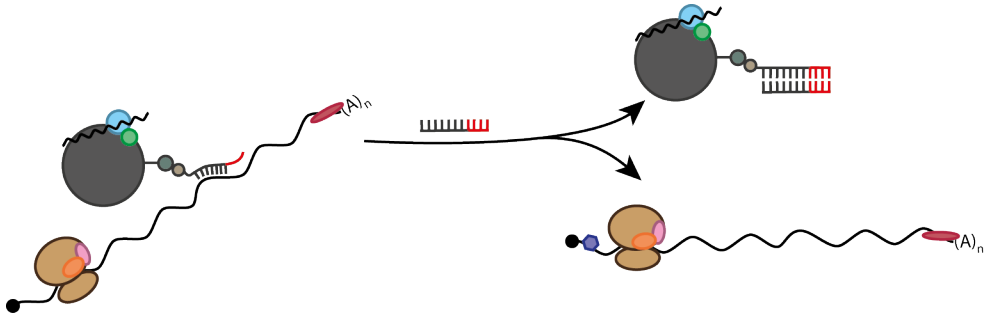


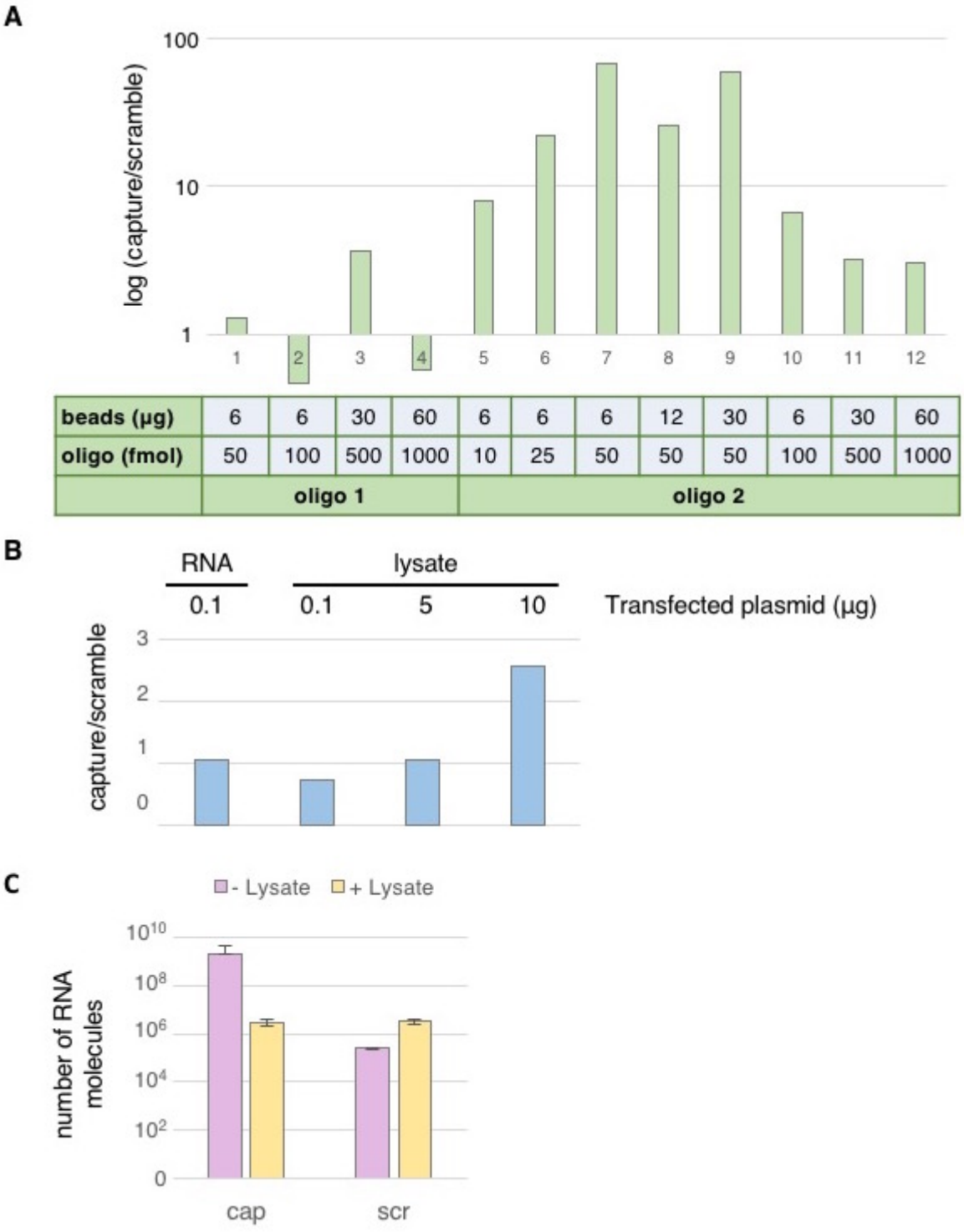
Figure 29: Elution of A3G mRNA using a competitor oligo. *A3G mRNA is pulled down using a complementary oligo which contains a toehold of 10 nts at its 3'-end that doesn't hybridize with A3G mRNA. Addition of a competitor oligonucleotide that hybridizes with the full-length of the capture oligo displaces the A3G mRNA from the beads.*

A3G mRNA has been captured from total cell lysates using a biotinylated capture oligo of a total length of 30 nucleotides, of which only the 20 nucleotides in 5' specifically hybridize with A3G mRNA and the 10 remaining nucleotides remain single-stranded as a so called "toehold". This toehold allows hybridization of a 30 nucleotides competitor oligo to the entire sequence of the capture oligo, thereby displacing the A3G mRNA; the duplex of the 30 complementary nucleotides of the capture and competitor oligos being more stable than the 20 nucleotides

complex of the capture oligo with A3G mRNA (Fig. 29). This allows specific elution of the A3G mRNA from the beads at the end of the pull-down. A scrambled oligo containing the same nucleotide composition as the capture oligo but rearranged in a random order that recognizes none of the cellular RNAs has been used as a negative control (108, 113).

3.4.2 Tests

In order to find the optimal conditions for the quantity of beads and oligos to use, the pull-down was first performed on an *in vitro* transcribed A3G mRNA. In order to be as close as possible to physiological conditions, A3G mRNA has first been quantified in cells using RT-qPCR. For each pull-down sample around 10^7 cells are used 24 h after transfection of 100 ng of the pCMV-A3G plasmid. These cells contain approximately 300 attomol of A3G mRNA, which is about 6 times more abundant in cells than the mRNA of the housekeeping gene GAPDH (data not shown). Given that the qPCR is not efficient at 100 %, this number is taken as a rough estimate. Subsequent tests have been performed with 1 fmol of *in vitro* transcribed A3G mRNA which was pulled down using different amounts of two different capture or scramble oligo pairs as well as different amounts of beads. Overall, the second capture and scramble oligo pair is more efficient in pulling down A3G mRNA than the first pair (Fig. 30A). At 6 µg of beads and 50 fmol of oligos, the best ratios of pull-down are achieved with 66 times more A3G mRNA that is pulled down in the presence of the capture compared to the scramble oligo (Fig. 30A). These conditions were then used to pull-down A3G mRNA from cellular RNAs as well as from total cellular lysates, however the pull-down efficiency was entirely lost (Fig. 30B lanes 1 and 2). This might have been caused by an insufficient amount of A3G mRNA available in cell extracts. Therefore, in an attempt of increasing the amount of A3G mRNA in cells, cells were transfected with an increasing amount of A3G expression plasmid. While an increase of 50 times of the original amount of plasmid showed no effect, a slight increase of pull-down efficiency seemed to be obtained at a 100 time increase of transfected plasmid (Fig. 30B lanes 3 and 4), however this could not be reproduced in subsequent experiments (data not shown). In order to explain the discrepancy between the results obtained by an *in vitro* transcribed RNA in the absence of lysate (Fig. 30A) and the pull-down of an RNA from a cellular extract (Fig. 30B), an experiment was performed where the *in vitro* transcribed A3G mRNA was bound to the beads first and subsequently incubated with or without cellular lysate. In the absence of cellular lysate, on average 826 times more RNA was bound by the capture compared to the scramble oligo (Fig. 30C). In the presence of cellular lysate however, the amount of RNA recovered in the presence of the capture oligo decreased to levels close to the scramble oligo (Fig. 30C). Indeed, it seemed that addition of cellular lysate induced the drop off of A3G mRNA



beads (µg)	6	6	30	60	6	6	6	12	30	6	30	60
oligo (fmol)	50	100	500	1000	10	25	50	50	50	100	500	1000
	oligo 1				**oligo 2**							

Figure 30: Pull-down of A3G mRNA followed by specific elution by a competitor oligo*. **(A)** An in vitro transcribed A3G mRNA was pulled down in the presence of different amounts of beads and oligos. Two different types of capture or corresponding scramble oligos ("oligo 1" - Cap1, Scr1; "oligo 2" - Cap2, Scr2) were used. Eluted A3G mRNA was quantified by RT-qPCR and the ratio of A3G mRNA obtained with the capture compared to the scramble oligo was calculated. **(B)** Pull-down of A3G mRNA from cellular RNAs extracted with TriReagent or from total cellular lysates using the Cap2 or Scr2 oligos and quantified by RT-qPCR. Cells were transfected with different amount of pCMV-A3G expression plasmids as indicated. **(C)** Pull-down of an in vitro transcribed A3G mRNA incubated with or without cellular lysate with the Cap2 or Scr2 oligos followed by quantification by RT-qPCR. n=3.*

from the beads. This strongly suggested that the hybridization of the RNA with the capture oligo was not sufficient to maintain the RNA on the beads. Cellular competitors, as for example RNA binding proteins, with affinity for the same A3G mRNA region might displace the RNA from the capture oligo.

3.5 Pull-down of cellular proteins on an in vitro transcribed, biotinylated A3G mRNA

3.5.1 Experimental strategy

Previous experiments have shown, that the interaction of A3G mRNA with a complementary, biotinylated oligo is not strong enough to reliably maintain the RNA bound to magnetic beads throughout the pull-down protocol. An alternative approach has therefore been designed where A3G mRNA is transcribed *in vitro* in the presence of biotinylated UTP. This leads to an incorporation of approximately 200 biotins per RNA molecule (half of the RNA's uracils are biotinylated) and allows direct binding of the A3G mRNA to magnetic streptavidin-covered beads. This bead-associated RNA is then used as a bait for pull-down of interacting proteins from a cellular lysate.

3.5.2 Retention of biotinylated A3G mRNA on streptavidin-coated beads

The biotinylated A3G mRNA is bound to streptavidin-coated beads and incubated in the presence or absence of cellular lysate. Following several washes, the amount of A3G mRNA retained on the beads is quantified by RT-qPCR. Indeed, the A3G mRNA is sucessfully retained on beads in the presence and absence of lysate (Fig. 31A). Only a very minor loss of material can be observed in the presence of lysate, but this is not significant. The binding of a biotinylated and a non-biotinylated A3G mRNA to streptavidin-coated beads has been compared. No signal is detected in the absence of bait RNA, which shows that there is no contamination of beads with cellular RNAs. a considerable amount of non-biotinylated A3G mRNA is able to bind non-specifically to the beads, however, this amount is significantly lower than the amount of biotinylated A3G mRNA (Fig. 31B). Analysis of the pulled-down proteins by western blot has revealed that both Vif and A3G bind to the A3G mRNA (Fig. 31C).

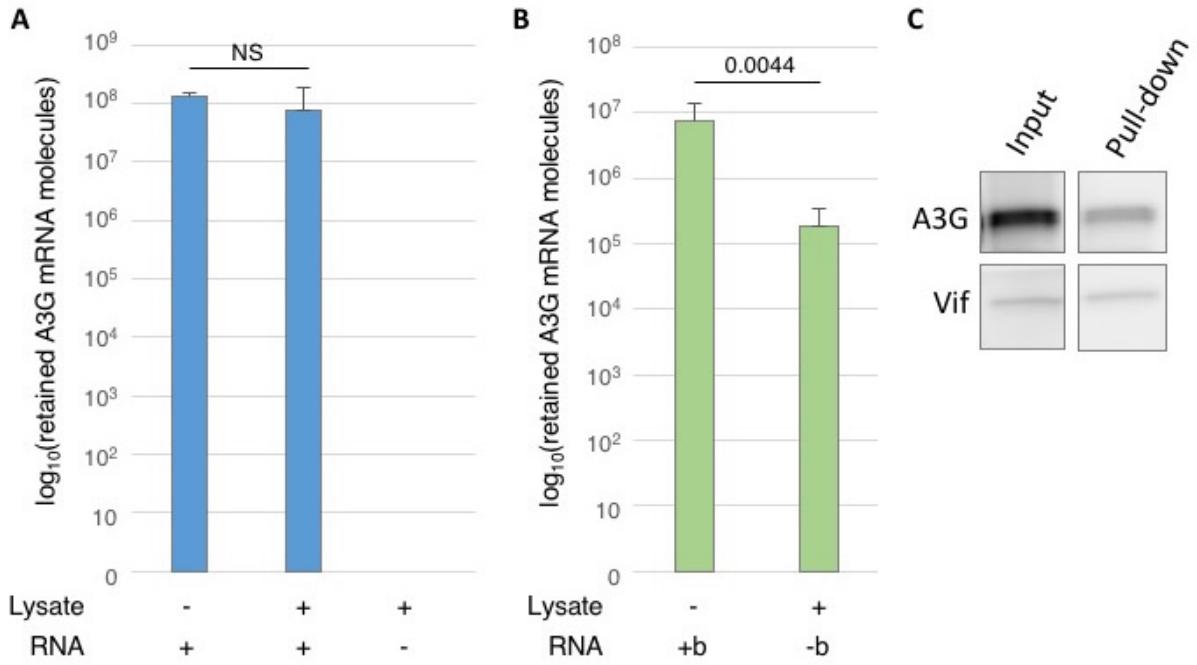


Figure 31: Retention of biotinylated A3G mRNA on streptavidin-coated beads. (A) *An in vitro transcribed, biotinylated A3G mRNA was bound to streptavidin-coated beads and subsequently incubated with cellular lysate ("+ lysate") or with the lysis buffer ("- lysate"). Streptavidin-coated beads without the A3G mRNA ("- RNA"), incubated with lysate were used as a control. A3G mRNA retained on the beads at the end of the protocol was quantified by RT-qPCR. The conditions were compared using a paired t-test (n=3; NS=not significant).* ***(B)*** *Retention of biotinylated ("+b") or non-biotinylated ("-b") A3G mRNA on streptavidin-coated beads in the presence or absence of lysate ("+/- lysate") was measured by RT-qPCR. Streptavidin-coated beads without the A3G mRNA ("- RNA"), incubated with lysate were used as a control. The conditions were compared using a paired t-test and the p-value is indicated (n=6).* ***(C)*** *Pull-down of proteins on a biotinylated A3G mRNA followed by analysis by western blotting.*

4. Discussion

In order to pull-down the full-length A3G mRNA from cellular lysates, several different protocols and various conditions have been tested, but none of them have given satisfactory results. The overall rationale of the approach was to pull-down full-length A3G mRNA from cellular lysates using a set of 1 to 10 complementary oligonucleotides coupled to magnetic beads. The main problem was the non-specific binding on the magnetic beads. While A3G mRNA was mainly used as a control for non-specific binding, some of the obtained results indicate that DNA, other RNAs and various proteins can also non-specifically bind to the beads. Multiple different approaches were tested in order to decrease non-specific binding on the beads, however, none of them showed a significant improvement. Different types of beads, saturated or not, different washing buffers and variations in lysate preparation all resulted in the same problem of non-specific binding. Overall, conditions that decreased the non-specific signal, like for example stringent washing or preclearing of the lysate also lead to a loss of specific A3G mRNA retention on the beads. In conclusion, the used oligos might not have been strong

enough to reliably retain a consistent amount of A3G mRNA on the beads. One possibility to improve this might be to use longer oligos that hybridize to A3G mRNA with more stability. While some protocols succeed in pull-down of mRNAs with oligos of around 20 nts (i.e. (105)) like the ones used in this study, others use oligos of up to 90 nts (140). However, a longer oligo could potentially disrupt secondary interactions in the A3G mRNA and it could also directly or indirectly disturb protein interactants. Moreover, it is not sure whether a longer oligo would be able to access A3G mRNA in the cellular lysate, as it seems to be highly complexed and even in some cases inaccessible for the small oligos used in our study.

Several different protocols for specific elution of A3G mRNA from the beads were also tested and failed to improve specificity.

The elution of desthiobiotinylated oligos from beads was not only poorly efficient even in the presence of high biotin concentrations, but moreover A3G mRNA also eluted even when it was not bound by the oligos. One possibility is that free biotin induced elution of non-specifically bound RNAs. However in this case, saturation of beads with free biotin before incubation with the RNA should have decreased non-specific binding, which was not the case. The other possibility is that the non-specific binding is very transitory and dynamic, with molecules detaching and reattaching constantly. In this case it might be interesting to increase the number of low-stringency washes in order to eliminate as much non-specific material as possible.

Elution of A3G mRNA using RNase H seemed to work rather nicely at first, but only when an *in vitro* transcribed RNA was used. The pull-down of A3G mRNA from cellular RNAs or from total lysates not only recovered a lot less material, but there was also a major problem of reproducibility. Indeed, this problem did not only concern the amount of eluted RNA, but also the amount of RNA bound on the beads across samples, which made it very difficult to compare replicates and generated high error bars. Moreover preliminary tests have shown that the RNase H digestion was not complete. This might have hampered elution efficiency, as one given RNA could be bound to beads by multiple oligos and if not all of them are digested, elution can't take place.

Finally, elution of A3G mRNA using a competitor oligo allowed recovery of significant quantities of an *in vitro* transcribed A3G mRNA. However, this was not at all reproducible with A3G mRNA expressed in cells. Indeed, even addition of lysate to *in vitro* transcribed RNA which had already previously been bound to beads lead to a complete loss of the specifically bound RNA. This suggests that interactants of A3G mRNA in the lysate might compete with the binding site of the capture oligo. To solve this problem, it might be possible to try another capture oligo.

Indeed, the efficiency of oligos seems to be very variable as the comparison of two capture oligos has shown. It might be necessary to test a set of oligos tiled all along the A3G mRNA to find those that bind to regions which are accessible both *in vitro* and in cellular lysates and are not competitively bound by other molecules present in cells.

In conclusion, despite all the different conditions tested, it was not possible to pull-down A3G mRNA directly from cellular lysates. The reason might be that A3G mRNA is too strongly complexed with proteins or other RNAs in the cell and might not be accessible for complementary oligos. Moreover, hybridization of the 20 nts oligos to A3G mRNA might not have been strong enough. A possibility to solve this problem would be to use longer oligos or peptide nucleic acid oligos (PNAs) which form stronger interactions with their targets (23, 231, 239). It is also possible that the oligos used here did not target the right regions of A3G mRNA. Some regions of the A3G mRNA might be more solicited by cellular interactants than others and certain regions might be more accessible for oligos. Nevetheless, 10 carefully selected oligos have been tested at the beginning and it is surprising that none of them seemed to be efficient enough to mediate pull-down of A3G mRNA.

Finally, an *in vitro* transcribed A3G mRNA has been used which incorporates biotinylated nucleotides in its sequence. This mRNA binds strongly to beads and is not lost even upon addition of cellular lysate. As binding of biotinylated A3G mRNA to beads was efficient and reproducible it is expected to reliably allow pull-down of A3G mRNA-interacting proteins. This protocol has therefore been used for the study of the A3G mRNA interactome in the presence and absence of Vif. While an effort was made to use conditions that reproduce as much as possible physiological conditions (gentle cell lysis and incubation with A3G mRNA in a buffer otherwise used for *in vitro* translation), the protocol is still a little less physiological than what was initially aimed for. The pull-down of A3G mRNA directly from cells would have allowed for the RNA to associate with its interactants inside the cell, while in the final protocol interactions are formed *in vitro* and on a modified RNA. The incorporated biotins could indeed perturb certain interactions, therefore it might be interesting for further experiments to decrease the amount of incorporated biotins gradually in order to find a threshold of a minimal number of incorporated biotins that still allows reliable binding to the beads. The advantage of the final protocol is that it can quickly be adapted and different mutants of A3G mRNA can easily be tested. Moreover, it is more reproducible and doesn't rely on expression efficiency of A3G mRNA in the cells, which indeed seemed to be rather low and can vary depending on the cell batch and age. The protocol can even be easily adapted to other RNAs, like for example

mRNAs of other APOBECs, without the tedious search for complementary oligos that are sufficiently specific.

Identification of cellular factors implicated in APOBEC3G translational inhibition by the HIV-1 Vif protein

VI Identification of cellular factors implicated in APOBEC3G translational inhibition by the HIV-1 Vif protein

Vif inhibits A3G translation by a yet unknown mechanism, that is dependent on the 5'-UTR of A3G mRNA. Our hypothesis is that Vif might interact with cellular factors that contribute to translational regulation of A3G. In order to study the impact of Vif on A3G mRNA-associated proteins I have used a protocol that I have developed during my thesis to pull-down proteins on an *in vitro* transcribed, biotinylated A3G mRNA, followed by their identification by mass spectrometry. This study has allowed us to characterize the proteic interactome of the full-length A3G mRNA in the presence and absence of Vif, the A3G uORF and the A3G protein. This has led to the identification of several cellular proteins that are modulated by Vif and therefore might play a role in Vif-mediated translational inhibition of A3G. Our preliminary results are presented in the following article. Further work is needed to confirm certain results and to validate the role of the obtained candidates in Vif-mediated translation regulation of A3G before submission of the article for publication.

Identification of cellular factors implicated in APOBEC3G translational inhibition by the HIV-1 Vif protein

Tanja Seissler[1], Cédric Verriez[1], Philippe Hammann[2], Béatrice Chane-Woon-Ming[1], Roland Marquet[1], Jean-Christophe Paillart[1]*

[1]Université de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, UPR 9002, F-67000 Strasbourg, France

[2]Plateforme protéomique Strasbourg Esplanade, CNRS, Université de Strasbourg, F-67000, France

*To whom correspondence should be addressed:

Tel: (+33) (0)3 88 41 70 35; Fax: (+33) (0)3 88 60 22 18

E-mail: jc.paillart@ibmc-cnrs.unistra.fr

Abstract

APOBEC3G (A3G) is one of the main restriction factors of HIV infection, which significantly reduces infectivity of virions. HIV counteracts this restriction factor using three different mechanisms. While Vif-mediated ubiquitination and degradation of A3G has been well characterized over the years, mechanisms aiming at reducing A3G transcription and translation of A3G mRNA are still poorly understood. With the objective to gain a better understanding of the translational inhibition and to identify potential cellular cofactors of Vif involved in this process, we pulled-down A3G mRNA-associated protein complexes from cellular lysates in the presence and absence of Vif and A3G protein. Analysis of these protein complexes by mass spectrometry allowed us first to identify a core interactome of A3G mRNA, which remains stable in all different conditions used in this study. We also identified a considerable number of proteins whose association with the A3G mRNA seems to be dependent on the presence of Vif, the A3G protein and a uORF in the 5'-UTR of A3G mRNA. Amongst these proteins, we identified RENT1 and CHIP and validated their interaction with Vif by co-immunoprecipitation. Overall, the identified proteins suggest a potential role of P-body and stress-granule components as well as the cytoskeleton in Vif-mediated translational inhibition of A3G. Moreover, most of the identified proteins are phosphoproteins, suggesting a potential regulation through phosphorylation.

Introduction

Restriction factors are IFN-responsive genes found amongst the first effectors of the intrinsic immune response against viral infections. Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G (APOBEC3G or A3G) is one of the best studied restriction factors, with an inhibitory effect on different retroviruses, retrotransposons and HBV replication. A3G restricts HIV-1 by C to U deamination during the reverse transcription of the viral genome. This leads to hypermutation of the provirus, resulting in production of truncated viral proteins as well as defects of the viral genomic RNA and ultimately arrest of the viral life cycle (18, 34, 45, 58). Restriction factors are often counteracted by viral proteins to allow viral replication. HIV-1 expresses Vif, which counteracts A3G. Vif-induced ubiquitination and subsequent degradation of A3G in the proteasome has been extensively studied. Vif interacts with A3G and recruits an E3-ubiquitin ligase complex composed of Cul5, EloB, EloC, Rbx2 and CBF-β. Vif thereby replaces the substrate recognition component of the E3 ubiquitin ligase, leading to ubiquitination of A3G. Ubiquitinated A3G is then recognized and degraded by the 26S proteasome (46, 57). Indeed, hijacking of the cellular ubiquitin-proteasome system is a common feature used by several viruses and especially HIV to mediate counter-defense to cellular restriction mechanisms (44). It has recently been shown that recruitment of the transcriptional cofactor CBF-β in the E3-ubiquitin ligase complex by Vif sequesters it away from the RUNX transcription factor and leads to a decrease in A3G transcription (3, 27). The third mechanism for A3G counteraction by Vif is inhibition of A3G translation, and was discovered early on but has been only poorly understood until recently (17, 37, 50). It has been shown that Vif binds to A3G mRNA and especially to the 5'- and 3'-UTRs (37). The 5'-UTR, but not the 3'-UTR has been shown to be necessary and sufficient to allow A3G translational inhibition by Vif (17). A very recent study identified a small uORF in the 5'-UTR of A3G mRNA and Vif-mediated translational inhibition has been shown to be dependent on this uORF. The uORF is also conserved in A3F, whose translation can also be inhibited by Vif (Libre, Seissler et al., in preparation). uORFs are present in almost 50 % of cellular mRNAs and constitute a negative regulatory element for the associated main ORFs (mORFs) which have to be translated by leaky scanning, re-initiation after translation of the uORF or through recruitment of the ribosome by an internal ribosome entry site (IRES) (6, 9). In the case of A3G, translation occurs through a mix of leaky scanning and re-initiation (Libre, Seissler et al., in preparation). uORFs allow regulation of mORF

expression in response to different stimuli, like for example stress or starvation, and can notably be regulated by different RNA binding proteins (RBPs). The sex-lethal (SXL) protein for example binds to a region in the 5'-UTR of the male-specific lethal 2 (msl-2) mRNA, downstream of a uORF and thereby decreases translation of msl-2 in favour of the uORF (8, 13, 36). In plants, translation of uORF containing transcripts can also be regulated by RBPs, like it is the case for RL24-containing ribosomes, which more efficiently re-initiate after uORF translation and might therefore be preferentially associated with translation of uORF-containing genes (38, 41). In human cells, the proteins DENR and MCT-1 have been associated with stimulation of re-initiation and their recruitment to particular uORF-containing mRNAs might increase their translation (32, 43).

The mechanism of Vif-mediated translation inhibition of A3G is yet unknown. As Vif binds to the 5'-UTR of A3G mRNA, it is possible that Vif interacts with the present protein complexes. Indeed, Vif could interact with different RBPs on the 5'-UTR of A3G mRNA and modulate their binding. A3G translational inhibition by Vif could for example be explained by drop-off of RBPs that stimulate translation or by the recruitment of translation inhibitory RBPs. In order to study the effect of Vif on A3G mRNA-associated protein complexes, we performed pull-down of proteins from cellular lysates on biotinylated A3G mRNA bound to magnetic beads, followed by their identification by mass spectrometry. This allowed us to identify over 400 A3G mRNA-associated proteins in different conditions of which at least 107 seem to be regulated positively or negatively by Vif. Interaction of some of these proteins with Vif has been validated by co-immunoprecipitation and their role in A3G translation has been studied by RNA silencing.

Material and methods

Plasmids. The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: HIV-1 NL4-3 Vif expression vector (pcDNA-hVif) from Dr. Stephan Bour and Dr. Klaus Strebel. For co-immunoprecipitation, a FLAG-tag was added at the C-terminus of hVif, resulting in the pcDNA-Vif-FLAG construct. The empty pcDNA vector was used as a negative control (pcDNA-Ø). The pCMV-T7-A3G-WT and pCMV-T7-A3G-ΔuORF contain the full-length A3G cDNA (both UTRs and the CDS) under the control of the T7 promoter for *in vitro* transcription.

The pCMV-A3G plasmid contains the full-length A3G cDNA (both UTRs and the CDS) under the control of the CMV promoter. The pCMV-A3G-ΔUTR contains the CDS of A3G without UTRs under the control of the CMV promoter.

***In vitro* transcription of biotinylated, capped, poly-adenylated A3G mRNA.** For transcription of a capped, poly-adenylated and biotinylated transcript, 1 µg of linearized pCMV-T7-A3G-WT or pCMV-T7-A3G-ΔuORF was used with the HiScribe T7 ARCA mRNA kit with tailing (NEB) according to manufacturer's instructions. The transcription mix was supplemented with 1.25 mM Bio-16-UTP (ThermoFisher) and incubated at 37 °C overnight. The plasmid was digested with 4 U DNase I (NEB) at 37 °C for 15 min followed by a tailing reaction at 37 °C for 30 min. The transcripts were precipitated by 2.5 M LiCl (NEB) according to manufacturer's instructions.

Preparation of cellular lysates using nitrogen cavitation. HEK293T cells were seeded into 15 cm dishes at $5.8*10^6$ cells per dish in a final volume of 20 ml DMEM medium (Gibco; supplemented with 10 % fetal calf serum (Pan Biotech), 100 U/ml penicillin and 100 µg/ml streptomycin (Gibco)). Cells were incubated at 37 °C and 5 % CO_2. Cells were transfected with expression plasmids around 16-18 h after their seeding. One µg of pCMV-A3G-ΔUTR and 8 µg of pcDNA-Vif were used per dish. The total amount of plasmid is adjusted to 9 µg using the pcDNA-Ø. The plasmid preparation was gently mixed with 27 µl of XtremeGene9 (Roche) and 373 µl of Opti-MEM (Gibco) and incubated at room temperature for 10-15 min, then pipetted onto the cells. Cells were detached in cold dPBS (Gibco) using cell scrapers 24 h after transfection. They were centrifuged at 300 xg for 5 min. The cell pellet was weighed and then resuspended in lysis buffer (20 mM HEPES, 100 mM KAc, 2 mM MgAc, 1 mM DTT, 100 U/ml RNasin; 1 x Halt Protease Inhibitor Cocktail (ThermoScientific)) at 2.5 ml per g of cells. Cells were then introduced into a nitrogen cell bomb (Parr) and incubated at 500 psi for 30 min at 4 °C. The resulting cell lysate was spun down twice at 1,000 xg for 5 min at 4 °C and the supernatant was recovered and stored at -80 °C.

Pull-down of cellular proteins on biotinylated A3G mRNA. SpeedBeads Magnetic streptavidin coated particles (GE Healthcare) were washed with wash buffer (20 mM HEPES, 100 mM KCl, 10 mM $MgCl_2$, 0.01 % NP40) and 40 µg of beads were incubated with 1 pmol of biotinylated A3G mRNA in wash buffer for 1 h at room temperature

under rotation. Beads were then washed with wash buffer, resuspended in cellular lysate (per sample, an equivalent of 1.5 cell dishes, diluted at 1/2 in lysis buffer was used) and incubated for 3 h at 4 °C under rotation. Beads were washed twice with wash buffer, once with no-detergent wash buffer (20 mM HEPES, 100 mM KCl, 10 mM $MgCl_2$) and transferred into a new tube. Beads were then resuspended in 15 µl H_20. Two µl were used for RT-qPCR and the rest was analysed by mass spectrometry.

RT-qPCR of A3G mRNA. The beads volume was adjusted to 16 µl with H_2O. Four µl of iScript Reverse Transcription Supermix for RT-qPCR (BioRad) were added and the mixture was incubated at 25 °C for 5 min, 42 °C for 30 min and 85 °C for 5 min. The RT-reaction was mixed with 0.2 µM of A3G sense (5'-TTC TCC AGA ATC AGG AAA AC-3') and antisense (5'-GTG TCT GTG ATC AGC TGG AG-3') primers, 1x Taq buffer, 0.25 mM of each dNTP, 2.5 U of Taq polymerase (homemade) and 1.25 µl of EvaGreen Dye (Biotium). The mixture was incubated at 95 °C for 3 min, then 40 cycles of 30 s at 95 °C, 30 s at 61 °C and 1 min at 72 °C were performed and the amount of DNA was monitored at each cycle using Evagreen fluorescence measurement. In parallel, a titration curve was generated using between 10^9 and 10^3 copies of the pCMV-A3G.

Mass Spectrometry. The beads suspension was mixed with 50 µl of elution buffer (Miltenyi) and incubated at 95 °C for 10 min. The beads were then discarded. Each sample was precipitated with 0.1 M ammonium acetate in 100% methanol, and proteins were resuspended in 50 mM ammonium bicarbonate. After a reduction-alkylation step (dithiothreitol 5 mM – iodoacetamide 10 mM), proteins were digested overnight with 1:25 (w/w) sequencing-grade porcine trypsin (Promega). One fifth of the peptide mixture was analyzed by nanoLC-MS/MS in an Easy-nanoLC-1000 system coupled to a Q-Exactive Plus mass spectrometer (ThermoFisher). Each sample was separated with an analytical C18 column (75 µm ID × 25 cm nanoViper, 3 µm Acclaim PepMap; ThermoFisher) with a 160 min 300 nL/min gradient of acetonitrile. The obtained data was searched against the Swissprot database with human taxonomy using the Mascot algorithm (version 2.5, Matrix Science). Mascot files were then imported into Proline v1.4 package (http://proline.profiproteomics.fr/) for post-processing. Proteins were validated with 1% FDR and the total number of MS/MS fragmentation spectra (Spectral Count) was used to quantify each protein in the

different samples.

Bioinformatic analyses. Mass spectrometry data obtained for each sample, including the proteins identified and their associated spectral counts (SpC), were stored in a local MongoDB database and several pairwise comparisons were then performed through a Shiny Application built upon the msmsEDA (16) and msmsTests (15) R/Bioconductor packages. The latter were respectively used to conduct exploratory data analyses of LC-MS/MS data by spectral counts and differential expression tests using a negative-binomial regression model. The p-values were adjusted with false discovery rate (FDR) control by the Benjamini-Hochberg method and 3 parameters were used (adjusted p-value < 0.1 or 0.05, a minimum of 2 SpC in the most abundant condition, and a minimum fold change of 2) to define differentially expressed proteins. Proteins significantly enriched ($p < 0.1$) in samples compared to the corresponding negative controls were analyzed using functional annotation clustering with medium classification stringency against the background of the total human proteome on the Database for Annotation, Visualization and Integrated Discovery (DAVID; https://david.ncifcrf.gov) (21). The Venn diagram was generated using Venny 2.1 (https://bioinfogp.cnb.csic.es/tools/venny) (39). String diagrams were generated using STRING 11.0 (https://string-db.org/) (52).

Co-immunoprecipitation without cross-linking. HEK293T cells were seeded in 10 cm dishes at 2.5 million cells per dish. After 16-18 h of seeding, cells were transfected with 0.4 µg of pCMV-A3G and 4 µg of pcDNA-Ø or pcDNA-Vif-FLAG. For the transfection, the plasmid preparation is gently mixed with 12 µl of XtremeGene9 (Roche) and 188 µl of Opti-MEM (Gibco), incubated at room temperature for 10-15 min, then pipetted onto the cells. Ten h after transfection, cells were treated with 25 µM ALLN. Twenty-four h after transfection, cells were washed once with dPBS (Gibco), then 250 µl of RIPA 1x (1x PBS, 1 % NP40, 0.5 % Na-DOC, 0.05 % SDS, 1x Halt Protease Inhibitor Cocktail (ThermoScientific)) were added and cells were incubated at 4 °C for 10 min. Cells were centrifuged at 18,188 x g for 1 h at 4 °C, then the supernatant was recovered.

