UNIVERSITÉ DE STRASBOURG

École Doctorale des Sciences de la Vie et de la Santé
Dynamic of Host-Pathogen Interactions – DHPI

THÈSE présentée par:
Marco De Rovere
Soutenue le: 02 Décembre 2020

Pour obtenir le grade de: Docteur de l'université de Strasbourg
Discipline: Science du Vivant
Spécialité: Aspects cellulaire et moléculaire de la Biologie

Identification of CTIP2 new proteins and ncRNAs interactants in microglial cells

THÈSE dirigée par:

M. ROHR Oliver	Professeur, Université de Strasbourg

RAPPORTEURS:

M. HERBEIN Georges	Professeur, Université de Franche-Comté
Mme. MARBAN Céline	Principal Scientist at F. Hoffmann-La Roche

EXAMINATEURS :

M. PAILLART Jean-Christophe	Directeur de recherche IBMC, Université de Strasbourg
Mme. KULA-PACURAR Anna	Assistant-professor, Jagiellonian University, Poland

Acknowledgements:

When you reach the end of a journey you need to take a moment to stop on your tracks, look back, and see the multitude of people that have assisted you to get till this point.

To the members of my jury: Professor Georges Herbein, Doctor Céline Marban, Professor Jean-Christophe Paillart and assistant-professor Anna Kula Pacurar. Time is the most precious resource humans have and you have used yours to read and correct my work. I will be forever grateful for this.

My biggest thanks go to my thesis director, Professor Olivier Rohr. As a leader you must be able to manage all kind of people, all kind of characters, to obtain the maximum output. I have my shortcoming, as a person and as a scientist; I am stubborn, and I dislike asking for help. But anytime I hit a wall in these past years he was there able to put me back on my feet. Every time I ended up in a dead end, he was ready to put me back on the right path.
You should not aim to be better than the others but to be better than who you were. I am now a better person of who I was 4 years ago. And it is especially thanks to him.

To Doctor Christian Schwartz, for his constant positivity and for never refusing to help me in whatever task regarding PCRs I was stuck in.

To Doctor Clémentine Wallet, for being the mature person in the lab we always needed. You are the sister to all of us and we can always count on you.

To Doctor Haitham Sobhy, which tireless and impeccable work made possible the analysis of my most complicated data.

To Fadoua Daouad, I named you the soulless soul of the lab years ago and I stand by it. We would all be lost without you managing the lab.

To Jeanne Van Assche, from my first student to full colleague, I am always proud to see you doing, at each time point of you PhD, better than how I was doing. Also, we may have disagreed sometimes, but it was always with the highest mutual respect.

To my father, who wanted a pilot, or a mechanical engineer, but by buying me as a child, microscopes, telescope, My little Chemist station and recording on VHS every episode of "*Il était une fois... la Vie"* formed the basis of a biologist.

To my mother, to best disciplinary model a boy could ask. To always showing me the paths but never forcing me upon anyone of them, and then to all the sacrifices she made for get me to this point.

To my girlfriend Yoseline, for being my best friend in these years abroad, and for relatively not stressing me too much while I was writing the thesis.

To my amazing Dungeons&Dragons group, David, Filipa, Guido, Renata, Rogeria e Sayali, I am an asocial bear, but you take out the best of me.

To Matteo, my last remaining friend from the original group of Italians, I cannot be abroad without a climbing companion, hopefully we will go again soon..or better to surf.

To my friends, Daniele, Eleonora, Franco, Francesco, Ped, Giulio, Lozy, Aaron, Roccia, Davide, Andrea, Ciullo, Gaz, Mattia, Batti, Kelly. You are the family I have chosen for myself, and this should be enough proof of the mental issues I have. I love you all, horrible people.

ABBREVIATIONS 7

1. INTRODUCTION 11

1.1. HIV and AIDS 11

1.1.1. HIV Taxonomy and Origin 11

1.1.2. Epidemiology 12

1.2. The HIV-1 retrovirus 14

1.2.1. HIV 1 structure 14

1.2.2. HIV-1 Genomic organization 16

1.2.2.1. Organization of the Genomic RNA 16

1.2.2.2. Non-Coding Regions 16

1.2.2.3. Coding Regions 17

1.2.3. HIV-1 replication cycle 19

1.2.3.1. HIV-1 targets of infection and entry strategies 20

1.2.3.2. Uncoating and reverse transcription 21

1.2.3.3. Viral gDNA Import in the nucleus 22

1.2.3.4. Integration 23

1.2.4. Post-Integration phase and viral gene expression 23

1.2.4.1. Tat independent phase 25

1.2.4.2. Tat-dependent phase 26

1.2.5. Production of new virions 27

1.2.5.1. Assembly 28

1.2.5.2. Budding 28

1.2.5.3. Maturation 28

1.3. HIV-1 infection 29

1.3.1. HIV-1 tropism 30

1.3.1.1. Infection of the CD4+ Lymphocytes T 30

1.3.1.2. Infection of the dendritic cells 31

1.3.1.3. Infection of monocytes and macrophages 32

1.3.1.4. Infection of the microglial cell 33

1.3.1.5. The infection of the Central Nervous System 34

1.4. Combination Anti-Retroviral Therapy 36

1.4.1. Success of cART 36

1.4.2. Failure of cART 37

1.4.3. HIV-1 Vaccine 39

1.5. HIV-1 latency 39

1.5.1. Type of latency 40

1.5.2. Types of cellular reservoirs 41

1.5.3. HIV-1 gene silencing and host transcription factors 42

1.5.4. Transcriptional interference 42

1.5.5. Site of integration and chromatin 43

1.5.6. Epigenetic silencing of the HIV-1 promoter 45

1.6. Strategies to target and remove the latently infected cells 47

1.6.1. Shock and Kill 48

1.6.2. Block and Lock 50

1.7. Bcl11b/CTIP2 52

1.7.1. CTIP2 as a transcription factor 54

1.7.2. CTIP2 seize the elongation factor P-TEFb 55

1.7.3. CTIP2 and HIV-1 56

1.8. PSF/SFPQ 57

1.8.1. SFPQ is a member of the DBHS family 57

1.8.2. PSF/SFPQ localization 58

1.8.3. PSF/SFPQ functions 58

1.8.4. Post-Translational modification of PSF/SFPQ 61

1.9. Paraspeckle 62

1.9.1. Paraspeckle components 62

2. MATERIAL AND METHODS 67

2.1. Cell culture 67

2.2. Plasmids 67

2.3. Immunoprecipitation assay (IP): 67

2.4. SDS-PAGE and Western Blot Analysis: 68

2.5. Mass spectrometry 68

2.6. RNA-Immunoprecipitation (RIP): 68

2.7. Cross-Linking Immunoprecipitation Sequencing (CLIP-Seq): 69

2.8. Luciferase Assay 70

2.9. CLIP sequencing and bioinformatics analysis 70

2.10. Primers: 71

2.11. Statistical Analysis: 72

3. OBJECTIVE OF THE THESIS 73

4. PUBLICATION 1 74

5. PUBLICATION 2 85

6. PUBLICATION 3 103

7. PUBLICATION 4 128

8. RESULTS 148

8.1. Identification of new protein partners of CTIP2 by quantitative mass spectrometry 149

8.2. CTIP2/SFPQ interaction 150

8.3. CTIP2/SFPQ/Tat interaction 153

8.4. Role of SFPQ in HIV-1 gene transcription and viral replication 155

8.5. CTIP2/FUS interaction 160

8.6. CTIP2/YTHDF family interaction 164

8.7. CTIP2/MATR3 interaction 166

8.8. Identification of new RNAs partners of CTIP2 by CLIP-Seq 167

9. DISCUSSION AND FUTURE PROSPECTIVE 174

9.1. CTIP2 interaction with paraspeckle components 175

9.2. CTIP2 interaction with the YTHDF family 179

9.3. CTIP2 interaction with the MATR3 180

9.4. CTIP2 newly discovered ncRNA partners 180

10. CONCLUSION 181

11. REFERENCES 182

Abbreviations

5' UTR 5' Untranslated Region
7SK snRNP 7SK small nuclear ribonucleoprotein
aa Amino acid
ADAR Adenosine Deaminase RNA Specific
ARF Alternative Reading Frame
BBB Blood–Brain Barrier
BCL11b B Cell Leukemia 11b
Bcl-2 B-Cell Lymphoma
BMI1 B lymphoma Mo-MLV (Moloney Murine Leukemia Virus) Insertion region 1 homolog
Brd4 Bromodomain containing protein 4
CA Capside
CBC Cap Binding Complex
CBP CREB Binding Protein
CBX4 Chromobox protein homolog 4
CCL19 C-C Motif Chemokine Ligand 19
CCR5 Chemokine (CC motif) Receptor 5
CDK9 Cyclin Dependent Kinase 9
cDNA Complementary Deoxyribonucleic acid
CLIP Cross-Linking Immunoprecipitation
CSF Cerebro Spinal Fluid
CTD C-Terminal Domain
CTIP2 Coup-TF (Chicken ovalbumin upstream promoter Transcription Factors) Interacting Protein 2
CTL Cytotoxic T Lymphocyte
CTLA-4 Cytotoxic T-Lymphocyte Associated Protein 4
CXCR4 Chemokine (CXC motif) Receptor 4
DBHS Drosophila behavior human splicing
DC-SIGN Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Nonintegrin
DDX3 DEAD Box Helicase 3
DNA Deoxyribonucleic acid

DNA-PKcs DNA-dependent Protein Kinase catalytic subunit

DNMT DNA Methyltransferase

dNTP Deoxynucleoside Triphosphate

DSIF DRB (5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole) Sensitivity Inducing Factor

eIF4G Eukaryotic Initiation Factor 4G

ELL Elongation factor for RNA polymerase II

ESCRT Endosomal Sorting Complexes Required for Transport

FUS FUS RNA Binding Protein

GALT Gut Associated Lymphoid Tissue

gp120/41 Glycoprotéine 120/41

HAART Highly Active Anti-Retroviral Therapy

HAND HIV-Associated Neurocognitive Disorder

HAT Histone Acetyl Transferase

HDAC Histone Deacetylase

HEXIM1 Hexaméthylène bis-acétamide inducible 1

HIC1 Hypermethylated In Cancer 1

HMGA1 High Mobility Group AT-hook 1

HMT Histone Methyl Transferase

HP1 Heterochromatin Protein 1

HTLV-1 Human T-Lymphotropic Virus type 1

HUSH Human Silencing Hub

IF Interferon

IN Integrase

IRES Internal Ribosome Entry Site

IRF7 Interferon Regulatory Factor 7

IRF3 Interferon Regulatory Factor 3

KMT7 Lysine methyltransferase 7

KRAB-ZFPs Krüppel-Associated Box Zinc Finger Proteins

LANA Latency-Associated Nuclear Antigen

LARP7 Lupus Antigen Related Protein 7

LPA Latency Promoting Agents

LPS Lipopolysaccharides

LRA Latency Reversing Agent

LRP Lipoprotein Receptor-related Protein

LSD1 Lysine-Specific histone Demethylase 1A

LTNP Long-Term Non-Progressors

LTR Long Terminal Repeat

MA Matrix

MALAT1 Metastasis associated lung adenocarcinoma transcript 1

MEG3 Maternally Expressed 3

miRNA Micro Ribonuclei acid

mRNA Messenger Ribonucleic acid

NC Nucleocapside

NCOR1 Nuclear receptor corepressor 1

NEAT1 Nuclear Enriched Abundant Transcript 1

Nef Negative regulatory factor

NELF Negative Elongation Factor

NEMO NF-κB Essential Modulator

NES Nuclear Export Signal

NFAT Nuclear Factor of Activated T-cells

NF-kB Nuclear Factor-kB

NLS Nuclear Localization Sequence

NNRTI Non-Nucleoside Reverse Transcriptase Inhibitor

NO Nitric oxide

NONO Non-POU Domain Containing Octamer Binding

NPC Nuclear Pore Complex

NRTI Nucleoside Reverse Transcriptase Inhibitor

NuRD Nucleosome Remodeling and Deacetylation

ORC Origin of Replication Complex

ORF Open Reading Frame

PABP Poly(A)-Binding Protein

PARP1 Poly(ADP-Ribose) Polymerase 1

PBMC Peripheral Blood Mononuclear Cell

PCAF p300/CBP-Associated Factor

PIC Pre-Integration Complex

PKR Protein Kinase R

PR Protease

PRC2 Polycomb Repressive Complex 2

PRKDC Protein Kinase DNA-activated Catalytic polypetide

PRMT6 Protein arginine methyltransferase 6

PSF PTB-Associated Splicing Factor

PSPC1 Paraspeckle Component 1

P-TEFb Positive Transcription Elongation Factor b

Rev Regulator of virion expression

RNA Ribonucleic acid

RNP Ribonucleoprotein Complex

RNF RING Finger protein

RT Reverse transcriptase

RTC Reverse Transcription Complex

SAMHD1 Sterile Alpha Motif domain and HD domain-containing protein 1

SENPs Sentrin-specific Protéases

SFPQ Splicing factor, proline- and glutamine-rich

SIRT1 Sirtuin 1

SNORD133 Small Nucleolar RNA, C/D Box 133

SP1 Specificity Protein 1

SUV39h1 Suppressor of Variegation 3-9 Homolog 1

TAFII250 TATA box binding protein Associated Factor RNA polymerase II 250 kDa

TAR Transactivation Response element

Tat Transactivator of Transcription

TCM Central memory CD4 T cells

TCR T Cell Receptor

TEM Effector memory CD4 T cells

TF Transcription factor

TGF Transforming Growth Factor

TLR Toll-Like Receptor

TNA Naïve CD4+ T cells

UTR Untranslated region

Vif Viral infectivity factor

Vpr Viral protein r

Vpu Viral protein u.

YTHDF YTH domain-containing family protein

1. Introduction

1.1. HIV and AIDS

The acquired immune deficiency syndrome (AIDS) is a disease that was discovered for the first time in the United States of America in 1981 by the Centre for Disease Control after hospitals and doctors noticed an alarming increases of cases of a rare skin cancer, the Kaposi sarcoma and a pulmonary infection from a yeast, *Pneumocystis jirovecii*. These two specific diseases are extremely rare in the healthy population and thrive only in immunodepressed organisms, also the main population exhibiting these symptoms were mostly homosexual men with multiple sexual partners and/or an history of use of injected drugs.

The first theory was that it should have been a sexually transmitted organism and later on this theory was reinforced when patients that had received infected blood started showing symptoms of the disease (CDC 1982). This led the researchers to start investigating the possible source of a what was now called AIDS. In 1983 almost at the same time, Professor Gallo in the USA isolated a virus named Human T-cell lymphotropic virus (HTLV-III) (Gallo et al., 1983) and Professors Montagnier and Barré-Sinoussi in France isolated another lymphotropic virus named Lymphadenophaty associated virus (LAV) (Barre-Sinoussi et al., 1983); soon after, they came to the conclusion that the virus isolated was indeed the same, and this organism was responsible for the destruction of the host organism immune defences and to be the etiological cause of the acquired immune deficiency syndrome. It was named Human Immunodeficiency virus, or HIV.