900 µg protein G Dynabeads (Invitrogen) per sample were washed 3 times with RIPA 1x, then resuspended in 30 µl RIPA 1x. Two µg of anti-FLAG M2 mouse monoclonal antibody (F1804, Sigma) were added and incubated for 2 h at room temperature under

rotation. Then beads were washed with RIPA 1x, resuspended in the cellular lysate (equivalent of 1.5 culture dishes per sample) and incubated at 4 °C for 2 h under rotation. Beads were washed 5 times with RIPA 1x and bound proteins were eluted in WB buffer (1x NuPAGE LDS Sample Buffer (Invitrogen), 1x NuPAGE Sample Reducing Agent (Invitrogen)) at 70 °C for 10 min. The supernatant was recovered and analysed by western blotting.

Co-immunoprecipitation with crosslinking. HEK293T cells were seeded into 10 cm dishes at 2.5 million cells per dish in a final volume of 5 ml DMEM medium (Gibco; supplemented with 10% fetal calf serum (Pan Biotech), 100 U/ml penicillin and 100 µg/ml streptomycin (Gibco)). Cells were incubated at 37 °C and 5 % CO_2. Cells were transfected with expression plasmids around 16-18 h after their seeding with 0.4 µg of pCMV-A3G and 4 µg of pcDNA-Vif-FLAG per dish. The total amount of plasmid is adjusted to 4.4 µg using the pcDNA-Ø. The plasmid preparation is gently mixed with 12 µl of XtremeGene9 (Roche) and 188 µl of Opti-MEM (Gibco) and incubated at room temperature for 10-15 min, then pipetted onto the cells. 10 h after transfection, cells were treated with 25 µM ALLN. 24 h after transfection, cells were washed twice with cold PBSCM buffer (dPBS (Gibco) supplemented with 0.1 mM $CaCl_2$ and 1 mM $MgCl_2$), and then incubated for 6 h at 4 °C in PBSCM supplemented with 1 mM dithiobis-succinimidyl-propionate (DSP, ThermoFisher). PBSCM-DSP was discarded and cells were incubated for 15 min at 4 °C in PBSCM supplemented with 20 mM Tris (pH 7.5). Cells were washed twice with cold PBSCM buffer and then incubated for 30 min at 4 °C in lysis buffer (50 mM Tris (pH 7.5), 300 mM NaCl, 5 mM EDTA, 1 % Triton x100, 1 x HALT Protease Inhibitor Cocktail (ThermoFisher)). Cells were spun down at 17,000 xg for 15 min at 4 °C and the supernatant was recovered. 60 µl of Pierce anti-DYKDDDDK Magnetic Agarose Beads (Invitrogen) were used per sample. Beads were washed 3 times with PBS, resuspended with the cell lysate and incubated over night at 4 °C under rotation. Beads were washed 5 times with wash buffer (50 mM Tris (pH 7.5), 300 mM NaCl, 5 mM EDTA, 0.1 % Triton x100) and once with PBS, then incubated at 37 °C for 15 min in 2x LDS Sample Buffer (NuPage) supplemented with 50 mM DTT. The supernatant was recovered and analysed by western blotting.

RNA silencing. On day 1, HEK293T cells were seeded into 6-well plates at $1*10^5$ cells per well in 1 ml DMEM. On day 2, cells were treated with a final concentration of 10

nM siRNAs. A TriFECTa RNAi Kit (Integrated DNA Technologies), containing the following siRNAs has been used: siRNA1 5'-AGGUUAUUGACGAUUCAUCUCUGA-3' and siRNA2 5'-CCCAACUUGGCUAUGAAGGUTA-3'. siRNAs were supplied as duplexes with a 2 nucleotides overhang at the 3' extremity of the complementary strand. siRNAs were first diluted in a final volume of 170 µl OptiMEM (Gibco), then mixed with 2 µl oligofectamin diluted in a final volume of 30 µl OptiMEM. This mix was incubated for 20 min at room temperature, then added to the cells in DMEM medium (non-supplemented). Around 4-6 h after siRNA treatment, 10 % fetal calf serum (Pan Biotech) was added. On day 4, cells were transfected with 0.1 µg of pCMV-A3G and 1 µg of pcDNA-Vif per dish. The total amount of plasmid was adjusted to 2 µg using the pcDNA-Ø. The plasmid preparation was gently mixed with 6 µl of XtremeGene9 (Roche) and 94 µl of Opti-MEM (Gibco) and incubated at room temperature for 10-15 min, then pipetted onto the cells. 10 h after transfection, cells were treated with 25 µM ALLN. On day 5, cells were washed with cold dPBS, then incubated in RIPA 1x (1x PBS, 1 % NP40, 0.5 % Na-DOC, 0.05 % SDS, 1x Halt Protease Inhibitor Cocktail (ThermoScientific)) for 10 min at 4 °C. Cells were centrifuged at 18,188 x g for 1 h at 4 °C, then the supernatant was recovered.

Name	Reference	Company	Dilution	Incubation
Polyclonal Rabbit @A3G	9968	NIH AIDS Reagent Program	1:10,000	ON
Polyclonal Rabbit @ARG1	E-AB-30551	Elabscience	1:1,000	ON
Polyclonal Rabbit @CALL5	GTX119159	GeneTex	1:2,000	ON
Polyclonal Rabbit @CHIP	GTX122827	GeneTex	1:10,000	ON
Polyclonal Rabbit @CLAP2	A302-155A-T	Bethyl	1:1,000	ON
Polyclonal Rabbit @DSG1	E-AB-31242	Elabscience	1:500	ON
HRP - Mouse @rabbit-GAPDH	MCA4739P	BioRad	1:10,000	ON
Monoclonal Mouse @LRC40	sc-515101	Santa Cruz Biotechnology	1:1,000	ON
Polyclonal Rabbit @MYPT1	E-AB-32148	Elabscience	1:2,000	ON
Polyclonal Rabbit @RENT1	GTX112303	GeneTex	1:2,000	ON
Monoclonal Mouse @Vif	319	NIH AIDS Reagent Program	1:10,000	ON
HRP - Goat @Rabbit-IgG	1706515	BioRad	1:10,000	1-2 h
HRP - Goat @Mouse-IgG	1706516	BioRad	1:10,000	1-2 h

Table 8: List of antibodies. *Antibody names, references and company are indicated. @GAPDH antibody as well as secondary antibodies are conjugated to horseradish peroxidase (HRP, indicated in the name). Primary antibodies are diluted in WB blocking solution and secondary antibodies in TNT at the indicated dilutions and incubated for the indicated time at 4 °C. The following reagents were obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: Anti-Human APOBEC3G Polyclonal from Dr. Warner C. Greene (50); HIV-1 Vif Monoclonal Antibody (#319) from Dr. Michael H. Malim (47, 48).*

Western blotting. The protein concentration of samples was determined by a Bradford assay using Bradford Assay Dye (BioRad) according to manufacturer's instructions. 80 µg of protein were used for siRNA samples and 20 µg for Co-IP input were denatured using NuPAGE LDS Sample Buffer and Reducing Agent (Invitrogen) at 70 °C for 10 min. For Co-IP samples, total proteins were used. Samples were migrated on a 4-15 % Criterion TGX Precast Midi Protein Gel (BioRad) in 1x TGS (25 mM Tris, 200 mM Glycine, 1 % SDS) for 35 min at 200 V. Then proteins were transferred onto a PVDF membrane using a Midi PVDF Transfer Pack (BioRad) at 25 V and 2.5 A for 10 min in a Trans-Blot Turbo (BioRad). The membrane was then incubated in WB blocking solution (50 mM Tris (pH 7.5), 150 mM NaCl, 1 % Triton x100, 5 % milk (Regilait)) for 1 h. Primary and secondary antibodies were added as described in table 1. After the primary antibody, the membrane was washed twice with TNT (50 mM Tris (pH 7.5), 150 mM NaCl, 1 % Triton x100). After the secondary antibody, the membrane was washed twice with TNT and once with TN (50 mM Tris (pH 7.5), 150 mM NaCl). One ml of Peroxide Solution and 1 ml Luminol Enhancer Solution (Amersham ECL Prime Western Blotting Detection Reagent, GE Healthcare) were mixed and added onto the membrane. Chemiluminescent signal was revealed using a Chemidoc (BioRad). For detection of multiple proteins on the same membrane, membranes were stripped for 25 min at room temperature under rotation with Antibody Stripping Buffer (Geba), then washed thoroughly with demineralized water.

Results

Determining the interactome of A3G mRNA. Cellular proteins have been pulled down on an *in-vitro* transcribed, capped, poly-adenylated and biotinylated full-length A3G mRNA coupled to streptavidin-covered magnetic beads, followed by identification by mass spectrometry. Retention of A3G mRNA at the end of the pull-down protocol has been verified by RT-qPCR, which showed a uniform retention in all samples (data not shown). Beads in the absence of A3G mRNA have been used as a negative control to evaluate non-specific binding of cellular proteins. Vif and A3G have been expressed in cells in order to study their effect on protein complexes binding to A3G mRNA. This allowed us to define the interactome of full-length A3G mRNA. 410 different proteins have been identified to significantly bind to A3G mRNA in the different conditions with a p-value under 5 % and 447 with a p-value under 10 % (Fig. 1A). The approach used

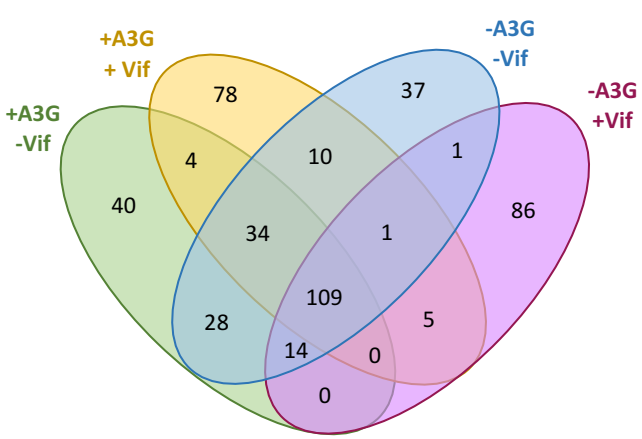
A

	+A3G -Vif		+A3G +Vif		-A3G -Vif		-A3G +Vif	
p <	0.1	0.05	0.1	0.05	0.1	0.05	0.1	0.05
	229	215	241	218	234	217	216	200

B

+A3G -Vif		+A3G +Vif		-A3G -Vif		-A3G +Vif	
Cluster	Enrichment	Cluster	Enrichment	Cluster	Enrichment	Cluster	Enrichment
Nucleic Acid Binding	38.17	Nucleic Acid Binding	39.67	Nucleic Acid Binding	42.38	Nucleic Acid Binding	29.55
Splicing	28.51	Keratin	22.51	Splicing	31.21	Cell-Cell Adhesion	24.88
RNA degradation	18.28	Splicing	20.39	RNP	19.05	Splicing	13.67
RNP	14.67	RNP	16.36	RNA degradation	15.33	RNA degradation	9.82
Translation initiation	10.33	RNA degradation	11.71	Translation initiation	10.49	RNP	9.39
Double straded RNA binding	7.54	mRNA 3' processing	9.24	mRNA 3' processing	9.19	Translation initiation	6.04
RNA helicase	7.52	Translation initiation	9.18	Zinc finger	6.45	Actin	4.52

C



D

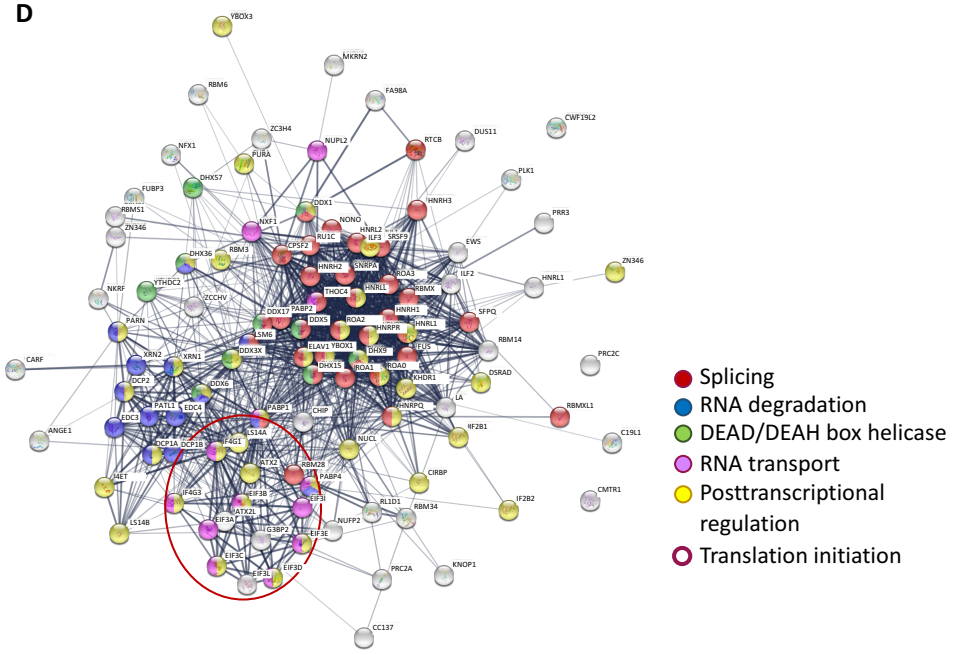


Figure 32: Proteins significantly associated with A3G mRNA in different conditions. *Pull-down of proteins from cell lysates expressing Vif and/or A3G on an in vitro-transcribed, capped, poly-adenylated and biotinylated A3G mRNA.* ***(A)*** *Number of proteins significantly associated with the A3G mRNA compared to the beads only control with a p-value < 0.1 or 0.05.* ***(B)*** *Functional annotation clustering of proteins associated with A3G mRNA in each condition at a p-value < 0.1 using the DAVID webservice. The 7 most enriched clusters are indicated for each condition.* ***(C)*** *Venn diagram of proteins identified with a p-value < 0.1 in the different conditions.* ***(D)*** *String diagram of the 109 proteins that represent the core interactome of A3G mRNA. The thickness of lines indicates the confidence of association of the linked proteins. Proteins involved in some of the most represented functional clusters are colored as indicated.*

here to identify A3G mRNA-binding proteins was very stringent, and we thus decided to use a p-value of 10 % as a threshold for further analyses in order to maximize our chances of identifying relevant cellular proteins. Between 200 and 241 proteins interact with A3G mRNA in any given condition (Fig. 1A). 109 of these proteins (24.4 % of the total interacting proteins) seem to constitute the core interactome of A3G mRNA, which is the same in all 4 conditions, and 158 (35.4 %) of the proteins are common in at least 3 conditions (Fig. 1C). The remaining 64.6 % of proteins are modulated depending on the presence of Vif and A3G protein (Fig. 1C). The cellular interactome of A3G mRNA seems to be composed of many proteins with known nucleic acid and RNA binding properties, as well as proteins containing domains that typically interact with RNAs, like for example zinc fingers (Fig. 1B). Amongst these proteins, we have identified many splicing factors, RNA helicases as well as proteins involved in RNA degradation, 3'-end processing and mRNA translation. Surprisingly, in the presence of Vif, proteins involved in keratinization and cell-cell adhesion appeared amongst the most enriched clusters (Fig. 1B). Amongst the 109 proteins that constitute the core interactome of A3G mRNA, we found proteins implicated in different post-transcriptional processes of gene regulation, including splicing factors, proteins involved in RNA degradation and transport, RNA helicases and translation initiation factors (Fig. 1D).

Effect of A3G protein on the interactome composition. A total of 83 proteins bound to A3G mRNA changed significantly when A3G was expressed in cells (Fig. 2A; Supplementary Table 1). The binding of 61 of these proteins to A3G mRNA seems to be negatively regulated by the addition of A3G protein. Amongst these proteins are many components of the proteasome and proteins implicated in cell cycle control (Fig. 2B). In turn, 22 proteins are upregulated upon addition of A3G protein and amongst these are many factors involved in ubiquitin conjugation as well as several RNA binding proteins involved in RNA degradation, post-transcriptional regulation and RNA metabolism (Fig. 2C). Some of these proteins are also components of RNP granules found in the cytoplasm like P-bodies or stress-granules. Overall, the A3G protein seems to significantly influence the interactome of its own mRNA, which might contribute to regulation of essential processes for its own expression.

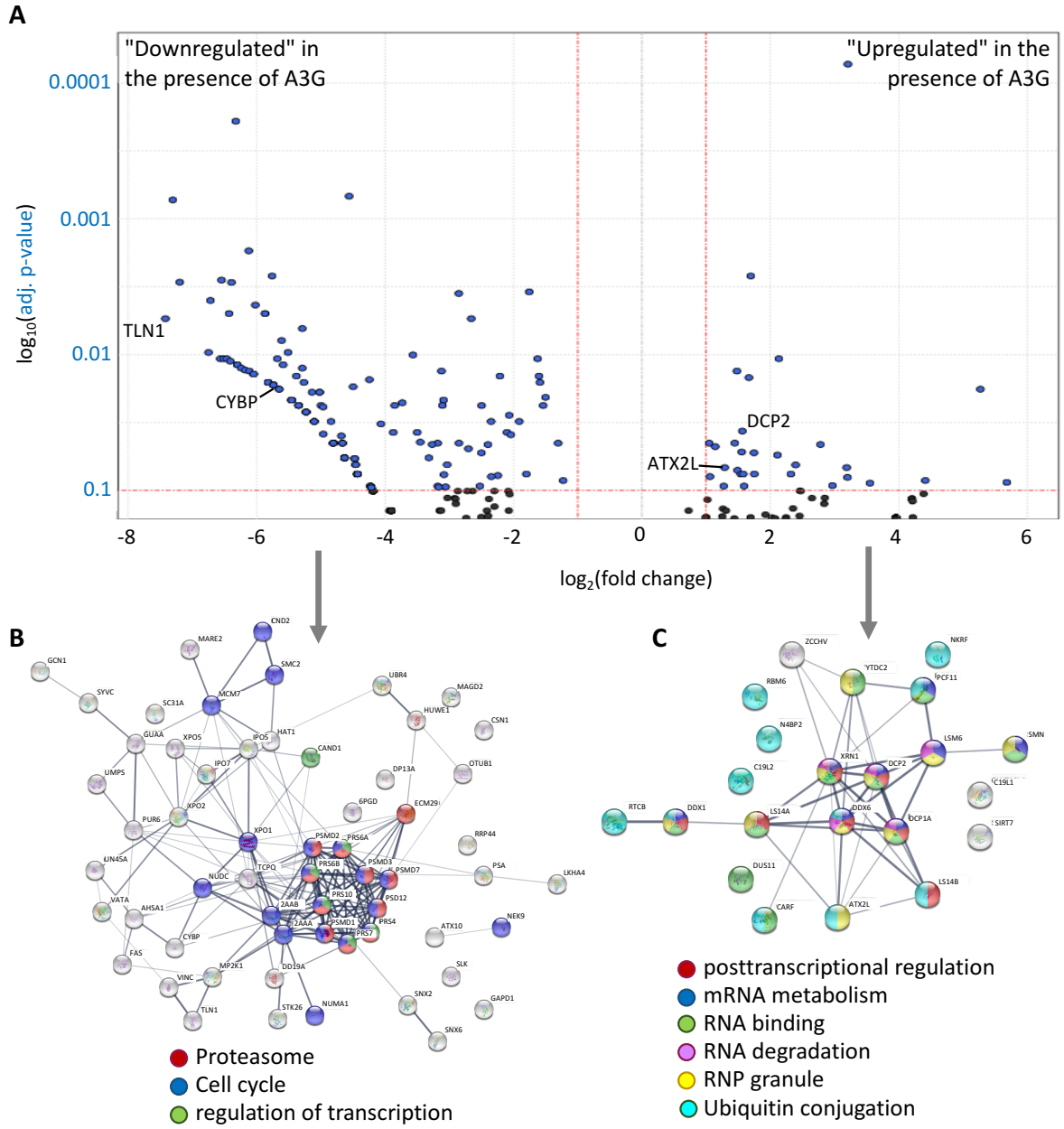


Figure 33: Effect of A3G protein on the interactome of its mRNA. (A) *Differential association of proteins with A3G mRNA in the presence and absence of the A3G protein analyzed by a generalized linear model of a negative-binomial distribution. Proteins with a positive log2(fold change) are upregulated in the presence of A3G, while a negative log2(fold change) indicates downregulation in the presence of A3G. P-values are adjusted using the Bejamini-Hochberg method and the threshold of a p-value of 0.1 is indicated as a red line. Proteins* ***(B)*** *downregulated or* ***(C)*** *upregulated in the presence of A3G are represented as a string diagram. The thickness of lines indicates the confidence of association of the linked proteins. Proteins involved in some of the most represented functional clusters are colored as indicated.*

Effect of Vif on the A3G mRNA interactome. In the absence of A3G protein, Vif seems to recruit 63 proteins to A3G mRNA and exclude 6 proteins from A3G mRNA-associated protein complexes (Fig. 3A; Supplementary table 2). Amongst the recruited proteins, a very large proportion is associated with the actin cytoskeleton and some proteins are implicated in tight junction formation or play a role in cell-cell

communication (Fig. 3C). Amongst the proteins downregulated by Vif, we mainly found proteins implicated in polyadenylation, as well as a couple of known RNA binding proteins (Fig. 3B). In the presence of A3G protein, a total of 65 proteins are upregulated on A3G mRNA in the presence of Vif, while 14 proteins are downregulated (Fig. 4A; Supplementary table 3). Amongst proteins upregulated in the presence of Vif, a very large proportion is associated with keratinization. Some of these proteins also play a role in amino acid biosynthesis and apoptosis. Some serpins have also been identified (Fig. 4C). Interestingly, of the total 147 proteins that are regulated by Vif, 88 are phosphoproteins, which means that they can be regulated through phosphorylation.

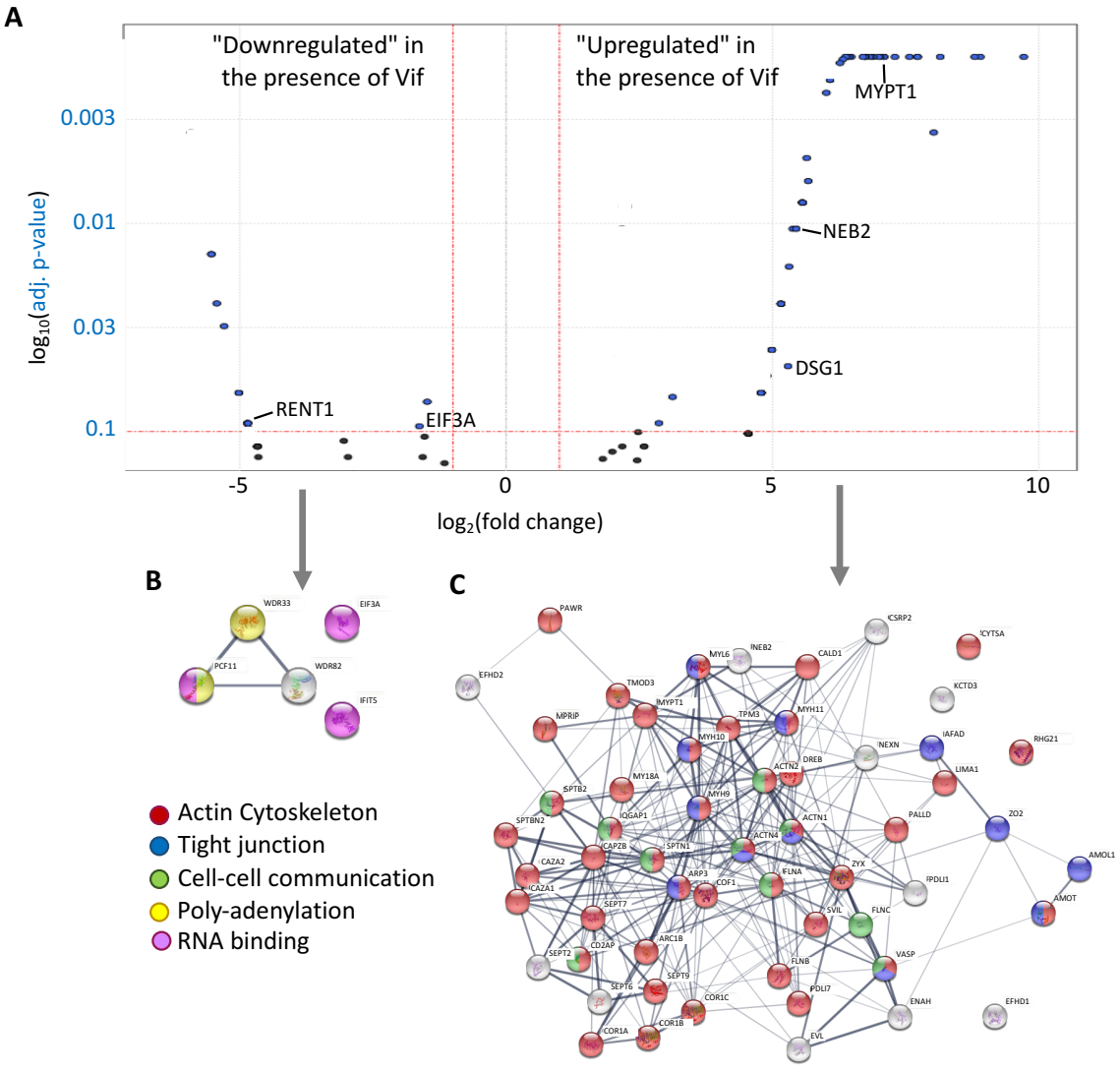


Figure 34: Effect of Vif on A3G mRNA-associated proteins in the absence of A3G protein. ***(A)*** *Differential association of proteins with A3G mRNA in the presence and absence of Vif. Proteins with a positive log_2(fold change) are upregulated in the presence of Vif, while a negative log_2(fold change) indicates downregulation in the presence of Vif. P-values are adjusted using the Bejamini-Hochberg method and the threshold of a p-value of 0.1 is indicated as a red line. Proteins* ***(B)*** *downregulated or* ***(C)*** *upregulated in the presence of Vif are represented as a string diagram. The thickness of lines indicates the confidence of association of the linked proteins. Proteins involved in some of the most represented functional clusters are colored as indicated.*

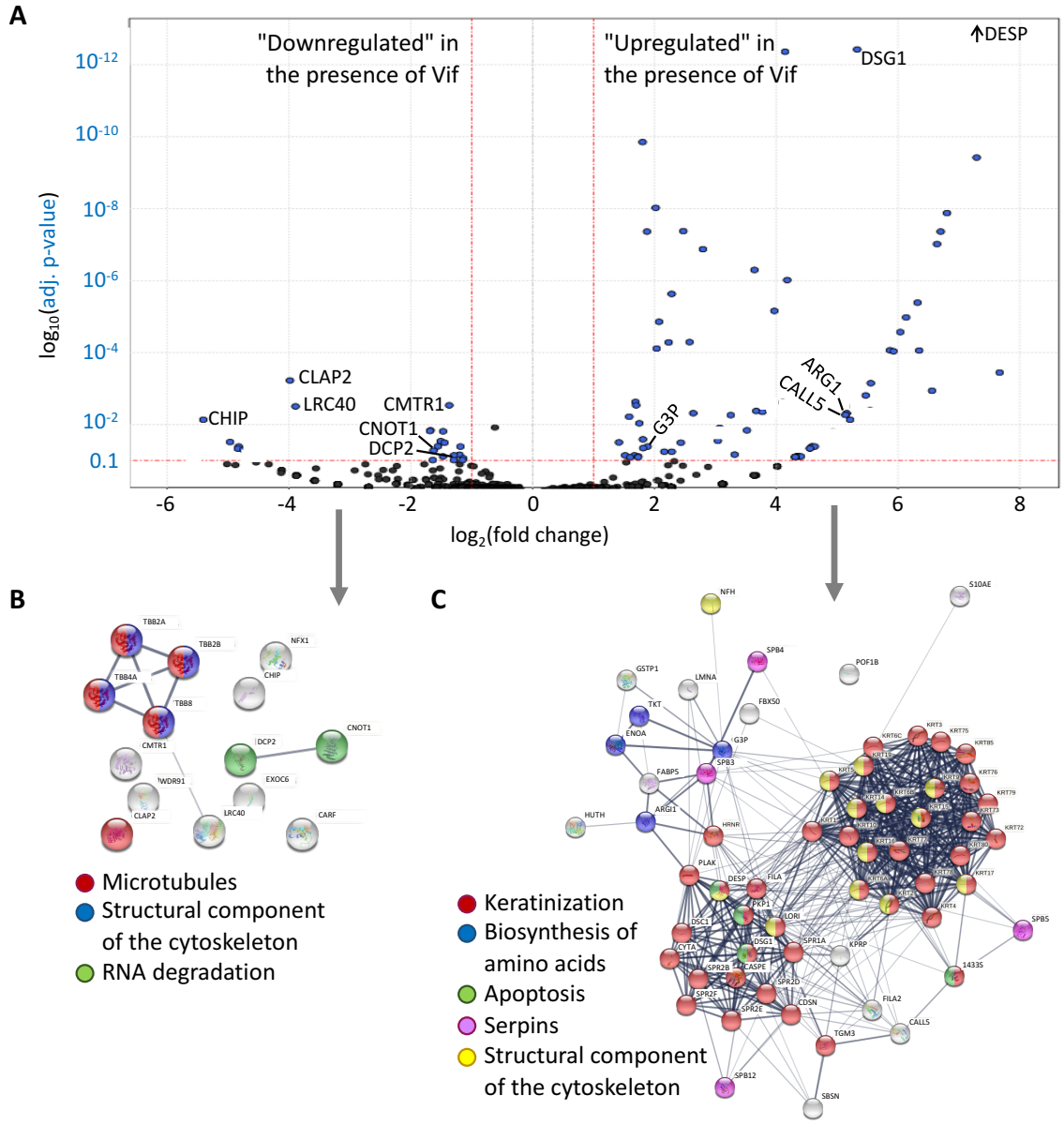
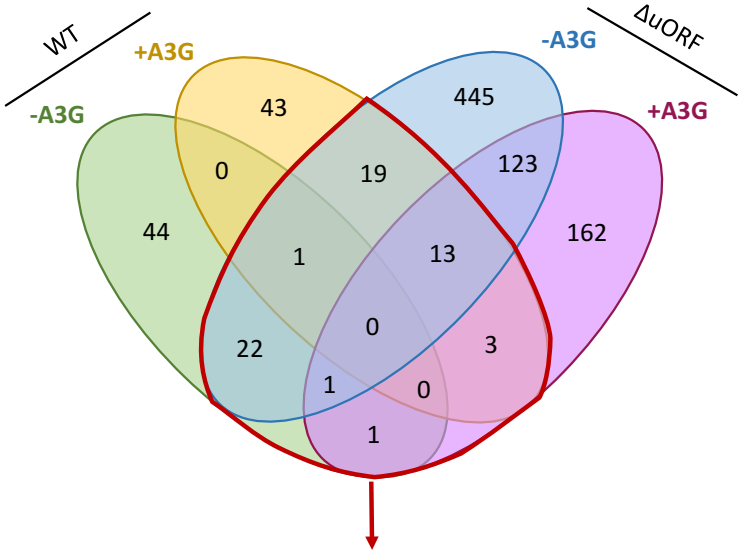


Figure 35: Effect of Vif on A3G mRNA-associated proteins in the presence of A3G protein. ***(A)*** *Differential association of proteins with A3G mRNA in the presence and absence of Vif. Proteins with a positive log_2(fold change) are upregulated in the presence of Vif, while a negative log_2(fold change) indicates downregulation in the presence of Vif. P-values are adjusted using the Bejamini-Hochberg method and the threshold of a p-value of 0.1 is indicated as a red line. Proteins* ***(B)*** *downregulated or* ***(C)*** *upregulated in the presence of Vif are represented as a string diagram. The thickness of lines indicates the confidence of association of the linked proteins. Proteins involved in some of the most represented functional clusters are colored as indicated.*

Effect of the uORF on proteins regulated by Vif. Previous studies have identified a short upstream ORF (uORF) in the 5'-UTR of A3G mRNA, and Vif-mediated translational inhibition of A3G has been shown to depend on the presence of this uORF (Libre, Seissler et al., in preparation). In order to gain a deeper insight into Vif-mediated translation regulation of A3G, the pull-down of cellular proteins in the presence and absence of Vif and A3G protein has therefore been repeated with a ΔuORF A3G mRNA. The proteins whose association with the ΔuORF A3G mRNA seemed to be

modulated by Vif were compared to those obtained for WT A3G mRNA. Preliminary results suggest that amongst the 147 proteins regulated by Vif on WT A3G mRNA, 60 were also found to be regulated in the absence of the uORF (Fig. 5, circled in red). Overall, the uORF seems to have a significant impact on the interactome of A3G mRNA.



1433S_HUMAN	DREB_HUMAN	PKP1_HUMAN
ACTN1_HUMAN	DSC1_HUMAN	PLAK_HUMAN
ACTN2_HUMAN	DSG1_HUMAN	SEPT2_HUMAN
ACTN4_HUMAN	FILA_HUMAN	SPB3_HUMAN
ARP3_HUMAN	FILA2_HUMAN	SPTB2_HUMAN
CAPZB_HUMAN	FLNA_HUMAN	SPTN1_HUMAN
CARF_HUMAN	FLNB_HUMAN	TPM3_HUMAN
CAZA1_HUMAN	FLNC_HUMAN	TPM4_HUMAN
CLAP2_HUMAN	IFIT5_HUMAN	WDR33_HUMAN
CNOT1_HUMAN	LIMA1_HUMAN	WDR91_HUMAN
COF1_HUMAN	LRC40_HUMAN	ZO2_HUMAN
CSRP2_HUMAN	MYH11_HUMAN	21 Keratins
DCP2_HUMAN	MYL6_HUMAN	
DESP_HUMAN	MYPT1_HUMAN	

Figure 36: Comparison of proteins whose association with WT and ΔuORF mRNA is regulated by Vif. *Pull-down of proteins on WT and ΔuORF A3G mRNA has allowed identification of proteins that seem to be regulated by Vif in the presence or absence of A3G. Proteins that are regulated both on WT and ΔuORF A3G mRNA are indicated in the table below.*

Interaction of selected proteins with Vif. Most of the proteins that significantly changed in the presence and absence of Vif are involved in the organization of the cytoskeleton. The cytoskeleton is an important player in HIV infection and seems to be modulated during infection in order to favour viral replication (26, 40, 53). Therefore, it is possible that expression of Vif leads to a modulation of the cytoskeleton, which might

explain accumulation of this class of proteins in our samples. Nevertheless, we concentrated further efforts on the other identified proteins. We selected a total of 10 candidates, which are significantly up- or downregulated by Vif in one of the conditions and which might have an effect on translational regulation (Table 2).

UniprotKB ID	UniProtKB Accession	Protein Name	Function
ARGI1_HUMAN	P05089	Arginase-1	Converts L-arginine to urea and L-ornithine
CALL5_HUMAN	Q9NZT1	Calmodulin-like protein 5	Binds calcium
DSG1_HUMAN	Q02413	Desmoglein-1	Component of desmosomes
G3P_HUMAN	P04406	Glyceraldehyde-3-phosphate dehydrogenase	Role in glycolysis, component of the GAIT complex
CNOT1_HUMAN	A5YKK6	CCR4-NOT transcription complex subunit 1	Component of the CCR4-NOT complex which is one of the major cellular mRNA deadenylases
CLAP2_HUMAN	O75122	CLIP-associating protein 2	Stabilization of microtubules
CHIP_HUMAN	Q9UNE7	E3 ubiquitin-protein ligase CHIP	E3-ubiquitin ligase
DCP2_HUMAN	Q8IU60	m7GpppN-mRNA hydrolase	Decapping of mRNAs for degradation
RENT1_HUMAN	Q92900	Regulator of nonsense transcripts 1	Recruited to stalled ribosomes as part of the SURF complex to induce NMD
MYPT1_HUMAN	O14974	Protein phosphatase 1 regulatory subunit 12A	Regulates PPP1C, binds to myosin, dephosphorylation of PLK1

Table 9: Selected candidates for a role in A3G translational inhibition by Vif. *The UniProtKB ID, Accession, full name and a summary of the function are indicated for each protein (59).*

Interaction of selected proteins with Vif has been validated by co-immunoprecipitation using a FLAG-tagged version of Vif. Amongst candidates, only CHIP and RENT1 have been shown to interact with Vif (Fig. 6). To a lower extent, CLAP2 might also be able to interact with Vif. However, we only detected this interaction after cross-linking of proteins in the cells, suggesting a weak or very dynamic interaction between Vif and these proteins, contrary to the interaction with A3G, which can also be detected without cross-linking (Fig. 6).

Role of CHIP and RENT1 in A3G translation. We have focussed further efforts on CHIP and RENT1, which seemed to be the most promising candidates, given their interaction with Vif in co-immunoprecipitation experiments. Their expression has been knocked down in cells using RNA silencing which allowed a mean decrease of 70-90

% of the expression level of CHIP (Fig. 7A) and 64 % of RENT1 (Fig. 7B). For both proteins, silencing seemed to have no visible effect on Vif-mediated counteraction of A3G nor on the expression level of A3G in the absence of Vif. Vif did not seem to have an impact on CHIP or RENT expression in the absence of siRNAs either (Fig. 7). However, these results have to be taken with caution since they are preliminary. Moreover, Vif-mediated translation inhibition was only very moderate in the control condition, which might have masked a possible effect of CHIP or RENT1 silencing.

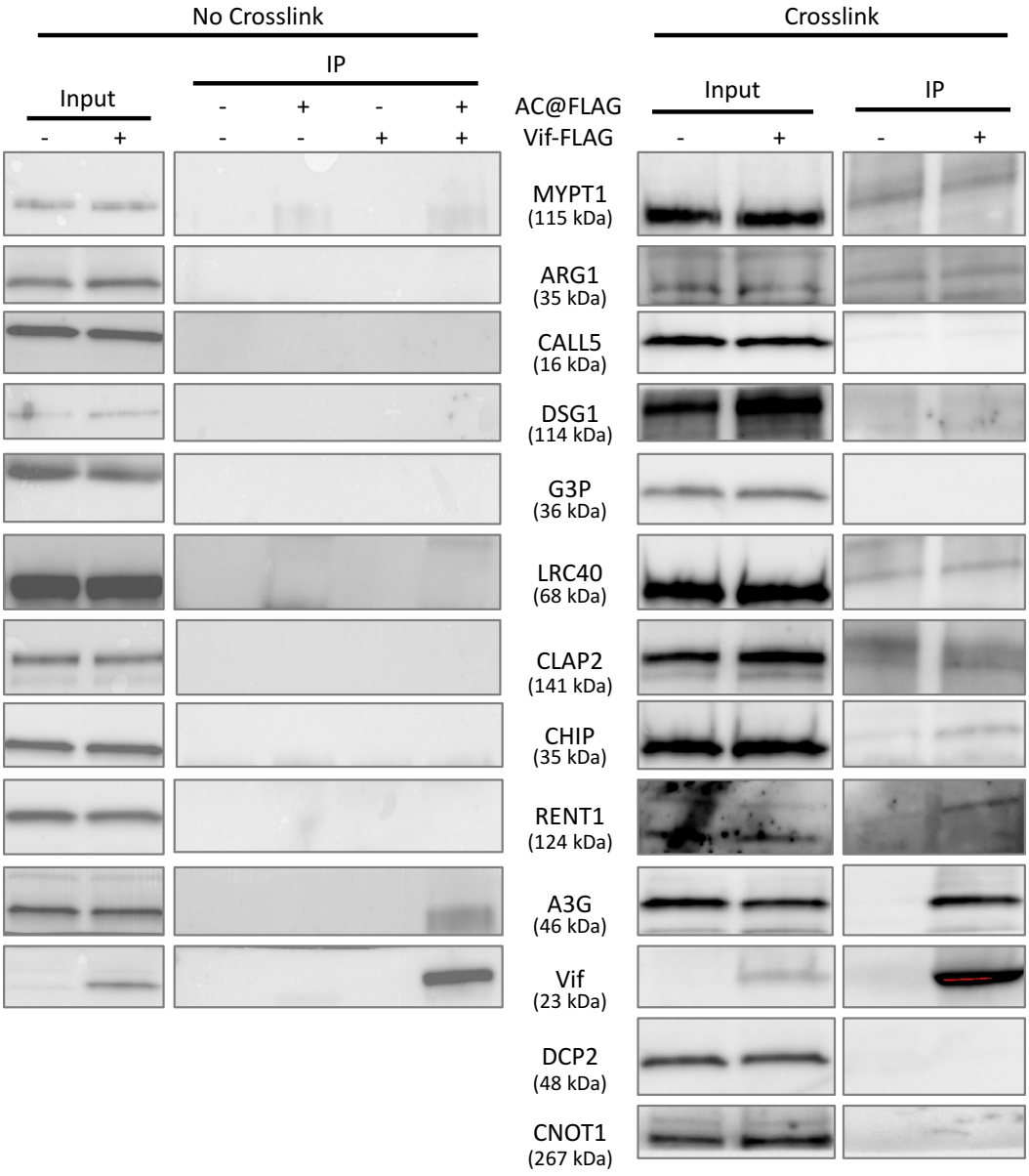


Figure 37: Interaction of Vif with selected proteins. *Cells were transfected or not to express a FLAG-tagged version of Vif as indicated. Cellular proteins were cross-linked or not and then immunoprecipitated using an @FLAG antibody. The input as well as the immunoprecipitated (IP) samples were analyzed by western blot.*

Discussion

In the present study we have for the first time identified the interactome of the A3G mRNA in different conditions. This has revealed a core interactome of 109 proteins which always remain associated with A3G mRNA independently of the experimental conditions, while 338 proteins are modulated depending on the presence of Vif and A3G proteins. Amongst the core interactome of A3G mRNA, we found proteins implicated in splicing, transport, translation and degradation. These proteins represent all of the typical interactants which play a role in the different steps of the life cycle of an mRNA.

Interestingly, the composition of A3G mRNA-associated protein complexes could be modulated by the presence of its own protein in the cellular lysate. This indicates that A3G protein influences the interactome of its own mRNA and might therefore play a role in the regulation of its expression. Indeed it has been shown previously that A3G protein localizes to P-bodies in cells (11, 12, 35, 55). Moreover, A3G can interact with and bind to RNAs through its zinc-coordinating domain (5, 23, 51), including its own mRNA (28). In our study, we observed an upregulation of RNP granule-associated proteins on A3G mRNA in the presence of A3G protein which might indicate that A3G recruits its own mRNA into P-bodies. It is possible that this allows A3G to regulate its expression.

In addition to the A3G protein itself, Vif could also modulate the A3G mRNA interactome. Vif has been shown to bind to the 5'-UTR of A3G mRNA (37). The 5'-UTR of A3G mRNA and more precisely a small uORF has been shown to be important for Vif-mediated inhibition of A3G translation (17; Libre, Seissler et al., in preparation). It is known that Vif downregulates A3G levels in cells by three different mechanisms, however the interplay of these different mechanisms in the cell has not been characterized yet. In the present study we have identified different proteins that are regulated by Vif in the presence and absence of the A3G protein. This suggests that Vif might not use the same mechanism of translational inhibition depending on the expression level of A3G. Indeed, when Vif is first expressed during an infection, the A3G protein is already present in the cell. At this stage, Vif has to quickly and efficiently downregulate A3G. When this is achieved, a low expression level of A3G has to be maintained, which might allow Vif to switch its mechanism. DSG1 (a desmosome

component with no known role in translation) is the only protein that seemed to be recruited on A3G mRNA by Vif independently of the A3G protein.

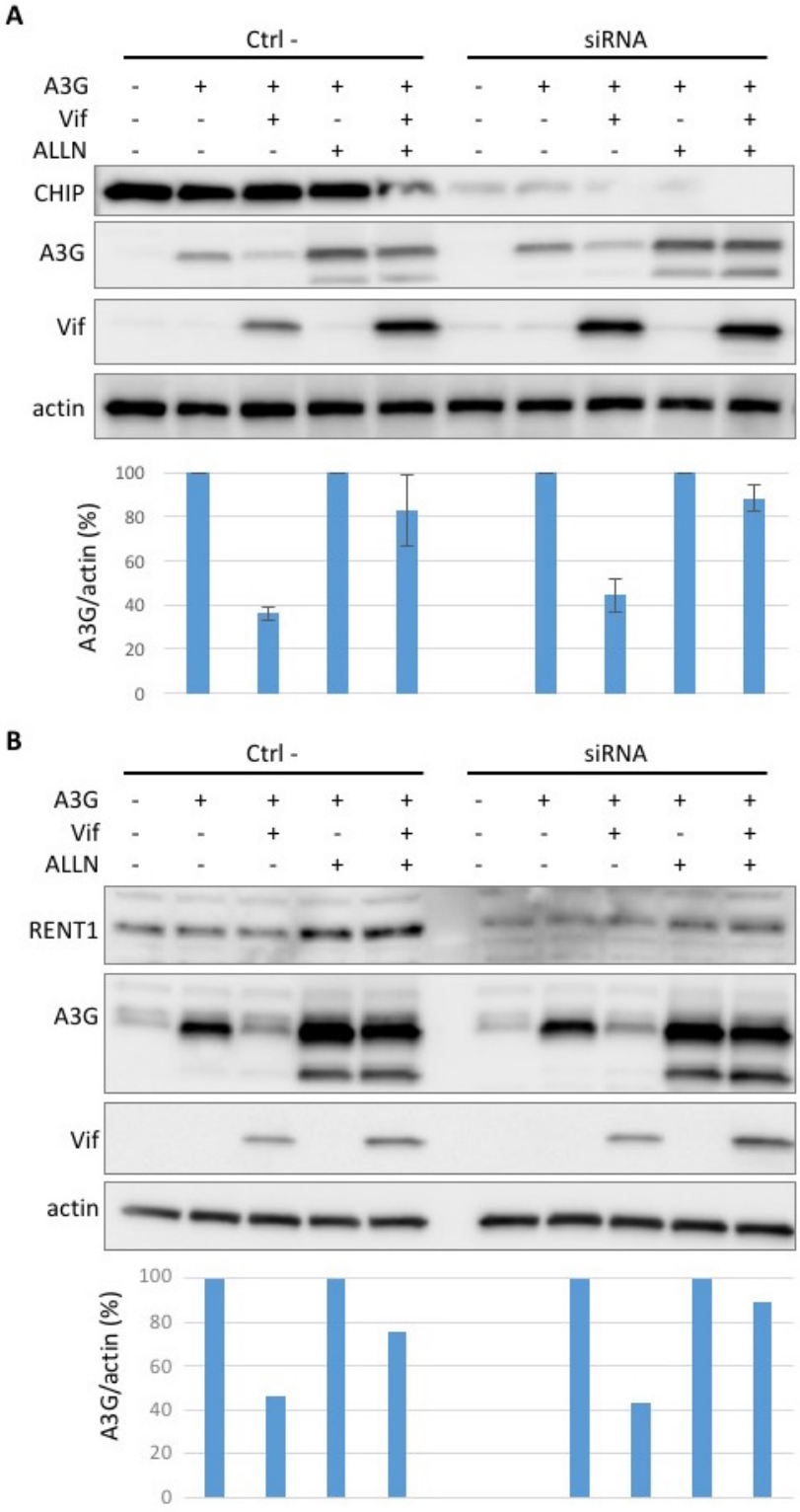


Figure 38: Effect of CHIP and RENT1 on Vif-mediated regulation of A3G expression levels. *Cells were treated with siRNAs directed against* ***(A)*** *CHIP,* ***(B)*** *RENT1 or with a negative control siRNA (Ctrl -). The effect of siRNA treatment on CHIP and RENT1 expression as well as on Vif-mediated regulation of A3G in the presence or absence of a proteasome inhibitor (ALLN) has been analysed by western blot and bands were quantified using ImageJ. The amount of A3G in the absence of Vif was set to 100 %.*

Surprisingly, a large proportion of cytoskeletal proteins have been identified as enriched in the presence of Vif. While keratins are typical contaminants in mass spectrometry (20), the actin cytoskeleton and intermediate filaments might play a role in the mechanism of Vif-mediated translational inhibition, for example by mediating transport of A3G mRNA into compartments like stress granules or P-bodies, where translation is inhibited. Indeed, it has been shown recently that in stress conditions, Vif relocalizes A3G mRNA into stress granules (Libre, Seissler et al., in preparation). Moreover, Vif is known to associate with intermediate filaments in the cell (26), which might explain their recruitment in our experiments. The actin cytoskeleton has also been shown to be regulated upon HIV infection, even though this has not yet been linked to Vif (40, 53).

Interestingly, Vif does not seem to directly interact with most of the proteins that it regulates. It is possible that Vif binds to A3G mRNA and induces changes to its secondary or even tertiary structure which lead to the drop-off or recruitment of different RBPs without an actual physical interaction with Vif. Indeed, Vif has been shown previously to possess an RNA chaperone activity (7, 19, 49). In order to confirm this hypothesis, it would be interesting to study the secondary structure of A3G mRNA *in cellula* in the presence and absence of Vif. Furthermore, most of the proteins regulated by Vif have been found to be phosphoproteins, which means that they can be regulated by phosphorylation. Vif is a known modulator of the phosphoproteome in the cell (14). Therefore, Vif might indeed induce changes in the phosphorylation status of A3G mRNA-associated proteins which could favour or hinder their interaction with the RNA. MYPT1 and NEB2 for example are phosphatase cofactors (2, 56) which seemed to be recruited onto A3G mRNA by Vif and are potentially involved in Vif-mediated modulation of the phosphoproteome. Further experiments will be needed to test this hypothesis.

CHIP and RENT1 were the only proteins whose interaction with Vif was confirmed in co-immunoprecipitation experiments. Amongst the proteins whose association with A3G mRNA seemed to be regulated by Vif, the E3-ubiquitin ligase CHIP (25) was particularly interesting because it has previously been described as a Vif-interacting protein (24). Moreover, CHIP has very recently been identified as a restriction factor of HIV infection, inducing ubiquitination and degradation of Tat (1). While the interaction

of RENT1 and CHIP with Vif would have rather suggested a recruitment of these proteins onto A3G mRNA by Vif, our analyses actually showed a Vif-mediated decrease of these proteins on A3G mRNA. One possible explanation would be that Vif binds to these proteins to induce their degradation. This would explain a direct interaction of these proteins with Vif while it also explains the downregulation of these proteins on A3G mRNA. Western blotting experiments showed however that CHIP and RENT1 expression did not change in the presence and absence of Vif. Moreover, silencing of CHIP and RENT1 had no effect, nor on the expression level of A3G, nor on Vif-mediated translational inhibition. These results are still preliminary and further experiments will be necessary to study the importance of the Vif-CHIP interaction.

While many proteins typically associated with mRNAs and the regulation of their life cycle have been identified as significantly enriched on A3G mRNA in our pull-down experiments, only few of them seem to be regulated by Vif. Indeed, most of the proteins whose association with A3G mRNA changed upon addition of Vif have no known role in regulation of translation. They might be multifunctional proteins whose role in translational regulation has not yet been identified. Therefore, it is difficult to propose a mechanism based on the present data. Silencing or overexpression of these proteins might allow to identify a potential effect on A3G translation in the presence and absence of Vif. On the other hand, some proteins like RENT1, DCP2 and CNOT1 have already been associated with mRNA degradation. RENT1 (better known as UPF1) is part of the nonsense-mediated decay pathway, which recruits decapping enzymes like DCP2 and deadenylases like CNOT1 to transcripts containing premature stop codons (22, 29, 30, 33, 54). These proteins can also be part of P-bodies and stress granules (4, 10, 42). Surprisingly, these proteins were downregulated from A3G mRNA in the presence of Vif. While this is consistent with previous results which have suggested that Vif does not induce degradation of A3G mRNA (17), it is nevertheless surprising because A3G mRNA also seems to be relocalized to stress-granules in the presence of Vif (Libre, Seissler et al., in preparation).

Finally, the results obtained for WT A3G mRNA have been compared to an A3G mRNA mutant where the uORF has been deleted (ΔuORF). Indeed the uORF has been shown previously to be required for Vif-mediated translational regulation (Libre, Seissler et al., in preparation). Our preliminary results showed that Vif-mediated regulation of many

proteins seems to depend on the presence of the uORF. Several proteins can't be regulated by Vif anymore when the uORF is deleted and amongst these, we have found for example CHIP and RENT1. A considerable number of proteins is regulated by Vif even in the absence of the uORF. Amongst these were for example CLAP2, CNOT1, DCP2, DSG1, LRC40 and MYPT1. For these proteins, two possibilities can be considered: (1) since Vif-mediated translational inhibition can't take place in the absence of the uORF, these proteins might not be valid candidates for a role in this mechanism; (2) in the absence of the uORF, Vif might still be able to bind to the 5'-UTR of A3G and regulate the association of these proteins with A3G mRNA, however, the mode of action of these proteins might depend on the presence of the uORF. Further studies will be necessary to discriminate between these possibilities. It would for example be interesting to evaluate whether Vif is still able to bind to the 5'-UTR of A3G in the absence of the uORF. Moreover, the results obtained with the ΔuORF A3G mRNA are still preliminary and will have to be confirmed.

In conclusion, this study allowed to characterize the impact of the A3G uORF, the A3G protein and Vif on the proteic interactome of A3G mRNA. It has also allowed identification of multiple potential cellular cofactors in Vif-mediated translational inhibition of A3G. The cytoskeleton, as well as P-bodies and stress granules, have been the main factors identified in this study which might play a central role in regulation of A3G translation. Nevertheless, several other proteins have been identified whose role in translation still has to be characterized.

Acknowledgements.

The following reagents were obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: A3G polyclonal antibody (#9968) from Dr. Warner Greene and Vif monoclonal antibody (#319) from Dr. Michael H. Malim. This work was supported by a grant from the French National Agency for Research on AIDS and Viral Hepatitis (ANRS) and SIDACTION to JCP, and by doctoral fellowships (TS and CV) from the French Ministry of Research and Higher Education.

Bibliography

1. **Ali A**, **Farooqui SR**, **Banerjea AC**. The host cell ubiquitin ligase protein CHIP is a potent suppressor of HIV-1 replication. *J Biol Chem* 294: 7283–7295, 2019.
2. **Allen PB**, **Ouimet CC**, **Greengard P**. Spinophilin, a novel protein phosphatase 1 binding protein localized to dendritic spines. *Proc Natl Acad Sci* 94: 9956–9961, 1997.
3. **Anderson BD**, **Harris RS**. Transcriptional regulation of APOBEC3 antiviral immunity through the CBF-β/RUNX axis. *Sci Adv* 1: e1500296, 2015.
4. **Anderson P**, **Kedersha N**. Stress granules: the Tao of RNA triage. *Trends Biochem Sci* 33: 141–150, 2008.
5. **Apolonia L**, **Schulz R**, **Curk T**, **Rocha P**, **Swanson CM**, **Schaller T**, **Ule J**, **Malim MH**. Promiscuous RNA Binding Ensures Effective Encapsidation of APOBEC3 Proteins by HIV-1. *PLOS Pathog* 11: e1004609, 2015.
6. **Barbosa C**, **Peixeiro I**, **Romão L**. Gene expression regulation by upstream open reading frames and human disease. *PLoS Genet* 9: e1003529, 2013.
7. **Batisse J**, **Guerrero S**, **Bernacchi S**, **Sleiman D**, **Gabus C**, **Darlix J-L**, **Marquet R**, **Tisné C**, **Paillart J-C**. The role of Vif oligomerization and RNA chaperone activity in HIV-1 replication. *Virus Res* 169: 361–376, 2012.
8. **Beckmann K**, **Grskovic M**, **Gebauer F**, **Hentze MW**. A Dual Inhibitory Mechanism Restricts msl-2 mRNA Translation for Dosage Compensation in Drosophila. *Cell* 122: 529–540, 2005.
9. **Calvo SE**, **Pagliarini DJ**, **Mootha VK**. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc Natl Acad Sci* 106: 7507–7512, 2009.
10. **Eulalio A**, **Behm-Ansmant I**, **Izaurralde E**. P bodies: at the crossroads of post-transcriptional pathways. *Nat Rev Mol Cell Biol* 8: 9–22, 2007.
11. **Gallois-Montbrun S**, **Holmes RK**, **Swanson CM**, **Fernández-Ocaña M**, **Byers HL**, **Ward MA**, **Malim MH**. Comparison of cellular ribonucleoprotein complexes associated with the APOBEC3F and APOBEC3G antiviral proteins. *J Virol* 82: 5636–5642, 2008.
12. **Gallois-Montbrun S**, **Kramer B**, **Swanson CM**, **Byers H**, **Lynham S**, **Ward M**, **Malim MH**. Antiviral protein APOBEC3G localizes to ribonucleoprotein complexes found in P bodies and stress granules. *J Virol* 81: 2165–2178, 2007.
13. **Gebauer F**, **Grskovic M**, **Hentze MW**. Drosophila sex-lethal inhibits the stable association of the 40S ribosomal subunit with msl-2 mRNA. *Mol Cell* 11: 1397–1404, 2003.
14. **Greenwood EJ**, **Matheson NJ**, **Wals K**, **van den Boomen DJ**, **Antrobus R**, **Williamson JC**, **Lehner PJ**. Temporal proteomic analysis of HIV infection reveals remodelling of the host phosphoproteome by lentiviral Vif variants. *eLife* 5, 2016.
15. **Gregori J**, **Sanchez A**, **Villanueva J**. msmsTests: LC-MS/MS Differential Expression Tests. *R Package Version 1120*: 2013.
16. **Gregori J**, **Sanchez A**, **Villanueva J**. msmsEDA: Exploratory Data Analysis of LC-MS/MS data by spectral counts. *R Package Version 1120*: 2014.
17. **Guerrero S**, **Libre C**, **Batisse J**, **Mercenne G**, **Richer D**, **Laumond G**, **Decoville T**, **Moog C**, **Marquet R**, **Paillart J-C**. Translational regulation of APOBEC3G mRNA by Vif requires its 5'UTR and contributes to restoring HIV-1 infectivity. *Sci Rep* 6: 39507, 2016.
18. **Harris RS**, **Bishop KN**, **Sheehy AM**, **Craig HM**, **Petersen-Mahrt SK**, **Watt IN**, **Neuberger MS**, **Malim MH**. DNA Deamination Mediates Innate Immunity to Retroviral Infection. *Cell* 113: 803–809, 2003.
19. **Henriet S**, **Sinck L**, **Bec G**, **Gorelick RJ**, **Marquet R**, **Paillart J-C**. Vif is a RNA chaperone that could temporally regulate RNA dimerization and the early steps of HIV-1 reverse transcription. *Nucleic Acids Res* 35: 5141–5153, 2007.
20. **Hodge K**, **Have ST**, **Hutton L**, **Lamond AI**. Cleaning up the masses: Exclusion lists to reduce contamination with HPLC-MS/MS. *J Proteomics* 88: 92–103, 2013.
21. **Huang DW**, **Sherman BT**, **Lempicki RA**. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44–57, 2009.
22. **Ito K**, **Takahashi A**, **Morita M**, **Suzuki T**, **Yamamoto T**. The role of the CNOT1 subunit of the CCR4-NOT complex in mRNA deadenylation and cell viability. *Protein Cell* 2: 755–763, 2011.
23. **Iwatani Y**, **Takeuchi H**, **Strebel K**, **Levin JG**. Biochemical activities of highly purified, catalytically active human APOBEC3G: correlation with antiviral effect. *J Virol* 80: 5992–6002, 2006.

24. **Jäger S**, **Cimermancic P**, **Gulbahce N**, **Johnson JR**, **McGovern KE**, **Clarke SC**, **Shales M**, **Mercenne G**, **Pache L**, **Li K**, **Hernandez H**, **Jang GM**, **Roth SL**, **Akiva E**, **Marlett J**, **Stephens M**, **D'Orso I**, **Fernandes J**, **Fahey M**, **Mahon C**, **O'Donoghue AJ**, **Todorovic A**, **Morris JH**, **Maltby DA**, **Alber T**, **Cagney G**, **Bushman FD**, **Young JA**, **Chanda SK**, **Sundquist WI**, **Kortemme T**, **Hernandez RD**, **Craik CS**, **Burlingame A**, **Sali A**, **Frankel AD**, **Krogan NJ**. Global landscape of HIV-human protein complexes. *Nature* 481: 365–370, 2011.
25. **Jiang J**, **Ballinger CA**, **Wu Y**, **Dai Q**, **Cyr DM**, **Höhfeld J**, **Patterson C**. CHIP is a U-box-dependent E3 ubiquitin ligase: identification of Hsc70 as a target for ubiquitylation. *J Biol Chem* 276: 42938–42944, 2001.
26. **Karczewski MK**, **Strebel K**. Cytoskeleton association and virion incorporation of the human immunodeficiency virus type 1 Vif protein. *J Virol* 70: 494–507, 1996.
27. **Kim DY**, **Kwon E**, **Hartley PD**, **Crosby DC**, **Mann S**, **Krogan NJ**, **Gross JD**. CBFβ stabilizes HIV Vif to counteract APOBEC3 at the expense of RUNX1 target gene expression. *Mol Cell* 49: 632–644, 2013.
28. **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. *J Biol Chem* 281: 29105–29119, 2006.
29. **Kurosaki T**, **Popp MW**, **Maquat LE**. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat Rev Mol Cell Biol* 20: 406–420, 2019.
30. **Lejeune F**, **Li X**, **Maquat LE**. Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. *Mol Cell* 12: 675–687, 2003.
31. **Libre C**, **Seissler T**, **Guerrero S**, **Batisse J**, **Stupfler B**, **Gilmer O**, **Weber M**, **Etienne L**, **Marquet R**, **Paillart J-C**. A conserved uORF in APOBEC3G mRNA is essential to its translational inhibition by the HIV-1 Vif protein. *Prep.*: [date unknown].
32. **Lomakin IB**, **Stolboushkina EA**, **Vaidya AT**, **Zhao C**, **Garber MB**, **Dmitriev SE**, **Steitz TA**. Crystal Structure of the Human Ribosome in Complex with DENR-MCT-1. *Cell Rep* 20: 521–528, 2017.
33. **Lykke-Andersen J**. Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay. *Mol Cell Biol* 22: 8114–8121, 2002.
34. **Mangeat B**, **Turelli P**, **Caron G**, **Friedli M**, **Perrin L**, **Trono D**. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* 424: 99–103, 2003.
35. **Marin M**, **Golem S**, **Rose KM**, **Kozak SL**, **Kabat D**. Human immunodeficiency virus type 1 Vif functionally interacts with diverse APOBEC3 cytidine deaminases and moves with them between cytoplasmic sites of mRNA metabolism. *J Virol* 82: 987–998, 2008.
36. **Medenbach J**, **Seiler M**, **Hentze MW**. Translational Control via Protein-Regulated Upstream Open Reading Frames. *Cell* 145: 902–913, 2011.
37. **Mercenne G**, **Bernacchi S**, **Richer D**, **Bec G**, **Henriet S**, **Paillart J-C**, **Marquet R**. HIV-1 Vif binds to APOBEC3G mRNA and inhibits its translation. *Nucleic Acids Res* 38: 633–646, 2010.
38. **Nishimura T**, **Wada T**, **Yamamoto KT**, **Okada K**. The Arabidopsis STV1 protein, responsible for translation reinitiation, is required for auxin-mediated gynoecium patterning. *Plant Cell* 17: 2940–2953, 2005.
39. **Oliveros JC**. Venny. An interactive tool for comparing lists with Venn's diagrams. *https://bioinfogp.cnb.csic.es/tools/venny/index.html*: 2007.
40. **Ospina Stella A**, **Turville S**. All-Round Manipulation of the Actin Cytoskeleton by HIV. *Viruses* 10, 2018.
41. **Park HS**, **Himmelbach A**, **Browning KS**, **Hohn T**, **Ryabova LA**. A plant viral "reinitiation" factor interacts with the host translational machinery. *Cell* 106: 723–733, 2001.
42. **Poblete-Durán N**, **Prades-Pérez Y**, **Vera-Otarola J**, **Soto-Rifo R**, **Valiente-Echeverría F**. Who Regulates Whom? An Overview of RNA Granules and Viral Infections. *Viruses* 8, 2016.
43. **Schleich S**, **Strassburger K**, **Janiesch PC**, **Koledachkina T**, **Miller KK**, **Haneke K**, **Cheng Y-S**, **Kuechler K**, **Stoecklin G**, **Duncan KE**, **Teleman AA**. DENR-MCT-1 promotes translation re-initiation downstream of uORFs to control tissue growth. *Nature* 512: 208–212, 2014.
44. **Seissler T**, **Marquet R**, **Paillart J-C**. Hijacking of the Ubiquitin/Proteasome Pathway by the HIV Auxiliary Proteins. *Viruses* 9, 2017.
45. **Sheehy AM**, **Gaddis NC**, **Choi JD**, **Malim MH**. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 418: 646–650, 2002.

46. **Sheehy AM**, **Gaddis NC**, **Malim MH**. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nat Med* 9: 1404–1407, 2003.
47. **Simon JH**, **Fouchier RA**, **Southerling TE**, **Guerra CB**, **Grant CK**, **Malim MH**. The Vif and Gag proteins of human immunodeficiency virus type 1 colocalize in infected human T cells. *J Virol* 71: 5259–5267, 1997.
48. **Simon JH**, **Southerling TE**, **Peterson JC**, **Meyer BE**, **Malim MH**. Complementation of vif-defective human immunodeficiency virus type 1 by primate, but not nonprimate, lentivirus vif genes. *J Virol* 69: 4166–4172, 1995.
49. **Sleiman D**, **Bernacchi S**, **Xavier Guerrero S**, **Brachet F**, **Larue V**, **Paillart J-C**, **Tisné C**. Characterization of RNA binding and chaperoning activities of HIV-1 Vif protein: Importance of the C-terminal unstructured tail. *RNA Biol* 11: 906–920, 2014.
50. **Stopak K**, **de Noronha C**, **Yonemoto W**, **Greene WC**. HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. *Mol Cell* 12: 591–601, 2003.
51. **Svarovskaia ES**, **Xu H**, **Mbisa JL**, **Barr R**, **Gorelick RJ**, **Ono A**, **Freed EO**, **Hu W-S**, **Pathak VK**. Human apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G (APOBEC3G) is incorporated into HIV-1 virions through interactions with viral and nonviral RNAs. *J Biol Chem* 279: 35822–35828, 2004.
52. **Szklarczyk D**, **Gable AL**, **Lyon D**, **Junge A**, **Wyder S**, **Huerta-Cepas J**, **Simonovic M**, **Doncheva NT**, **Morris JH**, **Bork P**, **Jensen LJ**, **Mering C von**. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 47: D607–D613, 2019.
53. **Usmani SM**, **Murooka TT**, **Deruaz M**, **Koh WH**, **Sharaf RR**, **Di Pilato M**, **Power KA**, **Lopez P**, **Hnatiuk R**, **Vrbanac VD**, **Tager AM**, **Allen TM**, **Luster AD**, **Mempel TR**. HIV-1 Balances the Fitness Costs and Benefits of Disrupting the Host Cell Actin Cytoskeleton Early after Mucosal Transmission. *Cell Host Microbe* 25: 73-86.e5, 2019.
54. **Wang Z**, **Jiao X**, **Carr-Schmid A**, **Kiledjian M**. The hDcp2 protein is a mammalian mRNA decapping enzyme. *Proc Natl Acad Sci U S A* 99: 12663–12668, 2002.
55. **Wichroski MJ**, **Robb GB**, **Rana TM**. Human retroviral host restriction factors APOBEC3G and APOBEC3F localize to mRNA processing bodies. *PLoS Pathog* 2: e41, 2006.
56. **Yamashiro S**, **Yamakita Y**, **Totsukawa G**, **Goto H**, **Kaibuchi K**, **Ito M**, **Hartshorne DJ**, **Matsumura F**. Myosin phosphatase-targeting subunit 1 regulates mitosis by antagonizing polo-like kinase 1. *Dev Cell* 14: 787–797, 2008.
57. **Yu X**, **Yu Y**, **Liu B**, **Luo K**, **Kong W**, **Mao P**, **Yu X-F**. Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* 302: 1056–1060, 2003.
58. **Zhang H**, **Yang B**, **Pomerantz RJ**, **Zhang C**, **Arunachalam SC**, **Gao L**. The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* 424: 94–98, 2003.
59. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47: D506–D515, 2019.

Supplementary material

Supplementary table 1: Proteins significantly regulated by the A3G protein. The protein Uniprot ID, Accession, common name and general function are indicated. The log2(fold change) of each protein in the presence of A3G protein as well as the associated adjusted p-value are indicated.