1.1.1. HIV Taxonomy and Origin

The Human Immunodeficiency virus 1 and 2 are Lentiviruses and are part of the family *Retrovidae.* While the disease, which the virus is responsible of, was first detected in the USA, the origin of the virus was still unknown. In 1999, (Gao at al., 1999) through mitochondrial DNA analysis, was able to locate the origin of HIV-1 in West Central Africa, more specifically in the *Pan troglodytes* population of Cameroon. Within these chimpanzee populations, are present different strains of Simian Immunodeficiency Viruses (SIV) and the phylogenetic analysis linked some of them to the HIV-1 strains M (Major), N (New) and O (Outlier) and the

newly discovered P which is also endemic in the region (Chaneau et al., 1994, Simon et al., 1998). This results supported the hypothesis that the origin of HIV-1 in humans was from not one but different zoonosis events with humans infected with the SIV; although this virus shows a low person to person transmission rate, a cause still to be identified must have exacerbated this transmission rate allowing the SIVcpz to evolve enough to be more resilient to the human immune system and take advantage of it (Marx et al., 2001). Other research conducted to pinpoint more precisely the origin of the virus also added a new link as an intermediate step between the chimpanzee and the humans in the form of the SIVgor infecting the Western lowland gorillas (*Gorilla gorilla gorilla*) (Takehisa at al., 2009). The most recent phylogenetic analysis has linked the strain M to have derived from the SIVcpz and the N and O from the SIVgor (Li et al., 2012, D'Arc et al., 2015). Once the origin has been discovered a process of molecular phylogenetic evidence has started to pinpoint the time in which this zoonosis has created the first HIV-1 in humans; the search for the first common ancestor for the strain M suggest that the zoonosis event may have happened between 1915 and 1940 (Korber et al., 2000). HIV-2, which is also represented in the population by different subtypes (A to H) and with each of them arised from an independent zoonosis event, is associated with lower infectivity and a lower rate of AIDS upon infection. (Silva et al., 2008; Sharp and Hahn, 2011; Campbell-Yesufu and Gandhi, 2011)

1.1.2. Epidemiology

Since HIV-1 discovery, the WHO has estimated that more than 34 million life were lost due to HIV-1 related diseases. While the number of people living with HIV-1 has increased steadily (14% compared to the 2010), with the last report of the WHO in 2018 estimating between 31.1 million and 43.9 million infected people, a comparison of these data with the 2010 shows a reduction of the new infections of 18% and a decrease in death of 34% (Figure 1). The distribution between genders regarding the infected is evenly split but the demographic changes based on the region taken in account. First of all, Eastern and Southern Africa remains the regions with the greatest number of infected people with 25.7 million people currently living with HIV.

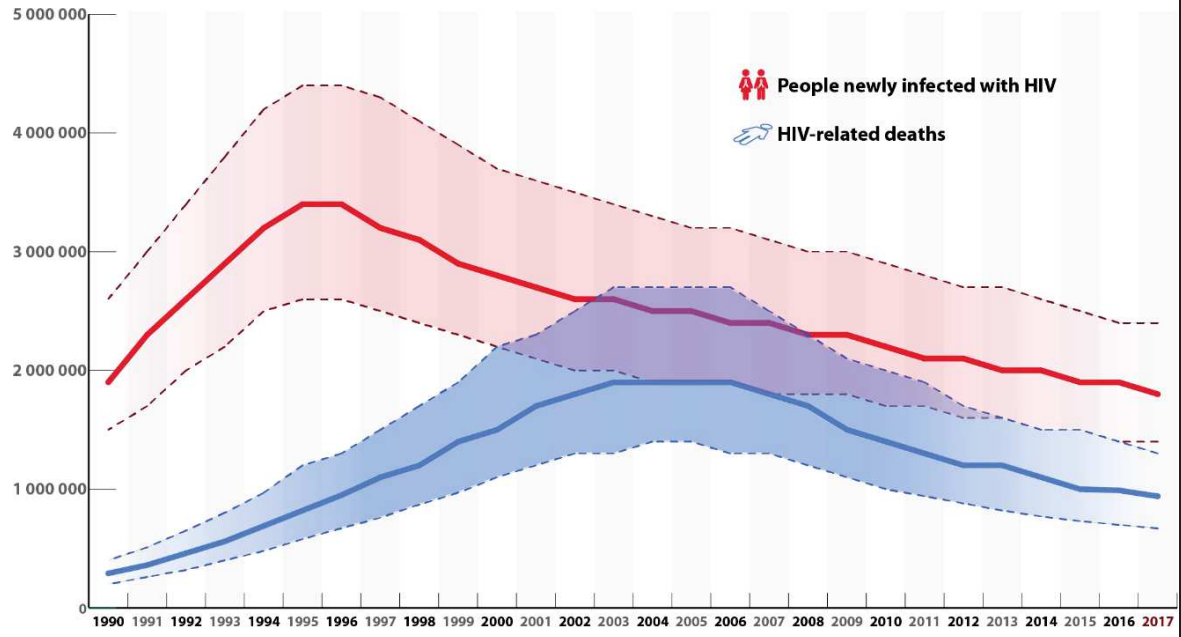


Figure 1: HIV incidence and mortality over time. Adapted from UNAIDS and WHO 2018

As noted, the steady increase of infected people should be seen as a positive data and can be totally linked to the increasing access of infected people to Anti-Retroviral therapies, which have led the decrease of the death toll worldwide. As seen in Figure 1 after the mortality peak of 2004, a steadily decrease of mortality is shown. This decrease can be linked to the cART coverage among the infected population worldwide that show how, from the last data of the WHO, almost 60% of the infected population is currently under cART regimen (Figure 2). The remaining 40% who are not under treatment, while seeming enormous, has to take in account how often, newly infected people are not aware of their status till later on when they start developing symptoms, but with more cheap and fast diagnostic tools there is the hope to be able to cover as much as 30 million people by 2020 (UNAIDS, 2018; WHO; 2018).

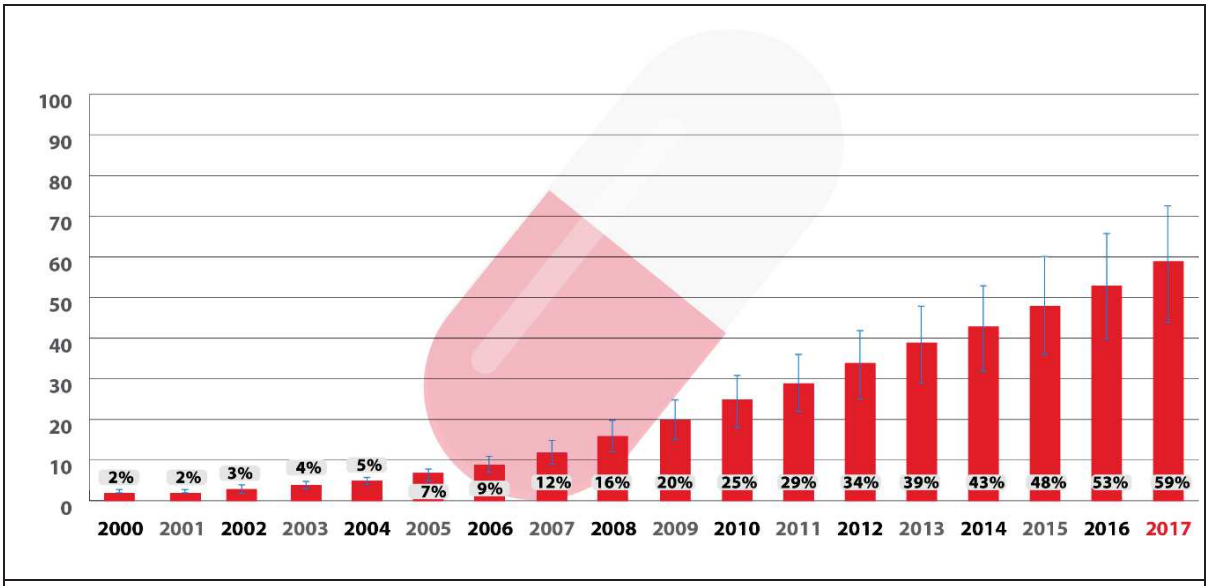


Figure 2: Cominbed Anti-Retroviral Theracpy (cART) over time. Adapted from from UNAIDS and WHO 2018.

1.2. The HIV-1 retrovirus

In the realm of the *Riboviria*, in the family *Retroviridae*, subclass *Orthoretrovirinae,* the HIV-1 is classified in the genera *Lentivirus* (Virus Taxonomy 2019). The *Retroviridae* is a family comprising many viruses infecting birds and mammals, they share an icosahedral geometry and their size range between 70 to 120nm of diameter. The viral genome is composed by two strands of positive RNA, identic and their size is around 9.2 kilobases. These strands of nucleic acids are tightly associated with the nucleocapsid protein forming a Ribonucleoprotein complex (RNP) (Coffin et al., 1997). One of the most distinguishing characteristics of these viruses is the presence of a protein named Reverse transcriptase, from which they take their name, able to copy the genomic RNA in a double strand of DNA. This process, going in reverse of the canonic DNA-RNA pathway, gave the "Retrovirus" name of this family.

1.2.1. HIV 1 structure

The morphology of the HIV-1 virion has been established for the first time in the late 1980 by electron microscopy and cryo-microscopy (Gelderblom et al., 1987). It is a roughly spherical virus (Figure 3), with a 120nm-140nm diameter on average, with protruding proteins projections included in the viral envelope. The core, visible at the electron microscopy, is cone shaped, formed by multimers of the p24 and capsid protein (CA) is contained by the Matrix formed by the (MA p17) proteins. The exterior envelope is formed by a lipid bilayer derived

from the host cell at the moment of the budding; this peculiarity is also one of the reason why till this day, a vaccine against HIV-1 has been so difficult to develop, due to the fact that the major part of the exterior surface of the virion is made of material from the previous cell infected from which the virion has originated and are not recognized by the host immune system (Aloia et al., 1993). From the envelope protrudes a trimer of transmembrane glycoproteins (TM, gp41) connected with the surface glycoprotein (SU, gp120) which has the function to dock the virion to the target cell before infecting it by interacting specifically with the CD4 receptor (Dalgleish et al., 1984; Lu M., 1995). Within the core, two copies of positive sense RNA of 9200 nucleotides are stabilized as an RNP complex by the nucleocapsid protein (Briggs et al., 2003). These molecules harbours the same post transcriptional modification of the normal messenger RNA of the cell with a polyA tail at the 3'end. In the core are present as well the viral enzymes necessary for its process of integration to continue its life cycle; these proteins with enzymatic function are the reverse Transcriptase (RT), the Protease (PR) and the Integrase (IN), and other accessory proteins such Vpr (Accola et al., 2000), Nef (Kotof et al., 1999) and Vif (Vasudevan et al., 2013) and p7 the nucleocapsid protein (NC) bound specifically the viral RNA. Figure 3 shows a schematic of the HIV-1 virion.

Within the virion are enclosed as well molecules originated in the host cell from which the virion has come from such actin molecules (Wilk et al., 1999) tRNA, $tRNA_{Lys3}$ and tRNA synthetase to support the retro transcription (Kleiman, L., 2002).

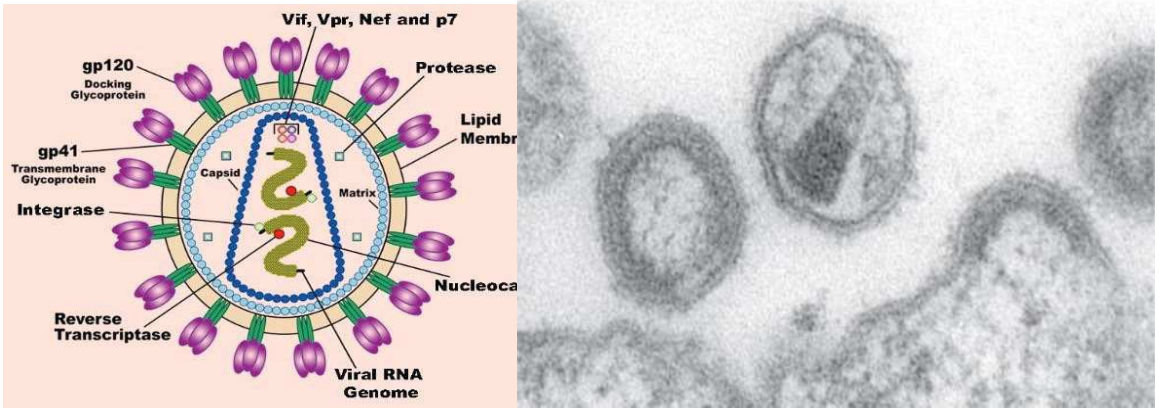


Figure 3: Schematic representation of the HIV-1 virion and electron microscopy of a mature (right) and immature (left) HIV-1 virion particles. Image credit: US National Institute of Health and (Adapted from Baumgartel et al., 2012)

1.2.2. HIV-1 Genomic organization

The HIV-1 genome can be found in two different forms. Within the virion particle as genomic RNA and then within the target host cell, during the replication process, this RNA becomes genomic DNA after it is reverse transcribed. The HIV-1 gDNA is then integrated withing the host cell own genome, taking the name of proviral DNA.

1.2.2.1. Organization of the Genomic RNA

When describing the genomic RNA, it is necessary to remind that it come from the already integrated proviral DNA. Once it is transcribed, the viral RNA forms a secondary structure which helps in the process of the viral replication; the trans-activation response (TAR) element is a key factor in the formation of this secondary structure. TAR is located at the 5' end of the viral RNA transcript and forms a hairpin structure which, through a lateral protuberance of three pyrimidines, act as binding site for another viral protein Tat to augment the activity on the 5' LTR promoter (Dingwall et al., 1989; Lu et al., 2011). Recently the TAR element has been also linked to its ability to dimerize the viral RNA, helping in the process of packaging within the newly formed virions as well as reverse transcribed it and integration of it in the host cell (Vrolijk et al., 2008; Mousseau et al., 2015). Another element involved in the RNA secondary structure is the Rev Response Element (RRE) which can be found in the Env coding region. It is present in unspliced and partially spliced viral mRNA. By interacting with the viral protein r*ev* and other host cofactors as MATR3, it contributes to the export of the unspliced viral RNA outside the nucleus (Cullen B. R., 2003; Sarracino et al., 2008). Other RNA secondary structures, as the SLIP, function as ribosomal frameshift motifs regulating for example, the shift between the *pol* and *gag* (in the case of the TTTTT motif) Open Reading frame (ORF). The function of acting as acceptor and donor sites for splicing has well been explored together with their function in offering polyadenylation signals, function in genomic RNA dimerization and translation initiation (Heng et al., 2012; Olson et al., 2015).

1.2.2.2. Non-Coding Regions

The genomic RNA of the HIV-1 could also be divided in coding and non-coding regions. The non-coding regions, while not responsible for the production of viral proteins, play important roles in different stages of the HIV-1 replication cycle. At both the 5' and 3' is present a Long

Terminal Repeat (LTR) region which is divided as well in several subregions, the U3, R, and U5 regions: The R region of HIV-1 has a length of 97 nt and encodes two well-conserved stem-loop structures, the TAR and poly(A) hairpins. The TAR with its role in activating the transcription thanks to Tat binding, the PolyA is responsible for dimerization and genome packaging. The U5 region constitutes the first module to be transcribed during reverse transcription. The PBS which binds the $tRNA_{Lys3}$ to initiate the reverse transcription, the Psi which regulates that packaging of the viral genome in the capsid and DIS with also roles in the packaging of the viral genome (Wu 2004; Johnson et al., 2010; Heng et al., 2012). Other non-coding regions have a more internal location; the RRE is located in the *rev* gene, and two sequences rich in purines, the Poly Purine Tract 3' (close to the 3' LTR) and the PolyPurine Tract central (close to the *pol* gene) have a role in the reverse transcription step, conferring a resistance to the activity of the ribonuclease H (Charneau P., 1992); the Central Termination Signal, which has the name implies has a role in interrupting the reverse transcription, is a AT rich region at the centre of the genome (Charneau P., 1994).

The UTR 3 'sequence includes the U3 region (Unique in 3') and the R sequence, identical to those found in 5'.

1.2.2.3. Coding Regions

The HIV-1 genome consists of nine genes *gag*, *pol*, and *env*, *tat*, *rev*, *nef*, *vif*, *vpr*, *vpu*, coding for 19 proteins (Figure 4). Three of these genes, *gag* (Group-specigic antigen), *pol* (polymerase) and *env* (envelope) are shared among all retroviruses.

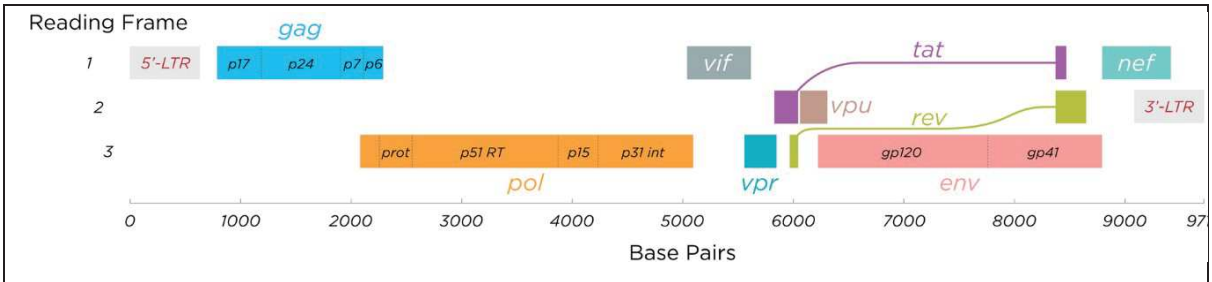


Figure 4: HIV-1 proviral genome organization showing the tat and rev exons in purple and yellow respectively and the different ORF.