UniprotKB ID	UniProtKB Accession	Protein Name	Function	Log(Fold change)	adj. P-value
2AAA_HUMAN	P30153	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	Scaffolding of protein phosphatase 2A	-2,85	4,5E-02
2AAB_HUMAN	P30154	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A beta isoform	Scaffolding of protein phosphatase 2A	-5,51	9,7E-03
6PGD_HUMAN	P52209	6-phosphogluconate dehydrogenase, decarboxylating	Catalyzes a step in the pentose phosphate metabolic pathway	-4,81	4,5E-02
AHSA1_HUMAN	O95433	Activator of 90 kDa heat shock protein ATPase homolog 1	Co-chaperone of HSP90AA1	-6,39	2,9E-03
ATX10_HUMAN	Q9UBB4	Ataxin-10	Role in survival of neurons	-4,56	6,8E-04
ATX2L_HUMAN	Q8WWM7	Ataxin-2-like protein	Regulation of stress granule and P-body formation	1,29	6,8E-02
C19L1_HUMAN	Q69YN2	CWF19-like protein 1	Potential role in splicing	1,54	7,6E-02
C19L2_HUMAN	Q2TBE0	CWF19-like protein 2	Potential role in splicing	1,75	5,3E-02
CAND1_HUMAN	Q86VP6	Cullin-associated NEDD8-dissociated protein 1	Assembly factor for SCF E3 ubiquitin ligases	-7,20	2,9E-03
CARF_HUMAN	Q9NXV6	CDKN2A-interacting protein	Regulates DNA damage response	1,55	5,2E-02
CND2_HUMAN	Q15003	Condensin complex subunit 2	Role in mitosis	-5,65	1,8E-02
CSN1_HUMAN	Q13098	COP9 signalosome complex subunit 1	Regulation of SCF-type E3 ubiquitin ligases	-4,81	4,5E-02
CYBP_HUMAN	Q9HB71	Calcyclin-binding protein	Molecular bridge in E3 ubiquitin ligases	-5,74	1,7E-02
DCP1A_HUMAN	Q9NPI6	mRNA-decapping enzyme 1A	Role in decapping and mRNA degradation	1,06	7,9E-02
DCP2_HUMAN	Q8IU60	m7GpppN-mRNA hydrolase	Mediates decapping, necessary for mRNA degradation	1,56	3,7E-02
DD19A_HUMAN	Q9NUU7	ATP-dependent RNA helicase DDX19A	Helicase involved in nulear mRNA export	-5,74	1,7E-02
DDX1_HUMAN	Q92499	ATP-dependent RNA helicase DDX1	Potential role in polyadenylation, sensor of dsRNA for antiviral immunity	1,48	1,3E-02
DDX6_HUMAN	P26196	Probable ATP-dependent RNA helicase DDX6	Role in decapping and mRNA degradation	1,49	7,1E-02
DP13A_HUMAN	Q9UKG1	DCC-interacting protein 13-alpha	Involved in multiple signaling pathways	-5,76	2,6E-03
DUS11_HUMAN	O75319	RNA/RNP complex-1-interacting phosphatase	Potential role in mRNA metabolism	1,67	1,5E-02
ECM29_HUMAN	Q5VYK3	Proteasome adapter and scaffold protein ECM29	Component of the 26S proteasome	-5,35	2,4E-02
FAS_HUMAN	P49327	Fatty acid synthase	Catalyzes formation of long-chain fatty acids	-4,50	1,7E-02
GAPD1_HUMAN	Q14C86	GTPase-activating protein and VPS9 domain-containing protein 1	Participates in endocytosis and trafficking	-6,33	1,9E-04
GCN1_HUMAN	Q92616	eIF-2-alpha kinase activator GCN1	Involved in repression of global protein synthesis and activation of gene-specific translation	-5,29	6,4E-03
GUAA_HUMAN	P49915	GMP synthase	Biosynthesis of GMP	-5,35	2,4E-02
HAT1_HUMAN	O14929	Histone acetyltransferase type B catalytic subunit	Acetylation of histones	-5,00	2,4E-02
HUWE1_HUMAN	Q7Z6Z7	E3 ubiquitin-protein ligase HUWE1	E3 ubiquitin ligase	-6,30	1,2E-02
IPO5_HUMAN	O00410	Importin-5	Nuclear import of NLS-containing proteins	-3,57	1,0E-02
IPO7_HUMAN	O95373	Importin-7	Nuclear import of NLS-containing proteins	-5,82	1,6E-02
LKHA4_HUMAN	P09960	Leukotriene A-4 hydrolase	Involved in synthesis of leukotriene B4	-4,81	4,5E-02
LS14A_HUMAN	Q8ND56	Protein LSM14 homolog A	Essential for formation of P-bodies	2,39	6,5E-02
LS14B_HUMAN	Q9BX40	Protein LSM14 homolog B	Potential role in translation	2,32	7,6E-02
LSM6_HUMAN	P62312	U6 snRNA-associated Sm-like protein LSm6	Involved in splicing, potential role in mRNA degradation	2,11	5,5E-02
MAGD2_HUMAN	Q9UNF1	Melanoma-associated antigen D2	Regulation of NaCl transporters SLC12A1 and A3	-2,21	1,4E-02

MARE2_HUMAN	Q15555	Microtubule-associated protein RP/EB family member 2	Role in microtubule stabilization	-4,81	4,5E-02
MCM7_HUMAN	P33993	DNA replication licensing factor MCM7	Role in DNA replication	-5,23	2,6E-02
MP2K1_HUMAN	Q02750	Dual specificity mitogen-activated protein kinase kinase 1	Component of the Map kinase signal transduction pathway	-5,02	1,9E-02
N4BP2_HUMAN	Q86UW6	NEDD4-binding protein 2	Potential role in DNA repair	5,27	1,8E-02
NEK9_HUMAN	Q8TD19	Serine/threonine-protein kinase Nek9	Role in cell cycle progression	-3,09	2,2E-02
NKRF_HUMAN	O15226	NF-kappa-B-repressing factor	Transcriptional repression of NF-kB responsive genes	1,45	4,5E-02
NUDC_HUMAN	Q9Y266	Nuclear migration protein nudC	Role in neurogenesis	-6,24	1,3E-02
NUMA1_HUMAN	Q14980	Nuclear mitotic apparatus protein 1	Binds to microtubules, involved in mitosis	-4,63	5,8E-02
OTUB1_HUMAN	Q96FW1	Ubiquitin thioesterase OTUB1	Mediates Lys-48-linked deubiquitination	-4,81	4,5E-02
PCF11_HUMAN	O94913	Pre-mRNA cleavage complex 2 protein Pcf11	Component of the pre-mRNA cleavage complex II	2,78	4,6E-02
PRS10_HUMAN	P62333	26S proteasome regulatory subunit 10B	Component of the 26S proteasome	-5,82	1,6E-02
PRS4_HUMAN	P62191	26S proteasome regulatory subunit 4	Component of the 26S proteasome	-5,46	2,2E-02
PRS6A_HUMAN	P17980	26S proteasome regulatory subunit 6A	Component of the 26S proteasome	-5,10	3,1E-02
PRS6B_HUMAN	P43686	26S proteasome regulatory subunit 6B	Component of the 26S proteasome	-5,82	1,6E-02
PRS7_HUMAN	P35998	26S proteasome regulatory subunit 7	Component of the 26S proteasome	-3,18	4,5E-02
PSA_HUMAN	P55786	Puromycin-sensitive aminopeptidase	Involved in antigen-processing for MHC-I	-5,82	1,6E-02
PSD12_HUMAN	O00232	26S proteasome non-ATPase regulatory subunit 12	Component of the 26S proteasome	-5,46	2,2E-02
PSD13_HUMAN	Q9UNM6	26S proteasome non-ATPase regulatory subunit 13	Component of the 26S proteasome	-5,74	1,7E-02
PSMD1_HUMAN	Q99460	26S proteasome non-ATPase regulatory subunit 1	Component of the 26S proteasome	-6,30	1,2E-02
PSMD2_HUMAN	Q13200	26S proteasome non-ATPase regulatory subunit 2	Component of the 26S proteasome	-6,18	1,3E-02
PSMD3_HUMAN	O43242	26S proteasome non-ATPase regulatory subunit 3	Component of the 26S proteasome	-6,05	1,4E-02
PSMD6_HUMAN	Q15008	26S proteasome non-ATPase regulatory subunit 6	Component of the 26S proteasome	-5,46	2,2E-02
PSMD7_HUMAN	P51665	26S proteasome non-ATPase regulatory subunit 7	Component of the 26S proteasome	-5,35	2,4E-02
PUR6_HUMAN	P22234	Multifunctional protein ADE2 [Includes: Phosphoribosylaminoimidazole-succinocarboxamide synthase	Catalyzes a step in the IMP biosynthesis pathway	-5,46	2,2E-02
RBM6_HUMAN	P78332	RNA-binding protein 6	Binds poly-G RNAs	1,59	9,3E-02
RRP44_HUMAN	Q9Y2L1	Exosome complex exonuclease RRP44	RNA degradation and processing	-6,43	5,0E-03
RTCB_HUMAN	Q9Y3I0	tRNA-splicing ligase RtcB homolog	tRNA maturation, may act as an RNA ligase on other RNAs	1,05	4,5E-02
SC31A_HUMAN	O94979	Protein transport protein Sec31A	Role in vesicle transport from the ER	-5,23	2,6E-02
SIR7_HUMAN	Q9NRC8	NAD-dependent protein deacetylase sirtuin-7	Histone deacetylation	2,97	9,2E-02
SLK_HUMAN	Q9H2G2	STE20-like serine/threonine-protein kinase	Involved in apoptosis	-5,23	2,6E-02
SMC2_HUMAN	O95347	Structural maintenance of chromosomes protein 2	Role in mitosis	-6,47	1,1E-02
SMN_HUMAN	Q16637	Survival motor neuron protein	Involved in splicing	1,27	9,3E-02
SNX2_HUMAN	O60749	Sorting nexin-2	Involved in intracellular trafficking	-5,29	1,3E-02
SNX6_HUMAN	Q9UNH7	Sorting nexin-6	Involved in intracellular trafficking	-5,38	1,4E-02
STK26_HUMAN	Q9P289	Serine/threonine-protein kinase 26	Involved in cell growth and apoptosis	-5,27	1,6E-02

SYVC_HUMAN	P26640	Valine--tRNA ligase	Ligates valine to its tRNA	-5,65	1,8E-02
TCPQ_HUMAN	P50990	T-complex protein 1 subunit theta	Protein chaperone	-6,72	4,0E-03
TLN1_HUMAN	Q9Y490	Talin-1	Connection of the cytoskeleton to the plasma membrane	-7,42	5,4E-03
UBR4_HUMAN	Q5T4S7	E3 ubiquitin-protein ligase UBR4	E3-ubiquitin ligase	-5,23	2,6E-02
UMPS_HUMAN	P11172	Uridine 5'-monophosphate synthase	Biosynthesis of UMP	-3,17	9,5E-02
UN45A_HUMAN	Q9H3U1	Protein unc-45 homolog A	Co-chaperone of HSP90	-5,35	2,4E-02
VATA_HUMAN	P38606	V-type proton ATPase catalytic subunit A	Involved in acidification of intracellular compartments	-5,10	3,1E-02
VINC_HUMAN	P18206	Vinculin	Actin binding, involved in cell-cell adhesion	-4,81	4,5E-02
XPO1_HUMAN	O14980	Exportin-1	Nuclear export of NES-containing proteins and certain RNAs	-2,66	5,4E-03
XPO2_HUMAN	P55060	Exportin-2	Mediates export of importin-alpha	-2,70	4,9E-02
XPO5_HUMAN	Q9HAV4	Exportin-5	Nuclear export of dsRNA and proteins containing a dsRNA-binding domain	-4,24	1,5E-02
XRN1_HUMAN	Q8IZH2	5'-3' exoribonuclease 1	mRNA decay	3,21	7,3E-05
YTDC2_HUMAN	Q9H6S0	3'-5' RNA helicase YTHDC2	Involved in degradation of m6A-containing RNAs	1,70	2,6E-03
ZCCHV_HUMAN	Q7Z2W4	Zinc finger CCCH-type antiviral protein 1	Recruits degradation machinery to viral RNAs	1,14	4,8E-02

Supplementary table 2: Proteins significantly regulated by Vif in the absence of the A3G protein. The protein Uniprot ID, Accession, common name and general function are indicated. The log_2(fold change) of each protein in the presence of Vif as well as the associated adjusted p-value are indicated.

UniprotKB ID	UniProtKB Accession	Protein Name	Function	Log(Fold change)	adj. P-value
ACTN4_HUMAN	O43707	Alpha-actinin-4	Involved in actin crosslinking and tight junction formation	9,72	1,6E-03
FLNB_HUMAN	O75369	Filamin-B	Anchoring of the cell membrane and proteins to actin	8,90	1,6E-03
ACTN1_HUMAN	P12814	Alpha-actinin-1	Involved in actin cross-linking	8,80	1,6E-03
DREB_HUMAN	Q16643	Drebrin	Recruitment of CXCR4 to immunological synapse, actin organization	8,15	1,6E-03
FLNA_HUMAN	P21333	Filamin-A	Anchoring of the cell membrane and proteins to actin	8,03	3,7E-03
LIMA1_HUMAN	Q9UHB6	LIM domain and actin-binding protein 1	Regulation of actin dynamics	7,72	1,6E-03
SPTN1_HUMAN	Q13813	Spectrin alpha chain, non-erythrocytic 1	Involved in secretion	7,72	1,6E-03
SPTB2_HUMAN	Q01082	Spectrin beta chain, non-erythrocytic 1	Involved in secretion	7,57	1,6E-03
CALD1_HUMAN	Q05682	Caldesmon	Interacts with ctin and myosin	7,30	1,6E-03
IQGA1_HUMAN	P46940	Ras GTPase-activating-like protein IQGAP1	Regulation of actin dynamics	7,09	1,6E-03
MYH10_HUMAN	P35580	Myosin-10	Involved in cytoskeleton reorganization and the stabilization of CO1A1 and CO1A2 mRNAs, antagonist of MYH9	7,09	1,6E-03
TPM4_HUMAN	P67936	Tropomyosin alpha-4 chain	Stabilization of actin filaments	7,03	1,6E-03
ACTN2_HUMAN	P35609	Alpha-actinin-2	Cross-linking of actin filaments	6,99	1,6E-03
ENAH_HUMAN	Q8N8S7	Protein enabled homolog	Cytoskeleton remodelling	6,89	1,6E-03
MYPT1_HUMAN	O14974	Protein phosphatase 1 regulatory subunit 12A	Regulates PPP1C, binds to myosin, dephosphorylation of PLK1	6,85	1,6E-03
COR1C_HUMAN	Q9ULV4	Coronin-1C	Cytoskeleton organization, fission of vesicles from endosomal transport	6,82	1,6E-03
COR1B_HUMAN	Q9BR76	Coronin-1B	Cell motility	6,74	1,6E-03
FLNC_HUMAN	Q14315	Filamin-C	Organization of the actin cytoskeleton	6,74	1,6E-03
MYH9_HUMAN	P35579	Myosin-9	Cytoskeleton reorganization, antagonist of MYH10	6,74	1,6E-03
TPM3_HUMAN	P06753	Tropomyosin alpha-3 chain	Stabilization of actin filaments	6,70	1,6E-03
PALLD_HUMAN	Q8WX93	Palladin	Organization of the actin cytoskeleton	6,48	1,6E-03
MY18A_HUMAN	Q92614	Unconventional myosin-XVIIIa	Cytoskeleton-associated protein that mediates transport of several proteins involved in immunity and inflammation	6,43	1,6E-03
TPM1_HUMAN	P09493	Tropomyosin alpha-1 chain	Stabilization of actin filaments	6,43	1,6E-03
TPM2_HUMAN	P07951	Tropomyosin beta chain	Stabilization of actin filaments	6,38	1,6E-03
VASP_HUMAN	P50552	Vasodilator-stimulated phosphoprotein	Cytoskeleton remodelling, link with ENAH ?	6,38	1,6E-03
PDLI1_HUMAN	O00151	PDZ and LIM domain protein 1	Recruits proteins like PALLD and ACTN1 to stress fibers	6,32	1,6E-03
SRC8_HUMAN	Q14247	Src substrate cortactin	Organization of the actin cytoskeleton	6,27	1,7E-03
SEPT9_HUMAN	Q9UHD8	Septin-9	Filament-forming GTPase	6,15	2,0E-03
CAPZB_HUMAN	P47756	F-actin-capping protein subunit beta	Caps and protects actin filament ends	6,08	2,0E-03
CSRP2_HUMAN	Q16527	Cysteine and glycine-rich protein 2	Role in vascular development	6,08	2,0E-03
ZYX_HUMAN	Q15942	Zyxin	Targets ENA/VASP proteins to focal adhesions	6,08	2,0E-03
AMOL1_HUMAN	Q8IY63	Angiomotin-like protein 1	Inhibits the Wnt/beta-catenin signaling	6,01	2,4E-03
EFHD1_HUMAN	Q9BUP0	EF-hand domain-containing protein D1	Calcium sensor	5,67	6,3E-03
COF1_HUMAN	P23528	Cofilin-1	Regulation of actin dynamics	5,64	4,9E-03
AIF1L_HUMAN	Q9BQI0	Allograft inflammatory factor 1-like	Promotes actin bundling	5,56	7,9E-03
CYTSA_HUMAN	Q69YQ0	Cytospin-A	Cytoskeleton organization	5,56	7,9E-03
MPRIP_HUMAN	Q6WCQ1	Myosin phosphatase Rho-interacting protein	Targets myosin phosphatase to the actin cytoskeleton	5,56	7,9E-03
PAWR_HUMAN	Q96IZ0	PRKC apoptosis WT1 regulator protein	Induces apoptosis and inhibits transcription (NF-KB pathway)	5,56	7,9E-03
AFAD_HUMAN	P55196	Afadin	Cell-cell junctions	5,56	7,9E-03

SEPT2_HUMAN	Q15019	Septin-2	Cytoskeleton organization	5,56	7,9E-03
NEB2_HUMAN	Q96SB3	Neurabin-2	Involved in different signalling pathways, binds to actin, links protein phosphatase 1 to actin	5,44	1,1E-02
ARP3_HUMAN	P61158	Actin-related protein 3	Mediates formation of branched actin filaments	5,44	1,1E-02
AMOT_HUMAN	Q4VCS5	Angiomotin	Role in tight junction maintenance	5,38	1,1E-02
SEPT7_HUMAN	Q16181	Septin-7	Organization of the actin cytoskeleton, GTPase	5,31	1,6E-02
CAZA2_HUMAN	P47755	F-actin-capping protein subunit alpha-2	Stabilization of actin filament extremities	5,16	2,4E-02
CD2AP_HUMAN	Q9Y5K6	CD2-associated protein	Mediates binding of actin to the plasma membrane	5,16	2,4E-02
PDLI7_HUMAN	Q9NR12	PDZ and LIM domain protein 7	Binding of proteins to actin	5,16	2,4E-02
RHG21_HUMAN	Q5T5U3	Rho GTPase-activating protein 21	Activates GTPase activity of RHOA and CDC42	5,16	2,4E-02
MYH11_HUMAN	P35749	Myosin-11	Muscle contraction	5,16	2,4E-02
EFHD2_HUMAN	Q96C19	EF-hand domain-containing protein D2	Potentially regulates apoptosis	4,99	4,1E-02
SPTN2_HUMAN	O15020	Spectrin beta chain, non-erythrocytic 2	Potential role in neuronal cytoskeleton	4,99	4,1E-02
SVIL_HUMAN	O95425	Supervillin	Links actin to the plasma membrane	4,99	4,1E-02
TMOD3_HUMAN	Q9NYL9	Tropomodulin-3	Role in actin dynamics	4,99	4,1E-02
ARC1B_HUMAN	O15143	Actin-related protein 2/3 complex subunit 1B	Mediates formation of branched actin filaments	4,78	6,5E-02
COR1A_HUMAN	P31146	Coronin-1A	Component of the cytoskeleton	4,78	6,5E-02
EVL_HUMAN	Q9UI08	Ena/VASP-like protein	cytoskeleton remodelling	4,78	6,5E-02
KCTD3_HUMAN	Q9Y597	BTB/POZ domain-containing protein KCTD3	Component of a K/Na channel	4,78	6,5E-02
NEXN_HUMAN	Q0ZGT2	Nexilin	Associated with the actin cytoskeleton	4,78	6,5E-02
SEPT6_HUMAN	Q14141	Septin-6	Organization of the actin cytoskeleton, GTPase	4,78	6,5E-02
ZO2_HUMAN	Q9UDY2	Tight junction protein ZO-2	Role in tight junctions	4,78	6,5E-02
MYL6_HUMAN	P60660	Myosin light polypeptide 6	Regulatory component of myosin	3,12	6,8E-02
CAZA1_HUMAN	P52907	F-actin-capping protein subunit alpha-1	Stabilization of actin filament extremities	2,87	9,1E-02
EIF3A_HUMAN	Q14152	Eukaryotic translation initiation factor 3 subunit A	RNA binding component of eIF3, can activate or repress translation	-1,63	9,4E-02
IFIT5_HUMAN	Q13325	Interferon-induced protein with tetratricopeptide repeats 5	RBP involved in innate immunity, recognizes 5'-ppp ends	-4,85	9,1E-02
PCF11_HUMAN	O94913	Pre-mRNA cleavage complex 2 protein Pcf11	Component of pre-mRNA cleavage complex II	-5,02	6,5E-02
WDR33_HUMAN	Q9C0J8	pre-mRNA 3' end processing protein WDR33	Maturation of mRNA 3' ends	-5,43	2,4E-02
WDR82_HUMAN	Q6UXN9	WD repeat-containing protein 82	Role in transcription regulation	-5,53	1,4E-02
DSG1_HUMAN	Q02413	Desmoglein-1	Component of desmosomes	5,29	4,9E-02
RENT1_HUMAN	Q92900	Regulator of nonsense transcripts 1	Recruited to stalled ribosomes as part of the SURF complex to induce NMD	-4,85	9,1E-02

Supplementary table 3: Proteins significantly regulated by Vif in the presence of the A3G protein. The protein Uniprot ID, Accession, common name and general function are indicated. The log_2(fold change) of each protein in the presence of Vif as well as the associated adjusted p-value are indicated.

UniprotKB ID	UniProtKB Accession	Protein Name	Function	Log(Fold change)	adj. P-value
1433S_HUMAN	P31947	14-3-3 protein sigma	Adapter protein in many signaling pathways. Regulates MDM2	3,52	1,4E-02
ARGI1_HUMAN	P05089	Arginase-1	Converts L-arginine to urea and L-ornithine	5,16	4,8E-03
CALL5_HUMAN	Q9NZT1	Calmodulin-like protein 5	Binds calcium	5,14	5,4E-03
CARF_HUMAN	Q9NXV6	CDKN2A-interacting protein	Involved in signaling pathways in response to DNA damage	-1,47	1,5E-02
CASPE_HUMAN	P31944	Caspase-14	Non-apoptotic caspase involved in epidermal differentiation.	5,93	9,2E-05
CDSN_HUMAN	Q15517	Corneodesmosin	Important for the epidermal barrier integrity.	5,47	1,5E-03
CHIP_HUMAN	Q9UNE7	E3 ubiquitin-protein ligase CHIP	E3-ubiquitin ligase	-5,41	7,3E-03
CLAP2_HUMAN	O75122	CLIP-associating protein 2	Stabilization of microtubules	-3,98	5,9E-04
CMTR1_HUMAN	Q8N1G2	Cap-specific mRNA (nucleoside-2'-O-)-methyltransferase 1	Methylation of the first nucleotide of mRNAs	-1,37	2,9E-03
CNOT1_HUMAN	A5YKK6	CCR4-NOT transcription complex subunit 1	Component of the CCR4-NOT complex which is one of the major cellular mRNA deadenylases	-1,63	5,3E-02
CYTA_HUMAN	P01040	Cystatin-A	Proteinase inhibitor important in desmosome-mediated cell-cell adhesion	4,59	4,1E-02
DCP2_HUMAN	Q8IU60	m7GpppN-mRNA hydrolase	Decapping of mRNAs for degradation	-1,29	7,2E-02
DESP_HUMAN	P15924	Desmoplakin	Component of desmosomes	6,82	1,3E-52
DSC1_HUMAN	Q08554	Desmocollin-1	Component of desmosomes	6,35	8,8E-05
DSG1_HUMAN	Q02413	Desmoglein-1	Component of desmosomes	5,33	3,8E-13
ENOA_HUMAN	P06733	Alpha-enolase	Role in glycolysis, growth control, hypoxia tolerance, allergic responses	2,29	5,7E-02
EXOC6_HUMAN	Q8TAG9	Exocyst complex component 6	Involved in docking of exocytic vesicles with fusion site on the plasma membrane	-4,97	3,1E-02
FABP5_HUMAN	Q01469	Fatty acid-binding protein 5	Transport of fatty acids, modulates inflammation	6,81	1,3E-08
FBX50_HUMAN	Q6ZVX7	F-box only protein 50	Role in cell proliferation	4,31	8,0E-02
FILA_HUMAN	P20930	Filaggrin	Organizes keratin filaments	7,67	3,6E-04
FILA2_HUMAN	Q5D862	Filaggrin-2	Cell-cell adhesion	4,18	9,7E-07
G3P_HUMAN	P04406	Glyceraldehyde-3-phosphate dehydrogenase	Role in glycolysis, component of the GAIT complex	1,89	4,1E-02
GSTP1_HUMAN	P09211	Glutathione S-transferase P	Metabolism	3,32	6,8E-02
HUTH_HUMAN	P42357	Histidine ammonia-lyase	L-histidine degradation into L-glutamate	4,31	8,0E-02
IGHG2_HUMAN	P01859	Immunoglobulin heavy constant gamma 2		4,37	7,6E-02
IGHG4_HUMAN	P01861	Immunoglobulin heavy constant gamma 4		4,37	7,6E-02
IGKC_HUMAN	P01834	Immunoglobulin kappa constant		4,41	7,7E-02
LMNA_HUMAN	P02545	Prelamin-A/C	Nuclear assembly and dynamics	2,64	4,8E-03
LRC40_HUMAN	Q9H9A6	Leucine-rich repeat-containing protein 40		-3,89	3,1E-03
NFH_HUMAN	P12036	Neurofilament heavy polypeptide	Intermediate filament in neurons	3,78	4,4E-03
NFX1_HUMAN	Q12986	Transcriptional repressor NF-X1	Regulates MHC class II expression	-1,18	4,1E-02
PKP1_HUMAN	Q13835	Plakophilin-1	Role in junctional plaques	6,64	9,6E-08
PLAK_HUMAN	P14923	Junction plakoglobin	Junction plaque	4,15	4,4E-13
POF1B_HUMAN	Q8WVV4	Protein POF1B	Regulates the actin cytoskeleton	5,28	3,1E-03
S10AE_HUMAN	Q9HCY8	Protein S100-A14	Regulates apoptosis by modulation of p53/TP53	4,59	4,1E-02
SBSN_HUMAN	Q6UWP8	Suprabasin		6,14	1,1E-05
SPB12_HUMAN	Q96P63	Serpin B12	Inhibitor of proteases	4,34	7,6E-02
SPB3_HUMAN	P29508	Serpin B3	Inhibitor of proteases, role in immune response against tumors	6,56	1,1E-03
SPB4_HUMAN	P48594	Serpin B4	Inhibitor of proteases, role in immune response against tumors	5,89	2,6E-03
SPB5_HUMAN	P36952	Serpin B5	Inhibitor of proteases, role in immune response against tumors	4,93	1,6E-02
SPR2B_HUMAN	P35325	Small proline-rich protein 2B	Keratinocyte envelope protein	6,04	2,7E-05
SPR2D_HUMAN	P22532	Small proline-rich protein 2D	Keratinocyte envelope protein	5,87	8,6E-05
SPR2E_HUMAN	P22531	Small proline-rich protein 2E	Keratinocyte envelope protein	6,32	4,1E-06
SPR2F_HUMAN	Q96RM1	Small proline-rich protein 2F	Keratinocyte envelope protein	5,56	7,1E-04
TBB2A_HUMAN	Q13885	Tubulin beta-2A chain	Component of tubulin	-1,68	1,5E-02
TBB2B_HUMAN	Q9BVA1	Tubulin beta-2B chain	Component of tubulin	-1,68	1,4E-02

TBB4A_HUMAN	P04350	Tubulin beta-4A chain	Component of tubulin	-1,55	4,0E-02
TBB8_HUMAN	Q3ZCM7	Tubulin beta-8 chain	Component of tubulin	-1,64	9,7E-02
TGM3_HUMAN	Q08188	Protein-glutamine gamma-glutamyltransferase E	Cross-links Glu and Lys, keratinocyte envelope formation	6,70	4,3E-08
TKT_HUMAN	P29401	Transketolase	Metabolism	5,22	7,3E-03
WDR91_HUMAN	A4D1P6	WD repeat-containing protein 91	Involved in cargo transport between early and late endosomes	-4,84	4,6E-02

Discussion

VII Discussion

A3G is one of the main restriction factors of HIV, which in the absence of Vif induces hypermutation of the viral genome and causes abortion of the viral life cycle. The HIV Vif protein counteracts A3G by 3 different mechanisms, one of which is inhibition of A3G translation. This translational inhibition is dependent on the 5'-UTR of A3G mRNA and more precisely on a small uORF which spans stem loops 2 and 3 of the 5'-UTR.

The first objective of my thesis was to better characterize the mode of translation of A3G in dependence of the uORF and to evaluate the importance of different characteristics of this uORF for Vif-mediated translational inhibition. This study has shown that the translation of A3G is negatively regulated by the presence of the uORF. A3G is indeed translated using a unique combination of leaky scanning and re-initiation after translation of the uORF, while no IRES activity has been detected. Furthermore, we have shown that translational inhibition of A3G by Vif can only take place when the uORF is intact. Upon mutation of the uORF start codon or deletion of the uORF, the Vif-mediated translation inhibition is abolished. The uORF has been found to be conserved in A3G and A3F, but none of the other A3s. As expected, A3F seemed to be regulated by Vif in the same manner as A3G, while this is the case for none of the other A3s, which is consistent with the absence of the uORF in their mRNAs. Finally, in stress conditions, it has been shown that A3G mRNA is relocalized to stress-granules in the presence of Vif (and not when the uORF is deleted), which might be either a consequence of the Vif-mediated translational block or part of the mechanism itself. In addition, the importance of the uORF for Vif-mediated counteraction of A3G has been studied using an H9 cell-line where the A3G uORF has been deleted using the CRISPR/Cas9 technology. The preliminary results obtained after infection of this cell line with HIV-1 seem to suggest that the uORF might have an impact on viral production, however this still has to be confirmed. This first part of my thesis has shown that the uORF plays a crucial role in the mechanism of A3G translational inhibition by Vif and might be equally important for counteraction of A3G in the context of a viral infection.

It has been shown previously that Vif binds to the stem loops 1 and 3 of A3G 5'-UTR. This binding brings Vif into close proximity with regulatory protein complexes, pre-initiation complexes and scanning ribosomes along the 5'-UTR, suggesting various possibilities for the mechanism of translational inhibition by Vif. Vif might for example hinder recruitment of the

ribosome, modulate leaky scanning, inhibit re-initiation after translation of the uORF or induce stalling of the ribosomes on the uORF. To do so, it is probable that Vif does not act on its own. Indeed, Vif might interact with A3G mRNA-associated protein complexes either to induce drop-off of cellular translation stimulatory factors or to recruit translation inhibitory cellular proteins to the A3G 5'-UTR. Indeed, numerous mechanisms exist where RBPs modulate translation by binding to an mRNA and it is possible that Vif takes advantage of one of these strategies.

To test this hypothesis, we have first studied the association of proteins with the 5'-UTR of A3G mRNA by RaPID. This has allowed us to identify 26 proteins whose association with the A3G 5'-UTR depends on the uORF and 21 proteins that are modulated by Vif. Many of these proteins are translation initiation factors, ribosomal proteins or other known regulators of translation or mRNA metabolism. Most interestingly, we have identified RL24, which is a known stimulator for re-initiation on uORF containing transcripts in plants (155, 163). In our case, RL24 is associated with the 5'-UTR of A3G only in the presence of the uORF, potentially to stimulate re-initiation, and is downregulated in the presence of Vif, which might lead to a decrease in A3G translation. The big downside of the RaPID protocol however was the low number of proteins that have been significantly enriched above the background of proteins associated with a scrambled RNA.

Therefore, we have decided to develop a more physiological protocol with a full-length A3G mRNA which is expressed and assembles with its interacting proteins in cells, before being pulled-down for analysis.

Pull-down of a non-modified, full-length A3G mRNA from cells would have allowed identification of the most physiological interactions and would have best represented the normal interactome of A3G mRNA in cells. However, it was not possible to specifically pull-down and elute A3G mRNA for different reasons: (i) A3G mRNA could non-specifically bind to beads; (ii) the amount of A3G mRNA pulled-down and eluted in the presence of complementary oligos was not sufficient to stand out from the unspecific background; (iii) even A3G mRNA that had previously been bound to beads was lost upon addition of cellular lysate. This might have been due to a tight interaction of A3G with cellular interactants, that hamper with the pull-down. Our subsequent results have suggested that A3G mRNA might be associated with the cytoskeleton, P-bodies and stress granules. The association of A3G mRNA with stress granules in the presence of Vif has also been identified by previous microscopy studies (Libre, Seissler et al., in preparation). If A3G mRNA is indeed associated in the cell

with these kinds of complexes, this might explain why A3G mRNA seemed inaccessible for our bead-bound complementary oligos. Despite many efforts to improve this protocol, problems with specificity of the pull-down have finally led us to use an alternative protocol based on an *in vitro* transcribed, biotinylated A3G mRNA which is used as a bait to capture interacting proteins *in vitro*.

Pull-down of proteins on a biotinylated RNA bait or their biotinylation in the cell by RaPID are both mediated by modified RNAs. The advantage of RaPID is that the RNA-protein interaction takes place in the physiological context of the cell while association of proteins with biotinylated A3G mRNA happens *in vitro*. The advantage of pull-down on a biotinylated bait is that in this case the entire interactome of the full-length A3G mRNA can be captured, while RaPID is performed only on the 5'-UTR of A3G and a large number of the 5'-UTR-interacting proteins are potentially not even biotinylated.

<table>
<tr><th></th><th>RaPID</th><th>Biotinylated A3G</th><th>Function</th></tr>
<tr><td>EIF3A</td><td>Enriched on A3G ΔuORF</td><td>Negatively regulated by Vif (in the absence of A3G)</td><td>Can bind to RNA and activate or repress translation</td></tr>
<tr><td>TLN1</td><td>Recruited by Vif, enriched on A3G ΔuORF</td><td>Negatively regulated by A3G</td><td>Connection of the cytoskeleton to the plasma membrane</td></tr>
<tr><td>ATX2L</td><td>Enriched on A3G ΔuORF</td><td>Positively regulated by A3G</td><td>Regulation of stress granule and P-body formation.</td></tr>
<tr><td>CYBP</td><td>Enriched on A3G ΔuORF</td><td>Negatively regulated by A3G</td><td>Potential component of ubiquitin E3 complexes.</td></tr>
<tr><td>DESP</td><td>Negatively regulated by Vif</td><td>Positively regulated by Vif (in the presence of A3G)</td><td>Component of desmosomes</td></tr>
<tr><td>DDX17</td><td rowspan="3">Enriched on the A3G 5'-UTR compared to the Scr (only in the presence of Vif)</td><td rowspan="3">Present on A3G mRNA in all conditions (part of the core interactome)</td><td>RNA helicase</td></tr>
<tr><td>DDX5</td><td>RNA helicase</td></tr>
<tr><td>RBM6</td><td>RNA binding protein</td></tr>
</table>

Table 10: Proteins that have been identified by RaPID and by pull-down of proteins on biotinylated A3G mRNA. *The name of each protein is indicated on the left. For each protocol, the condition in which the protein has been identified is indicated. The known function of each protein is indicated on the right as found on the Uniprot Knowledgebase.*

Despite these differences, it is surprising that only very few A3G mRNA-interacting protein have been found in common between the two protocols. Indeed, only 8 proteins appear in both screens (Table 8), but the conditions in which they have been identified are not consistent between the two experiments. Amongst these proteins, the most interesting is the translation initiation factor EIF3A, which can act as a translational activator (122). RaPID has shown that this protein is enriched on the A3G 5'-UTR in the absence of the uORF, which might be explained by the increased translation of A3G in the absence of the uORF (even though

translation might not be able to take place on the constructs used for these experiments). Pull-down of proteins on biotinylated A3G mRNA however seemed to indicate that Vif might cause a drop-off of this protein from A3G mRNA, which appears logical considering that Vif inhibits translation of A3G and that eIF3A can act as a translation stimulator. In this context it is surprising that the negative regulation of eIF3A by Vif has not been confirmed by RaPID.

Overall, in all mass spectrometry analyses performed during this thesis, we have mainly identified RNA binding proteins and many proteins with a known role in the mRNA metabolism. This shows the strength of the approaches used for identification of an mRNA interactome. However, the analysis process has also shown some difficulties. First of all, Vif is a known interactant of A3G mRNA, which has been characterized in a previous study (147), and has been shown by western blot to be enriched in our samples (Fig. 31). Nevertheless, it has never been identified by mass spectrometry, probably because no suitable peptides could be generated by trypsin digestion. Moreover, none of the known Vif interactants like A3G, CBF-β or components of the E3-ubiquitin ligase complex (CUL5, EloB, EloC) were identified either. Furthermore, proteins significantly enriched in different conditions have been analysed by differential expression tests using a negative-binomial regression model. These tests allowed to compare only two conditions at a time, which was not very well adapted to our experimental setup and did not allow to take into account the negative control and compare samples obtained in different conditions at the same time. To obtain more exact results it might be necessary to re-analyse the data with a multifactorial model, however, this is not currently available in the lab.

Despite these technical difficulties, pull-down of proteins on a biotinylated A3G mRNA has allowed us to identify a large number of proteins specifically associated with this mRNA and several of these proteins seemed to be regulated by Vif. This regulation might be through a Vif-induced downregulation of the expression of these proteins or through a direct interaction between Vif and the protein to mediate its recruitment on the mRNA. Recruitment and drop-off of cellular proteins from the A3G mRNA might also be mediated by conformational changes of the A3G mRNA in response to Vif binding. Moreover, Vif might also modulate the phosphorylation status of these proteins in order to mediate regulation.

Conclusion
&
Perspectives

VIII Conclusion and Perspectives

In conclusion, this thesis has contributed to show the importance of a small uORF in the 5'-UTR of A3G mRNA for Vif-mediated translational inhibition of A3G and A3F, while the translation of the other A3 proteins is not regulated by Vif. A T-cell line has also been generated where the A3G uORF is deleted, and this T-cell line will allow to confirm the importance of the uORF for HIV infection. My thesis has also allowed to study the interactome of A3G mRNA using two different techniques. These studies have shown that protein complexes that associate with A3G mRNA vary depending on the presence of the uORF, the A3G protein and the Vif protein. These three parameters have been shown to significantly alter the interactome of A3G mRNA and all three might play an important role in the regulation of expression and stability of this mRNA.

The obtained results have given rise to several new perspectives to further study Vif-mediated translational inhibition of A3G:

Single nucleotide polymorphisms (SNPs) have been shown to affect uORFs and some even play important roles in human disease (34). In order to further study the importance of the A3G uORF on HIV infection, it would therefore be interesting to look for polymorphisms in the A3 locus in the human population. Comparison of A3G transcripts in patient cohorts could allow identification of potential uORF-altering SNPs and evaluation of their association with viral control and mutation rates.

Different mechanisms can be considered for Vif-mediated translational inhibition of A3G. Vif could for example inhibit recruitment of ribosomes on the 5'-UTR of A3G mRNA or induce ribosome stalling during translation of the uORF. Our results for instance suggest the Vif-mediated downregulation of certain ribosomal proteins from A3G mRNA-associated protein complexes. In order to evaluate this hypothesis and to study what happens to ribosomes in the presence of Vif, it would be interesting to do ribosomal toeprinting assays on A3G mRNA in the presence and absence of Vif.

In our different experiments, the association of several proteins with the A3G mRNA has been found to be modulated by Vif. On the one hand, it will be necessary to study the role of these

proteins in A3G translation and most importantly in Vif-mediated regulation of A3G translation. This might be achieved by silencing expression of these proteins or by their overexpression in cells followed by analysis of the impact on A3G expression. On the other hand, it would be interesting to study how Vif modulates the presence of these proteins on A3G mRNA, knowing that for the most part, a direct interaction with Vif could not be shown.

Firstly, Vif has been shown previously to be a modulator of the phosphoproteome in infected cells (74) and many of the proteins identified in our study are phosphoproteins. In order to evaluate the possibility that Vif regulates the association of these proteins with A3G mRNA through modification of their phosphorylation status, it would be interesting to study phosphorylation of these proteins in the presence and absence of Vif.

Vif is also a known RNA chaperone, which plays an important role in virion assembly and regulation of reverse transcription (20, 86, 214). This RNA chaperone activity might allow Vif to induce changes in the secondary structure of A3G mRNA, which in turn might modulate binding of different cellular factors. In order to study this possibility, it would be interesting to study the secondary structure of the 5'-UTR of A3G mRNA in the presence and absence of Vif to see whether Vif induces any changes susceptible to impact the associated proteins.

Some of our results indicate that Vif modulates the interactome of A3G mRNA differently in the presence and absence of the A3G protein. This might suggest that Vif can use different mechanisms for translational regulation of A3G and that a switch might exist that induces transfer from one mechanism to the other. In this context, it would be interesting to better study the timeline of Vif-mediated A3G counteraction during an infection. Cells could for example be harvested at different time points after an infection in order to evaluate the contribution of Vif-induced proteasomal degradation and translational inhibition over the course of the infection.

Summary in french
-
Résumé en français

IX. Summary in french / Résumé en français

1. Introduction

Le virus de l'immunodéficience humaine (VIH) a été décrit pour la première fois en 1983 comme l'agent étiologique du syndrome d'immunodéficience acquise (SIDA). Ce virus est aujourd'hui responsable d'une pandémie mondiale avec 37,9 millions de personnes infectées en 2018 et près de 0,8 millions de personnes qui meurent encore chaque année suite à l'infection. Le virus a été introduit dans la population humaine au cours d'au moins 13 évènements de transmission distincts à partir de singes à l'homme, ce qui a donné lieu à différents types, groupes et sous-types du virus. Il existe aujourd'hui un nombre considérable d'agents antirétroviraux qui permettent de maitriser la charge virale pendant très longtemps et ont ainsi permis d'améliorer considérablement la durée de vie des personnes infectées. Néanmoins, le virus peut échapper à ces médicaments, car il évolue rapidement grâce à une reverse transcriptase peu fidèle et un taux élevé de production virale.

Le VIH infecte des cellules du système immunitaire dont principalement les lymphocytes T (LT) $CD4^+$ activées, mais aussi les LT mémoires et quiescents ainsi que les macrophages. Ces cellules sont caractérisées par la présence du récepteur CD4 ainsi qu'un des deux corécepteurs, CCR5 ou CXCR4 à leur surface. Le cycle viral se déroule comme suit et est également représenté en figure F1. La protéine virale de l'enveloppe (Env) se fixe d'abord au CD4 et ensuite à un des corécepteurs, ce qui induit des changements de conformation qui permettent l'insertion d'un peptide de fusion dans la membrane cellulaire. Les membranes virales et cellulaires se rapprochent et fusionnent, permettant la libération du core viral dans le cytoplasme de la cellule hôte. Le VIH est un rétrovirus et de ce fait doit reverse transcrire son génome à ARN (ARNg) simple brin de polarité positive en un ADN double brin. Cette reverse transcription est effectuée par la reverse transcriptase (RT) du virus dans le

cytoplasme de la cellule hôte au sein d'un complexe de protéines virales qui protègent l'ADN néoformé de la détection par les senseurs de l'immunité innée et permettent en même temps l'entrée des nucléotides nécessaires à l'intérieur du complexe. L'ADN viral ainsi formé est ensuite importé dans le noyau ou il est intégré dans le génome de la cellule hôte sous forme d'un provirus. L'intégration est catalysée par l'intégrase virale et s'effectue de préférence au niveau de l'euchromatine à proximité des pores nucléaires et dans des gènes fortement transcrits. Les ARN viraux mono- ou multi-épissés ainsi que l'ARNg non-épissé peuvent ensuite être produits et sont exportés dans le cytoplasme où a lieu la production des protéines virales. L'ARNg est

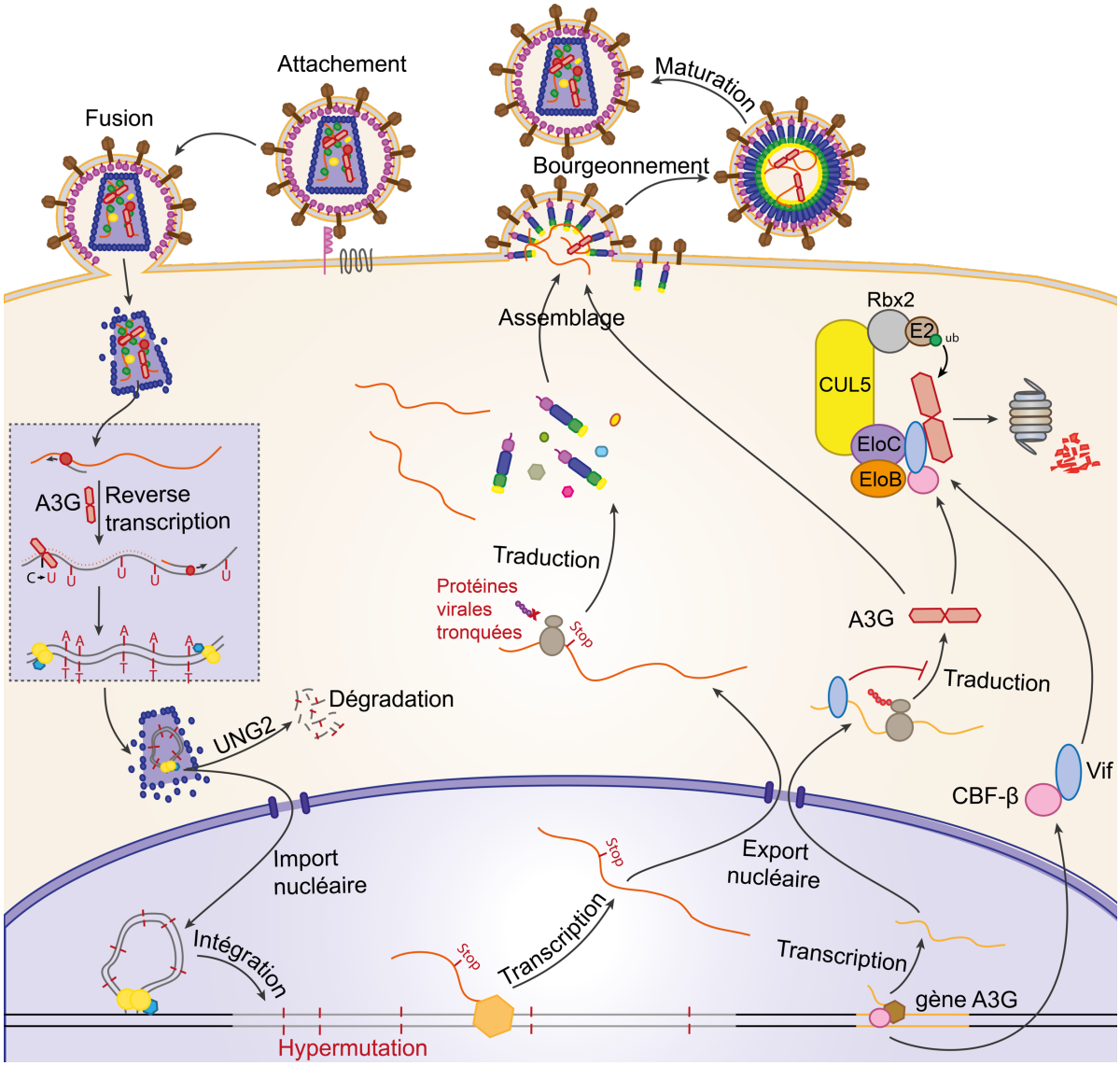


Figure F1 : Cycle de vie du VIH, restriction par A3G et contre-action par Vif*. A3G (rouge) est incorporé dans les virions et induit l'hypermutation de l'ADN proviral ce qui mène soit à sa dégradation ou à la production de protéines virales tronquées et l'arrêt du cycle. Vif (bleu) diminue la transcription d'A3G (1), inhibe sa traduction (2) et induit sa dégradation par le protéasome (3).*

sélectionné par les protéines virales structurales qui s'assemblent au niveau de la membrane plasmique de la cellule et bourgeonnent afin de former de nouvelles particules virales.

Les facteurs de restriction font partie de l'immunité innée et constituent un des premiers mécanismes de défense de l'organisme contre les pathogènes. L'expression de ces protéines est universelle, mais peut être stimulée par l'interféron. Il existe une panoplie de facteurs de restriction qui sont capables d'inhiber les différentes étapes du cycle de vie de multiples virus. APOBEC3G *(Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G*) ou A3G est un des facteurs de restriction majeurs du VIH. A3G est encapsidé dans les nouvelles particules virales et lors de l'étape de reverse transcription au cours du cycle suivant peut induire des mutations de C en U sur le brin d'ADN (-) grâce à son activité désaminase de cytosine. Ceci conduit à l'apparition de mutations de G en A sur le brin d'ADN (+), mutations qui peuvent impacter des séquences régulatrices du génome viral ou bien les séquences codantes, notamment par apparition de codons stop prématurés (Fig. F1). La production d'ARNg défectueux et de protéines virales tronquées mène finalement à l'arrêt du cycle viral. A3G fait partie d'une famille de désaminases de cytosine constituée de 7 membres (A3A, B, C, D, F, G, H), dont 4 montrent une activité de restriction du VIH (D, F, G, H). Ces protéines sont constituées de 1 (A, C, H) ou de 2 (B, D, F, G) domaines conservés à coordination de zinc. Elles exercent leur activité de désamination de cytosine principalement sur l'ADN sb au niveau de sites spécifiques (CC pour A3G, TC pour les autres). Les protéines A3 constituent en même temps une protection et une menace pour la cellule. D'un côté, ils protègent la cellule contre les effets délétères des transposons et contre les virus. De l'autre côté, leur activité représente également un pouvoir mutagène pour le génome de la cellule et doit être strictement contrôlé, notamment au niveau de l'expression de ces protéines ou de leur localisation cellulaire. De plus, les protéines A3 peuvent aussi stimuler l'évolution lorsqu'un faible nombre de mutations est introduit dans l'ADN cible.

Le VIH peut se répliquer efficacement dans les cellules qui expriment A3G grâce à l'expression de la protéine Vif qui diminue fortement le niveau d'expression d'A3G en utilisant 3 mécanismes différents. Premièrement, Vif interagit avec A3G et recrute une

E3 ubiquitine ligase, composée de Cul5, EloB, EloC, CBF-β et Rbx2. Ceci conduit à l'ubiquitination d'A3G suivi de sa reconnaissance et dégradation par le protéasome 26S. La transcription de l'ARNm d'A3G se fait sous le contrôle du facteur de transcription RUNX en association avec son cofacteur CBF-β. L'interaction de Vif avec CBF-β conduit à sa séquestration dans le complexe E3-ubiquitine ligase, ce qui diminue la transcription d'A3G. Finalement, il a été montré que Vif se fixe à l'ARNm d'A3G et inhibe sa traduction. Ce dernier mode d'action de Vif est encore peu étudié et l'objectif de ma thèse a été de contribuer à une meilleure compréhension du mécanisme mis en œuvre par Vif pour inhiber la traduction d'A3G.

L'initiation de la traduction chez les eucaryotes est fortement régulée. L'ARNm est complexé à un certain nombre de protéines, dont le complexe eIF4F qui se fixe à la coiffe et les protéines PABP qui se fixent à la queue poly-A. La petite sous-unité du ribosome est elle-même liée à plusieurs facteurs régulant l'initiation de la traduction (appelés eIFs). La petite sous-unité du ribosome est d'abord recrutée sur la 5'-UTR de l'ARNm par interaction de eIF3 lié au ribosome avec eIF4F lié à l'ARN. A ce stade le ribosome est déjà chargé avec l'ARNt initiateur qui est lié à son acide aminé (Met-ARNt$_i^{Met}$) formant le complexe de pré-initiation (PIC). Le ribosome entame ensuite le balayage de la 5'-UTR de l'ARNm à la recherche du codon initiateur. Lorsque celui-ci est reconnu par appariement avec le Met-ARNt$_i^{Met}$, les eIFs se dissocient et la grande sous-unité du ribosome rejoint le complexe pour démarrer la traduction. Les facteurs d'élongation de la traduction (eEFs) apportent peu à peu les aminoacyl-ARNt correspondants aux codons suivants et la chaine peptidique grandit par catalyse de la formation d'une liaison peptidique au niveau du centre peptidyl transférase du ribosome. Lorsque le ribosome atteint un codon stop (UAA, UGA ou UAG), celui-ci est reconnu par des facteurs de terminaison (eRFs), qui stimulent la libération de la chaine peptidique. Finalement, des facteurs de recyclage rejoignent le ribosome et permettent la libération des deux sous-unités du ribosome ainsi que de l'ARNm.

La traduction est un processus fortement régulé qui fait intervenir de nombreux éléments, comme d'un côté les éléments de séquence et de structure de l'ARNm et de l'autre côté les protéines se liant à l'ARN (RBPs). Parmi ces éléments de séquence

régulateurs on trouve notamment les *upstream ORF* (uORF), c'est à dire des petits cadres de lecture en amont du cadre de lecture principal (mORF). Environ 50 % des transcrits humains contiennent des uORF qui ajoutent un niveau de régulation supplémentaire à ces ARNm avec en général une régulation négative du mORF. Pour la traduction du mORF en présence d'un uORF, différents mécanismes existent : le premier, appelé *leaky scanning* est basé sur le fait que le ribosome ne reconnait pas systématiquement les codons d'initiation qu'il rencontre lors du balayage. De ce fait il est possible que le ribosome n'initie pas la traduction de l'uORF mais continue sa progression jusqu'à atteindre le mORF. Le deuxième mécanisme consiste en la réinitiation après traduction de l'uORF. Ceci est possible lorsque les eIFs ne se dissocient pas entièrement du ribosome lors de la traduction de l'uORF et permettent ainsi au ribosome de recommencer le balayage, de réacquérir un nouvel ARNt initiateur et d'initier par la suite la traduction du mORF. Finalement, le ribosome peut être recruté directement au niveau du mORF par un site d'entrée interne du ribosome (IRES), un élément structural qui permet la traduction coiffe- et parfois aussi scanning-indépendante.

2. Objectifs de cette thèse

Il a été montré précédemment que Vif inhibe la traduction d'A3G. Ce mécanisme en lui-même est responsable d'environ 50 % de la diminution du niveau d'expression d'A3G et il est suffisant pour augmenter significativement l'infectivité du virus. Il a été montré que les tiges-boucles SL2 et SL3 dans la 5'-UTR de l'ARNm d'A3G sont nécessaires pour permettre l'inhibition traductionnelle par Vif. De plus, Vif se fixe sur la 5'-UTR de l'ARNm d'A3G au niveau des tiges-boucles SL1 et SL3. Malgré l'importance de l'inhibition traductionnelle dans la contre-action d'A3G par Vif, le mécanisme mis en œuvre n'est pas encore connu. L'objectif de ma thèse a donc été de contribuer à une meilleure compréhension de ce mode d'action de Vif.

La première partie de ma thèse s'est basée sur la découverte d'un uORF dans la 5'-UTR de l'ARNm d'A3G qui s'étend des tiges boucles SL2 à SL3. L'objectif dans un premier temps était de caractériser l'impact de cet uORF sur le mode de traduction

d'A3G ainsi que sur la régulation traductionnelle par Vif. Dans un deuxième temps nous avons voulu générer une lignée de lymphocytes T où l'uORF d'A3G est supprimée afin d'étudier son effet sur l'infectivité virale.

La deuxième partie de ma thèse avait pour objectif d'identifier des facteurs cellulaires qui pourraient jouer un rôle dans l'inhibition traductionnelle d'A3G par Vif. En effet il a été montré précédemment que Vif peut se fixer sur l'ARNm d'A3G. Il est donc possible que Vif induise par exemple le recrutement de protéines inhibitrices ou qu'il décroche des protéines stimulatrices de la traduction. L'association des protéines en présence et absence de Vif a d'abord été analysée en utilisant la 5'-UTR de l'ARNm d'A3G. Ensuite, l'objectif a été la mise en place d'un protocole qui permette l'identification des partenaires de l'ARNm d'A3G entier.

3. Rôle d'une uORF conservée chez A3G et A3F dans leur régulation traductionnelle par Vif

Il avait été montré précédemment que la 5'-UTR de l'ARNm d'A3G et plus précisément les tiges boucles SL2 et SL3 sont importantes pour permettre l'inhibition traductionnelle d'A3G par Vif. Une analyse *in silico* a révélé la présence d'un petit cadre de lecture ouvert (*upstream ORF*, uORF) qui s'étend de SL2 à SL3 en amont du cadre de lecture codant pour A3G (mORF). Sachant que les uORF sont des éléments régulateurs de la traduction, nous avons généré des mutants afin d'analyser l'impact de l'uORF sur la traduction d'A3G (Fig. F2A). En effet, l'uORF semble négativement réguler la traduction d'A3G, car lorsque l'uORF est supprimé ou que son codon d'initiation est muté, l'expression d'A3G augmente sensiblement. Lorsque le codon stop de l'uORF est muté et que l'uORF est placée en phase avec l'ORF codant A3G, une forme longue, issue de l'initiation de la traduction à partir du codon AUG de l'uORF, et une forme courte d'A3G, issue de l'initiation à partir du codon AUG du mORF, sont produits. Ceci suggère une traduction d'A3G par leaky scanning car le ribosome semble pouvoir initier la traduction à partir des deux codons d'initiation. Néanmoins, des changements de la taille de l'uORF et de sa distance par rapport au mORF impactent également la traduction d'A3G, ce qui suggère qu'un mécanisme de

réinitiation de la traduction est également mis en œuvre. Par contre, aucun IRES n'a été identifié dans la 5'-UTR de l'ARNm d'A3G. Ces résultats montrent qu'A3G est traduit en présence de l'uORF par un mélange de leaky scanning et de réinitiation.

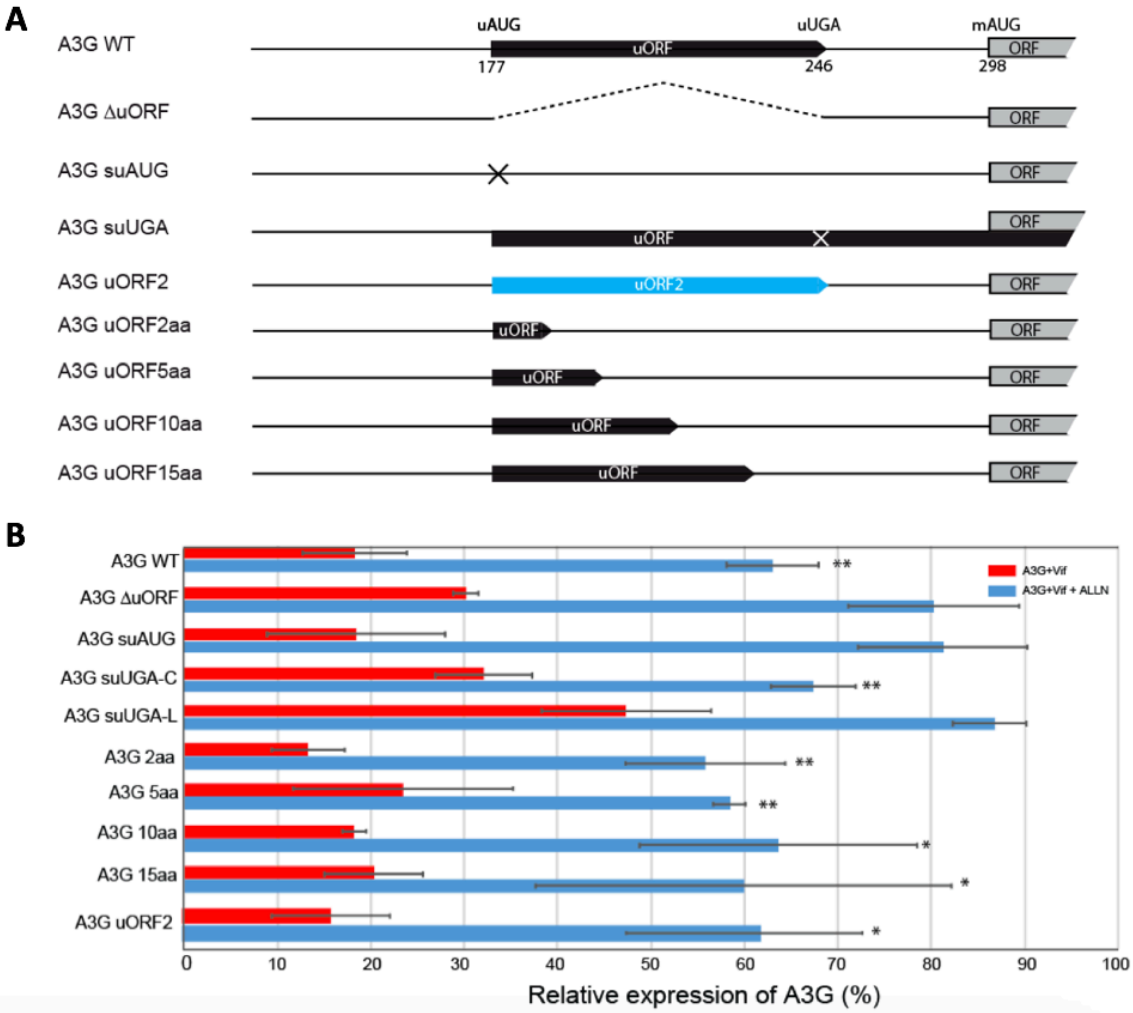


Figure F2 : Etude de l'importance de l'uORF contenu dans la 5'-UTR de l'ARNm d'A3G pour son inhibition traductionnelle par Vif. (A) *La 5'-UTR de l'ARNm d'A3G ainsi que les différents mutants utilisés sont représentés schématiquement. ΔuORF : délétion de l'uORF ; suAUG : mutation du codon d'initiation de l'uORF ; suUGA : mutation du codon stop de l'uORF (C : forme courte ; L : forme longue) ; uORF2 : changement de la séquence de l'uORF ; uORF2-15aa : raccourcissement de l'uORF à la taille indiquée.* ***(B)*** *Analyse de l'expression d'A3G à partir des différentes constructions en présence de Vif et d'un inhibiteur du protéasome (ALLN). L'expression d'A3G a été analysée par western blot, les bandes obtenues ont été quantifiées par ImageJ. Les barres d'erreurs sont représentatives de trois expériences. Les p-valeurs sont indiquées avec des étoiles : * < 0,05 ; ** < 0,01.*

L'uORF est conservée chez A3G et A3F mais chez aucun des autres A3. Nous avons montré qu'A3G et A3F sont régulés par Vif au niveau traductionnel, néanmoins cette inhibition traductionnelle n'a pas lieu chez les autres A3. Ceci suggère que l'inhibition traductionnelle d'A3G par Vif est dépendante de l'uORF. En effet nous avons montré par la suite que lorsque l'uORF d'A3G est supprimé ou que son codon d'initiation est

muté, Vif ne peut plus inhiber la traduction d'A3G (Fig. F2B, ΔuORF et suAUG), tandis qu'aucun effet significatif n'a été observé sur l'inhibition traductionnelle d'A3G par Vif pour les autres mutants de l'uORF. Lorsque le codon stop de l'uORF est muté, Vif est capable d'inhiber la traduction de la forme courte mais non de la forme longue (Fig. F2B, suUGA-C et suUGA-L), ce qui suggère que Vif n'agit pas sur la traduction à partir du codon AUG de l'uORF.

Finalement, la localisation de l'ARNm d'A3G a été étudiée dans la cellule en présence et absence de l'uORF par microscopie confocale. Cette étude a confirmé qu'en conditions de stress induites par traitement des cellules à l'arsénite ou à la chaleur, la protéine A3G est localisée dans les granules de stress. L'ARNm d'A3G est également localisée dans les granules de stress, mais uniquement en présence de Vif. Cette localisation est perdue lorsque l'uORF est supprimé.

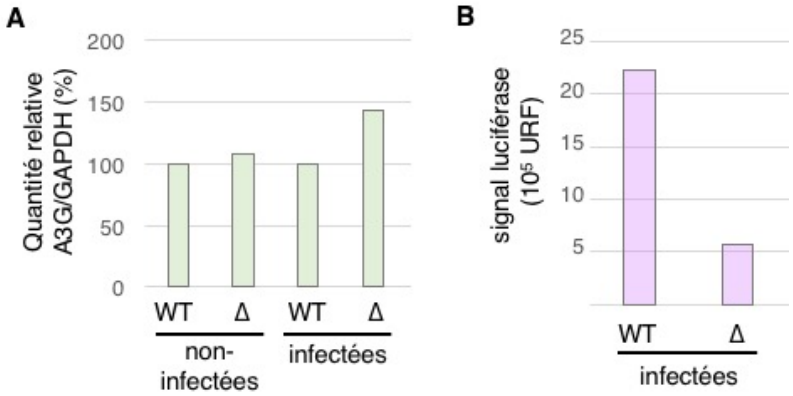


Figure F3 : Infection de cellules H9-WT et ΔuORF par le VIH-1. *Les cellules H9 sauvages (WT) et ΔuORF ont été infectées avec le VIH-1 (isolat LAI).* ***(A)*** *L'expression d'A3G dans les cellules infectées a été analysée par western blot et les bandes ont été quantifiées par ImageJ. La GAPDH a été utilisée comme témoin de charge. La quantité d'A3G exprimée dans les cellules WT non-infectées a été fixée à 100 %.* ***(B)*** *Les virus produits dans les cellules WT et ΔuORF ont été utilisés pour l'infection de cellules TZM-bl. Le signal luciférase, qui est représentatif du nombre de cellules infectées, a été mesuré et est indiqué en unités relatives de fluorescence (URF).*

Afin de pouvoir étudier l'effet de l'uORF sur l'infection par le VIH, nous avons cherché à supprimer l'uORF dans le génome de cellules H9, une lignée cellulaire de lymphocytes T. Pour ce faire nous avons utilisé la technologie CRISPR/Cas9, ce qui nous a permis l'obtention d'un clone où l'uORF d'A3G est supprimé (ΔuORF) tandis que celle d'A3F reste inchangé. Le clone obtenu est hétérozygote, c'est à dire qu'il possède un allèle sauvage et un allèle ΔuORF du gène A3G. Néanmoins, au niveau

des transcrits nous avons détecté quasi exclusivement l'allèle ΔuORF. Une première expérience d'infection de ces cellules a montré une légère augmentation du niveau d'A3G dans les cellules ΔuORF infectées comparé au WT. En même temps, une forte baisse de la quantité de virus produits a pu être observée (Fig. F3). Ces résultats, bien que préliminaires, suggèrent un impact de l'uORF sur l'infection virale.

4. Identification de protéines qui interagissent avec la 5'-UTR de l'ARNm d'A3G

Afin de mieux comprendre le mécanisme utilisé par Vif pour inhiber la traduction d'A3G, nous avons voulu identifier les protéines qui se fixent à sa 5'-UTR afin d'étudier si Vif modifie l'interactome protéique de cet ARN. Pour cela j'ai utilisé un protocole récemment publié pour la détection des interactions ARN-protéines (RaPID). Ce protocole se base sur la biotinylation de protéines à une distance de 10 nm de la biotine ligase bactérienne BirA. BirA est fusionnée à la protéine λN qui se fixe sur les tiges boucles BoxB. La 5'-UTR de l'ARNm d'A3G est exprimée en fusion avec 3 tiges-boucles BoxB de part et d'autre. Le recrutement de BirA sur cet ARN grâce à l'interaction λN-BoxB permet la biotinylation des protéines qui interagissent avec la 5'-UTR de l'ARNm d'A3G. Ces protéines sont ensuite isolées sur des billes magnétiques recouvertes de streptavidine suivi de leur identification par spectrométrie de masse.

Cette approche nous a permis d'identifier entre 581 et 916 protéines dans chaque échantillon. Parmi ces protéines, ce sont surtout des RBP (RNA Binding Protein) qui sont enrichies, comme par exemple des facteurs d'épissage et de traduction. La comparaison des protéines identifiées en présence et absence de l'uORF a permis l'identification de 11 protéines enrichies en la présence et 15 protéines enrichies en l'absence de l'uORF (Fig. F4A). Parmi ces protéines, certaines ont une fonction connue dans la régulation de la traduction, la stabilité des ARNm, l'export nucléaire des ARN et l'épissage. La comparaison des protéines identifiées en présence et absence de Vif a permis l'identification de 14 protéines dont l'association avec l'ARNm d'A3G semble diminuer en présence de Vif (Fig. F4B). De façon intéressante, 4 de ces protéines (RL24, RL27, AP3D1 et THOC2) sont aussi spécifiquement associées avec

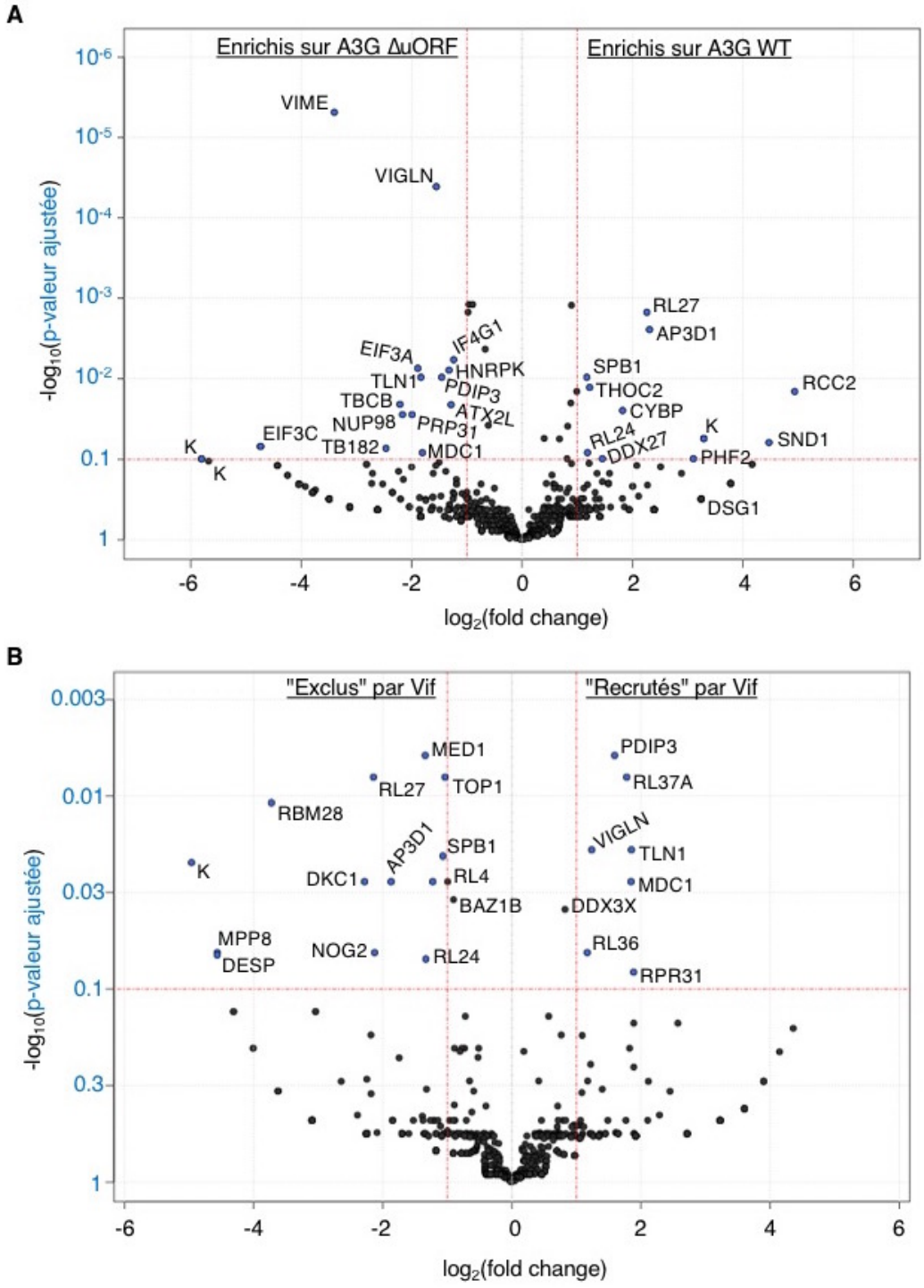


Figure F4 : ***Influence de Vif et de l'uORF sur les protéines interagissant avec la 5'-UTR de l'ARNm d'A3G.*** *La comparaison des protéines fixées à la 5'-UTR de l'ARNm d'A3G dans deux conditions différentes a été effectuée en utilisant un modèle de régression binomiale négative. Une p-value ajustée par la méthode Benjamini-Hochberg < 0,1 ainsi qu'une diminution ou augmentation d'au moins deux fois (fold change > 2) ont été utilisés comme valeurs limites et sont indiquées avec des lignes rouges sur l'axe des y et des x respectivement. Les noms des protéines sont indiqués à côté des points correspondants, les kératines sont marquées "K".* ***(A)*** *Comparaison de protéines fixées à la 5'-UTR de l'ARNm d'A3G WT et ΔuORF en l'absence de Vif. Les protéines enrichies en présence de l'uORF sont indiquées à droite et celles enrichies en l'absence de l'uORF à gauche.* ***(B)*** *Comparaison des protéines fixées à la 5'-UTR de l'ARNm d'A3G en présence et absence de Vif. Les protéines enrichies en présence de Vif sont indiquées à droite et celles enrichies en l'absence de Vif à gauche.*

l'ARNm d'A3G WT en comparaison du ΔuORF. Sept protéines semblent recrutées sur l'ARNm d'A3G par Vif, dont par exemple PDIP3, RL37A et MDC1. L'interaction de Vif avec ces trois protéines a été étudiée par co-immunoprécipitation sur des lysats cellulaires cross-linkés ou non. Cette expérience n'a pas pu confirmer l'interaction entre Vif et ces protéines.

Tandis que le présent protocole nous a donné des indications intéressantes sur la dépendance de l'interactome de l'ARNm d'A3G de l'uORF et de Vif, le nombre de protéines identifiées spécifiquement était plutôt bas. Ceci pourrait être dû à des contraintes imposés par BirA qui est capable de biotinyler des protéines à seulement 10 nm de distance. Ce protocole étant optimal pour de petits ARN d'intérêt, la 5'-UTR de l'ARNm d'A3G est bien plus grande et de ce fait la méthode a probablement permis d'identifier qu'une petite partie des protéines qui y sont associées. Pour cela nous avons voulu mettre en place un protocole pour pourvoir étudier l'ARNm d'A3G entier.