The *gag* gene code for a polyprotein which is cleaved by the viral protease to create the matrix protein (MA p17), the capsid protein (CA, p24), the nucleocapsid protein (NC, p7), and the p6 which share its function by binding the Vpr (Solbak et al., 2013)

The *pol* gene code as well for a polypeptide which through a process autocatalysis produce the protease (PR, p12) which cleaves and release the reverse transcriptase (RT, p51) and the integrase (IN, p31) (Debouck C., 1987).

The *env* gene code for another polypeptide which is cleaved thanks to the support of a protease of the host cell which then produce the 2 sole exterior viral proteins in the virion, the transmembrane glycoprotein (TM, gp41) and the surface glycoprotein (SU, gp120).

The *tat* gene code for a 101 amino acids protein with essential functions for viral gene expression. Tat binds the TAR element and recruits an active transcription elongation complex to the HIV-1 promoter. This protein is coded by two exons where the first part is composed by 72 aa and the remaining amino acids in the latter. While being rare, a stop codon in the second exon can give the arise of an 86 amino acids Tat form (Jones and Peterlin, 1994; Li er al., 2012).

The *rev* (regulator of virion expression) code for the protein Rev. This protein, by interacting with the RRE motif, is capable of bypassing the natural nuclear retention of unspliced RNAs allowing the export of the single spliced and unspliced viral RNA in the cytoplasm. (Hadzopoulou-Cladaras et al., 1989)

The *nef* gene (Negative Regulatory Factor) code for the Nef protein which role is to establish a persistent state of infection within the host cell. Nef does so by reducing the threshold of activation of the infected cells by downregulation the surface receptor CD4 and Lck. As a result the cells start producing Interleukin 2 which stimulates growth and proliferation in CD4 T cells creating a new population that the virus can infect (Geyer et al., 2001) and has also being linked to be able to stop apoptotic signals within the cell by inactivating Bad (Wolf et al., 2011).

The *vif* gene (viral infectivity factor) code for the protein Vif. Vif has a specific role in protecting the HIV-1 reverse transcription process by targeting the human enzyme APOBEC for ubiquitination and follow up degradation by the proteasome. APOBEC is a cytidine deaminase enzyme that bind and causes hypermutation on the negative strand of the newly retrotranscribed viral DNA. These modifications allow the cell to recognize the aberrant DNA and target it for degradation (Conticello et al., 2003; Lecossier et al., 2003; Mangeat et al., 2003).

The gene *vpr* (viral protein r) code for the protein vpr. Vpr is required for viral replication in non-proliferating cells such as macrophages and is also responsible for the import of the pre-integration complex (Popov et al., 1998), the arrest of the cellular cycle in the G2 phase (Muthumani et al., 2006), the reduction of the random mutations during the retro-transcription (Chen et al., 2004) and the elimination of CTIP2 to prevent gene silencing (Forouzanfar et al., 2019).

The gene *vpu* (viral protein u) code for the protein Vpu which in combination with Nef causes the internalization and degradation of the CD4 receptor enhancing HIV-1 infectivity (Bour et al., 1995).

1.2.3. HIV-1 replication cycle

The HIV-1 life cycle is divided in two phases: the pre-integration phase or early phase involves the virion infecting a new cells and comprise all the steps from binding to the target cell, entry, reverse transcription of the viral RNA and integration in the host genome; the post-integration phase or late phase cover the viral gene expression after a latency period, replication, virion assembly, budding and maturation (Figure 5).

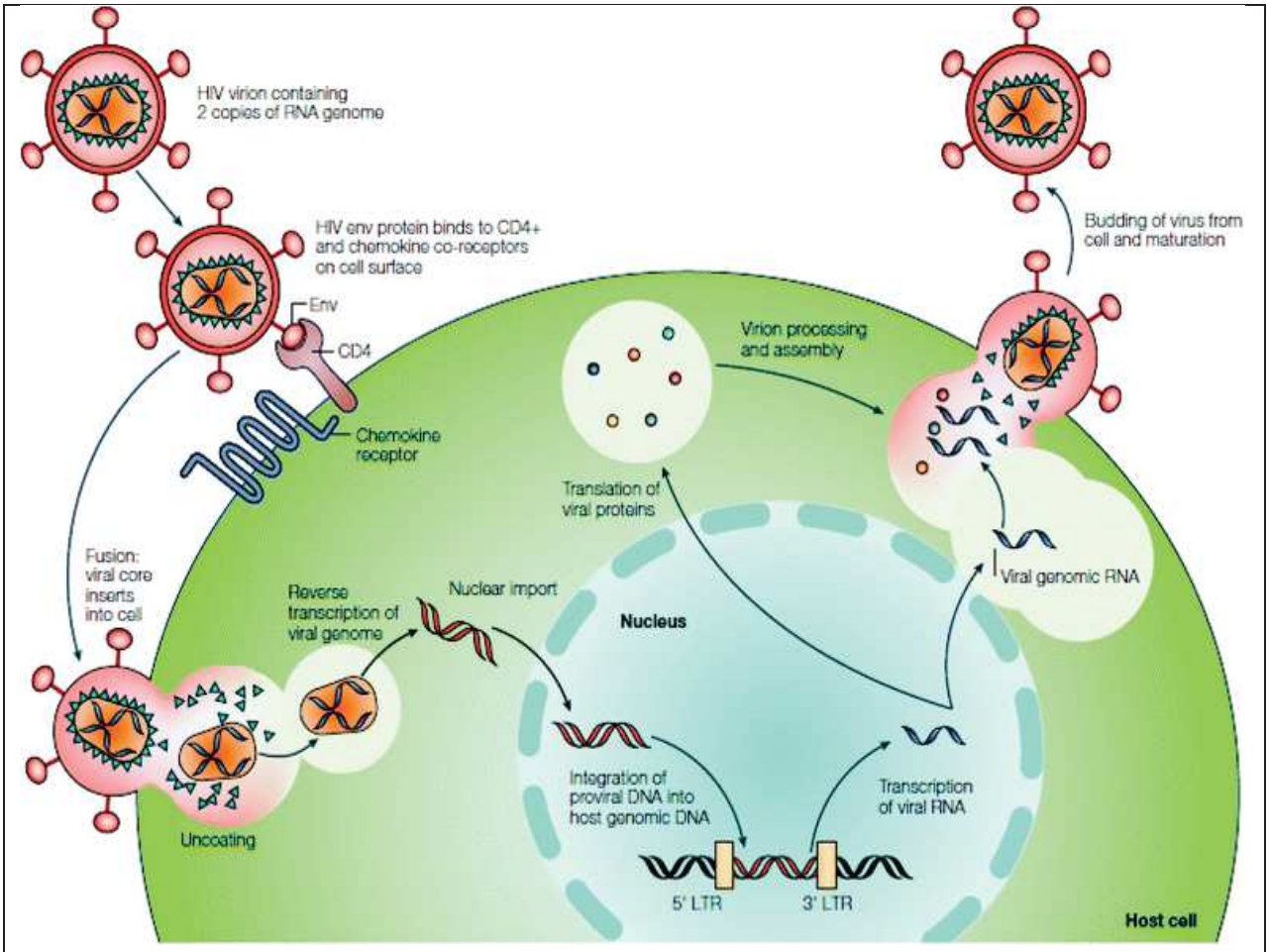


Figure 5: HIV-1 replication cycle (Rambaut et al., 2004)

1.2.3.1. HIV-1 targets of infection and entry strategies

The main targets of HIV-1 are cells expressing the CD4 receptor and associated chemokine co-receptor and include CD4+T cells, Dendritic cells, Microglial and the Monocyte/Macrophages lineage. The infection starts with the interaction of the viral surface glycoprotein gp120 which binds the CD4 with high affinity; this interaction is aided by the chemokine receptor CCR5 and CXCR4.

The main route for the virus to infect the target cell is by fusion. When the gp120 protein binds the receptor CD4, gp120 undergoes a conformational change exposing a hydrophobic peptide which penetrates the cell membrane and deliver the capsid core, with all associated proteins, in the cell cytoplasm (Weiss, 2002). Different viral strains exhibit a different cellular tropism which is found by looking at which co-receptor they prefer; viral strains which binds to the CCR5 are labelled R5 and prefer T lymphocytes, macrophages and dendritic cells, the strain

labelled X4 binds the CXCR4 on T lymphocytes and the strain X4R5 bind both of them with no particularly preference (Doms and Moore, 2000; Lama and Planelles, 2007).

Another entry strategy is the one through the endocytosis pathway; recent studies have discovered how the first cells to entry in contact with the virus upon entry in the organism may be the dendritic cells (DC) (Padilla-Parra and Dustin, 2016). The interaction of these cell with the virus uses the DC-SIGN lectin which trigger the endocytosis response; here the virus escapes the degradation and uses the dendritic cells as a Trojan horse. These cells, especially when come in contact with pathogens, migrate to the lymph nodes where the virus can have easily access to lymphocytes and macrophages to spread the infection. In the lymph nodes it is believed the virus adapt a trans-infection mechanism which uses synapses and conjunction point between dendritic cell and other leukocytes to pass through and infect them. This mechanism is now believed to be one of the most important way of virus spreading (Sewald et al., 2015; Sattentau, and Stevenson 2016).

1.2.3.2. Uncoating and reverse transcription

The viral capsid is internalized in the cell cytoplasm and rapidly destabilized to release the gRNA and the viral proteins. Cyclophiline A, a human protein which is sequestered in the newly formed virion by Gag, facilitate the creation of the retro transcription complex that is formed by the viral RNA, tRNALys3, the reverse transcriptase and different other proteins of both viral and cellular origin (Braaten D., 1996), such as p17, p24 and p7 (all coded in the *gag* gene), the Integrase and Vpr.

The reverse transcription allows the synthesis of a double strand proviral DNA from the source of a strand of viral RNA. Retroviral Retro Transcriptase has three sequential biochemical activities: RNA-dependent DNA polymerase activity, ribonuclease H (RNAse H), and DNA-dependent DNA polymerase activity (Hu and Hughes, 2012). In Figure 6 the mains steps of the retro transcription process are shown.

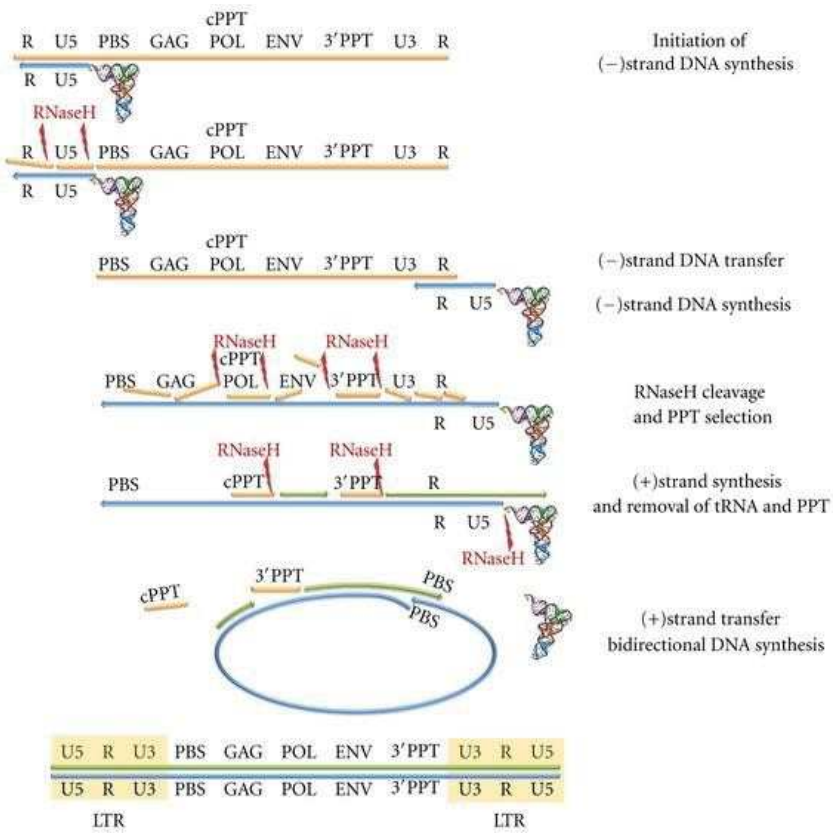


Figure 6: The steps of the retro transcription in HIV-1 (Espostio et al., 2012)

The first step is the synthesis of a strand of DNA with negative polarity which is started by the binding of the 18 nucleotides at the 3' terminal of the tRNALys3 to the PBS steam loop at the 5' on the viral genome (Hu and Hugues, 2012). RT synthesize then the region R to U5 and the original viral RNA is hydrolyzed by the RNAse H activity of the RT. The DNA strand just formed allow the synthesis of a full DNA strand till the PBS sequence and the RNA is hydrolyzed apart from two regions rich in purines called cPPT and 3'PPT. This two oligos will act as primers for the initiation of the synthesis of the (+) sense DNA and will be later degraded. Last, the strand switch of the (+) PBS sequence to the complementary PBS site on the negative DNA strand form a bidirectional, circular DNA be the base for the synthesis of the complete linearized gDNA with the 5' and 3' flanking LTR sequences (Espostio et al., 2012).

1.2.3.3. Viral gDNA Import in the nucleus

Once the viral RNA has been retrotranscribed in DNA the Retro Transcription Complex (RTC) evolve in the Pre-Integration Complex (PIC) formed by the viral DNA newly synthetized, the RT, the Integrase (IN), Vpr, matrix proteins (MA) and numerous cell proteins (Farnet C.M. at

al., 1997; Miller M.D. et al., 1997). At this point the viral genetic material must cross the nuclear pore complex (NPC), a series of channels that allow the transport of molecules between the nucleus and the cytoplasm. The problem here is that the size limit to allow passive diffusion is of only 9nm while the PIC is roughly 55nm and so, to succeed in this task, the CA viral protein binds components of the NPC and involve nuclear pore proteins and karypherins in this effort (Gallay et al., 1997; Bukrinsky and Adzhubei, 1999; Suzuki and Craigie, 2007; Endsley et al., 2014).

1.2.3.4. Integration

Once within the nucleus the integrase catalyses the insertion of the viral DNA in the genome of the host, this is a crucial step because from now on the cell is irreversibly infected (Van Maele and Debyser, 2005). The Integrase cleaves a dinucleotide from the 3' LTR and the cellular DNA as well, then through a process of trans-esterification the 3' free hydroxy group is attached to a 5' free phosphate of the cellular DNA. The non-complementary regions are recognized and fixed by the cellular DNA repairing machinery (Van Maele et al., 2006). This step, like many others before, is aided by cellular cofactors like the protein HMG-1 and BAF (Chen and Engelman, 1998; Hindmarsh et al., 1999). The site of integration has been subject of many studies and it has been proven to not be random, nor sequence specific but to be determined of the interaction of the IN with the Lens Epithelium-Derived Growth Factor (LEDGF/p75) which target for integration actively transcribed genes away from the start of transcription of cellular genes (Maertens et al., 2003, Craigie and Bushman, 2014).

1.2.4. Post-Integration phase and viral gene expression

Once integrated within the genome of the host cells the virus take the name of provirus and its genes expression is determined by the host cells state of activation and proliferation as well as the site in which the integration happened (Jordan et al., 2001). The provirus at this step can start to express its genome for the production of new viral particle or can enter a phase of post-integration latency in which the provirus is considered silent. This two different destiny of the provirus is connected to the location of its integration; while it prefer (and statistically with higher change) to integrate in the actively transcribed regions of the host cell genome (euchromatin), a smaller portion of the virus will indeed integrate in regions which will shift to an heterochromatin state, making them silenced; these cells, infected, but not actively producing

new viral particles, take the names of viral reservoir and are refractory to the anti-retroviral therapy making them the main reason why still at this day we manage to control but not eradicate an HIV infection form a patient (Schröder et al., 2002). In the next chapters we will discuss more in deep the formation of the reservoir.

After integration a nucleosomal structure is formed (Figure 7), where the nucleosome 0 and 1 play the most important roles since are positioned at the proviral promoter (-145 to -255) and the site where the transcription will start (-3 to +141). The intermediate region is called DHS (DNase 1 Hypersensitive Sites), this region, free of nucleosomes have the majority of the sequences able to bind transcription factors (Verdin et al., 1993; Ne et al., 2018). The LTR U3 region can be divided in three different parts (Figure 7) called modulatory, enhancer and core regions; the modulatory (-454 to -104) has a Negative Regulatory Element (NRE) to which follows the enhancer region (-105 to -79). The core promoter (-78 to -1) is composed by three SP1 sites, a TATA box and an initiator sequence (van Opijnen et al., 2004; Karn, 2011).

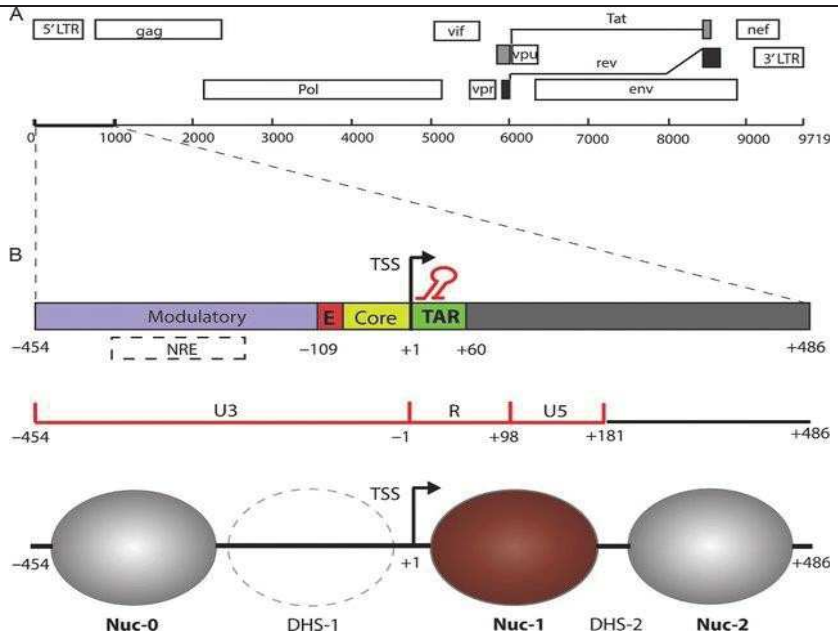


Figure 7: Representation of the HIV-1 5' LTR and its position among the nucleosomes. In vivo observed chromatin landscape of the transcriptionally inactive HIV-1 5'LTR is defined by the presence of two strictly positioned nucleosomes Nuc-0 and Nuc-1 as indicated connected by a region of accessible DNA, hypersensitive to nucleases, DHS-1 which contains the majority of cellular transcription factor consensus binding sites and which has high propensity for nucleosome formation (Ne et al., 2018)

Once the transcription as started, three different types of viral transcripts are generated, differentiated among them by the splicing process they undergo. Multi-spliced transcripts will lead to the production of Tat, Rev and Nef, the mono-spliced code for Vif, Vpr, Vpu and Env while the non-spliced mRNAs are translated in the structural proteins of the virions and the viral enzymes or incorporated into the newly formed virions. Viral transcription can be divided in two phases, an early phase which is Tat independent and a later one, more productive, which is Tat-dependent.

1.2.4.1. Tat independent phase

As normal cellular gene would do, the provirus need the transcription machinery of the host cell as well, but this process is aided, mediated and enhanced by both viral and host proteins. The gene expression starts on the 5' LTR promoter but, in the absence of the viral protein Tat, the transcription activity is low. The LTR functions like an eukaryotic promoter. It makes use of the RNApol II (DNA dependent RNA polymerase II) which is recruited on the promoter as soon as the NF-κB binds together with the Specificity Protein 1 (SP1) at the enhancer and core regions respectively, of the U3 promoter (Perkins et al., 1993)(Figure 7). NF-κB and SP1 recruit Histone Acetyl-Transferase enzymes (HATs) which remodel the epigenetic modification on the nucleosomes, remodelling the chromatin and remove the steric block of transcription from the promoter. At this point the Transcription Factor IIH (TFIIH) forms the pre-initiation complex and phosphorylates the C-terminal domain of the RNApolII to allow the transcription of the TAR RNA (Parada and Roeder, 1996; Tyagi et al., 2001). During this phase, the transcription is often paused by the presence on the RNApolII of the negative elongation factors (NELF) and DSIF (DRB sensitivity inducing factor) which results in the production of mostly short and abortive transcripts (Peterlin and Price, 2001). Also the production of non-spliced and single-spliced transcripts is inhibited by the presence of the CRS sequences (Cis-acting Repression Sequence) in the *gag*, *pol*, and *env* genes that acts as nuclear retention signals (Maldarelli et al., 1991) and only multi-spliced mRNA (which code for proteins as Tat, Rev and Nef) manage to reach the cytoplasm and be translated; the protein Tat and Rev then return to the nucleus and start the Tat dependent phase.

1.2.4.2. Tat-dependent phase

Just few copies of Tat are necessary to start a positive feedback loop which will boost viral gene transcription and enhances its own expression (Stauber and Pavlakis 1998). Tat binds on the TAR RNA and recruits a heterodimer composed by the Cyclin T1 and CDK9 (Cyclin-Dependent Kinase 9) which form the positive transcription factor P-TEFb. Originally it is found in an inactive form as a Ribo-NucleoProtein Complex (RNP) with the small nucleolar RNA 7SK acting as scaffold, Bcl11b/CTIP2 (B-cell lymphoma/leukemia 11B), LaRP7 (La-related protein 7), MePCE (7sk specific methyl phosphate-capping enzyme) and HEXIM1 (Hexamethylene Bis-Acetamide Inducible 1) (Schonichen et al., 2010; Cherrier et al., 2013)(Figure 8).

Tat is able to disrupt this complex in two ways: it shifts the binding of the pTEFb by competitive pushing the binding of pTEFb with the TAR (D'Orso et al., 2012) and by interacting with pTEFb, via its C-rich domain releases HEXIM1 from the complex (Contreras et al., 2007).

Once the Tat-pTEFb complex is formed on the TAR, P-TEFb induces phosphorylation of Serine 2 within heptapeptide repeats (YSPTSPS) of the C-terminal domain (CTD) of RNA Pol II (Zhou et al., 2012). In addition to the RNA Pol II CTD, DSIF and NELF are also phosphorylated by P-TEFb. Phosphorylation antagonizes their inhibitory actions, leading to the release of RNA Pol II paused from the proximal promoter, and the transition to productive elongation (Fujinaga et al., 2004; Ivanov et al., 2000). These events allow the production of full-length viral transcripts. At the same time, in the nucleus, Rev form a multimer on the RRE sequence of the non-spliced and single spliced mRNA. Rev Nuclear Export Signal (NES), with MATR3 as cofactor, recruits the nuclear export cellular machinery to shuttle the viral mRNA to the cytoplasm (Neville et al., 1997; Kula et al., 2011).

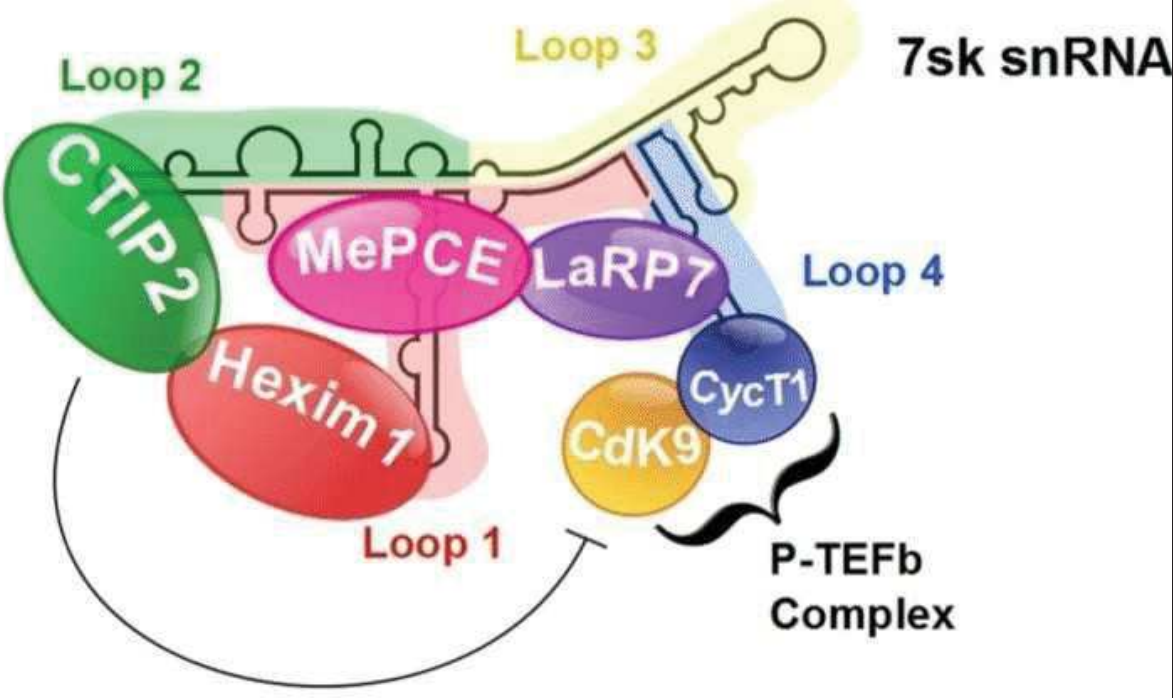


Figure 8: The inactive form of P-TEFb in the Ribonucleo Protein Complex (Adapted from Cherrier et al., 2013)

1.2.5. Production of new virions

In the cytoplasm the viral mRNA transcripts undergo the process of being translated in proteins by the host cell ribosomal machinery; while the multi-spliced RNA are translated without issues, the full-length mRNA translation has to overcome some difficulties. These mRNA are polyadenylated and capped and, on the 5' is present an untranslated region (UTR), and it is necessary, for the ribosomal complex to bind, that the secondary RNA structure is remodelled. To do so, an RNA helicase called DDX3 is recruited which then allow the elongation initiation factor 4F (elF4F) and the 43S pre-initiation complex to bind the cap (Rojas-Araya et al., 2015; Jan et al., 2016). Once all the viral proteins are able to be translated in the cytoplasm, the assembly and production of the new virions can begin.

1.2.5.1. Assembly

The assembly of the different viral proteins is located at the plasmatic membrane of the cell, this phenomenon happens after the gag precursors Pr55 and Pr160 are modified through a process of Myristoylation, allowing them to anchor themselves the lipidic bilayer of the cellular membrane (Ono, 2010). Another important detail to keep in mind is that the full length, unspliced viral mRNA can have two mutually exclusive fates: To be translated as we have already stated and described or to be packed within the newly formed virions. It is necessary then to have a selecting mechanism to decide when and which transcripts will be led to one or the other destiny. A higher level of gag protein within the cytoplasm switch the translation to packaging by initiating a conformational change in the secondary structure of the gRNA exposing two specific structure of the 5'UTR, SL3 and DIS which trigger the dimerization and its association with the gag protein (Paillart et al., 2004).

The envelope precursor protein pr160 is cleaved by a host cell's protease in proximity of the Golgi to obtain the gp120 and gp41 (Cosson, 1996); the precursors pr160 and pr55 allow as well the recruitment of molecules necessary for the virion future replication cycle like the tRNALys3 through their RT domain (Mak et al., 1997), cyclophilin A through the CA domain (Luban et al., 1993) and Vpr thanks to the domain p6 (Kondo et al., 1995).

1.2.5.2. Budding

At this step HIV-1 is able to hijack the cellular exocytosis machinery to start the process of releasing the immature virions outside the cell. The p6 domain of the precursors viral protein of gag, pr55 and pr160, recruits the cellular protein TSG101 and ALIX which are components of the Endosomal Sorting Complexes Required for Transport (ESCRT) (Garrus et al., 2001; Martin Serrano et al., 2003) and will close the necks of the budding vesicles allowing the exocytosis of the immature particles.

1.2.5.3. Maturation

Maturation and budding happen almost at the same time; the maturation process is visible at the electron microscopy due to morphological changes in the virion where there is a concentration of the viral genome in the characteristic cone shape of the core. The maturation process start when the protease (PR) is released by an autocatalytic mechanism (Debouck et al.,

1987) and cleaves the gag and pol precursors pr160 and pr55 to release both the viral enzymes (RT and IN) and the structural proteins (MA,CA, p2, NC, p1, p6). At this step, the newly formed virion particles are considered almost mature and infectious (Tang et al., 2008)

1.3. HIV-1 infection

HIV-1 infection follows a three phases pattern (Figure 9):

- **The primary infection phase** main event is an exponential increase of the viremia or the copies of viral RNA detected in the plasma of the infected individual. This phase is also named acute phase and due to the intense viral replication, allows HIV-1 to disseminate in the whole organism and seeding the lymphoid organs. This phase shows a drop as well in the population of the lymphocytes T CD4+ followed up by a recovery when the organism starts to develop an immune response against HIV-1. At this point a drop in the plasma viremia is detected and the organism enter in the asymptomatic phase. During the acute HIV syndrome, which last for few weeks, often are noted some clinical symptoms as fever and recurring headache due to a cytokine burst and the infected individual is highly contagious.
- **Clinical latency** is the longest phase and can last for years after the first infection event. A constant fight between HIV-1 and the immune system keeps the patient mostly asymptomatic and the T cells destroyed are compensated by newly formed ones. This phase can last up to a decade and in case of patient aware of their condition and actively under anti-retroviral therapy, this phase can be pushed to several decades. The end of this phase is centred on the steady drop in the CD4+ T Lymphocytes till a point in which the patient starts to develop clinical signs of opportunistic infections (Ho et al., 1995).
- **Symptomatic phase and Acquired immunodeficiency syndrome (AIDS).** At this stage, the immune system is extremely damaged and a positive diagnosis for AIDS is given when CD4 T-cell count drops below 200 cells/μL.

  At the same time the patient often develops opportunistic infections such as *Pneumocystis pneumonia*, tuberculosis, toxoplasmosis and cryptosporidiosis, rare form of cancers such as the Kaposi's sarcoma plus a variety of symptoms such as swollen lymph glands, chills, fever, sweats, weakness, and weight loss. The patients are highly infectious and as the AIDS progress dementia and wasting syndrome lead an untreated patient to death within 3 years form the beginning of this phase.