5. Mise en place d'un protocole pour identifier les protéines associées avec l'ARNm d'A3G entier.

Afin d'étudier l'interactome protéique de l'ARNm d'A3G entier, nous avons voulu mettre au point un protocole de pull-down. Pour cela, l'ARNm d'A3G a été exprimé dans les cellules HEK293T. L'objectif était d'isoler l'ARNm d'A3G ainsi que les protéines qui y sont associées à partir d'un lysat cellulaire en utilisant des oligonucléotides complémentaires, biotinylés et couplés à des billes magnétiques recouvertes de streptavidine.

Tout d'abord nous avons cherché à optimiser le protocole en utilisant différents types de billes. Pour étudier la rétention de l'ARNm d'A3G sur les billes en fin de protocole de pull-down, une détection par RT-PCR a été effectuée. Ces expériences ont révélé une très mauvaise spécificité de rétention de l'ARNm d'A3G par les oligos complémentaires. En effet, le ratio de l'ARNm d'A3G retenu en présence des oligos comparé aux billes nues était proche de 1 pour tous les types de billes testés (Fig. F5), ce qui signifie qu'autant d'ARNm d'A3G était retenu en présence qu'en absence des oligos spécifiques. La bande détectée en l'absence d'oligos a été purifiée et

séquencée, et a montré qu'il s'agissait bien de l'ARNm d'A3G et non d'un autre contaminant aspécifique. Différentes approches ont été mis en œuvre afin d'éliminer les ARN se fixant aspécifiquement aux billes. Ni des lavages de haute stringence ni la saturation des billes ou le pré-clearing du lysat n'a montré d'effet.

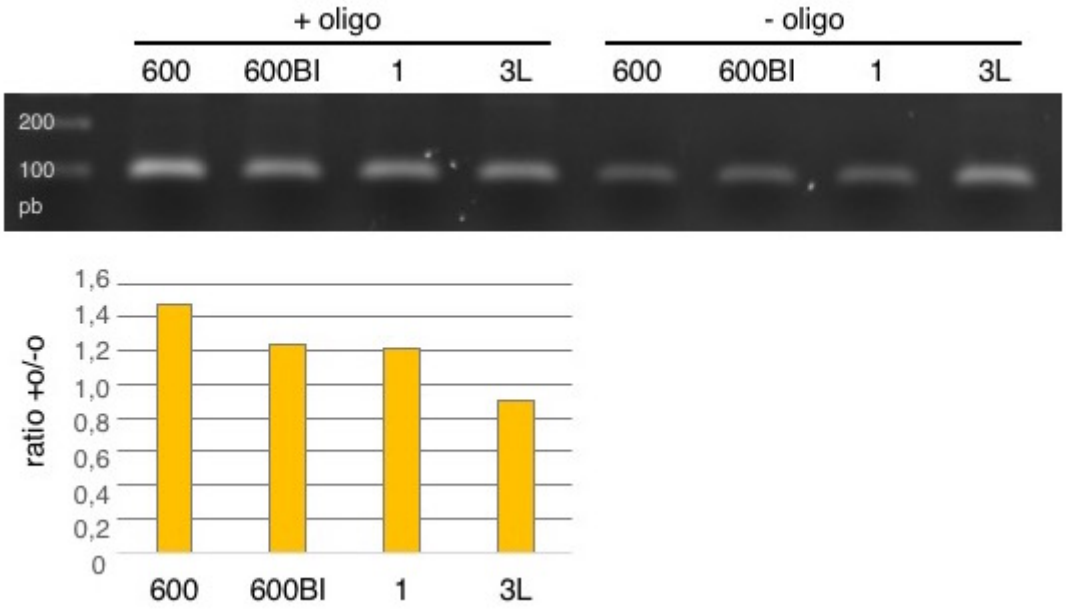


Figure F5 : Spécificité de rétention de l'ARNm d'A3G sur les billes magnétiques*. Le protocole de pull-down a été effectué avec un ARNm d'A3G transcrit in vitro sur différents types de billes (600, 600BI, 1 et 3L du kit MagSi-STA (MagnaMedics)) couplées aux oligos complémentaires à l'ARNm d'A3G (+ oligo) ou sur les billes seules (- oligo). A la fin du protocole, l'ARNm d'A3G a été amplifié par RT-PCR, les amplicons ont été chargés sur un gel d'agarose 0,8 % (TBE 0,5 x) et les bandes obtenues ont été quantifiées par ImageJ. Le ratio entre la bande obtenue en présence de l'oligo comparé au contrôle billes seules a été calculé.*

Une variation du ratio billes-lysat n'a pas non plus permis d'obtenir une amélioration du problème. Différents protocoles d'élution spécifique de l'ARNm d'A3G des billes en fin de pull-down ont donc été testés. Tout d'abord nous avons utilisé des oligos desthiobiotinylés, qui se fixent aux billes avec une moindre affinité et peuvent être décrochés des billes par ajout de biotine libre. Or nos résultats ont montré que l'élution de ces oligos était très inefficace même à de très fortes concentrations en biotine. De plus, une quantité comparable d'ARNm d'A3G se décrochait des billes même en absence des oligos. Nous avons ensuite essayé d'éluer l'ARNm d'A3G en utilisant la RNase H qui dégrade spécifiquement l'ARN hybridé à de l'ADN. Ainsi des fragments de l'ARNm d'A3G hybridés aux oligos complémentaires devraient être libérés des billes. Tandis que ce protocole semblait fonctionner sur un ARNm d'A3G transcrit *in vitro*, une grande variabilité a été constatée lorsque le protocole a été réalisé sur des

ARN cellulaires ou un lysat cellulaire total. Finalement, un protocole d'élution a été mis en œuvre qui utilise un oligo compétiteur. En effet l'ARNm d'A3G est fixé sur les billes par un oligo de capture qui est composé de 20 nucléotides complémentaires à l'ARNm d'A3G ainsi que de 10 nucléotides supplémentaires à son extrémité 5'. L'ajout d'un oligo compétiteur qui s'hybride sur toute la longueur des 30 nucléotides provoque le déplacement de l'ARNm d'A3G des billes. Tandis que ce protocole fonctionne très efficacement sur un ARN transcrit *in vitro* et a permis d'isoler jusqu'à 80 fois plus d'ARN avec l'oligo capture comparé à un oligo à séquence aléatoire (scramble), seule une très faible quantité d'ARNm d'A3G, identique entre l'oligo spécifique et l'oligo scramble, a pu être isolée d'un lysat cellulaire. En effet nous avons montré par la suite que même l'ajout de lysat sur un ARNm transcrit *in vitro* qui était déjà préalablement fixé aux billes induisait le décrochage de cet ARN des billes (Fig. F6). Ceci suggère que l'hybridation entre l'oligo et l'ARNm d'A3G n'est pas assez forte pour maintenir cet ARN fixé aux billes en présence des autres composants cellulaires qui probablement entrent en compétition avec l'oligo pour la fixation à l'ARNm d'A3G.

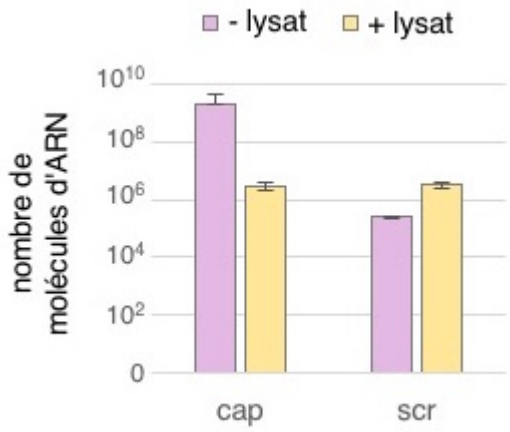


Figure F6 : Pull-down de l'ARNm d'A3G suivi de l'élution spécifique par un oligo compétiteur. *Un ARNm d'A3G transcrit in vitro a d'abord été fixé aux billes magnétiques (Speedbeads, Thermo Fisher) grâce à un oligo complémentaire, puis a été incubé en présence d'un lysat cellulaire (+ lysat) ou non (- lysat). L'ARNm d'A3G a ensuite été élué en utilisant un oligo compétiteur suivi de sa quantification par RT-qPCR. n=3.*

Par conséquent nous avons décidé d'effectuer le pull-down avec un ARNm d'A3G transcrit *in vitro*, coiffé, poly-adénylé et biotinylé. Ceci permet la fixation de l'ARN sur les billes par interaction directe des nucléotides biotinylés, intégrés dans la séquence de l'ARN, avec les billes. Cet ARN fixé aux billes est ensuite incubé avec un lysat cellulaire afin de permettre aux interactants de s'y fixer. Avec ce protocole, une

quantité considérable de l'ARNm d'A3G reste fixée sur les billes jusqu'à la fin du protocole même après ajout de lysat. Ce protocole a donc été choisi pour les analyses de l'interactome protéique de l'ARN.

6. Identification de facteurs cellulaires impliqués dans l'inhibition traductionnelle d'A3G par Vif

Des protéines cellulaires ont été isolées par pull-down sur un ARNm d'A3G transcrit *in vitro*, coiffé, poly-adénylé et biotinylé, suivi de leur identification par spectrométrie de masse. Des billes en l'absence de l'ARNm d'A3G ont été utilisées comme contrôle négatif. Nous avons ainsi déterminé l'interactome protéique de l'ARNm d'A3G en présence et absence de la protéine A3G et de Vif. Parmi les protéines identifiées, 109, soit 24,4 % des protéines totales, sont associées avec l'ARNm d'A3G indépendamment de la condition utilisée tandis que les autres protéines semblaient modulées par la présence des protéines A3G et Vif (Fig. F7). Parmi les 109 protéines qui constituent l'interactome de base de l'ARNm d'A3G, nous avons retrouvé des protéines impliquées dans toutes les différentes étapes du cycle de vie d'un ARNm.

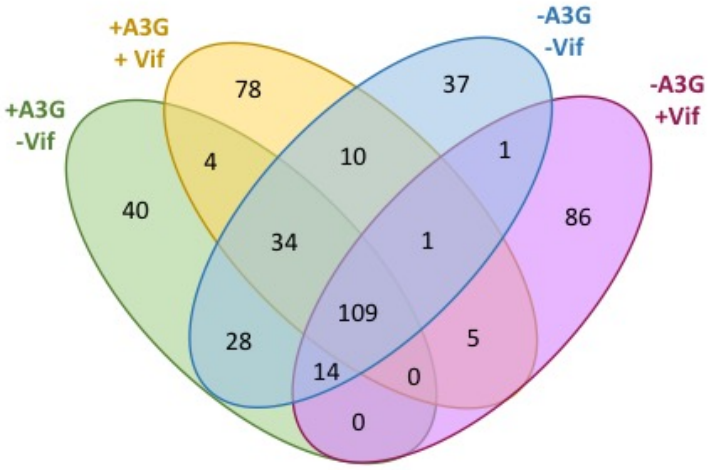


Figure F7 : Protéines significativement associées avec l'ARNm d'A3G dans des conditions différentes. *Le pull-down de protéines a été effectué à partir de lysat cellulaire exprimant les protéines Vif et/ou A3G. Les protéines enrichies au moins 2 fois en présence de l'ARNm d'A3G comparé au contrôle de billes seules avec une p-value < 0,1 sont représentées dans un diagramme de Venn.*

Une analyse d'expression différentielle nous a permis l'identification de 83 protéines dont la présence sur l'ARNm d'A3G change significativement en présence de la

protéine A3G dans les cellules. Parmi les protéines qui semblent recrutées par A3G sur son propre ARNm, nous avons notamment identifié des composants des P-bodies et des granules de stress. Ces résultats suggèrent qu'A3G régule l'interactome de son propre ARNm, ce qui pourrait contribuer à la régulation de son expression.

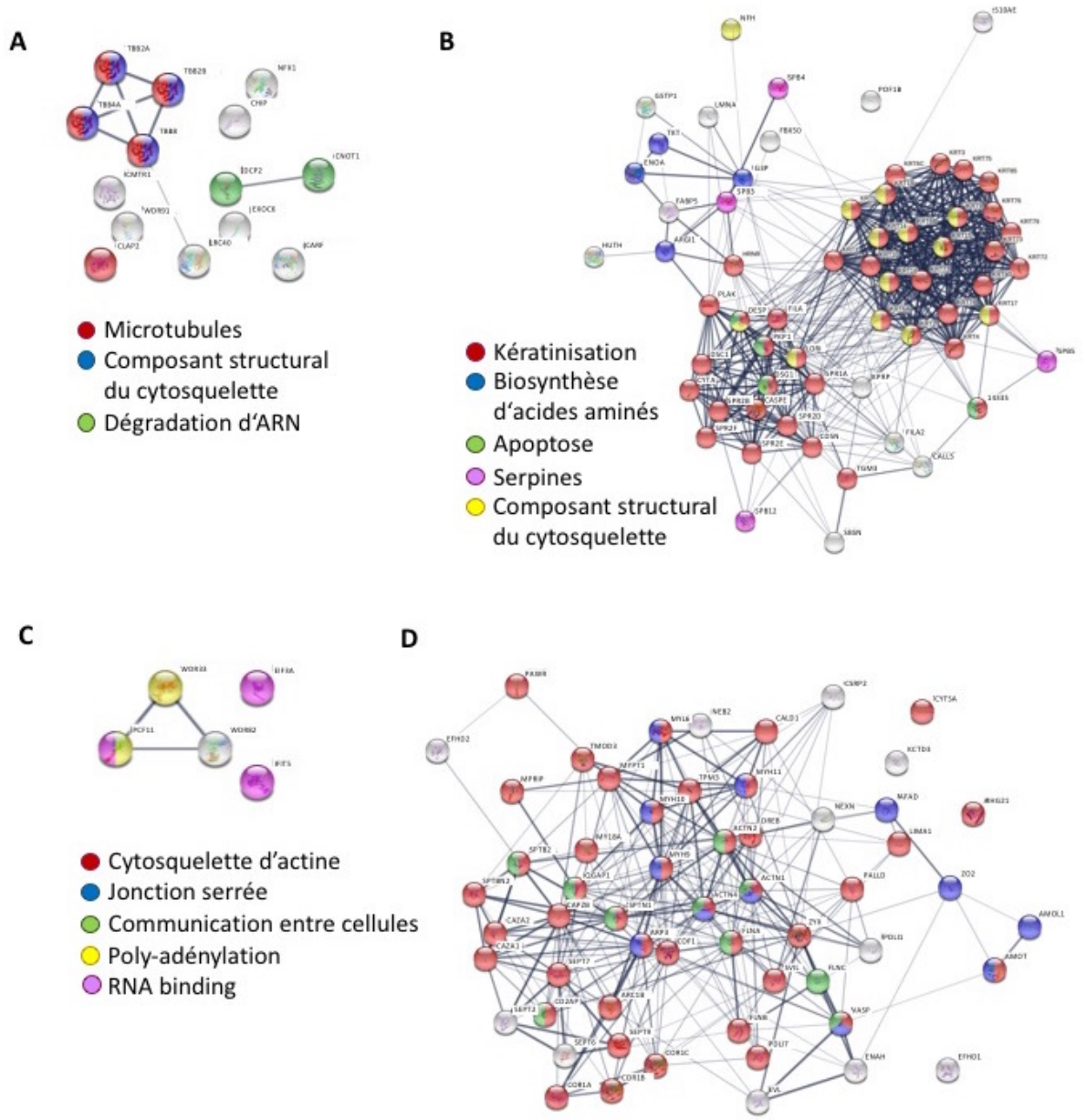


Figure F8 : Protéines régulées sur l'ARNm d'A3G par Vif. *Les protéines associées avec l'ARNm d'A3G dans différentes conditions ont été identifiées par spectrométrie de masse suivi d'un test d'expression différentiel en fonction de la présence de la protéine Vif. (A) Protéines régulées négativement par Vif en présence de la protéine A3G. (B) Protéines régulées positivement par Vif en présence de la protéine A3G. (C) Protéines régulées négativement par Vif en absence de la protéine A3G. (D) Protéines régulées positivement par Vif en absence de la protéine A3G. Les protéines sont représentées sous forme d'un diagramme string. L'épaisseur des traits représente la force de l'association entre deux protéines. Les principales fonctions biologiques sont indiquées par un code couleur. La légende de C et D est commune.*

Nous avons également identifié 146 protéines dont la présence sur l'ARNm d'A3G est régulée par Vif (Fig. F8). Parmi ces protéines, 68 sont régulées par Vif exclusivement en absence de la protéine A3G (Fig. F8C et D) tandis que 78 sont régulées en présence d'A3G (Fig. F8A et B) et une seule (DSG1) est régulée dans les deux conditions. Ceci suggère que Vif pourrait utiliser des mécanismes différents pour réguler la traduction d'A3G en présence et absence de la protéine A3G. De façon intéressante, 88 des 147 protéines régulées par Vif sont des phosphoprotéines et peuvent donc être régulées par phosphorylation.

Il a été montré précédemment qu'une petite uORF dans la 5'-UTR de l'ARNm d'A3G est nécessaire pour permettre à Vif d'inhiber la traduction d'A3G. Afin d'évaluer l'impact de cette uORF sur l'interactome de l'ARNm d'A3G, l'expérience a été répétée avec un ARNm d'A3G où l'uORF est supprimé. Des résultats préliminaires indiquent que parmi les 147 protéines régulées par Vif sur l'ARNm d'A3G WT, 60 sont également régulées en l'absence de l'uORF. Dans l'ensemble, l'uORF semble avoir un impact majeur sur l'interactome de l'ARNm d'A3G ainsi que sur les protéines dont la présence sur l'ARNm d'A3G peut être influencée par Vif.

Finalement, 10 protéines parmi celles qui sont régulées par Vif sur l'ARNm d'A3G ont été sélectionnées pour des analyses supplémentaires. Tout d'abord nous avons évalué l'interaction de ces protéines avec Vif grâce à des expériences de co-immunoprécipitation. Nos résultats ont montré que parmi les protéines sélectionnées, seules CHIP et RENT1 sont capables d'interagir avec Vif. Pour étudier le rôle de ces protéines dans la régulation de la traduction d'A3G, leur expression a été diminuée par RNA silencing. Cependant, nos résultats préliminaires n'ont montré aucun impact du silencing de ces protéines sur l'expression d'A3G ni en présence, ni en absence de Vif.

7. Discussion

A3G est un facteur de restriction du VIH qui induit l'hypermutation du génome viral et provoque l'arrêt du cycle viral en absence de Vif. Vif contrecarre A3G de 3 manières différentes, dont une est l'inhibition de sa traduction. Il avait été montré précédemment que la 5'-UTR d'A3G et plus particulièrement les tiges-boucles SL2 et SL3 qui contiennent une petite uORF sont importantes pour permettre à Vif d'inhiber la traduction d'A3G.

Le premier objectif de ma thèse a été de caractériser le mode de traduction d'A3G en présence de l'uORF et d'étudier l'impact de l'uORF sur l'inhibition traductionnelle par Vif. Cette étude nous a permis de définir qu'A3G est traduit par un mélange de leaky scanning et de réinitiation de la traduction. Il a également été montré que l'uORF est essentielle pour l'inhibition traductionnelle d'A3G par Vif et que la mutation de son codon d'initiation est suffisante pour abolir ce mode d'action de Vif. Parmi les autres A3, l'uORF a été conservée uniquement chez A3F, et celui-ci est le seul à part A3G qui subit une régulation traductionnelle par Vif. Nous avons également montré qu'en conditions de stress, Vif induit la relocalisation de l'ARNm d'A3G dans les granules de stress de façon uORF-dépendante.

Le mécanisme utilisé par Vif pour inhiber la traduction d'A3G est encore inconnu. Sachant que Vif se fixe à la 5'-UTR de l'ARNm d'A3G, il est possible que Vif interagisse ainsi avec les complexes protéiques qui y sont associés. Vif pourrait par exemple recruter des facteurs inhibiteurs ou induire le décrochage de facteurs stimulateurs de la traduction de la 5'-UTR d'A3G. Afin d'étudier cette hypothèse, nous avons étudié les protéines associées avec l'ARNm d'A3G en présence et absence de Vif par deux protocoles différents.

Dans un premier temps, le protocole RaPID nous a permis d'identifier un nombre considérable de protéines qui semblent régulés par Vif ou par l'uORF. Nous pouvons en déduire que les deux sont importants pour définir l'interactome protéique de la 5'-UTR de l'ARNm d'A3G. De façon intéressante, deux protéines ribosomales, RL24 et

RL27 sont enrichis sur l'ARNm d'A3G en présence de l'uORF. Ceci indique que ces protéines pourraient être importantes pour permettre la traduction d'A3G en présence de l'uORF. En effet, il a été décrit précédemment que les ribosomes changent de composition en fonction des transcrits qu'ils traduisent. RL24 et RL27 pourraient donc être des protéines dont la présence dans le ribosome est modulée de cette façon. De plus, RL24 a été associé chez les plantes avec la stimulation de la traduction de transcrits qui contiennent une uORF, en augmentant le taux de réinitiation de la traduction. De surcroit, nous avons également montré que l'association de RL24 avec la 5'-UTR d'A3G diminue en présence de Vif. RL24 pourrait donc être une protéine nécessaire à la traduction de l'ARNm d'A3G en présence de l'uORF en favorisant la réinitiation. En diminuant la présence de RL24 sur l'ARNm d'A3G, Vif pourrait induire une diminution de réinitiation qui se répercute finalement dans une baisse de la traduction d'A3G.

Le protocole de pull-down de protéines sur un ARNm d'A3G biotinylé nous a permis d'identifier un grand nombre de protéines dont la présence semble régulée par Vif, par la protéine A3G ou par l'uORF. Vif semble recruter certaines protéines sur l'ARNm d'A3G et diminuer la présence de certaines autres protéines. Parmi les protéines régulées par Vif, beaucoup sont des composants du cytosquelette. En effet il a été montré auparavant que le cytosquelette joue un rôle important dans l'infection par le VIH et que celui-ci semble moduler le cytosquelette en sa faveur. Il est donc possible que Vif interagisse en effet avec le cytosquelette. De plus, de nombreuses phosphoprotéines ont été identifiées. En effet Vif est connue pour provoquer des changements dans le phosphoprotéome de la cellule et ceci pourrait permettre à Vif d'impacter l'activité de facteurs cellulaires impliqués dans la traduction d'A3G. De plus, nos résultats suggèrent que Vif ne régule pas les mêmes protéines en présence et absence de la protéine A3G, ce qui suggère que Vif pourrait utiliser différents mécanismes en fonction du niveau d'expression d'A3G. Lorsque Vif est exprimée dans une cellule nouvellement infectée, il y a un grand pool de protéines A3G qui doit être rapidement diminué par Vif. Ensuite, un niveau faible d'expression d'A3G doit être maintenu. En effet, il serait possible qu'un switch s'opère entre les deux conditions qui fait que Vif adopte un mode d'action différent. Parmi les protéines régulées par Vif,

une interaction avec Vif n'a pu être confirmée que pour une petite partie des protéines. Il est possible que Vif régule l'association de ces protéines avec l'ARNm d'A3G sans interaction directe, par exemple en induisant des changements structuraux de l'ARN ou encore en modifiant l'état de phosphorylation des protéines ciblées.

8. Conclusion et Perspectives

Cette thèse a contribué à montrer l'importance d'une petite uORF dans la régulation de la traduction d'A3G et d'A3F en présence et absence de Vif. La lignée de lymphocytes T A3G ΔuORF générée au cours de cette thèse permettra d'étudier plus en détail l'impact de cet uORF sur l'infection par le VIH. Cette thèse a également permis l'identification de l'interactome protéique de l'ARNm d'A3G entier et de sa 5'-UTR en particulier. Un impact important de Vif, de la protéine A3G ainsi que de l'uORF sur cet interactome a été montré et pourrait jouer un rôle significatif dans la régulation de l'expression d'A3G. Un nombre considérable de protéines a été identifié et elles pourraient être récrutées par Vif sur l'ARNm d'A3G ou être exclus par Vif des complexes associés avec l'ARNm d'A3G. Afin d'étudier l'effet de ces protéines sur l'inhibition traductionnelle d'A3G par Vif, nous envisageons de diminuer l'expression de ces protéines par RNA silencing ou de l'augmenter par surexpression. De plus, Vif est un chaperon d'ARN connu et de ce fait pourrait réguler l'association des protéines sur l'ARNm d'A3G en modifiant la structure de ces ARN. Afin d'étudier cette hypothèse il serait intéressant d'étudier la structure secondaire de l'ARNm d'A3G en présence et absence de Vif. Finalement, il serait intéressant de connaitre plus en détail l'étape dans l'initiation de la traduction d'A3G sur laquelle Vif agit. Par conséquent, des expériences de toeprinting du ribosome sur la 5'-UTR de l'ARNm d'A3G en présence et absence de Vif pourraient permettre de définir à quel niveau les ribosomes sont impactés par Vif.

Bibliography

X Bibliography

1. **Adachi S**, **Natsume T**. Purification of noncoding RNA and bound proteins using FLAG peptide-conjugated antisense-oligonucleotides. *Methods Mol Biol Clifton NJ* 1262: 265–274, 2015.
2. **Altfeld M**, **Fadda L**, **Frleta D**, **Bhardwaj N**. DCs and NK cells: critical effectors in the immune response to HIV-1. *Nat Rev Immunol* 11: 176–186, 2011.
3. **Altfeld M**, **Gale M**. Innate immunity against HIV-1 infection. *Nat Immunol* 16: 554–562, 2015.
4. **Ambrose Z**, **Aiken C**. HIV-1 Uncoating: Connection to Nuclear Entry and Regulation by Host Proteins. *Virology* 0: 371–379, 2014.
5. **Anderson BD**, **Harris RS**. Transcriptional regulation of APOBEC3 antiviral immunity through the CBF-β/RUNX axis. *Sci Adv* 1: e1500296, 2015.
6. **Anderson P**, **Kedersha N**. Stress granules: the Tao of RNA triage. *Trends Biochem Sci* 33: 141–150, 2008.
7. **Aoshi T**, **Koyama S**, **Kobiyama K**, **Akira S**, **Ishii KJ**. Innate and adaptive immune responses to viral infection and vaccination. *Curr Opin Virol* 1: 226–232, 2011.
8. **Archin NM**, **Sung JM**, **Garrido C**, **Soriano-Sarabia N**, **Margolis DM**. Eradicating HIV-1 infection: seeking to clear a persistent pathogen. *Nat Rev Microbiol* 12: 750–764, 2014.
9. **Arias JF**, **Koyama T**, **Kinomoto M**, **Tokunaga K**. Retroelements versus APOBEC3 family members: No great escape from the magnificent seven. *Front Microbiol* 3: 275, 2012.
10. **Arif A**, **Yao P**, **Terenzi F**, **Jia J**, **Ray PS**, **Fox PL**. The GAIT translational control system. *Wiley Interdiscip Rev RNA* 9, 2018.
11. **Arts EJ**, **Hazuda DJ**. HIV-1 antiretroviral drug therapy. *Cold Spring Harb Perspect Med* 2: a007161, 2012.
12. **Aylett CHS**, **Ban N**. Eukaryotic aspects of translation initiation brought into focus. *Philos Trans R Soc Lond B Biol Sci* 372, 2017.
13. **Bai Q**, **Bai Z**, **Sun L**. Detection of RNA-binding Proteins by In Vitro RNA Pull-down in Adipocyte Culture. *J Vis Exp JoVE* , 2016.
14. **Barbosa C**, **Peixeiro I**, **Romão L**. Gene expression regulation by upstream open reading frames and human disease. *PLoS Genet* 9: e1003529, 2013.
15. **Barnes C**, **Kanhere A**. Identification of RNA-Protein Interactions Through In Vitro RNA Pull-Down Assays. *Methods Mol Biol Clifton NJ* 1480: 99–113, 2016.
16. **Barra J**, **Leucci E**. Probing Long Non-coding RNA-Protein Interactions. *Front Mol Biosci* 4, 2017.
17. **Barré-Sinoussi F**, **Chermann JC**, **Rey F**, **Nugeyre MT**, **Chamaret S**, **Gruest J**, **Dauguet C**, **Axler-Blin C**, **Vézinet-Brun F**, **Rouzioux C**, **Rozenbaum W**, **Montagnier L**. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220: 868–871, 1983.
18. **Barré-Sinoussi F**, **Ross AL**, **Delfraissy J-F**. Past, present and future: 30 years of HIV research. *Nat Rev Microbiol* 11: 877–883, 2013.
19. **Barton K**, **Winckelmann A**, **Palmer S**. HIV-1 Reservoirs During Suppressive Therapy. *Trends Microbiol* 24: 345–355, 2016.
20. **Batisse J**, **Guerrero S**, **Bernacchi S**, **Sleiman D**, **Gabus C**, **Darlix J-L**, **Marquet R**, **Tisné C**, **Paillart J-C**. The role of Vif oligomerization and RNA chaperone activity in HIV-1 replication. *Virus Res* 169: 361–376, 2012.
21. **Beckmann BM**, **Castello A**, **Medenbach J**. The expanding universe of ribonucleoproteins: of novel RNA-binding proteins and unconventional interactions. *Pflugers Arch* 468: 1029–1040, 2016.
22. **Beckmann K**, **Grskovic M**, **Gebauer F**, **Hentze MW**. A Dual Inhibitory Mechanism Restricts msl-2 mRNA Translation for Dosage Compensation in Drosophila. *Cell* 122: 529–540, 2005.
23. **Bell TJ**, **Eiríksdóttir E**, **Langel Ü**, **Eberwine J**. PAIR Technology: Exon-Specific RNA-Binding Protein Isolation in Live Cells. In: *Cell-Penetrating Peptides: Methods and Protocols*, edited by Langel Ü. Humana Press, p. 473–486.
24. **Bennett RP**, **Salter JD**, **Smith HC**. A New Class of Antiretroviral Enabling Innate Immunity by Protecting APOBEC3 from HIV Vif-Dependent Degradation. *Trends Mol Med* 24: 507–520, 2018.
25. **Bergantz L**, **Subra F**, **Deprez E**, **Delelis O**, **Richetta C**. Interplay between Intrinsic and Innate Immunity during HIV Infection. *Cells* 8, 2019.

26. **Bierhoff H**. Analysis of lncRNA-Protein Interactions by RNA-Protein Pull-Down Assays and RNA Immunoprecipitation (RIP). *Methods Mol Biol Clifton NJ* 1686: 241–250, 2018.
27. **Brar GA**. Beyond the Triplet Code: Context Cues Transform Translation. *Cell* 167: 1681–1692, 2016.
28. **Brelot A**, **Chakrabarti LA**. CCR5 Revisited: How Mechanisms of HIV Entry Govern AIDS Pathogenesis. *J Mol Biol* 430: 2557–2589, 2018.
29. **de Breyne S**, **Ohlmann T**. Focus on Translation Initiation of the HIV-1 mRNAs. *Int J Mol Sci* 20, 2018.
30. **de Breyne S**, **Soto-Rifo R**, **López-Lastra M**, **Ohlmann T**. Translation initiation is driven by different mechanisms on the HIV-1 and HIV-2 genomic RNAs. *Virus Res* 171: 366–381, 2013.
31. **Brierley I**, **Dos Ramos FJ**. Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Res* 119: 29–42, 2006.
32. **Briggs JAG**, **Kräusslich H-G**. The Molecular Architecture of HIV. *J Mol Biol* 410: 491–500, 2011.
33. **Buchan JR**, **Parker R**. Eukaryotic Stress Granules: The Ins and Outs of Translation. *Mol Cell* 36: 932–941, 2009.
34. **Calvo SE**, **Pagliarini DJ**, **Mootha VK**. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc Natl Acad Sci* 106: 7507–7512, 2009.
35. **Campbell EM**, **Hope TJ**. HIV-1 capsid: the multifaceted key player in HIV-1 infection. *Nat Rev Microbiol* 13: 471–483, 2015.
36. **Castello A**, **Fischer B**, **Eichelbaum K**, **Horos R**, **Beckmann BM**, **Strein C**, **Davey NE**, **Humphreys DT**, **Preiss T**, **Steinmetz LM**, **Krijgsveld J**, **Hentze MW**. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 149: 1393–1406, 2012.
37. **Castelló A**, **Franco D**, **Moral-López P**, **Berlanga JJ**, **Alvarez E**, **Wimmer E**, **Carrasco L**. HIV-1 protease inhibits Cap- and poly(A)-dependent translation upon eIF4GI and PABP cleavage. *PloS One* 4: e7997, 2009.
38. **Cen S**, **Peng Z-G**, **Li X-Y**, **Li Z-R**, **Ma J**, **Wang Y-M**, **Fan B**, **You X-F**, **Wang Y-P**, **Liu F**, **Shao R-G**, **Zhao L-X**, **Yu L**, **Jiang J-D**. Small Molecular Compounds Inhibit HIV-1 Replication through Specifically Stabilizing APOBEC3G. *J Biol Chem* 285: 16546–16552, 2010.
39. **Chen E**, **Joseph S**. Fragile X Mental Retardation Protein: A Paradigm for Translational Control by RNA-Binding Proteins. *Biochimie* 114: 147–154, 2015.
40. **Chen K**, **Zhao BS**, **He C**. Nucleic Acid Modifications in Regulation of Gene Expression. *Cell Chem Biol* 23: 74–85, 2016.
41. **Chew G-L**, **Pauli A**, **Schier AF**. Conservation of uORF repressiveness and sequence features in mouse, human and zebrafish. *Nat Commun* 7, 2016.
42. **Chiang H-S**, **Liu HM**. The Molecular Basis of Viral Inhibition of IRF- and STAT-Dependent Immune Responses. *Front Immunol* 9: 3086, 2018.
43. **Coleman CM**, **Wu L**. HIV interactions with monocytes and dendritic cells: viral latency and reservoirs. *Retrovirology* 6: 51, 2009.
44. **Collier AC**, **Coombs RW**, **Schoenfeld DA**, **Bassett RL**, **Timpone J**, **Baruch A**, **Jones M**, **Facey K**, **Whitacre C**, **McAuliffe VJ**, **Friedman HM**, **Merigan TC**, **Reichman RC**, **Hooper C**, **Corey L**. Treatment of human immunodeficiency virus infection with saquinavir, zidovudine, and zalcitabine. AIDS Clinical Trials Group. *N Engl J Med* 334: 1011–1017, 1996.
45. **Colomer-Lluch M**, **Ruiz A**, **Moris A**, **Prado JG**. Restriction Factors: From Intrinsic Viral Restriction to Shaping Cellular Immunity Against HIV-1. *Front Immunol* 9: 2876, 2018.
46. **Cook KB**, **Hughes TR**, **Morris QD**. High-throughput characterization of protein–RNA interactions. *Brief Funct Genomics* 14: 74–89, 2015.
47. **Council OD**, **Joseph SB**. Evolution of Host Target Cell Specificity During HIV-1 Infection. *Curr HIV Res* 16: 13–20, 2018.
48. **Couso J-P**, **Patraquim P**. Classification and function of small open reading frames. *Nat Rev Mol Cell Biol* 18: 575–589, 2017.
49. **Craigie R**. Nucleoprotein Intermediates in HIV-1 DNA Integration: Structure and Function of HIV-1 Intasomes. *Subcell Biochem* 88: 189–210, 2018.
50. **Crowe JE**. 122 - Host Defense Mechanisms Against Viruses. In: *Fetal and Neonatal Physiology (Fifth Edition)*, edited by Polin RA, Abman SH, Rowitch DH, Benitz WE, Fox WW. Elsevier, p. 1175-1197.e7.
51. **D Urbano V**, **De Crignis E**, **Re MC**. Host Restriction Factors and Human Immunodeficiency

Virus (HIV-1): A Dynamic Interplay Involving All Phases of the Viral Life Cycle. *Curr HIV Res* 16: 184–207, 2018.
52. **Daugherty MD**, **Malik HS**. Rules of engagement: molecular insights from host-virus arms races. *Annu Rev Genet* 46: 677–700, 2012.
53. **Deeks SG**, **Overbaugh J**, **Phillips A**, **Buchbinder S**. HIV infection. *Nat Rev Dis Primer* 1: 15035, 2015.
54. **Dever TE**, **Dinman JD**, **Green R**. Translation Elongation and Recoding in Eukaryotes. *Cold Spring Harb Perspect Biol* 10, 2018.
55. **Dever TE**, **Green R**. The Elongation, Termination, and Recycling Phases of Translation in Eukaryotes. *Cold Spring Harb Perspect Biol* 4: a013706, 2012.
56. **Dick RA**, **Mallery DL**, **Vogt VM**, **James LC**. IP6 Regulation of HIV Capsid Assembly, Stability, and Uncoating. *Viruses* 10, 2018.
57. **Doyle T**, **Goujon C**, **Malim MH**. HIV-1 and interferons: who's interfering with whom? *Nat Rev Microbiol* 13: 403–413, 2015.
58. **Dubois N**, **Khoo KK**, **Ghossein S**, **Seissler T**, **Wolff P**, **McKinstry WJ**, **Mak J**, **Paillart J-C**, **Marquet R**, **Bernacchi S**. The C-terminal p6 domain of the HIV-1 Pr55Gag precursor is required for specific binding to the genomic RNA. *RNA Biol.* (June 28, 2018). doi: 10.1080/15476286.2018.1481696.
59. **Esparza J**. A brief history of the global effort to develop a preventive HIV vaccine. *Vaccine* 31: 3502–3518, 2013.
60. **Eulalio A**, **Behm-Ansmant I**, **Izaurralde E**. P bodies: at the crossroads of post-transcriptional pathways. *Nat Rev Mol Cell Biol* 8: 9–22, 2007.
61. **Falkenhagen A**, **Joshi S**. HIV Entry and Its Inhibition by Bifunctional Antiviral Proteins. *Mol Ther Nucleic Acids* 13: 347–364, 2018.
62. **Fauci AS**, **Pantaleo G**, **Stanley S**, **Weissman D**. Immunopathogenic mechanisms of HIV infection. *Ann Intern Med* 124: 654–663, 1996.
63. **Fernandes JD**, **Booth DS**, **Frankel AD**. A structurally plastic ribonucleoprotein complex mediates post-transcriptional gene regulation in HIV-1. *Wiley Interdiscip Rev RNA* 7: 470–486, 2016.
64. **Fischl MA**, **Richman DD**, **Grieco MH**, **Gottlieb MS**, **Volberding PA**, **Laskin OL**, **Leedom JM**, **Groopman JE**, **Mildvan D**, **Schooley RT**, **Jackson GG**, **Durack DT**, **King D**. The Efficacy of Azidothymidine (AZT) in the Treatment of Patients with AIDS and AIDS-Related Complex. *N Engl J Med* 317: 185–191, 1987.
65. **Freed EO**. HIV-1 replication. *Somat Cell Mol Genet* 26: 13–33, 2001.
66. **Freed EO**. HIV-1 assembly, release and maturation. *Nat Rev Microbiol* 13: 484–496, 2015.
67. **Gabuzda DH**, **Lawrence K**, **Langhoff E**, **Terwilliger E**, **Dorfman T**, **Haseltine WA**, **Sodroski J**. Role of vif in replication of human immunodeficiency virus type 1 in CD4+ T lymphocytes. *J Virol* 66: 6489–6495, 1992.
68. **Gallois-Montbrun S**, **Kramer B**, **Swanson CM**, **Byers H**, **Lynham S**, **Ward M**, **Malim MH**. Antiviral protein APOBEC3G localizes to ribonucleoprotein complexes found in P bodies and stress granules. *J Virol* 81: 2165–2178, 2007.
69. **Galloway A**, **Cowling VH**. mRNA cap regulation in mammalian cell function and fate. *Biochim Biophys Acta Gene Regul Mech* 1862: 270–279, 2019.
70. **Gebauer F**, **Grskovic M**, **Hentze MW**. Drosophila sex-lethal inhibits the stable association of the 40S ribosomal subunit with msl-2 mRNA. *Mol Cell* 11: 1397–1404, 2003.
71. **Ghimire D**, **Rai M**, **Gaur R**. Novel host restriction factors implicated in HIV-1 replication. *J Gen Virol* 99: 435–446, 2018.
72. **Gorle S**, **Pan Y**, **Sun Z**, **Shlyakhtenko LS**, **Harris RS**, **Lyubchenko YL**, **Vuković L**. Computational Model and Dynamics of Monomeric Full-Length APOBEC3G. *ACS Cent Sci* 3: 1180–1188, 2017.
73. **Gottlieb MS**, **Schroff R**, **Schanker HM**, **Weisman JD**, **Fan PT**, **Wolf RA**, **Saxon A**. Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. *N Engl J Med* 305: 1425–1431, 1981.
74. **Greenwood EJ**, **Matheson NJ**, **Wals K**, **van den Boomen DJ**, **Antrobus R**, **Williamson JC**, **Lehner PJ**. Temporal proteomic analysis of HIV infection reveals remodelling of the host phosphoproteome by lentiviral Vif variants. *eLife* 5, 2016.
75. **Gregori J**, **Sanchez A**, **Villanueva J**. msmsTests: LC-MS/MS Differential Expression Tests. *R Package Version 1120*: 2013.