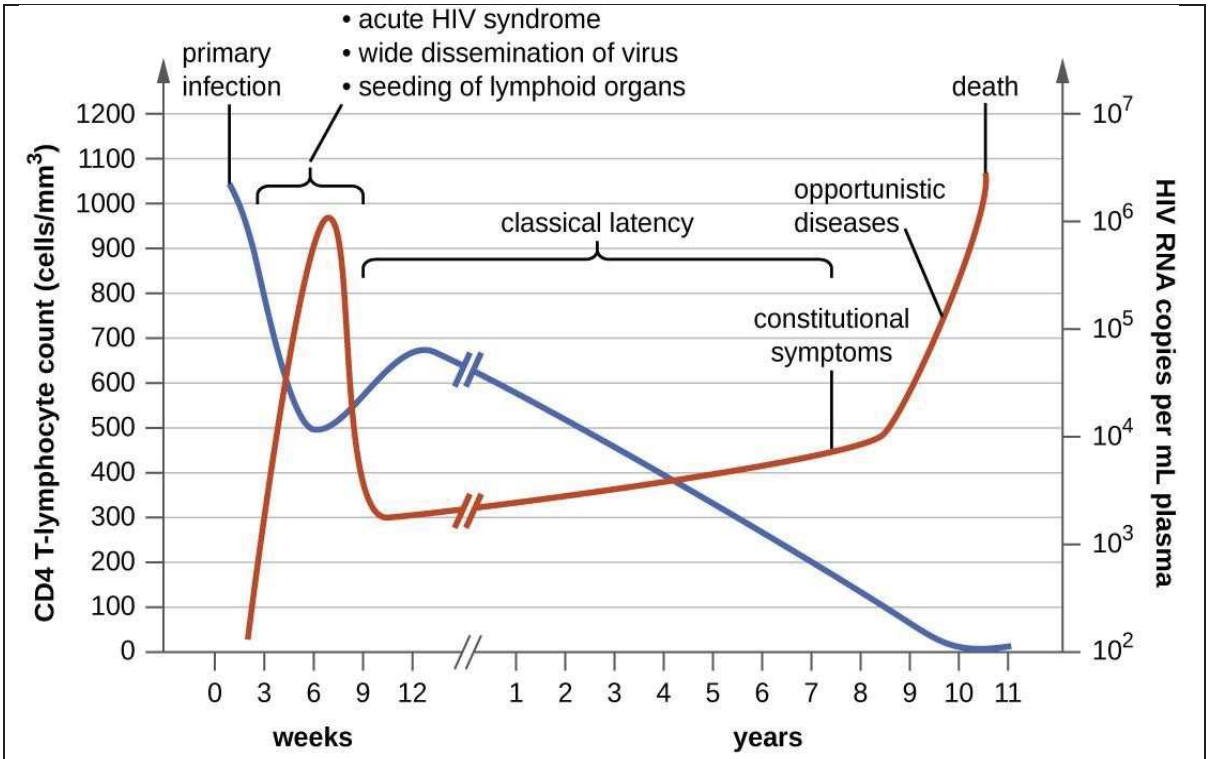


Figure 9: HIV-1 infection phases in rapport to the circulating CD4 T cells (blue line), and viral RNA (red line). Adapted from Kogan, M., & Rappaport, 2011

1.3.1. HIV-1 tropism

While we discussed in the previous chapter the count of the lymphocytes CD4+, which remains the main target of HIV-1, the virus can infect and replicate within many other cells of the immune systems. During the infection, an essay is performed on patients called the Trofile Assay, which assess which main strain is currently the predominant in the organism (R5 or X4). Usually the first viruses are mainly CCR5 after the virus then derives to CXCR4 and double tropism. (Whitcomb et al., 2007).

1.3.1.1. Infection of the CD4+ Lymphocytes T

Lymphocytes CD4+, also known as T helper cells, play a critical role in the adaptive immune response by exerting an activation role upon presented with foreign antigens by Antigen Presenting Cells (APCs). Once activated CD4+ activate other cell of the immune systems both through cell-to-cell contact and through chemokine release (Figure 10)

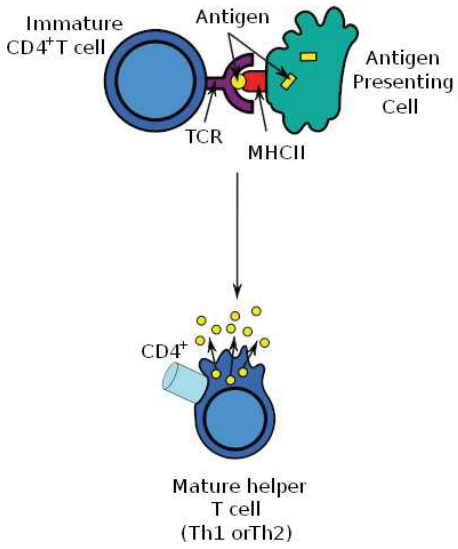


Figure 10: Activation of CD4+ T Helper lymphocytes activation upon contact with foreign antigen presented by an APC (Adapted from Wikipedia)

Once infected, the activation state of these cell plays a crucial role in the expression of the provirus. A non-activated CD4+ cell will carry a latent HIV-1 provirus (Schnittman et al., 1989) usually in these cells is present a so called pre-integration latency, where the virus replication cycle is halted at the retro transcription and integration step (Zack et al., 1990). Once activated these cells become the source of an abundant and rapid viral replication which, due to the virus cytotoxic effect, will lead rapidly to the cell death.

1.3.1.2. Infection of the dendritic cells

Dendritic cells (DC) main localization is at the peripheral tissues and are usually the first cells of the immune system to encounter invading pathogens. These cells take also the more general role (even if other cell can cover this function) of Antigen Presenting Cells (APCs) and due to this they play a pivotal role in the immune response. Dendritic cells internalize pathogens via the endocytic pathway an after a process of intracellular lysis they deal with the antigenic peptides presenting them on their surface to activate T lymphocytes (Steinman, 1991). HIV-1 has evolved a way to hijack these cells to further propagate in the organism and reach their main target of infection, the CD4+ T lymphocytes. In the early 1990 it was noticed how DC infected with HIV-1 were responsible of an incredibly potent cytopathic infection of CD4+ lymphocytes (Cameron et al., 1992). From these discoveries it was proposed the theory that HIV-1 not only could survive the endocytosis and lysis when internalized by the DC, but would use them as "Trojan Horses" to be brought in other lymphoid tissues to then have easy

access to other cells to infect. Interesting also that DCs do not need to be actively infected to spread the infection (like CD4+ and macrophages need to be) but they are able to propagate the virus via a process called trans-infection (Geijtenbeek et al., 2000; Figdor et al., 2002; van Kooyk and Geijtenbeek, 2003) well represented in Figure 11. Different types of dendritic cells shows different levels of infection rates and recently active replication of HIV-1 was shown also within these cells.

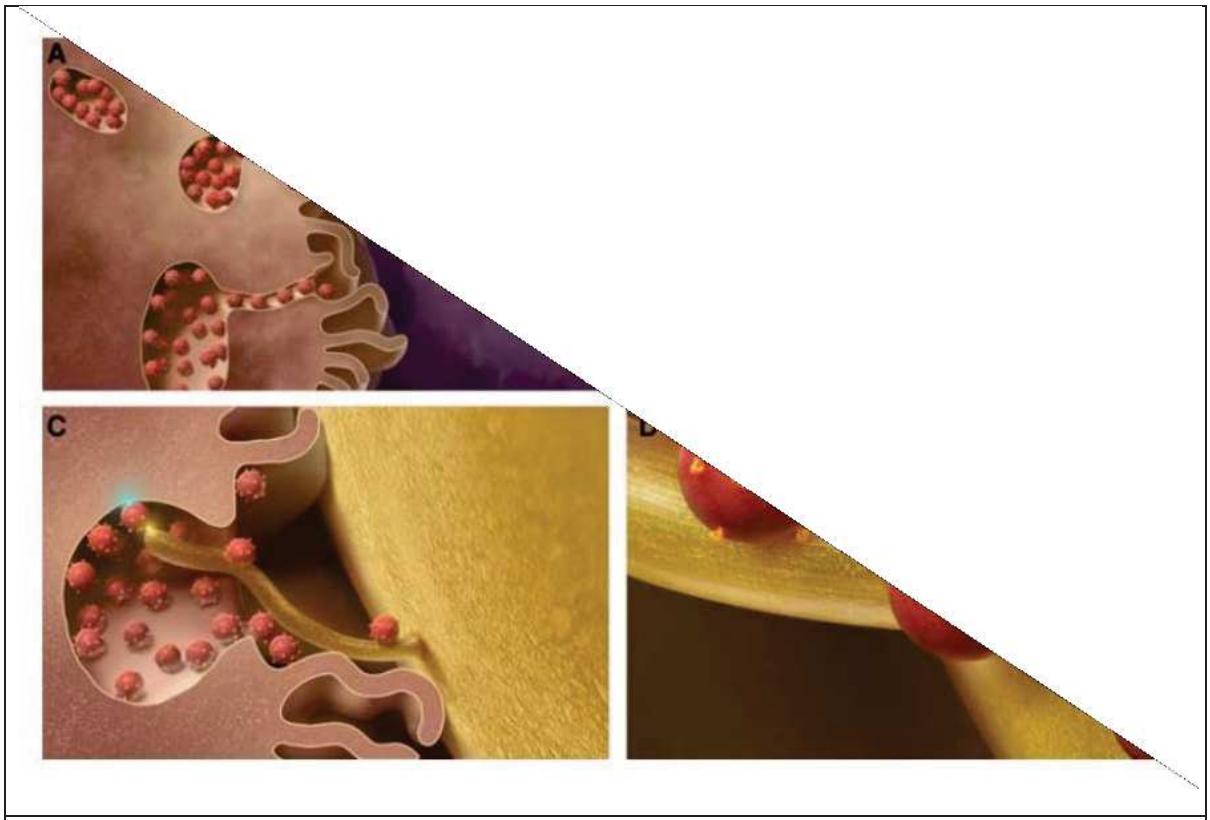


Figure 11: Antigen presentation and trans-infection of a CD4+ T cell mediated by a DC. (A) DCs capture and concentrate virions. (B) CD4 + T lymphocytes, containing membrane protrusions, bind DCs. (C) The protrusions of CD4 + T cells invade the compartments containing the virus and bind effectively HIV. (D) The virus then migrates to the cell body to trigger the infection. (Wilen et al., 2012)

1.3.1.3. Infection of monocytes and macrophages

Macrophages, and their precursor, the monocytes, are cell of the immune system which contrary to the lymphocytes, are originated form the myeloid line. These cells take the role of phagocytes with the active role of capturing, internalizing and destroying microbes, pathogens, cancer cells, debris and anything else lacking the proteins identifying as part of the healthy

organism. Macrophages are located all around the body and specific macrophages found in specific organs or tissues gain specific names (as an example, the Microglia in the brain, Kuppfer cells in the liver etc.). An important key characteristic of the mature macrophages is their inability to replicate, these cells always originate from a monocytes precursor; due to this characteristic the HIV-1 infection of these cells has adapted and especially the assembly of new virions differ from the one in the other cells. Within the macrophages the assembly of the immature virions is located in proximity of the cytoplasmic vesicles which are enriched with the major histocompatibility complex class II (MHCII). Within these vesicles the virions can meet three different fates: degradation, stabilization within the cytoplasm or release after the fusion of the cytoplasmic vesicles with the cellular membrane (Raposo et al., 2002). This process of exiting the cells is more delicate then the normal budding happening in the CD4+ cells, and causes a lower cytopathic effect to the cells; for this reason macrophages infected with HIV-1 have an incredibly longer life span compared to infected CD4+ cells and act as one of the major viral reservoirs. Also a study performed on humanized irradiated mice generated by transplanting human $CD34^{+}$ hematopoietic stem cells into NOD/SCID mice, which are unable to support human T cell development, showed how even in the absence of the main target of the virus, HIV-1 was still able to persist within the myeloid lineage (Schultz et al., 2007;

1.3.1.4. Infection of the microglial cell

Microglial cells are the main immune defence of the Central Nervous System (CNS) consisting primarily of the brain and the spinal cord. Microglial count to 10 to 15% of all the cells in the brain and play the role of keepers of the brain maintenance by removing plaques, damaged and unnecessary neurons, and synapses and of course, pathogens (Lawson et al., 1992).

While these cells express many different co-receptors, HIV-1 especially uses the CCR5 to invade and infect them (Bajetto et al., 2002) making these cells the primary source of infection within the CNS. The infection of the microglial is extremely important in the study of HIV-1 infection for other two reasons, the first one is connected to one of the main clinical symptoms of an advanced HIV-1 infection which is the HIV Associated Dementia (HAD); in the late phase of the infection, especially in untreated patients ($\cong 20\%$), patients infected with HIV-1 start to develop motor and cognitive neurological symptoms which nowadays are all grouped under the HAD classification (Navia et al., 1986). These symptoms arise as the results of the activation of the microglial upon infection which start the produce and secrete different

molecules toxic for neuron and astrocytes such as Tumor Necrosis Factor α (TNF α), nitrogen monoxide, quinolinic and arachidonic acids, Platelet Activating Factor (PAF) and viral proteins which cytotoxic effect such as Tat and Vpr (González-Scarano and Martín-García, 2005). The effect on the overall CNS will be discussed further in another chapter. Another important role of the infected microglial within the CNS is the fact that these cells will form a reservoir (like the macrophages normally do) but in a anatomically sanctuary protected by the blood brain barrier, making them extremely refractory to drug access and effect. Moreover, microglial have an extremely long life, and they harbours different block mechanisms that favours latency, notably CTIP2, and once infected became long lived reservoirs.

1.3.1.5. The infection of the Central Nervous System

Due to its anatomical importance the CNS is separated and almost isolated by the rest of the body by a structure called the Blood Brain Barrier (BBB). This anatomical barrier is formed by a layer of endothelial cells tightly connected to block and select the passage of substances through it and maintaining the brain homeostasis and filtering the exchange between blood and the cerebrospinal fluid. Nevertheless HIV-1 can penetrate this barrier in three different ways (Albright et al; 2003) (Figure 12):

- A “Trojan Horse” mechanism in which an infected monocyte or lymphocytes (which are able to pass through the BBB through the process of diapedesis) “squeeze” themselves between the endothelial cells granting access to the virus to this anatomical sanctuary.
- By infecting the endothelial cells and then budding from the other side of the BBB.
- The passage of free circulation virion across the BBB by transcytosis.

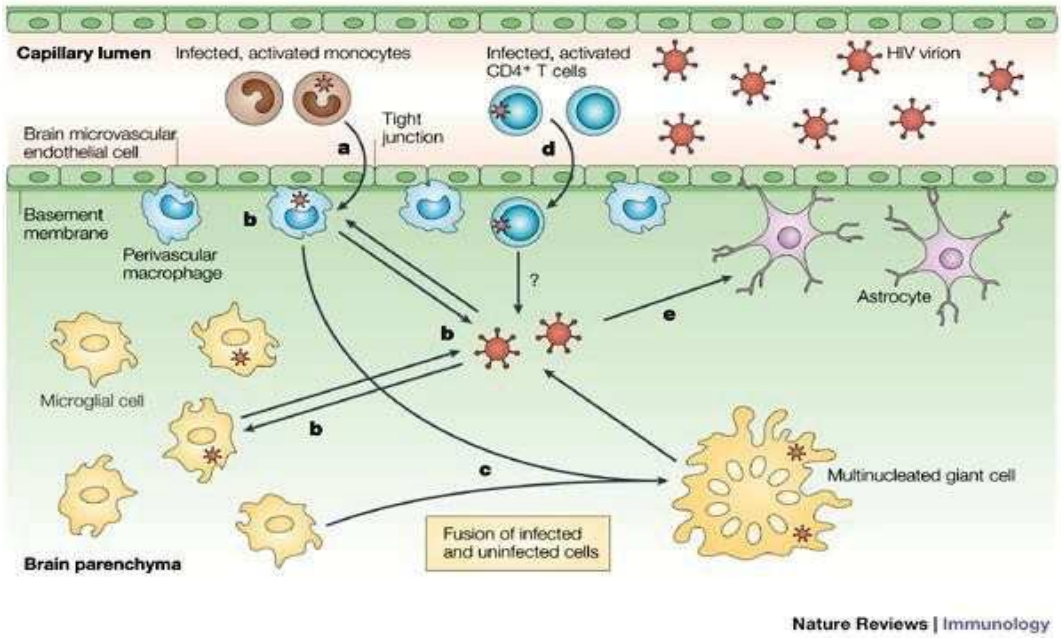


Figure 12: HIV-1 invasion of the Central Nervous System (González-Scarano and Martín-García, 2005)

As mentioned before, 1/3rd of the patients in the late phase of the AIDS, especially if they have not been treated with an anti-retroviral therapy, will develop HAD. Among the symptoms described in this phase are short-term memory loss and lack of concentration, apathy, mood swing and change of personality. These symptoms are correlated with a disrupt of the homeostasis and the health of the CNS. As described above once the virus is able to cross the BBB it starts its infection of the resident macrophages, the microglial, which once activated will produce toxic cytokines and molecules which harm the neurons. Another effect of the microglial activation is the followed activation of the astrocytes. Astrocytes are another type of cell of the CNS with a role in maintaining the homeostasis. HIV-1 can infect these cells as well even if the infection is extremely weak and almost non-productive. Anyhow the activation of these cells will cause a signalling cascade which cause the release of massive quantity of glutamate in the extracellular fluid causing a destabilization of the BBB increasing its permeability and allowing an easier access to this sanctuary for infected cells and circulating virions (Kaul et al., 2001).

The activation of the microglial and astrocytes upon infection, with the correlated production of cytokines, have of course, a deleterious effect on the health of the neuron cells. The glutamate activates the p38 MAPK pathway, for example, which leads to the apoptosis of the cell (González-Scarano and Martín-García, 2005). Last but not least, the astrocytes, responsible on maintaining the myelin coating around the axons, suffer greatly by this environment, and

with damage on the myelin and a lower production of it, the nervous signal are lost or becomes fragmented.

1.4. Combination Anti-Retroviral Therapy

1.4.1. Success of cART

Almost 40 years have passed since HIV discovery, and the search for a cure or a treatment for HIV-1 positive patients has made great strides. In the 80' and 90', being HIV-1 positive was equal to a death sentence, nowadays with the current developed antiretroviral drugs, an infected patient can look forward to a life expectancy similar to a healthy person. Also, an HIV-1 positive patient, under a regimen of Anti-retroviral therapy, has is viral tier reduced to a point in which is hardly considered infectious and its viremia can only be detected with the most accurate and sensible tools available; nevertheless the treatment has to be considered lifelong since a cure has not yet being found (Cihlar and Fordyce, 2016).

As the name suggest, this therapy does not avail of only one class of drugs but a combination of them. Currently five classes of drugs are used, each of them targeting a specific and crucial step of the HIV-1 replication cycle:

- Fusion inhibitors (or entry inhibitors) prevent the virus to fuse with the cell membrane and release its core in the cytoplasm. Maraviroc and enfuvirtide are currently the two drugs used where the first one target the CCR5 co-receptor the second block the entry by binding the gp41 viral protein and forming an in-active hetero six-helix bundle (Bai et al., 2016).
- Nucleoside/nucleotide reverse-transcriptase inhibitors (NRTIs) such as zidovudine, abacavir, lamivudine, emtricitabine, and tenofovir are chain terminators. Once incorporated in the nascent newly DNA strand, they prevent other nucleosides from also being incorporated into the DNA chain because of the absence of a 3' OH group (Das & Arnold, 2013)
- Non-nucleoside reverse-transcriptase inhibitors (NNRTIs) such as nevirapine, efavirenz, etravirine and rilpivirine affect the handling of substrate (nucleotides) by reverse transcriptase by binding to an allosteric site of the enzyme near the active site (Das & Arnold, 2013)

- Integrase inhibitors such as raltegravir, elvitegravir and dolutegravir inhibits the viral protein integrase.
- Protease inhibitors such as lopinavir, indinavir, nelfinavir, amprenavir, ritonavir, darunavir and atazanavir prevent the formation of mature virions by blocking the action of the viral proteasome. They especially block the cleavage of the gag and gag/pol precursors proteins (Wensing et al., 2010).

Currently a cART regimen consists of 3 drugs; a "back bone" of two NRTIs plus a "base" of one of the Integrase inhibitors/Protease inhibitors/ NNRTIs. This combination is due to the fact that HIV-1 has a very short replication cycle and no proof-reading enzyme. This give HIV a high rate of mutation which, under the evolutionary pressure of a single drug, lead to the quick appearing of antiviral resistant mutants. To avoid this, the combination of three drugs significantly reduce the chance since if one mutation appear it cannot propagate since that virus will still be blocked by other two drugs. Currently, before starting a cART regimen, newly infected individuals are genotypically tested to assess the base line resistance, after of which the best cART regimen is decided. For example, in the UK, there is 11.8% medium to high-level resistance at baseline to the combination of efavirenz + zidovudine + lamivudine, and 6.4% medium to high level resistance to stavudine + lamivudine + nevirapine (Cane et al., 2005).

1.4.2. Failure of cART

The virus in a patient currently under cART is unable to infect new cells; this result gave the optimistic idea that if a patient was kept under treatment for long enough, the organism could be cleared from HIV-1 within a span of 3 years. This idea come from the fact that infected activated T cells shows a half-life lower than a day and analysing the viremia of patient undergoing cART (Figure13) the 1st phase shows a drastic drop in the presence of HIV-1 RNA copies in the plasma; this contraction phased happens upon death of the infected T cell. Following there is a second phase where the more resilient infected macrophages and dendritic cells start dying and last between one and four weeks. At this point the plasma viremia is low enough that can be detected only by the most sophisticated diagnostic tools since the amount of circulating HIV-1 RNA is lower than the average diagnostic limit of detection of 50 copies/ml. However, this third continuous phase, is the result of the stable latent reservoir of infected cells which persist despite the cART treatment and reactivates or slowly produce

infectious virions (Van Lint et al., 2013). As a testimony of the presence of these reservoir are presents “blips” which may be the effect of latently infected cells reactivating, producing high numbers of virions and then quickly dying out without the virus being able to infect other cells (Martinez-Picado & Deeks, 2016). The partial failure of cART is depicted in the viral rebound phase upon interruption of treatment, where the viremia level, within a span a few weeks, return to pre-cART level or even higher levels, definitely giving the proof of the existence of a population of cells acting as reservoirs hindered, in their viral production, making them resistant to cART, but able to reactivate and start the infection process once more. (Ho et al., 2013).

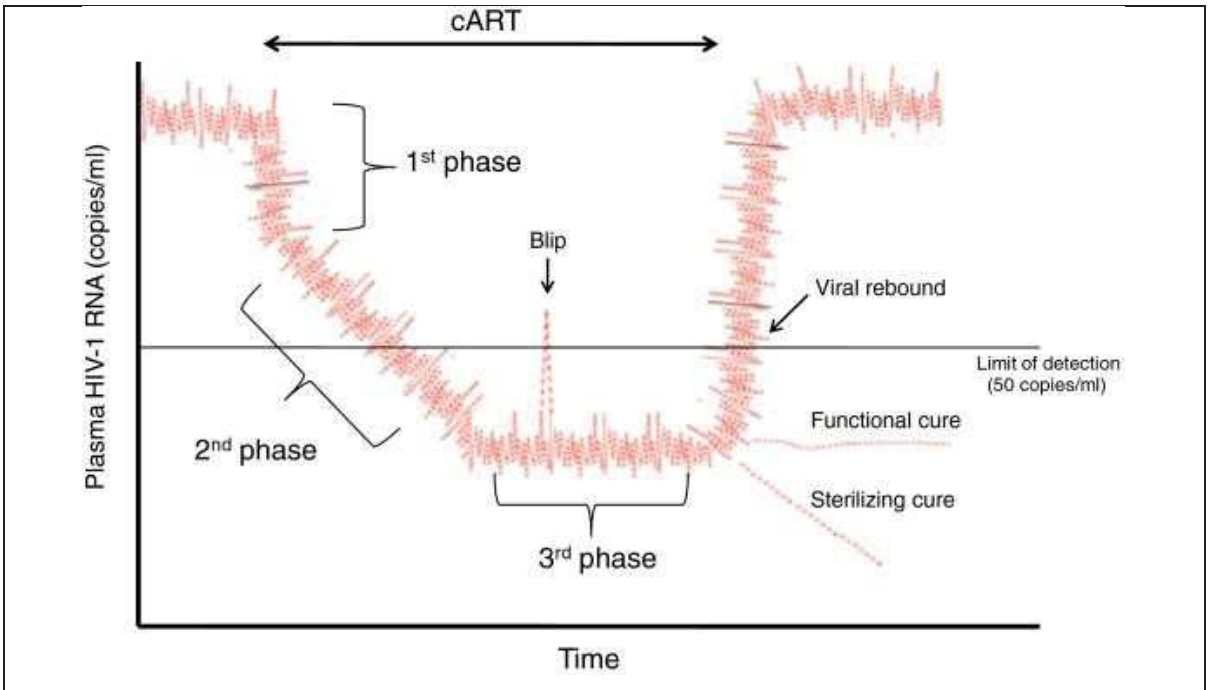


Figure 13: Plasma viremia progression in patient under cART regimen. (Van Lint et al., 2013)

Due to this rapid rebound of the viremia a HIV-1 positive patient has to take in account that the cART regimen, in the current state of the therapies, is a lifelong commitment. As stated before cART have changed the life expectancy of an infected individual from an almost sure death sentence to a life expectancy almost equal to an healthy individual living in the same environment, but this bring us to another problem with the current cART regimen. A lifelong therapy of whichever drug brings with it side effects more or less problematic. Patients undergoing cART therapy will develop co-morbidities as nephrotoxicity, hepatotoxicity, osteoporosis, lipodystrophy, hyperlipidaemia and a vast array or cardiovascular diseases. It is

also important to keep in mind the taxing cost of a life-long treatment for the patients and the public health system that in France is around at 10.000€ per year per patient.(Mallon 2007; Mallon, 2014; Cotter and Mallon 2014, Milburn et al., 2017).

1.4.3. HIV-1 Vaccine

Despite decades of research, a tolerable and effective HIV-1 vaccine is still unavailable. One of the most challenging aspects to develop a vaccine is overcoming HIV-1 envelope (Env) sequence diversity to lead to the production of protective broadly neutralizing antibodies (bNAbs) and non-neutralizing humoral and cellular responses (Bonsignori et al., 2017)
The history of HIV-1 vaccine development started immediately after the virus discovery; at the beginning vaccine development followed the traditional pathways that focused on the humoral immune response and the oldest, classic approach of inactivating, and injecting whole or partial virions were initially attempted for HIV-1. In 1999 this approach produced evidence of immunologic damage from the nef viral protein and was rapidly abandoned (Learmont et al., 1999). In the early 2000s it was completed a 2 Phase IIb/III trials in 2 different groups of patients but the protein-based vaccine (AIDSVAX) that contained gp120 proteins from various HIV-1 subtypes gave disappointing results, with an HIV-1 infection rate of 6.7% in the vaccinated group compared with 7.0% in the placebo group. After other failure in looking to obtain an adequate humoral response, the vaccinology field shifted their attention to create a vaccine able to stimulate protective CD4+ and CD8+ T-cell responses (Liang et al, 2005; Wilson et al., 2006). In 2009 the RV144 trial brought some optimism in the field after years of showing a modest vaccine efficacy of 31% at 3.5 years after vaccination, with a vaccine efficacy as high as 61% during the first year (Rerks-Engarm et al., 2009; Robb et al., 2012)

1.5. HIV-1 latency

As was described in the previous chapter, patients undergoing cART treatment remain infected and HIV-1 persist in their organism in reservoir which are protected from both the immune system and the anti-retroviral drugs. These cells, while being almost to non-productive, are extremely long-lived and the understanding on the mechanism behind the virus ability to hide in these cells and how it reactivates has drawn the attention of the scientific community. A shared idea is that by tackling this problem we could reach the full purge of the virus form an

infected individual (Deeks et al., 2016). HIV-1 reservoirs are specific cells in which the provirus is silence, but there are also specific tissues were the cells while still being active, are protected by the specific characteristics of the surrounding environment. These places are named Anatomical Reservoirs or anatomical sanctuaries. The Gut-Associated Lymphoid Tissue (GALT) is one of these, were a viral replication persist even under cART regiment. Another site, which we have discussed before, is the CNS. This site, protected by the BBB, is a site recalcitrant to the absorption of foreign molecules from the blood stream. The BBB acts as a selecting filter and ends up protecting the population of cells (microglial and peripheral macrophages) which end up being an extremely well protected reservoir, from the cART treatment, but also from other immune cells able to clear infected cells (Gras and Kaul, 2010; Wallet et al., 2019). The testis are protected as well, in a similar fashion of the CNS, by the Sertoli cell barrier.

1.5.1. Type of latency

Latency can come in two different forms. The more rare and unstable pre-integration latency and the more stable, and important from a clinical point of view, post-integration latency.
The pre-integration latency arises when the virus is unable to integrate within the host genome; this can happen form many different reasons and at the different stages of the replication cycle. In naïve CD4 T cells for example, the intracellular environment is extremely poor of nucleotides which will prolong the reverse transcriptase process to up three days. Even if the retro transcription is successful, if the cells is energetically starved, the lack of ATP will prevent the Pre-Integration Complex (PIC) to pass through the nucleus. The PIC is utterly unstable and will be degraded within few days making the pre-integration latency not the main problem to be addressed when facing the viral reservoir issue. (Van Lint et al., 2013).
The more important is the post-integration latency, where the virus has managed to integrate withing the host genome, but its gene are not expressed. This provirus is transcriptionally silent and allow these cells to persist for longer periods of time, making them a ticking bomb ready to restart the infection process. Due to major impact of these reservoirs, the post-integrational latency will be discussed further and in detail in the next chapters.

1.5.2. Types of cellular reservoirs

Naïve T cells

These cells are not the most preponderant latent reservoirs; infection of these cell by HIV-1 is considered difficult and shows many restrictions but the current accepted theory is that these cells may be infected during the thymopoiesis step of their maturation and then return to a quiescent state with a latent provirus within (Van Lint et al., 2013; Churchill et al., 2016).

Hematopoietic progenitor cells

HPC are known to be infected by HIV-1. There has been some debate if they also can become latent reservoir, but recent studies have pinpointed specific HPC like multipotent hematopoietic progenitor cells in the bone marrow and circulating mast cell progenitors to be possible targets. Of these two types, only CD34+ bone marrow cells have been shown to harbor latent provirus in HIV+ individuals with undetectable viral loads on HAART (McNamara & Collins, 2011; Pace & O'Doherty, 2013).

Memory CD4+ T cells

Since CD4+ T cells are the main target of HIV-1 it does not come as a surprise that they are also the main source of resting reservoirs. Specifically, the central (Tcm), transitional (Tt,) and effectory (Tem) memory T cells form these reservoirs. As we mentioned before, the state of these cells render the infection by HIV-1 a difficult process, so the most accepted idea is that these cell are infected once activated and survive long enough to revert to their quiescent state (Murray et al., 2016). Recently a specific cellular receptor, CD32a, was proposed as the marker for latently infected CD4+ lymphocytes; this discovery lead to insights that could facilitate the specific targeting and elimination of this reservoir (Descours et al., 2017). Sadly, the excitement for this discovery was short lived since other groups have tried to replicate this finding without success (Osuna et al., 2018). A renewed hope was instilled when again another group manage to in part replicate these finding, discovering an underlying technical difficulty to obtain a sufficiently pure population of *bona fide* $CD32^{+}CD4^{+}$ T cells. The enrichment for replication-competent proviruses in these cells was probably caused by the inability to separate the CD32a+ lymphocytes by also the APC carrying the same markers. The introduction of extra round of selection manage to obtain a more pure population of these demonstrating a progressive increase in HIV DNA enrichment in $CD32^{+}CD4^{+}$ cells with further purification of these cells, and in the purified population they found a very prominent enrichment for HIV

DNA (Darcis et al., 2019). While this discovery is still under debate, the research for specific markers for latently infected cells continues.

Monocyte/Macrophage/Microglial

Monocytes cells become infected in the acute phase of the infection and, in patients treated with cART, it is possible to detect viral DNA in the circulating monocytes. Macrophages, especially the resident one of the CNS, the microglial, have a long half-life compared to other cells of the immune system. They can survive for several years to the whole life of the patient. Another important characteristic of the microglial is their ability replicate *in-situ.* They also get infected during the acute phase of the infection (He et al., 1997), to then become silent reservoir rapidly when the viral replication is halted. During the final stage of the disease, the overall inflammation response of the organism to the pathogen will cause the reactivation of these cell, drastically increasing the number of infected cells in the CNS leading to the already described final neurological symptoms of a late stage AIDS (Glass et al., 1995).

1.5.3. HIV-1 gene silencing and host transcription factors

One of the way viral gene may or may not be expressed depends on the availability of host transcription factors. The presence of these factors, and their quantity, is directly correlated with the state of activation of the cell and, as we discussed in the previous chapters, non-activated cell will show a reduced presence of these factors in the nucleus. For example resting memory CD4+ T cells, in their inactive state, have the necessary transcription factors for HIV-1 (NFAT and NF-kB) sequestered in the cytoplasm (Archin et al., 2014) and microglial cells instead have their transcription factors pTEFb locked in an inactive form (Cherrier et al., 2013).