76. **Gregori J**, **Sanchez A**, **Villanueva J**. msmsEDA: Exploratory Data Analysis of LC-MS/MS data by spectral counts. *R Package Version 1120*: 2014.
77. **Guerrero S**, **Libre C**, **Batisse J**, **Mercenne G**, **Richer D**, **Laumond G**, **Decoville T**, **Moog C**, **Marquet R**, **Paillart J-C**. Translational regulation of APOBEC3G mRNA by Vif requires its 5'UTR and contributes to restoring HIV-1 infectivity. *Sci Rep* 6: 39507, 2016.
78. **Gunišová S**, **Hronová V**, **Mohammad MP**, **Hinnebusch AG**, **Valášek LS**. Please do not recycle! Translation reinitiation in microbes and higher eukaryotes. *FEMS Microbiol Rev* 42: 165–192, 2018.
79. **Guo H**. Specialized ribosomes and the control of translation. *Biochem Soc Trans* 46: 855–869, 2018.
80. **Guo Y**, **Dong L**, **Qiu X**, **Wang Y**, **Zhang B**, **Liu H**, **Yu Y**, **Zang Y**, **Yang M**, **Huang Z**. Structural basis for hijacking CBF-β and CUL5 E3 ligase complex by HIV-1 Vif. *Nature* 505: 229–233, 2014.
81. **Hammer SM**, **Squires KE**, **Hughes MD**, **Grimes JM**, **Demeter LM**, **Currier JS**, **Eron JJ**, **Feinberg JE**, **Balfour HH**, **Deyton LR**, **Chodakewitz JA**, **Fischl MA**. A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. *N Engl J Med* 337: 725–733, 1997.
82. **Harris RS**, **Dudley JP**. APOBECs and virus restriction. *Virology* 479–480: 131–145, 2015.
83. **Harvey RF**, **Smith TS**, **Mulroney T**, **Queiroz RML**, **Pizzinga M**, **Dezi V**, **Villenueva E**, **Ramakrishna M**, **Lilley KS**, **Willis AE**. Trans-acting translational regulatory RNA binding proteins. *Wiley Interdiscip Rev RNA* 9: e1465, 2018.
84. **Hellen CUT**. Translation Termination and Ribosome Recycling in Eukaryotes. *Cold Spring Harb Perspect Biol* 10, 2018.
85. **Hemelaar J**, **Elangovan R**, **Yun J**, **Dickson-Tetteh L**, **Fleminger I**, **Kirtley S**, **Williams B**, **Gouws-Williams E**, **Ghys PD**, **WHO–UNAIDS Network for HIV Isolation Characterisation**. Global and regional molecular epidemiology of HIV-1, 1990-2015: a systematic review, global survey, and trend analysis. *Lancet Infect Dis* 19: 143–155, 2019.
86. **Henriet S**, **Sinck L**, **Bec G**, **Gorelick RJ**, **Marquet R**, **Paillart J-C**. Vif is a RNA chaperone that could temporally regulate RNA dimerization and the early steps of HIV-1 reverse transcription. *Nucleic Acids Res* 35: 5141–5153, 2007.
87. **Hidalgo L**, **Swanson CM**. Regulation of human immunodeficiency virus type 1 (HIV-1) mRNA translation. *Biochem Soc Trans* 45: 353–364, 2017.
88. **Hilditch L**, **Towers GJ**. A model for cofactor use during HIV-1 reverse transcription and nuclear entry. *Curr Opin Virol* 4: 32–36, 2014.
89. **Hinnebusch AG**, **Lorsch JR**. The mechanism of eukaryotic translation initiation: new insights and challenges. *Cold Spring Harb Perspect Biol* 4, 2012.
90. **Hirsch JD**, **Eslamizar L**, **Filanoski BJ**, **Malekzadeh N**, **Haugland RP**, **Beechem JM**, **Haugland RP**. Easily reversible desthiobiotin binding to streptavidin, avidin, and other biotin-binding proteins: uses for protein labeling, detection, and isolation. *Anal Biochem* 308: 343–357, 2002.
91. **Hladik F**, **McElrath MJ**. Setting the stage: host invasion by HIV. *Nat Rev Immunol* 8: 447–457, 2008.
92. **Hogg JR**, **Collins K**. RNA-based affinity purification reveals 7SK RNPs with distinct composition and regulation. *RNA N Y N* 13: 868–880, 2007.
93. **Hu W-S**, **Hughes SH**. HIV-1 reverse transcription. *Cold Spring Harb Perspect Med* 2, 2012.
94. **Huang DW**, **Sherman BT**, **Lempicki RA**. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44–57, 2009.
95. **Hughes SH**. Reverse Transcription of Retroviruses and LTR Retrotransposons. *Microbiol Spectr* 3: MDNA3-0027–2014, 2015.
96. **Inamdar K**, **Floderer C**, **Favard C**, **Muriaux D**. Monitoring HIV-1 Assembly in Living Cells: Insights from Dynamic and Single Molecule Microscopy. *Viruses* 11, 2019.
97. **Ingolia NT**, **Lareau LF**, **Weissman JS**. Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. *Cell* 147: 789–802, 2011.
98. **Jackson RJ**, **Hellen CUT**, **Pestova TV**. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* 11: 113–127, 2010.
99. **Jackson RJ**, **Hellen CUT**, **Pestova TV**. Termination and post-termination events in eukaryotic translation. *Adv Protein Chem Struct Biol* 86: 45–93, 2012.
100. **Jacques DA**, **McEwan WA**, **Hilditch L**, **Price AJ**, **Towers GJ**, **James LC**. HIV-1 uses dynamic

capsid pores to import nucleotides and fuel encapsidated DNA synthesis. *Nature* 536: 349–353, 2016.
101. **Johnstone TG**, **Bazzini AA**, **Giraldez AJ**. Upstream ORFs are prevalent translational repressors in vertebrates. *EMBO J* 35: 706–723, 2016.
102. **Jones CE**, **McKnight Á**. Retroviral restriction: nature's own solution. *Curr Opin Infect Dis* 29: 609–614, 2016.
103. **Joseph SB**, **Swanstrom R**, **Kashuba ADM**, **Cohen MS**. Bottlenecks in HIV-1 transmission: insights from the study of founder viruses. *Nat Rev Microbiol* 13: 414–425, 2015.
104. **Karn J**, **Stoltzfus CM**. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harb Perspect Med* 2: a006916, 2012.
105. **Kedersha N**, **Anderson P**. Regulation of translation by stress granules and processing bodies. *Prog Mol Biol Transl Sci* 90: 155–185, 2009.
106. **Kennedy EM**, **Bogerd HP**, **Kornepati AVR**, **Kang D**, **Ghoshal D**, **Marshall JB**, **Poling BC**, **Tsai K**, **Gokhale NS**, **Horner SM**, **Cullen BR**. Posttranscriptional m(6)A Editing of HIV-1 mRNAs Enhances Viral Gene Expression. *Cell Host Microbe* 19: 675–685, 2016.
107. **Kennedy EM**, **Courtney DG**, **Tsai K**, **Cullen BR**. Viral Epitranscriptomics. *J Virol* 91, 2017.
108. **Kennedy-Darling J**, **Holden MT**, **Shortreed MR**, **Smith LM**. Multiplexed programmable release of captured DNA. *Chembiochem Eur J Chem Biol* 15: 2353–2356, 2014.
109. **Kim DI**, **Jensen SC**, **Noble KA**, **Kc B**, **Roux KH**, **Motamedchaboki K**, **Roux KJ**. An improved smaller biotin ligase for BioID proximity labeling. *Mol Biol Cell* 27: 1188–1196, 2016.
110. **Kim DY**. The assembly of Vif ubiquitin E3 ligase for APOBEC3 degradation. *Arch Pharm Res* 38: 435–445, 2015.
111. **Kim DY**, **Kwon E**, **Hartley PD**, **Crosby DC**, **Mann S**, **Krogan NJ**, **Gross JD**. CBFβ stabilizes HIV Vif to counteract APOBEC3 at the expense of RUNX1 target gene expression. *Mol Cell* 49: 632–644, 2013.
112. **Knisbacher BA**, **Gerber D**, **Levanon EY**. DNA Editing by APOBECs: A Genomic Preserver and Transformer. *Trends Genet* 32: 16–28, 2016.
113. **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. *Sci Rep* 7: 16965, 2017.
114. **Koito A**, **Ikeda T**. Apolipoprotein B mRNA-editing, catalytic polypeptide cytidine deaminases and retroviral restriction. *Wiley Interdiscip Rev RNA* 3: 529–541, 2012.
115. **Kozak M**. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Res* 9: 5233–5252, 1981.
116. **Kozak M**. Effects of intercistronic length on the efficiency of reinitiation by eucaryotic ribosomes. *Mol Cell Biol* 7: 3438–3445, 1987.
117. **Kozak M**. Constraints on reinitiation of translation in mammals. *Nucleic Acids Res* 29: 5226–5232, 2001.
118. **Kozak M**. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* 361: 13–37, 2005.
119. **Krummheuer J**, **Johnson AT**, **Hauber I**, **Kammler S**, **Anderson JL**, **Hauber J**, **Purcell DFJ**, **Schaal H**. A minimal uORF within the HIV-1 vpu leader allows efficient translation initiation at the downstream env AUG. *Virology* 363: 261–271, 2007.
120. **Kurosaki T**, **Popp MW**, **Maquat LE**. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat Rev Mol Cell Biol* 20: 406–420, 2019.
121. **LaRue RS**, **Andrésdóttir V**, **Blanchard Y**, **Conticello SG**, **Derse D**, **Emerman M**, **Greene WC**, **Jónsson SR**, **Landau NR**, **Löchelt M**, **Malik HS**, **Malim MH**, **Münk C**, **O'Brien SJ**, **Pathak VK**, **Strebel K**, **Wain-Hobson S**, **Yu X-F**, **Yuhki N**, **Harris RS**. Guidelines for naming nonprimate APOBEC3 genes and proteins. *J Virol* 83: 494–497, 2009.
122. **Lee ASY**, **Kranzusch PJ**, **Cate JHD**. eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. *Nature* 522: 111–114, 2015.
123. **Lee K-M**, **Chen C-J**, **Shih S-R**. Regulation Mechanisms of Viral IRES-Driven Translation. *Trends Microbiol* 25: 546–561, 2017.
124. **Leppek K**, **Stoecklin G**. An optimized streptavidin-binding RNA aptamer for purification of ribonucleoprotein complexes identifies novel ARE-binding proteins. *Nucleic Acids Res* 42: e13, 2014.
125. **Lichinchi G**, **Gao S**, **Saletore Y**, **Gonzalez GM**, **Bansal V**, **Wang Y**, **Mason CE**, **Rana TM**. Dynamics of the human and viral m(6)A RNA methylomes during HIV-1 infection of T cells. *Nat Microbiol* 1: 16011, 2016.

126. **Ling C**, **Ermolenko DN**. Structural insights into ribosome translocation. *Wiley Interdiscip Rev RNA* 7: 620–636, 2016.
127. **Lomakin IB**, **Stolboushkina EA**, **Vaidya AT**, **Zhao C**, **Garber MB**, **Dmitriev SE**, **Steitz TA**. Crystal Structure of the Human Ribosome in Complex with DENR-MCT-1. *Cell Rep* 20: 521–528, 2017.
128. **Lusic M**, **Siliciano RF**. Nuclear landscape of HIV-1 infection and integration. *Nat Rev Microbiol* 15: 69–82, 2017.
129. **Ma L**, **Zhang Z**, **Liu Z**, **Pan Q**, **Wang J**, **Li X**, **Guo F**, **Liang C**, **Hu L**, **Zhou J**, **Cen S**. Identification of small molecule compounds targeting the interaction of HIV-1 Vif and human APOBEC3G by virtual screening and biological evaluation. *Sci Rep* 8: 8067, 2018.
130. **Ma XM**, **Yoon S-O**, **Richardson CJ**, **Jülich K**, **Blenis J**. SKAR links pre-mRNA splicing to mTOR/S6K1-mediated enhanced translation efficiency of spliced mRNAs. *Cell* 133: 303–313, 2008.
131. **Mailler E**, **Bernacchi S**, **Marquet R**, **Paillart J-C**, **Vivet-Boudou V**, **Smyth RP**. The Life-Cycle of the HIV-1 Gag–RNA Complex. *Viruses* 8, 2016.
132. **Mailliot J**, **Martin F**. Viral internal ribosomal entry sites: four classes for one goal. *Wiley Interdiscip Rev RNA* 9, 2018.
133. **Malim MH**, **Bieniasz PD**. HIV Restriction Factors and Mechanisms of Evasion. *Cold Spring Harb Perspect Med* 2, 2012.
134. **Mangeot PE**, **Risson V**, **Fusil F**, **Marnef A**, **Laurent E**, **Blin J**, **Mournetas V**, **Massouridès E**, **Sohier TJM**, **Corbin A**, **Aubé F**, **Teixeira M**, **Pinset C**, **Schaeffer L**, **Legube G**, **Cosset F-L**, **Verhoeyen E**, **Ohlmann T**, **Ricci EP**. Genome editing in primary cells and in vivo using viral-derived Nanoblades loaded with Cas9-sgRNA ribonucleoproteins. *Nat Commun* 10: 45, 2019.
135. **Martínez-Salas E**, **Lozano G**, **Fernandez-Chamorro J**, **Francisco-Velilla R**, **Galan A**, **Diaz R**. RNA-binding proteins impacting on internal initiation of translation. *Int J Mol Sci* 14: 21705–21726, 2013.
136. **Mathé G**, **Pontiggia P**, **Orbach-Arbouys S**, **Triana K**, **Ambetima N**, **Morette C**, **Hallard M**, **Blanquet D**. AIDS therapy with two, three or four agent combinations, applied in short sequences, differing from each other by drug rotation. I. First of two parts: a phase I trial equivalent, concerning five virostatics: AZT, ddI, ddC, acriflavine and an ellipticine analogue. *Biomed Pharmacother Biomedecine Pharmacother* 50: 220–227, 1996.
137. **Matreyek KA**, **Engelman A**. Viral and cellular requirements for the nuclear entry of retroviral preintegration nucleoprotein complexes. *Viruses* 5: 2483–2511, 2013.
138. **Mbonye U**, **Karn J**. The Molecular Basis for Human Immunodeficiency Virus Latency. *Annu Rev Virol* 4: 261–285, 2017.
139. **McCormick C**, **Khaperskyy DA**. Translation inhibition and stress granules in the antiviral immune response. *Nat Rev Immunol* 17: 647–660, 2017.
140. **McHugh CA**, **Guttman M**. RAP-MS: A Method to Identify Proteins that Interact Directly with a Specific RNA Molecule in Cells. *Methods Mol Biol Clifton NJ* 1649: 473–488, 2018.
141. **McHugh CA**, **Russell P**, **Guttman M**. Methods for comprehensive experimental identification of RNA-protein interactions. *Genome Biol* 15: 203, 2014.
142. **McMichael AJ**, **Borrow P**, **Tomaras GD**, **Goonetilleke N**, **Haynes BF**. The immune response during acute HIV-1 infection: clues for vaccine development. *Nat Rev Immunol* 10: 11–23, 2010.
143. **Medenbach J**, **Seiler M**, **Hentze MW**. Translational Control via Protein-Regulated Upstream Open Reading Frames. *Cell* 145: 902–913, 2011.
144. **Mehellou Y**, **De Clercq E**. Twenty-six years of anti-HIV drug discovery: where do we stand and where do we go? *J Med Chem* 53: 521–538, 2010.
145. **Menéndez-Arias L**, **Sebastián-Martín A**, **Álvarez M**. Viral reverse transcriptases. *Virus Res* 234: 153–176, 2017.
146. **Meng B**, **Lever AM**. Wrapping up the bad news: HIV assembly and release. *Retrovirology* 10: 5, 2013.
147. **Mercenne G**, **Bernacchi S**, **Richer D**, **Bec G**, **Henriet S**, **Paillart J-C**, **Marquet R**. HIV-1 Vif binds to APOBEC3G mRNA and inhibits its translation. *Nucleic Acids Res* 38: 633–646, 2010.
148. **Merrick WC**, **Pavitt GD**. Protein Synthesis Initiation in Eukaryotic Cells. *Cold Spring Harb Perspect Biol* 10, 2018.
149. **Meyer KD**. m6A-mediated translation regulation. *Biochim Biophys Acta BBA - Gene Regul Mech* 1862: 301–309, 2019.
150. **Meyer KD**, **Patil DP**, **Zhou J**, **Zinoviev A**, **Skabkin MA**, **Elemento O**, **Pestova TV**, **Qian S-B**,

Jaffrey SR. 5’ UTR m(6)A Promotes Cap-Independent Translation. *Cell* 163: 999–1010, 2015.
151. **Moore MJ**. From birth to death: the complex lives of eukaryotic mRNAs. *Science* 309: 1514–1518, 2005.
152. **Morse M**, **Huo R**, **Feng Y**, **Rouzina I**, **Chelico L**, **Williams MC**. Dimerization regulates both deaminase-dependent and deaminase-independent HIV-1 restriction by APOBEC3G. *Nat Commun* 8: 597, 2017.
153. **Muckenthaler M**, **Gray NK**, **Hentze MW**. IRP-1 binding to ferritin mRNA prevents the recruitment of the small ribosomal subunit by the cap-binding complex eIF4F. *Mol Cell* 2: 383–388, 1998.
154. **Ne E**, **Palstra R-J**, **Mahmoudi T**. Transcription: Insights From the HIV-1 Promoter. *Int Rev Cell Mol Biol* 335: 191–243, 2018.
155. **Nishimura T**, **Wada T**, **Yamamoto KT**, **Okada K**. The Arabidopsis STV1 protein, responsible for translation reinitiation, is required for auxin-mediated gynoecium patterning. *Plant Cell* 17: 2940–2953, 2005.
156. **Ocwieja KE**, **Sherrill-Mix S**, **Mukherjee R**, **Custers-Allen R**, **David P**, **Brown M**, **Wang S**, **Link DR**, **Olson J**, **Travers K**, **Schadt E**, **Bushman FD**. Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule enrichment and long-read sequencing. *Nucleic Acids Res* 40: 10345–10355, 2012.
157. **Ohlmann T**, **Mengardi C**, **López-Lastra M**. Translation initiation of the HIV-1 mRNA. *Transl Austin Tex* 2: e960242, 2014.
158. **Okada A**, **Iwatani Y**. APOBEC3G-Mediated G-to-A Hypermutation of the HIV-1 Genome: The Missing Link in Antiviral Molecular Mechanisms. *Front Microbiol* 7, 2016.
159. **Olson ME**, **Harris RS**, **Harki DA**. APOBEC Enzymes as Targets for Virus and Cancer Therapy. *Cell Chem Biol* 25: 36–49, 2018.
160. **Ostareck DH**, **Ostareck-Lederer A**, **Shatsky IN**, **Hentze MW**. Lipoxygenase mRNA silencing in erythroid differentiation: The 3’UTR regulatory complex controls 60S ribosomal subunit joining. *Cell* 104: 281–290, 2001.
161. **Paillart JC**, **Marquet R**, **Skripkin E**, **Ehresmann B**, **Ehresmann C**. Mutational analysis of the bipartite dimer linkage structure of human immunodeficiency virus type 1 genomic RNA. *J Biol Chem* 269: 27486–27493, 1994.
162. **Parekh BS**, **Ou C-Y**, **Fonjungo PN**, **Kalou MB**, **Rottinghaus E**, **Puren A**, **Alexander H**, **Hurlston Cox M**, **Nkengasong JN**. Diagnosis of Human Immunodeficiency Virus Infection. *Clin Microbiol Rev* 32, 2019.
163. **Park HS**, **Himmelbach A**, **Browning KS**, **Hohn T**, **Ryabova LA**. A plant viral “reinitiation” factor interacts with the host translational machinery. *Cell* 106: 723–733, 2001.
164. **Parker R**, **Sheth U**. P bodies and the control of mRNA translation and degradation. *Mol Cell* 25: 635–646, 2007.
165. **Peeters M**, **Jung M**, **Ayouba A**. The origin and molecular epidemiology of HIV. *Expert Rev Anti Infect Ther* 11: 885–896, 2013.
166. **Pereira-Montecinos C**, **Valiente-Echeverría F**, **Soto-Rifo R**. Epitranscriptomic regulation of viral replication. *Biochim Biophys Acta Gene Regul Mech* 1860: 460–471, 2017.
167. **Perreau M**, **Levy Y**, **Pantaleo G**. Immune response to HIV. *Curr Opin HIV AIDS* 8: 333–340, 2013.
168. **Pery E**, **Sheehy A**, **Nebane NM**, **Brazier AJ**, **Misra V**, **Rajendran KS**, **Buhrlage SJ**, **Mankowski MK**, **Rasmussen L**, **White EL**, **Ptak RG**, **Gabuzda D**. Identification of a Novel HIV-1 Inhibitor Targeting Vif-dependent Degradation of Human APOBEC3G Protein. *J Biol Chem* 290: 10504–10517, 2015.
169. **Poblete-Durán N**, **Prades-Pérez Y**, **Vera-Otarola J**, **Soto-Rifo R**, **Valiente-Echeverría F**. Who Regulates Whom? An Overview of RNA Granules and Viral Infections. *Viruses* 8, 2016.
170. **Pollpeter D**, **Parsons M**, **Sobala AE**, **Coxhead S**, **Lang RD**, **Bruns AM**, **Papaioannou S**, **McDonnell JM**, **Apolonia L**, **Chowdhury JA**, **Horvath CM**, **Malim MH**. Deep sequencing of HIV-1 reverse transcripts reveals the multifaceted antiviral functions of APOBEC3G. *Nat Microbiol* 3: 220–233, 2018.
171. **Pyndiah N**, **Telenti A**, **Rausell A**. Evolutionary genomics and HIV restriction factors. *Curr Opin HIV AIDS* 10: 79–83, 2015.
172. **Ramanathan M**, **Majzoub K**, **Rao DS**, **Neela PH**, **Zarnegar BJ**, **Mondal S**, **Roth JG**, **Gai H**, **Kovalski JR**, **Siprashvili Z**, **Palmer TD**, **Carette JE**, **Khavari PA**. RNA-protein interaction detection in living cells. *Nat Methods* 15: 207–212, 2018.

173. **Ramanathan M**, **Porter DF**, **Khavari PA**. Methods to study RNA-protein interactions. *Nat Methods* 16: 225–234, 2019.
174. **Rankovic S**, **Varadarajan J**, **Ramalho R**, **Aiken C**, **Rousso I**. Reverse Transcription Mechanically Initiates HIV-1 Capsid Disassembly. *J Virol* 91, 2017.
175. **Rausch JW**, **Le Grice SFJ**. HIV Rev Assembly on the Rev Response Element (RRE): A Structural Perspective. *Viruses* 7: 3053–3075, 2015.
176. **Re A**, **Joshi T**, **Kulberkyte E**, **Morris Q**, **Workman CT**. RNA-protein interactions: an overview. *Methods Mol Biol Clifton NJ* 1097: 491–521, 2014.
177. **Rey-Cuillé MA**, **Berthier JL**, **Bomsel-Demontoy MC**, **Chaduc Y**, **Montagnier L**, **Hovanessian AG**, **Chakrabarti LA**. Simian immunodeficiency virus replicates to high levels in sooty mangabeys without inducing disease. *J Virol* 72: 3872–3886, 1998.
178. **Rhee H-W**, **Zou P**, **Udeshi ND**, **Martell JD**, **Mootha VK**, **Carr SA**, **Ting AY**. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* 339: 1328–1331, 2013.
179. **Riquelme-Barrios S**, **Pereira-Montecinos C**, **Valiente-Echeverría F**, **Soto-Rifo R**. Emerging Roles of N6-Methyladenosine on HIV-1 RNA Metabolism and Viral Replication. *Front Microbiol* 9: 576, 2018.
180. **Roberts JW**. Phage lambda and the regulation of transcription termination. *Cell* 52: 5–6, 1988.
181. **Rodrigues V**, **Ruffin N**, **San-Roman M**, **Benaroch P**. Myeloid Cell Interaction with HIV: A Complex Relationship. *Front Immunol* 8, 2017.
182. **Rojas-Araya B**, **Ohlmann T**, **Soto-Rifo R**. Translational Control of the HIV Unspliced Genomic RNA. *Viruses* 7: 4326–4351, 2015.
183. **Roux KJ**, **Kim DI**, **Raida M**, **Burke B**. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol* 196: 801–810, 2012.
184. **Rowland-Jones SL**. AIDS pathogenesis: what have two decades of HIV research taught us? *Nat Rev Immunol* 3: 343–348, 2003.
185. **Ruan X**, **Li P**, **Cao H**. Identification of Transcriptional Regulators That Bind to Long Noncoding RNAs by RNA Pull-Down and RNA Immunoprecipitation. *Methods Mol Biol Clifton NJ* 1783: 185–191, 2018.
186. **Rustagi A**, **Gale M**. Innate antiviral immune signaling, viral evasion and modulation by HIV-1. *J Mol Biol* 426: 1161–1177, 2014.
187. **Sadowski I**, **Hashemi FB**. Strategies to eradicate HIV from infected patients: elimination of latent provirus reservoirs. *Cell. Mol. Life Sci.* (May 25, 2019). doi: 10.1007/s00018-019-03156-8.
188. **Salter JD**, **Bennett RP**, **Smith HC**. The APOBEC Protein Family: United by Structure, Divergent in Function. *Trends Biochem Sci* 41: 578–594, 2016.
189. **Salter JD**, **Smith HC**. Modeling the Embrace of a Mutator: APOBEC Selection of Nucleic Acid Ligands. *Trends Biochem Sci* 43: 606–622, 2018.
190. **Sauter D**, **Kirchhoff F**. Multilayered and versatile inhibition of cellular antiviral factors by HIV and SIV accessory proteins. *Cytokine Growth Factor Rev* 40: 3–12, 2018.
191. **Sauter D**, **Kirchhoff F**. Key Viral Adaptations Preceding the AIDS Pandemic. *Cell Host Microbe* 25: 27–38, 2019.
192. **Schaller T**, **Herold N**. The Early Bird Catches the Worm--Can Evolution Teach us Lessons in Fighting HIV? *Curr HIV Res* 14: 183–210, 2016.
193. **Schleich S**, **Strassburger K**, **Janiesch PC**, **Koledachkina T**, **Miller KK**, **Haneke K**, **Cheng Y-S**, **Kuechler K**, **Stoecklin G**, **Duncan KE**, **Teleman AA**. DENR-MCT-1 promotes translation re-initiation downstream of uORFs to control tissue growth. *Nature* 512: 208–212, 2014.
194. **Schmit JC**, **Weber B**. Recent advances in antiretroviral therapy and HIV infection monitoring. *Intervirology* 40: 304–321, 1997.
195. **Schmitt E**, **Coureux P-D**, **Monestier A**, **Dubiez E**, **Mechulam Y**. Start Codon Recognition in Eukaryotic and Archaeal Translation Initiation: A Common Structural Core. *Int J Mol Sci* 20, 2019.
196. **Schneider CA**, **Rasband WS**, **Eliceiri KW**. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9: 671–675, 2012.
197. **Schoggins JW**. Interferon-Stimulated Genes: What Do They All Do? *Annu. Rev. Virol.* (July 5, 2019). doi: 10.1146/annurev-virology-092818-015756.
198. **Schuller AP**, **Green R**. Roadblocks and resolutions in eukaryotic translation. *Nat Rev Mol Cell Biol* 19: 526–541, 2018.
199. **Schwartz C**, **Bouchat S**, **Marban C**, **Gautier V**, **Van Lint C**, **Rohr O**, **Le Douce V**. On the way

to find a cure: Purging latent HIV-1 reservoirs. *Biochem Pharmacol* 146: 10–22, 2017.
200. **Seissler T**, **Marquet R**, **Paillart J-C**. Hijacking of the Ubiquitin/Proteasome Pathway by the HIV Auxiliary Proteins. *Viruses* 9, 2017.
201. **Sertznig H**, **Hillebrand F**, **Erkelenz S**, **Schaal H**, **Widera M**. Behind the scenes of HIV-1 replication: Alternative splicing as the dependency factor on the quiet. *Virology* 516: 176–188, 2018.
202. **Sharma A**, **Yilmaz A**, **Marsh K**, **Cochrane A**, **Boris-Lawrie K**. Thriving under stress: selective translation of HIV-1 structural protein mRNA during Vpr-mediated impairment of eIF4E translation activity. *PLoS Pathog* 8: e1002612, 2012.
203. **Sharma S**, **Patnaik SK**, **Taggart RT**, **Baysal BE**. The double-domain cytidine deaminase APOBEC3G is a cellular site-specific RNA editing enzyme. *Sci Rep* 6: 39100, 2016.
204. **Sharp PM**, **Hahn BH**. Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med* 1: a006841, 2011.
205. **Shaw GM**, **Hunter E**. HIV transmission. *Cold Spring Harb Perspect Med* 2, 2012.
206. **Sheehy AM**, **Gaddis NC**, **Choi JD**, **Malim MH**. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 418: 646–650, 2002.
207. **Sheehy AM**, **Gaddis NC**, **Malim MH**. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nat Med* 9: 1404–1407, 2003.
208. **Shirokikh NE**, **Preiss T**. Translation initiation by cap-dependent ribosome recruitment: Recent insights and open questions. *Wiley Interdiscip Rev RNA* 9: e1473, 2018.
209. **Simon JH**, **Fouchier RA**, **Southerling TE**, **Guerra CB**, **Grant CK**, **Malim MH**. The Vif and Gag proteins of human immunodeficiency virus type 1 colocalize in infected human T cells. *J Virol* 71: 5259–5267, 1997.
210. **Simon JH**, **Southerling TE**, **Peterson JC**, **Meyer BE**, **Malim MH**. Complementation of vif-defective human immunodeficiency virus type 1 by primate, but not nonprimate, lentivirus vif genes. *J Virol* 69: 4166–4172, 1995.
211. **Simon MD**. Capture hybridization analysis of RNA targets (CHART). *Curr Protoc Mol Biol* Chapter 21: Unit 21.25., 2013.
212. **Simon V**, **Bloch N**, **Landau NR**. Intrinsic host restrictions to HIV-1 and mechanisms of viral escape. *Nat Immunol* 16: 546–553, 2015.
213. **Skripkin E**, **Paillart JC**, **Marquet R**, **Ehresmann B**, **Ehresmann C**. Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. *Proc Natl Acad Sci U S A* 91: 4945–4949, 1994.
214. **Sleiman D**, **Bernacchi S**, **Xavier Guerrero S**, **Brachet F**, **Larue V**, **Paillart J-C**, **Tisné C**. Characterization of RNA binding and chaperoning activities of HIV-1 Vif protein: Importance of the C-terminal unstructured tail. *RNA Biol* 11: 906–920, 2014.
215. **Somers J**, **Pöyry T**, **Willis AE**. A perspective on mammalian upstream open reading frame function. *Int J Biochem Cell Biol* 45: 1690–1700, 2013.
216. **Spector C**, **Mele AR**, **Wigdahl B**, **Nonnemacher MR**. Genetic variation and function of the HIV-1 Tat protein. *Med Microbiol Immunol (Berl)* 208: 131–169, 2019.
217. **Stahl RE**, **Friedman-Kien A**, **Dubin R**, **Marmor M**, **Zolla-Pazner S**. Immunologic abnormalities in homosexual men. Relationship to Kaposi's sarcoma. *Am J Med* 73: 171–178, 1982.
218. **Stake M**, **Singh D**, **Singh G**, **Marcela Hernandez J**, **Kaddis Maldonado R**, **Parent LJ**, **Boris-Lawrie K**. HIV-1 and two avian retroviral 5' untranslated regions bind orthologous human and chicken RNA binding proteins. *Virology* 486: 307–320, 2015.
219. **Stopak K**, **de Noronha C**, **Yonemoto W**, **Greene WC**. HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. *Mol Cell* 12: 591–601, 2003.
220. **Strebel K**. HIV accessory proteins versus host restriction factors. *Curr Opin Virol* 3: 692–699, 2013.
221. **Sundquist WI**, **Kräusslich H-G**. HIV-1 assembly, budding, and maturation. *Cold Spring Harb Perspect Med* 2: a006924, 2012.
222. **Tebit DM**, **Arts EJ**. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. *Lancet Infect Dis* 11: 45–56, 2011.
223. **Tirumuru N**, **Zhao BS**, **Lu W**, **Lu Z**, **He C**, **Wu L**. N(6)-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 Gag protein expression. *eLife* 5, 2016.
224. **Tough RH**, **McLaren PJ**. Interaction of the Host and Viral Genome and Their Influence on HIV