1.5.4. Transcriptional interference

In other situation, it is not the absence of transcriptional factors the key factor preventing viral genes to be expressed but a phenomenon that goes under the name of transcriptional interference. Three different interference have been described (Figure 14):

- Promoter competition, “a host promoter, close to the integration site, hijacks the enhancer region of the 5′LTR, leading to transcription interferences. Enhancers interacting with the LTR preferentially induce the host gene at the expense of the provirus” (Redel et al., 2010)
- Promoter occlusion, in a similar fashion to the promoter competition, a stronger cell host promoter can hijack the RNA polymerase II at the expense of the viral promoter
- RNA polymerase complex collision, when the host and viral promoter face each other, a collision between the two RNA polymerase II can bring a stop to the transcription

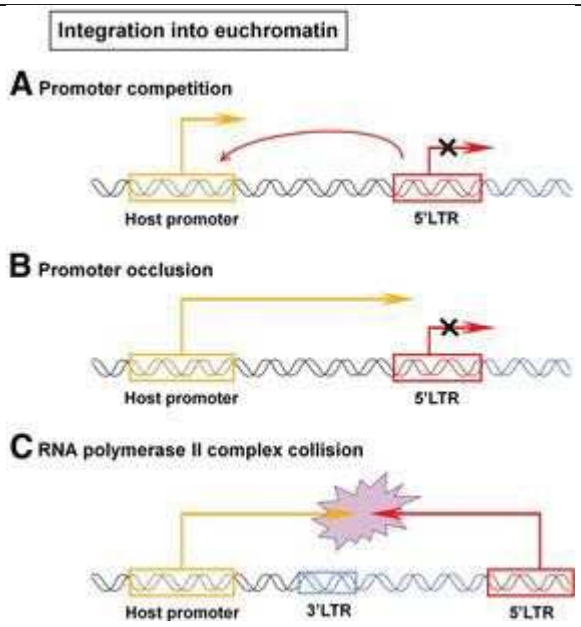


Figure 14: Mechanisms of latency of the HIV-1 provirus in the euchromatin: (A) Promoter competition, (B) Promoter occlusion, (C) RNA polymerase II complex collision: (Adapted from Redel et al., 2010)

1.5.5. Site of integration and chromatin

While the majority of the provirus (93%) integrates within the introns of highly transcribed regions of the host genome (Han et al., 2008), an important source of HIV-1 latency happens when the provirus is integrated within the highly condensed part of the host genome named heterochromatin. Chromatin shows a high level of plasticity, being able to shift from a relax, easy to access for the transcriptional factors state, the euchromatin, to a more packed, and due to steric interference, difficult to transcribe and so considered silent (Workman and Kingston, 1998; Berger et al., 2007; Yadav et al., 2018).

Chromatin structure, and state, is the result of the interaction of its fundament unit, the nucleosome. The nucleosome is formed by an octamer of proteins called histones, which form the scaffold around which the DNA is coiled, and upon specific chemical modification on these protein (and the DNA as well), this structure a more relax or packed configuration. These modifications take the name of epigenetic modifications and have inserted a great deal of complexity in the gene transcription (Kouzarides, 2007). These post-translational modifications are carried by enzymes which add specific chemical group to residues of the histone, which can be modified by acetylation, methylation, phosphorylation, SUMOylations, and ubiquitinations. To make the study of epigenetic even more complicated, these modifications will carry different effect depending on their numbers, and locations. In a brief summary, SUMOylation is connected to the heterochromatin (Maison et al., 2011; Maison et al., 2016) state while acetylation is usually connected with euchromatin; this modification is carried by histone acetyl-transferases (HAT) while it is removed (hence shifting to the heterochromatin state) by Histone de-acetyalses (HDAC).

In the case of the remaining modifications the process is more complex, and each modification will have a different effect depending on the residue that carries it. For example methylations of lysines 4, 36 and 79 of the 3rd histone (H3) will are typical marks of euchromatin state while methylations on lysines 9 and 27 of H3 and lysine 20 of the 4th histone (H4) are marks of heterochromatin state (Nakayama et al., 2001; Le Douce et al., 2012) (Figure 15).

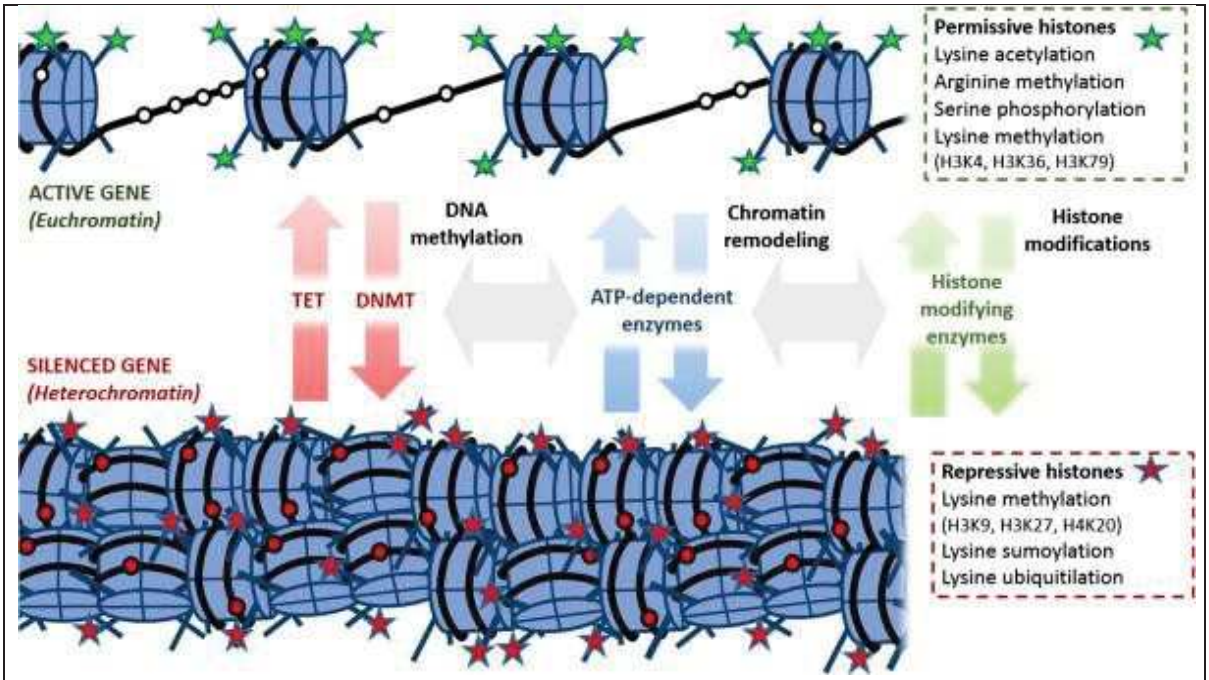


Figure 15: Summary of post-translational modifications on histones and their effect on the chromatin state. (Bégin and Nadeau, 2014)

1.5.6. Epigenetic silencing of the HIV-1 promoter

As we stated, the level of chromatin condensation determines the accessibility of transcription factors to the viral promoter. Once integrated, nucleosome 0 is precisely positioned at the HIV-1 promoter (Van Lint et al., 1996) and nucleosome 1 at the start site of transcription. Epigenetic modifications are prerequisites for the activation or the silencing of the viral promoter. These modifications are carried out by specific enzymes. Histone Methyl-Transferases (HMTs) are among these enzymes that regulates gene transcription. Under these class of enzymes, we can find G9a, SUV39H1 (Suppressor Of Variegation 3-9 Homolog 1) and EZH2 (Enhancer Of Zeste 2 Polycomb Repressive Complex 2). The first two are responsible of the H3K9 methylation and the latter methylates H3K27. All of them promote a repressive heterochromatin environment (Friedman et al., 2011).

Histone Deacetylases are also a class of enzymes connected with the silencing of HIV-1 silencing by removing the acetylations and allowing the free lysins to be further modified. After deacetylation, the HMTs will methylate H3K9 (He and Margolis, 2002) to which Heterochromatin Protein 1 (HP1) will bind to promote gene silencing (Figure 16) (Bannister et al., 2001).

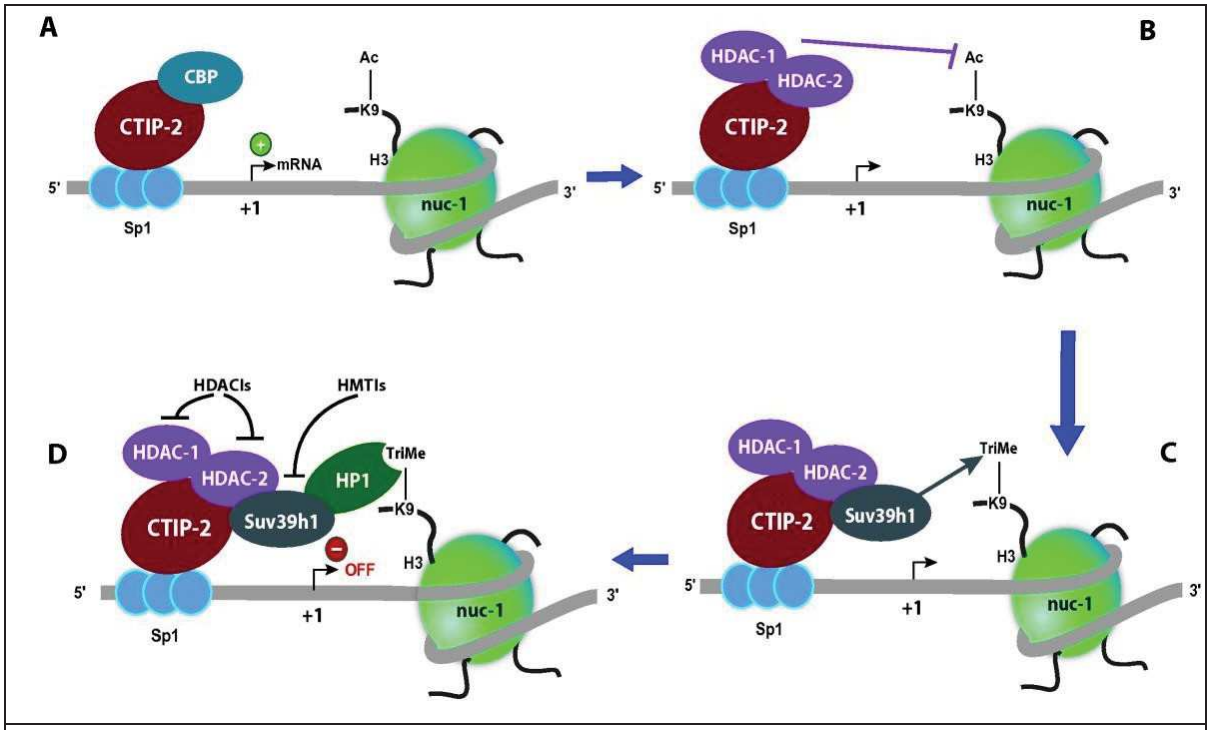


Figure 16: CTIP-2-mediated establishment of a heterochromatin environment at the HIV-1 promoter region. (Colin and Van Lint, 2009)

To confirm the key role of these proteins, a new class of drugs is being studied under the name of Latency Reverse Agents (LRAs) which should be playing a critical role in reactivating the latently infected cells to purge them. Some of these LRAs are HDAC inhibitors (HDACi) and specifically target the members of the complex responsible for the HIV-1 silencing. While showing the ability to reactivate the latent provirus, many of these LRAs have fallen short from the ability to reactivate all of the latent HIV-1 infected cells, proving how these reservoirs are kept, and maintained, by a vast array of players which need to be investigated further to find a proper treatment to reactivates all of these cells (Archin et al., 2012).

Epigenetic modifications can also occur directly on the DNA nucleotides. DNA methylation, carried on by the DNA Methyl Transferases (DNAMTs) enzymes, is associated as well with gene silencing. These modifications appear on the CpG islands and will act as biding sites for methyl-CpG-binding proteins which will recruit the histone modifying enzymes discussed before to remodel the chromatin. In latently infected CD4+ T cells, the HIV-1 promoter regions present different hypermethylated CpG islands (Blazkova et al., 2009; Kauder et al., 2009). Currently three DNMTs have been identified in mammals : DNMT1, DNMT3a, DNMT3b and a fourth enzyme DNMT3c found to be specifically expressed in the male germline (Bareu et al., 2016)

1.6. Strategies to target and remove the latently infected cells

To the current day we are not yet able to completely eradicate the viral infection form an organism with the sole utilization of the cART due to the presence of the latently infected cells acting as viral reservoirs. The two main strategies currently under study could be summarized as a sterilizing approach, aiming to the complete elimination of the virus from the organism and a functional one which aim to reduce and control the viral latency. Since the discovery of HIV-1 there have been extremely rare cases of patients that under specific circumstances have manage to control the infection, and the sterilizing approach was born from the idea to replicate what happened to the now famous Berlin patient. The Berlin patient is currently the only known HIV-1 positive person who has managed to completely eradicate the virus from its organism. This, of course, has not happened naturally but after the patient was diagnosed with an acute myeloid leukemia and was treated for it. After the destruction of his bone marrow due to the leukemia therapy he received the transplant form a patient harbouring a specific homozygote mutation of the receptor CCR5 named CCR5 Δ32/Δ32, a mutation which grant the T CD4+ cells resistance to the HIV-1 infection. While the Berlin patient has been living without cART treatment for more than 13 years, this approach is linked to an high mortality rate and it is ill advised to be used as a therapeutic alternative only to cure HIV-1 infections (Yukl et al., 2013b). Other cases worth mentioning are the one of the so called Mississippi baby, born from a mother infected with HIV-1 and not in treatment at the moment of birth, which was treated with cART for 30 hours after being born and kept in therapy till the 18 month of life. The viremia was undetectable at 30 months, but the baby experienced a rebound at two years since the stop of the treatment (Henrich et al., 2013; Persaud et al., 2013). The Boston patient was a similar case to the Berlin patient, diagnosed with a Hodgkin lymphoma, received a transplant from a donor without the CCR5 Δ32/Δ32 mutation and after the therapy he experienced a viral rebound, most likely caused by an extremely small portion of survived cells in the peripheral blood; more recently the now named London patient underwent allogeneic stem-cell transplantation with cells that did not express CCR5 (CCR5Δ32/Δ32). HIV-1 viral load in plasma remained undetectable in the London patient up to 30 months (last tested on March 4, 2020), using an assay with a detection limit of 1 copy per mL. The patient's CD4 count was 430 cells per μL (23·5% of total T cells) at 28 months. A very low-level positive signal for HIV-1 DNA was recorded in peripheral CD4 memory cells at 28 months. The viral load in

semen was undetectable in both plasma (lower limit of detection [LLD] <12 copies per mL) and cells (LLD 10 copies per 10^6 cells) at 21 months (Gupta et al., 2020).

Among the patients that manage to show a immune control of the viral infection, the VISCONTI cohort (Viro-Immunological Sustained Control after Treatment Interruption) is a group of fourteen patients who received early therapy for the virus and were considered to be in "long-term virological remission," meaning that they still harbour the virus within their bodies but HIV viral loads are low or undetectable despite being off antiretroviral therapy (Sáez-Cirión et al., 2013). Due to the many obstacles (and risks) to pursue a sterilization strategy, the functional ones has started to focus his attention of a set of patients named HIV Elite controllers, rare patients which are able to keep in check the HIV-1 infection thanks to a more reactive immune response carried on by their dendritic cells, Natural Killer cells and T CD8+ which are more successful in eliminating the infected CD4+, plus the presence of neutralizing antibodies. These patients also show a slow progression of the disease due to a higher degree of methylation of the viral promoter and a lower level of latently infected cells (Autran et al., 2011). Two strategies to pursue the functional healing are currently under study, named “Shock and Kill” and “Block and Lock”.

1.6.1. Shock and Kill

This strategy is based on three different mechanism of action and the utilization of specific compounds that goes under the name of Latency Reversing Agents (LRAs). The first target is to release the transcription factors (NF-kB, NFAT, PTEFb) from their block to allow the initiation of the transcription of the viral genes. The second is to target the epigenetic environment around the viral promoter, which in the case of the latently infected cells is usually blocked within the heterochromatin part of the genome and refractory to the access of the transcription factors. The third is to activate the immune system to target the reactivated cells and kill them.

For the first target, within the “shock” part of the therapy, different compounds are currently under study such a prostratine and bryostatine which are agonists of the Protein Kinase C (PKC) and lead to the activation of the NF-kB and causes a strong activation of the viral transcription; these compounds have shown, sadly, a high level of toxicity. P-TEFb was able to be released and the AKT pathway activated in vitro by two compounds, HMBA and Disulfram showing an inversion of the latency. Sadly, experiment in humans has shown only partial success with a modest reactivation of the T cells and a non-detected reduction of the

size of the reservoirs. Agonist to the TRL receptor have shown a good efficiency in reactivating the viral gene expression with the follow up reactivation of the specific immunity in patients under cART (Thorlund et al., 2017).