Disease. *Front Genet* 9: 720, 2018.
225. **Towers GJ**, **Noursadeghi M**. Interactions between HIV-1 and the Cell-Autonomous Innate Immune System. *Cell Host Microbe* 16: 10–18, 2014.
226. **UNAIDS**. Global HIV & AIDS statistics — 2019 fact sheet [Online]. 2019. https://www.unaids.org/en/resources/fact-sheet.
227. **Weber M**, **Weber F**. RIG-I-like receptors and negative-strand RNA viruses: RLRly bird catches some worms. *Cytokine Growth Factor Rev* 25: 621–628, 2014.
228. **Wichroski MJ**, **Robb GB**, **Rana TM**. Human retroviral host restriction factors APOBEC3G and APOBEC3F localize to mRNA processing bodies. *PLoS Pathog* 2: e41, 2006.
229. **Wilen CB**, **Tilton JC**, **Doms RW**. HIV: cell binding and entry. *Cold Spring Harb Perspect Med* 2, 2012.
230. **Wong RW**, **Mamede JI**, **Hope TJ**. Impact of Nucleoporin-Mediated Chromatin Localization and Nuclear Architecture on HIV Integration Site Selection. *J Virol* 89: 9702–9705, 2015.
231. **Xi C**, **Balberg M**, **Boppart SA**, **Raskin L**. Use of DNA and Peptide Nucleic Acid Molecular Beacons for Detection and Quantification of rRNA in Solution and in Whole Cells. *Appl Environ Microbiol* 69: 5673–5678, 2003.
232. **Yamashita M**, **Engelman AN**. Capsid-Dependent Host Factors in HIV-1 Infection. *Trends Microbiol* 25: 741–755, 2017.
233. **Yu X**, **Yu Y**, **Liu B**, **Luo K**, **Kong W**, **Mao P**, **Yu X-F**. Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* 302: 1056–1060, 2003.
234. **Zhang H**, **Wang Y**, **Lu J**. Function and Evolution of Upstream ORFs in Eukaryotes. *Trends Biochem Sci* 0, 2019.
235. **Zhang H**, **Yang B**, **Pomerantz RJ**, **Zhang C**, **Arunachalam SC**, **Gao L**. The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* 424: 94–98, 2003.
236. **Zhao G**, **Perilla JR**, **Yufenyuy EL**, **Meng X**, **Chen B**, **Ning J**, **Ahn J**, **Gronenborn AM**, **Schulten K**, **Aiken C**, **Zhang P**. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* 497: 643–646, 2013.
237. **Zheng X**, **Cho S**, **Moon H**, **Loh TJ**, **Jang HN**, **Shen H**. Detecting RNA-Protein Interaction Using End-Labeled Biotinylated RNA Oligonucleotides and Immunoblotting. *Methods Mol Biol Clifton NJ* 1421: 35–44, 2016.
238. **Zhou M**, **Luo R-H**, **Hou X-Y**, **Wang R-R**, **Yan G-Y**, **Chen H**, **Zhang R-H**, **Shi J-Y**, **Zheng Y-T**, **Li R**, **Wei Y-Q**. Synthesis, biological evaluation and molecular docking study of N-(2-methoxyphenyl)-6-((4-nitrophenyl)sulfonyl)benzamide derivatives as potent HIV-1 Vif antagonists. *Eur J Med Chem* 129: 310–324, 2017.
239. **Zielinski J**, **Kilk K**, **Peritz T**, **Kannanayakal T**, **Miyashiro KY**, **Eiríksdóttir E**, **Jochems J**, **Langel Û**, **Eberwine J**. In vivo identification of ribonucleoprotein-RNA interactions. *Proc Natl Acad Sci* 103: 1557–1562, 2006.
240. **Zuker M**. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31: 3406–3415, 2003.
241. **Zuo T**, **Liu D**, **Lv W**, **Wang X**, **Wang J**, **Lv M**, **Huang W**, **Wu J**, **Zhang H**, **Jin H**, **Zhang L**, **Kong W**, **Yu X**. Small-Molecule Inhibition of Human Immunodeficiency Virus Type 1 Replication by Targeting the Interaction between Vif and ElonginC. *J Virol* 86: 5497–5507, 2012.
242. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47: D506–D515, 2019.

Appendices

XI Appendices

BoxB-A3Gwt:

GCCCTGAAAAAGGGCAAGCTTGGCCCTGAAAAAGGGCAAGCTTGGCC
CTGAAAAAGGGCAAGCTTGGAGACGCTCTTTCCCTTTGCAATTGCCTT
GGGTCCTGCCGCACAGAGCGGCCTGTCTTTATCAGAGGTCCCTCTGC
CAGGGGGAGGGCCCCAGAGAAAACCAGAAAGAGGGTGAGAGACTGA
GGAAGATAAAGCGTCCCAGGGCCTCCTACACCAGCGCCTGAGCAGGA
AGCGGGAGGGGCCATGACTACGAGGCCCTGGGAGGTCACTTTAGGGA
GGGCTGTCCTAAAACCAGAAGCTTGGAGCAGAAAGTGAAACCCTGGT
GCTCCAGACAAAGATCTTAGTCGGGACTAGCCGGCCAAGGCGTCTCTA
GCTCTTGCCCTGAAAAAGGGCAGCTCTTGCCCTGAAAAAGGGCAGCT
CTTGCCCTGAAAAAGGGC

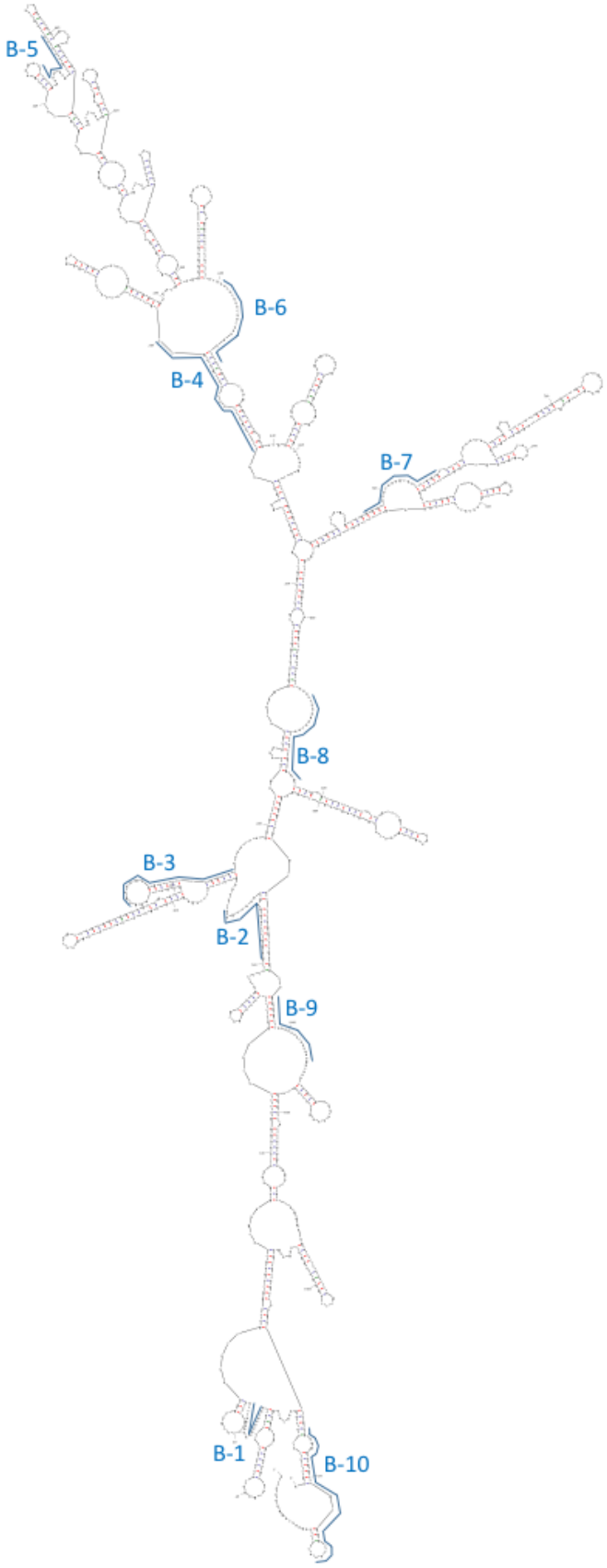
BoxB-A3GΔuORF:

GCCCTGAAAAAGGGCAAGCTTGGCCCTGAAAAAGGGCAAGCTTGGCC
CTGAAAAAGGGCAAGCTTGGAGACGCTCTTTCCCTTTGCAATTGCCTT
GGGTCCTGCCGCACAGAGCGGCCTGTCTTTATCAGAGGTCCCTCTGC
CAGGGGGAGGGCCCCAGAGAAAACCAGAAAGAGGGTGAGAGACTGA
GGAAGATAAAGCGTCCCAGGGCCTCCTACACCAGCGCCTGAGCAGGA
AGCGGGAGGGGCC*AACCCTGGTGCTCCAGACAAAGATCTTAGTCGG
GACTAGCCGGCCAAGGCGTCTCTAGCTCTTGCCCTGAAAAAGGGCAG
CTCTTGCCCTGAAAAAGGGCAGCTCTTGCCCTGAAAAAGGGC

BoxB-Scr:

GCCCTGAAAAAGGGCAAGCTTGGCCCTGAAAAAGGGCAAGCTTGGCC
CTGAAAAAGGGCAAGCTTGGAGACGAATCGCGCGAAAAGTGAATTGGT
AACGGCCCGCTAGAGTACATGGCGGTGTAGCAAACAGATCACCTACAG
AAGCTCCCGGTAGAACAGGCAAGGACCCTGAATGTGACCAGACTGCAA
CGACGGGCGCTAGTTGGTAGATGCGAAGGGTGAGAGAGTAAATGACG
GGAGATTCCGACCCGCCGGGTGTGAAAATGCTTTCGGACCAAACCGG
AATCACCAGCGCCGTCAGGTCACGTGGCCGTGTCGGTGTCACCGCAG
CGCAAGGATCCACCACCACTGAAAACCCTGATACGATCGTCTCTAGCTC
TTGCCCTGAAAAAGGGCAGCTCTTGCCCTGAAAAAGGGCAGCTCTTGC
CCTGAAAAAGGGC

Appendix 1: Sequences of the chimeric RNAs of BoxB-A3Gwt, BoxB-A3GΔuORF and BoxB-Scr. *BoxB stem loops are indicated in blue, the A3Gwt, A3GΔuORF and Scr sequences in purple, the uORF in red and the place where the uORF is deleted is marked with a red star.*



Appendix 2: Mfold predicted structure of the CDS of A3G mRNA and the hybridization sites of biotinylated oligos.

Name	Composition
Lysis buffer S	20 mM HEPES (pH 7.5); 100 mM KCl; 10 mM $MgCl_2$; 1 mM DTT; 0.5 % NP40; 0.1 u/ml RNAsin (Promega); 5 mM EDTA; 1x Halt protease inhibitor cocktail (ThermoScientific)
Hypotonic Lysis Buffer	10 mM HEPES (pH 7.5), 1.5 mM $MgCl_2$, 10 mM KCl, protease inhibitor mix, RNasin
Lysis Buffer K	469 mM LiCl ; 62.5 mM Tris HCl (pH 7.5); 1.25 % LiDS ; 1.25 % Triton X100 ; 12.5 mM RVC ; 12.5 mM DTT ; 125 U/mL RNasin ; 1.25 x Halt Protease Inhibitor Cocktail (ThermoScientific)
Lysis Buffer P	20 mM HEPES, 100 mM KAc, 2 mM MgAc, 1 mM DTT, 100 U/mL RNasin; 1 x Halt Protease Inhibitor Cocktail (ThermoScientific)
Lysis Buffer B	50 mM Tris (pH 7.5), 500 mM NaCl, 0.2 % SDS, 1 mM DTT
RIPA 1x	1x PBS, 1 % NP40, 0.5 % Na-DOC, 0.05 % SDS, 1x Halt Protease Inhibitor Cocktail (ThermoScientific)
Saturation Buffer	0.2 µg/µl total yeast tRNA; 0.5 µg/µl heparin; 0.64 U/µl RNasin; 20 mM HEPES (pH 7.5); 100 mM KCl; 10 mM $MgCl_2$; 0.01 % NP40
Formamide Blue	95 % formamide, 0.025 % xylene cyanol, 0.025 % bromophenol blue
Taq Buffer	75 mM Tris (pH 8.8), 20 mM ammonium sulfate, 0.01 % Tween20, 2 mM $MgCl_2$
RNase H Buffer	20 mM HEPES (pH 7.5); 50 mM KCl; 5 mM $MgCl_2$; 1 mM DTT ; 50 µg/µl BSA (ThermoScientific)
100-KCl Buffer	20 mM HEPES (pH 7.5); 100 mM KCl; 10 mM $MgCl_2$; 0.01 % NP40
500-KCl Buffer	20 mM HEPES (pH 7.5); 500 mM KCl; 10 mM $MgCl_2$; 0.01 % NP40
1000-KCl Buffer	20 mM HEPES (pH 7.5); 1 M KCl; 10 mM $MgCl_2$; 0.01 % NP40
Bead Washing Buffer	375 mM LiCl; 50 mM Tris HCl (pH 7.5); 1 % LiDS; 1 % Triton X100
Washing Buffer K	100 mM LiCl; 50 mM Tris-HCl (pH 7.5); 0.2 % LiDS; 0.2 % Triton X100
Wash Buffer B1	2 % SDS
Wash Buffer B2	50 mM HEPES, 500 mM NaCl, 0.1 % Na-DOC, 1 % Triton x100, 1 µM EDTA
Wash Buffer B3	10 mM Tris (pH 7.5), 250 µM LiCl, 0.5 % Na-DOC, 0.5 % NP40, 1 µM EDTA
Wash Buffer B4	50 mM Tris (pH 7.5)
ND Buffer	20 mM HEPES, 100 mM KCl, 10 mM $MgCl_2$
WB Buffer	1x NuPAGE LDS Sample Buffer (Invitrogen), 1x NuPAGE Sample Reducing Agent (Invitrogen)
1x TGS	25 mM Tris, 200 mM Glycine, 1 % SDS
WB Blocking Solution	50 mM Tris (pH 7.5), 150 mM NaCl, 1 % Triton x100, 5 % milk (Regilait)
TNT	50 mM Tris (pH 7.5), 150 mM NaCl, 1 % Triton x100
TN	50 mM Tris (pH 7.5), 150 mM NaCl

Appendix 3: List of all buffers. *All of the buffers used in this manuscript are indicated with their name and their composition.*

Translational inhibition of APOBEC3G by the HIV-1 Vif protein

Effect of Vif on A3G mRNA-associated protein complexes

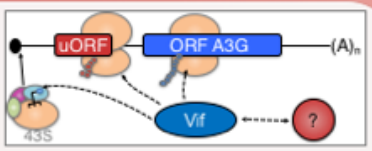
Tanja Seissler, Camille Libre, Roland Marquet & Jean-Christophe Paillart

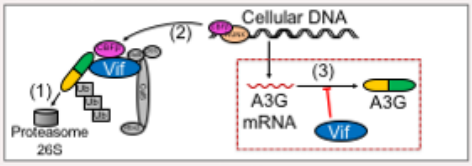
Université de Strasbourg, CNRS-UPR 9002, Architecture & Réactivité de l'ARN, IBMC, F-Strasbourg

Background

In order to defend themselves against infection by HIV-1, cells express the restriction factor APOBEC3G (A3G) which induces hypermutation of the viral genome thereby blocking viral replication.

HIV-1 expresses the protein Vif which counteracts A3G using 3 different strategies: (1) Vif induces A3G degradation by the proteasome; (2) Vif downregulates A3G transcription; (3) Vif inhibits A3G translation.





Translational inhibition of A3G by Vif is dependent on the 5' UTR of A3G mRNA and more precisely on a small ORF (uORF) upstream of the main A3G ORF.

The mechanism of this translational inhibition being still unknown, we have hypothesized that Vif might act on A3G mRNA-associated protein complexes like for example initiating ribosomes.

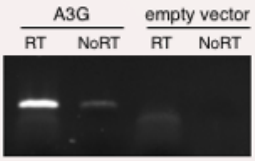
Method

Pull-down of A3G mRNPs from total cell lysates :

- HEK 293T cells are transfected with vectors allowing the expression of:
  A3G mRNA WT or ΔuORF
  + / - Vif

- Biotinylated oligos complementary to A3G mRNA allow retention of A3G mRNPs on streptavidin beads

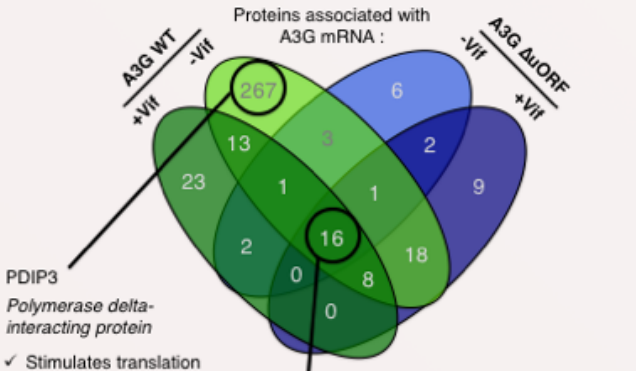
- Detection of A3G mRNA using RT-PCR



- Identification of proteins associated with A3G mRNA using mass spectrometry

Protein category	Enrichment Score
Nucleus	43,98
Splicing	40,83
Translation	37,08
RNA-binding	25,5
Nucleotide binding	22,62
Ribosomal biogenesis	18,21

Results



PDIP3 *Polymerase delta-interacting protein*

- ✓ Stimulates translation
- ✓ Present only on WT mRNA and in absence of Vif

SRSF1 *Serine/arginine-rich splicing factor*

- ✓ Implicated in splicing
- ✓ Enriched in the absence of Vif

Ratio -Vif/+Vif :	Replicate	A3G WT	ΔuORF
	1	10	0,6
	2	5	0,8

Other interesting proteins : KHDR1, ELAV1, RBM27 …

Conclusion

This study indicates that Vif might modify the composition of A3G mRNA-associated protein complexes. It might for example decrease the association of certain translation-stimulating factors (i.e. PDIP3, SRSF1) with A3G mRNA, thereby leading to a decrease in A3G translation.

Perspectives :

- ✓ Overexpression of candidate proteins
- ✓ Silencing of candidate protein expression

} Is Vif still able to inhibit A3G translation in these conditions?

→ Targeting translational inhibition of A3G by Vif (i.e. by inhibiting Vif interaction with necessary cellular proteins) is a promising strategy for new antiviral treatments.

Reference:
Guerrero S. X. et al., *Scientific Reports*, 2016 **6**, 39507

Acknowledgements:
Thank you to Philippe Hammann and his team for mass spectrometry analyses.

Appendix 4: Poster presented at the XIX Journées Francophones de Virologie (Paris, France, 30-31/03/2017) and at the Sidaction University for Young Researchers (Carry-le-Rouet, France, 14-20/10/2017).

Translational inhibition of APOBEC3G by the HIV-1 Vif protein

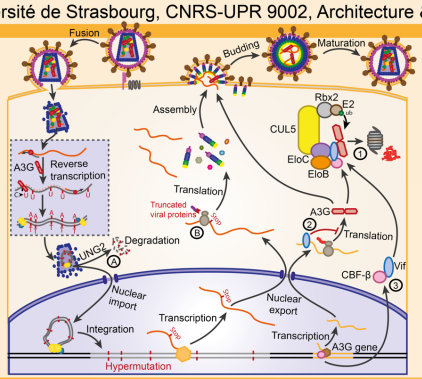
Effect of Vif on APOBEC3G mRNA-associated protein complexes

Tanja Seissler, Camille Libre, Roland Marquet & Jean-Christophe Paillart

Université de Strasbourg, CNRS-UPR 9002, Architecture & Réactivité de l'ARN, IBMC

Introduction

In order to defend themselves against infection by HIV-1, cells express the restriction factor **APOBEC3G (A3G)** which induces hypermutation of the viral genome.
This either leads to:
(A) degradation of viral DNA or
(B) accumulation of truncated viral proteins.



HIV-1 expresses the protein **Vif** which counteracts A3G using 3 different strategies:

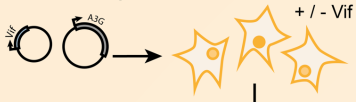
(1) Vif induces A3G degradation by the proteasome;
(2) **Vif inhibits A3G translation**;
(3) Vif downregulates A3G transcription.

Translational inhibition of A3G by Vif is dependent on the 5' UTR of A3G mRNA. The mechanism of this translational inhibition being still unknown, we have hypothesized that Vif might act on A3G mRNA-associated protein complexes like for example initiating ribosomes.

Methods

Pull-down of A3G mRNPs from total cell lysates

① HEK 293T cells are transfected with vectors allowing the expression of: A3G mRNA + / - Vif

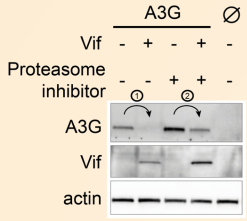


② Biotinylated oligos complementary to A3G mRNA allow retention of A3G mRNPs on streptavidin beads

③ Identification of proteins associated with A3G mRNA using mass spectrometry (RP-nanoLC-ESI-MS/MS)

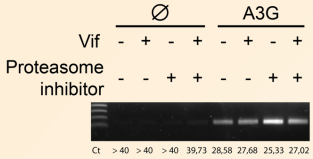
Results

Western Blot of total cell lysates



① Vif induces A3G degradation by the proteasome;
② Vif inhibits A3G translation;

Verification of A3G mRNA retention on beads using RT-qPCR



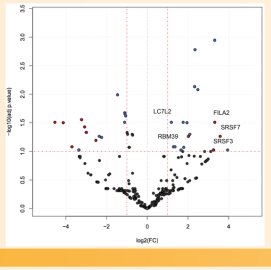
➔ A3G mRNA is pulled down successfully in all samples

Functional annotation clustering of total identified proteins (DAVID - https://david.ncifcrf.gov/)

Functional Cluster	Enrichment Score
translation	54,6
splicing	39,89
cadherin	35,79
nucleotide binding	26,56
RNA binding	19,96
mRNA export	14,48

➔ Identification of several mRNA associated proteins

Proteins identified by MS that interact with A3G mRNA



Proteins identified by MS that interact with A3G mRNA and increase / decrease in the presence of Vif

Proteins: Uniprot ID	Description	- Proteasome Inhibitor Replicate 1	Replicate 2	Replicate 3	+ Proteasome Inhibitor Replicate 1	Replicate 2	Replicate 3
SURF6_HUMAN	Surfeit locus protein 6	0,50	0,07		0,33	0,05	
TCP4_HUMAN	Activated RNA polymerase II transcriptional coactivator p15	10,07	0,12		0,07		
PR38B_HUMAN	Pre-mRNA-splicing factor 38B	12,67	0,31		0,24		
SRRM1_HUMAN	Serine/arginine repetitive matrix protein 1	1,50	0,04		0,17	0,03	
RBM39_HUMAN	RNA-binding protein 39	6,13	0,10		0,56	0,05	
U2AF2_HUMAN	Splicing factor U2AF 65 kDa subunit	2,50	0,09	0,10	0,10		
DDX46_HUMAN	Probable ATP-dependent RNA helicase DDX46		0,05		0,50	0,30	
NKAP_HUMAN	NF-kappa-B-activating protein	5,00	0,20		0,17	0,01	0,33
LCTL3_HUMAN	Luc7-like protein 3	3,00	0,13		0,25	0,01	
SRSF5_HUMAN	Serine/arginine-rich splicing factor 5	1,00	0,18	0,29	0,33		
SCLT1_HUMAN	Bcl-2-associated transcription factor 1	1,25	0,15	0,05	0,50		
IF4A3_HUMAN	Eukaryotic initiation factor 4A-III		0,20	0,33	0,50		
KPBP4_HUMAN	ADP-ribosylation factor-like protein 6-interacting protein 4		0,33		0,28		
SRSF2_HUMAN	Serine/arginine-rich splicing factor 2	2,00	0,01	0,05			
SRSF4_HUMAN	Serine/arginine-rich splicing factor 4	8,17	0,24		0,06		
RL23A_HUMAN	60S ribosomal protein L23a	5,33	0,10		0,42		
RL8_HUMAN	60S ribosomal protein L8	2,00	0,07	0,25			
RL29_HUMAN	60S ribosomal protein L29	2,00	0,06	0,43			
HNRPM_HUMAN	Heterogeneous nuclear ribonucleoprotein M		0,39		0,20		
K7EJV9_HUMAN	60S ribosomal protein L23a	9,00	0,17		0,33		
RS27A_HUMAN	Ubiquitin-40S ribosomal protein S27a		0,60	0,25	0,40		
RL40_HUMAN	Ubiquitin-60S ribosomal protein L40		0,60	0,25	0,40		
MMTA2_HUMAN	Multiple myeloma tumor-associated protein 2		0,06		1,00	0,25	
RNPS1_HUMAN	RNA-binding protein with serine-rich domain 1		0,10		0,50		
RBM14_HUMAN	RNA-binding protein 14		0,25	0,20			
RSRC2_HUMAN	Arginine/serine-rich coiled-coil protein 2	4,00	0,20		0,13	0,10	
B4DES9_HUMAN	cDNA FLJ60424, highly similar to Junction plakoglobin	2,00	7,00				
K7C1R_HUMAN		2,00	3,00				

Ratio of protein spectral counts +Vif / -Vif:
Red: proteins decreased by Vif
Green: proteins increased by Vif

SURF6: potentially involved in ribosomal assembly

IF4A3: RNA helicase, core component of the exon junction complex

Conclusion

This study suggests that Vif might modify the composition of A3G mRNA-associated protein complexes. It might for example decrease the association of certain translation-stimulating factors (i.e. SURF6, IF4A3) with A3G mRNA, thereby leading to a decrease in A3G translation.

Perspectives:
➔ Silencing of candidate protein expression
➔ Overexpression of candidate proteins

Is Vif still able to inhibit A3G translation in these conditions?

Targeting translational inhibition of A3G by Vif (i.e. by inhibiting Vif interaction with necessary cellular proteins) is a promising strategy for new antiviral treatments.

References:
Guerrero S. X. et al., Scientific Reports, 2016 6, 39507
Seissler T. et al., Viruses, 2017 9, 9110322

Acknowledgements:
Thank you to Philippe Hammann and his team for mass spectrometry analyses and to Béatrice Chane-Woon-Ming for bioinformatics.

Appendix 5: Poster presented at the Viruses 2018 Conference on Breakthroughs in Viral Replication (Barcelona, Spain, 7-9/02/2018).

Translational inhibition of APOBEC3G by the HIV-1 Vif protein

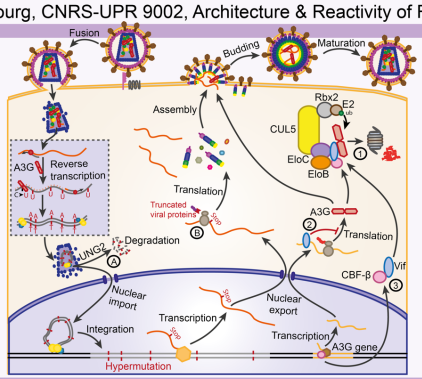
Effect of Vif on APOBEC3G mRNA-associated protein complexes

Tanja Seissler, Camille Libre, Roland Marquet & Jean-Christophe Paillart

University of Strasbourg, CNRS-UPR 9002, Architecture & Reactivity of RNA, IBMC, Strasbourg, France

Introduction

In order to defend themselves against infection by HIV-1, cells express the restriction factor **APOBEC3G (A3G)** which induces hypermutation of the viral genome.
This either leads to:
(A) degradation of viral DNA or
(B) accumulation of truncated viral proteins.



HIV-1 expresses the protein **Vif** which counteracts A3G using 3 different strategies:

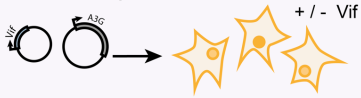
(1) Vif induces A3G degradation by the proteasome
(2) **Vif inhibits A3G translation**
(3) Vif downregulates A3G transcription

Translational inhibition of A3G by Vif is dependent on the 5' UTR of A3G mRNA. The mechanism of this translational inhibition being still unknown, we have hypothesized that Vif might act on regulatory A3G mRNA-associated protein complexes.

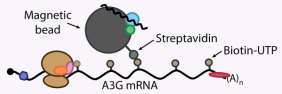
Methods

Pull-down of A3G mRNPs from total cell lysates

① HEK 293T cells are transfected with vectors allowing the expression of: + / - A3G
+ / - Vif



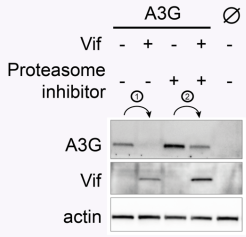
② Pull-down of proteins from cellular lysate on an in vitro-transcribed, biotinylated A3G mRNA.



③ Identification of proteins associated with A3G mRNA using mass spectrometry (RP-nanoLC-ESI-MS/MS).

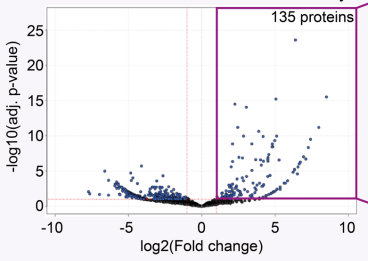
Results

Western Blot of total cell lysates

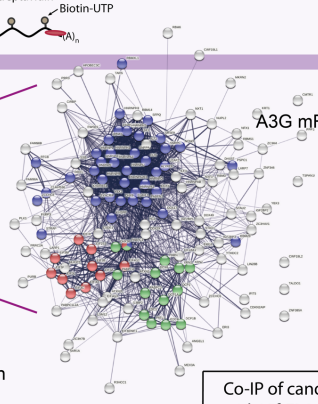


① Vif induces A3G degradation by the proteasome
② Vif inhibits A3G translation

Proteins on A3G mRNA / Beads only

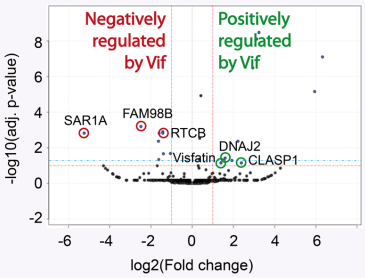


➔ 135 proteins enriched on A3G mRNA with a p-value < 0,1 and a fold change > 2



A3G mRNA interactome
- Splicing
- Translation initiation
- mRNA degradation

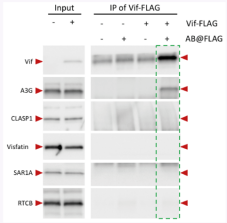
Proteins on A3G mRNA in the presence / absence of Vif



SAR1A	Involved in transport from the endoplasmic reticulum to the Golgi
FAM98B	Positively stimulates arginine methylation
RTCB	Catalytic subunit of the tRNA-splicing ligase complex
Visfatin	Catalyses an intermediate step in the NAD biosynthesis pathway
DNAJ2	Stimulates ATP hydrolysis and the folding of unfolded proteins
CLASP1	Promotes the stabilization of dynamic microtubules

➔ Proteins potentially recruited on A3G mRNA by Vif or dropped off from A3G mRNA by Vif

Co-IP of candidate proteins with Vif-FLAG:



➔ None of the identified proteins directly interact with Vif

Conclusion

This study suggests that Vif might modify the composition of A3G mRNA-associated protein complexes. Vif might for example decrease the association of certain translation-stimulating factors with A3G mRNA, or increase association of inhibitory factors, thereby leading to a decrease in A3G translation.

Perspectives:
➔ Silencing of candidate protein expression
➔ Overexpression of candidate proteins

Is Vif still able to inhibit A3G translation in these conditions?

Targeting translational inhibition of A3G by Vif (i.e. by inhibiting Vif interaction with necessary cellular proteins) is a potential strategy for new antiviral treatments.

References:
Guerrero S. X. et al., Scientific Reports, 2016 6, 39507
Seissler T. et al., Viruses, 2017 9, 9110322

Acknowledgements:
Thank you to Philippe Hammann and his team for mass spectrometry analyses and to Béatrice Chane-Woon-Ming for bioinformatics.

Appendix 6: Poster presented at the Retroviruses 2019 conference (Cold Spring Harbor Laboratory, USA, 20-25/05/2019).

Tanja SEISSLER

Translational inhibition of the Restriction Factor APOBEC3G (A3G) by the HIV-1 Vif protein: Role of a uORF in the 5'-UTR of A3G mRNA and identification of cellular factors

The HIV-1 Vif protein counteracts the restriction factor APOBEC3G (A3G) by downregulating its expression level in infected cells. This is achieved in different ways, one of which is translational inhibition, a mechanism that is still poorly understood. The first part of my thesis contributes to the characterization of a small upstream ORF (uORF), that is found in the 5'-UTR of A3G and A3F mRNAs. This uORF has been found to be crucial for regulation of A3G translation and is necessary to allow Vif-mediated translational inhibition. In the second part of this thesis, different protocols have been set up in order to identify A3G mRNA-associated cellular proteins which might play a role in the mechanism of Vif-mediated translational inhibition. Several proteins, whose presence on A3G mRNA seems to be modulated by Vif have been identified.

Keywords: APOBEC3G, HIV, Vif, uORF, translation, RNA-protein interaction, RNA pull-down,

Inhibition traductionnelle du facteur de restriction APOBEC3G (A3G) par la protéine Vif du VIH-1 : Rôle d'une uORF dans la 5'-UTR de l'ARNm d'A3G et identification de facteurs cellulaires

La protéine Vif du VIH-1 contrecarre le facteur de restriction APOBEC3G (A3G) en diminuant son niveau d'expression dans les cellules infectées. Ceci est mis en œuvre entre autres par l'inhibition de sa traduction, un mécanisme encore peu compris. La première partie de ma thèse contribue à la caractérisation d'une petite ORF (uORF) qui se situe dans la 5'-UTR de l'ARNm d'A3G et d'A3F en amont de leurs ORF respectives. Cette uORF s'est révélée cruciale pour la régulation de la traduction d'A3G en présence et absence de Vif. Dans la deuxième partie de cette thèse, différents protocoles ont été mis en œuvre pour identifier les protéines associées avec l'ARNm d'A3G, qui pourraient jouer un rôle dans le mécanisme d'inhibition traductionnelle d'A3G par Vif. Ainsi, plusieurs protéines ont été identifiées dont la présence sur l'ARNm d'A3G semble modulée par Vif.

Mots clés : APOBEC3G, VIH, Vif, uORF, traduction, interaction ARN-protéine, pull-down d'ARN