Agents to reshape the chromatin environment are also been tested and among them the inhibitors of the HDACs are the most studied both in vivo and in vitro. Disulfiram,

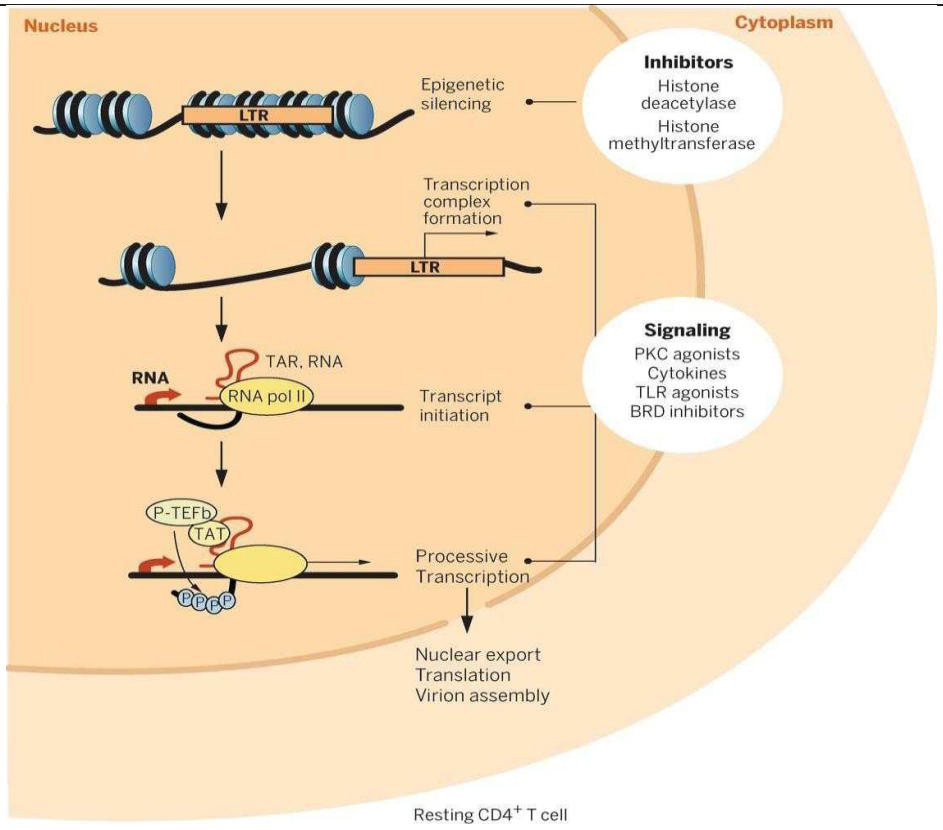


Figure 17: The "Shock" part of the "Shock and Kill" strategy. The Shock aim to both free the blocked transcription factor and reshape the chromatin environment around the viral promoter (Clutton & Jones, 2018)

Panobinostat and vorinostat have already been clinically tested showing activation of infected cells but no reduction in the size of the reservoirs (Archin et al., 2012; Elliott et al., 2015). Inhibitors of the Histone Methyltransferase such as chaetocine and BIX-01294 and DNA demethylations agents have shown to be able to reverse the latency in vitro (Bouchat et al., 2012; Bouchat et al., 2015). BET (Bromodomain and Extra-Terminal motif) inhibitors such as RVX-208 and PFI-1 reactivated HIV-1 transcription by up-regulation of P-TEFb by increasing CDK9 Thr-186 phosphorylation in latently infected Jurkat T cells *in vitro* and in CD4+ cell from HAART patients in *ex vivo* (Darcis et al., 2015).

The step of clearance or "kill" is actuated by the immune system carried by the NK cells and HIV-specific CD8+ T-cells and by the viral cytopathic effect. Recent studies have shown how a the use of a single compound as LRA showed suboptimal effects in reactivating the latent cells and a combination of LRAs is required, moreover, a prolonged treatment with cART have the side effect of reducing the numbers of HIV-specific CD8+ T-cells (Perreau et al.,2017) and introduced the necessity to stimulates as well these cells population to effectively killed the reactivated cells. A *study in vitro* as shown also how latent cell reactivated by Vorinostat were not dying for cytopathic effect, but they were efficiently cleared by the CD8+ lymphocytes (Shan et al., 2012). The current strategy has evolved then to include a stimulation of the CD8+ T cells which can detect and kill with astonishingly sensitivity and can as well form long-lived "memory" populations capable of rapidly act in case of future encounter with the virus (Wherry & Ahmed, 2004). During the acute phase of the infection, the emergence of these CD8+ cells coincides with the decline in the virus load. These cells target the HIV proteome in such a fashion that the virus cannot escape without evolving with a fitness cost (Pereyra et al., 2014). A combination of Romidepsin with a specific vaccine designed to stimulates CD8+ T cell have reported, in two different studies, to significant delay in the viral rebound after cART interruption (Mothe et al., 2017) and to reduce the size of the reservoir with a viremia level below the level of detection (Leth et al., 2016).

Also, the use of specific neutralizing antibodies is under study and some interesting result were obtained in humanized mice treated with Varinostat that showed a delay in the viral rebound after cART interruption (Halper-Stromberg et al., 2014).

1.6.2. Block and Lock

In parallel to the "Shock and Kill" strategy, a new approach is currently being tested with the aim to completely shut down the replication in the latently infected cells. As we described before, latently infected cells show an extremely low level of viral replication, which makes them responsible of the viral rebound once the cART regimen is interrupted. The "Block and Lock" strategy aim directly at utilizing compound that suppress completely this basal replication and then lock them in this state. Many viral and cellular proteins are involved in HIV transcription and silencing, and hence represent potential targets for future block-and-lock approaches (Darcis et al., 2017). Here is a list these different block-and-lock strategies (Figure 18):

- Tat Inhibition by Didehydro-Cortistatin A (dCA) is currently the most advanced block-and-lock approach and employs a Tat inhibitor. dCA has shown the ability in *ex vivo* of inhibiting the residual viral replication and to delay the reactivation after cART is interrupted. It also has been reported to reduce the size of the viral reservoir in the lymphoid tissue and brain of humanized mice (Kessing et al., 2017)
- LEDGINs are small molecule inhibitors of the interaction between HIV integrase (IN) and the cellular chromatin-thethering factor LEDGF/p75 (Christ et al., 2010)
- FACT Inhibition by Curaxin CBL0100. FACT acts as a histone chaperone and promotes RNAPII driven transcription by destabilizing the nucleosomal structure. A 2011 study showed that the anticancer compounds named curaxins inhibit FACT and suppress NF-κB mediated transcription (Gasparian et al., 2011). in 2017 HIV replication and reactivation was blocked by using curaxin CBL0100. Curaxin CBL0100 inhibited RNAPII mediated transcription elongation in a Tat-dependent manner (Jean et al., 2017)
- RNA-Induced Epigenetic Silencing. By using short interfering (si) or short hairpin (sh) RNA to maintain the repressive heterochromatic landscape at the HIV 5' LTR promoter. Two RNA were designed (siRNA143 and PromA) to target the transcription factors binding site on the LTR promoter reducing the reactivation level on J-Lat cells (Suzuki et al., 2008).
- Levosimendan is a drug used to treat heart failure and was shown to strongly inhibit the HIV-1 replication in latently infected primary CD4+ T-cells (Hayashi et al., 2017)
- HSP90 Inhibitors Heat shock protein 90 (HSP90) is a cellular chaperone protein that helps folding and stabilizing other proteins. It has been shown that during the HIV-1 infection, the expression of this chaperon protein increases in both T lymphocytes and mononuclear cells (Ringrose et al., 2008) and inhibitors of HSP90 suppress HIV transcription and replication.
- Triptolide is a compound that inhibits viral transcription by targeting Tat for degradation by the proteasomal pathway (Wan & Chen, 2014)
- Jak-STAT Inhibitors. The Jak-STAT pathway was shown to be involved in HIV persistence and reactivation in primary CD4+ T cells (Gavegnano et al., 2014). Two inhibitors of this pathway have been already tested (Ruxolitinib and tofacitinib) showing a strong inhibition of the reactivation. Moreover, the anti-inflammatory effect

of these compounds reduced activation of T cells limiting transmission of HIV to other cells.

- BRD4 Modulators. The bromodomain-containing protein 4 is an epigenetic reader that interacts with various proteins to stimulate gene expression (Jang et al., 2015), but by competing with Tat to bind the P-TEFb, inhibits HIV-1 transcription. ZL0580, a BRD4 modulator, has also been shown to promote a repressive chromatin environment at the LTR promoter in addition to inhibiting Tat transactivation and transcription (Niu et al., 2019).
- mTOR Inhibitors. Inhibition of mTOR suppressed HIV reactivation in primary CD4+ T cells and patient cells by downregulating CDK9 phosphorylation and blocking NF-κB signalling (Besnard et al., 2016)

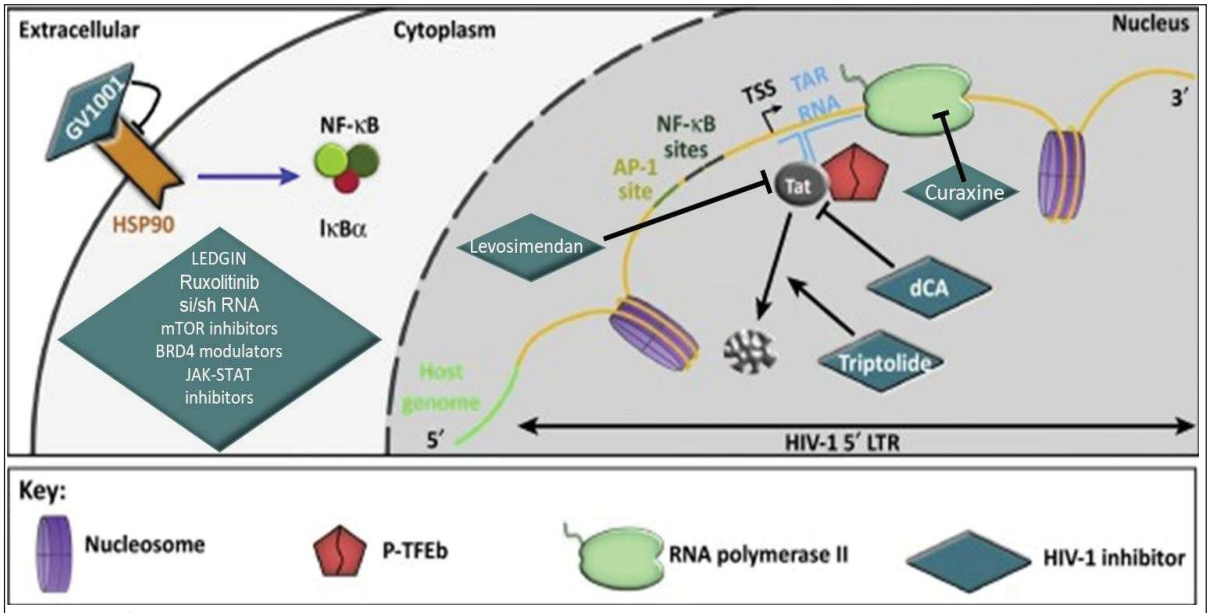


Figure 18: The recently discovered compound acting as HIV-1 inhibitors used in the Block and Lock approach (Adapted from Darcis et al., 2017)

1.7. Bcl11b/CTIP2

The transcriptional factor Bcl11b (B-cell lymphoma/leukemia 11b) also known as CTIP2 (COUP-TF interacting protein 2) act as a transcriptional repressor and it is mainly expressed in the brain and the immune system (Avram et al., 2000; Leid et al., 2004). The gene coding for this protein is localized on the 14p32.2 human chromosome and in the murine chromosome 12 with an 88% similarity between the two genes (Kominami, 2012). In humans this gene, formed by 4 exons, encodes for a zinc-finger protein (Cys2-His2) that has two different

isoforms of 812 or 884 aminoacids. This protein carries 6 different zinc-finger domains, required to elicit it binding on the DNA and other domains through which it interacts with other proteins (Figure 17).

CTIP2 has different functions. Required for a correct development of the CNS, CTIP2 plays a key role in the development of corticospinal motor neuron axonal projections to the spinal cord and has been linked as well to the formation of brain tumours (Arlotta et al., 2005; Ganguli-Indra et al., 2009). CTIP2 is necessary also for the development of the immune system where it play a crucial role in the differentiation of the thymocytes. In addition, CTIP2 has recently been linked to play a function in both the odontogenesis and the formation of the permeable protective barrier of the derma (Wakabayashi et al , 2003; Golonzhka et al., 2009a; Golonzhka et al., 2009b). Due to his wide range of functionality, the gene coding for CTIP2 was investigated in other model animals and a corresponding gene of the mammalian version was found in creature phylogenetically quite distant from us such as the common fruit fly, *Drosophila melanogaster*, the zebra fish, *Danio rerio*, and the nematode *Caenorhabditis elegans*. The presence of a similar protein in animal so distant from us, from the evolutionary point of view, shows already how this protein is responsible of many essential functions. As an example, a mutation in this gene in humans is responsible of immunodeficiency caused by the lack of T lymphocytes. This disease called Severe Combined Immunodeficiency carries with it other clinical onsets all connected with the already described roles of CTIP2 as teeth deformities, dermatitis and impaired intellectual development (Punwani et al., 2016; Lessel et al., 2018). Knock-down mice lacking CTIP2 will not survive once born.

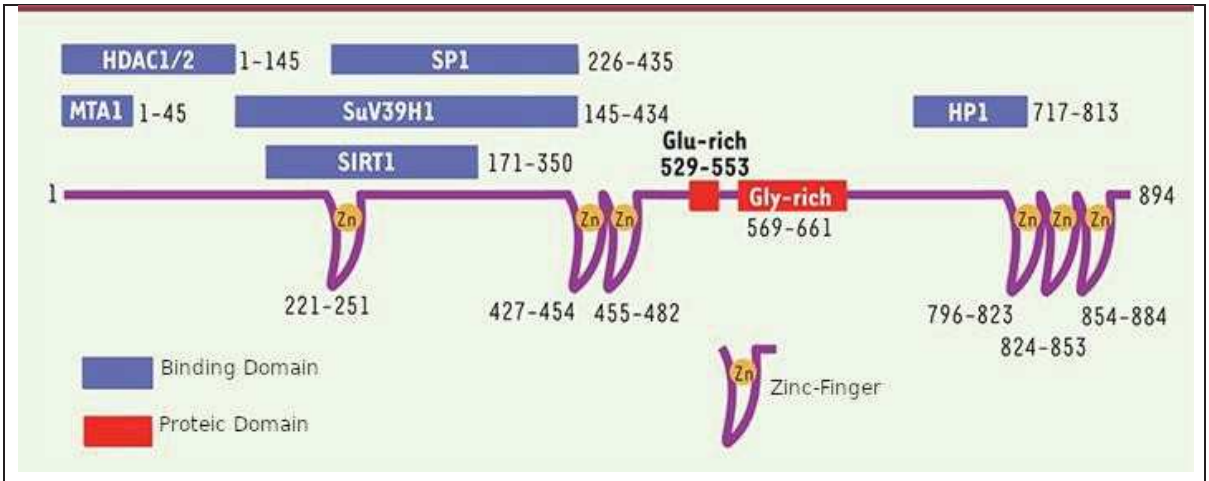


Figure 17: Schematic representation of the protein CTIP2. In blue are shown the location of the binding domain of CTIP2 with other proteins like the Histone Deacetylases 1 and 2 and the Methyltransferases SUV39H1. In red is the Glutamate and Glycine rich proteic domain and in purple the six zinc-finger motifs. (Adapted from Le Douce et al., 2014).

There have been different studies, on the role of CTIP2 in the regulation of gene expression and its ability to bind directly or indirectly on specific sequences on gene promoter (Avram et al., 2000; Marban et al., 2005). Nevertheless CTIP2 play an essential role in the expression of a multitude of genes, both as a repressor and an activator, and its presence in many different systems and proteins complexes make it the key factor in different pathologies if its function is altered.

1.7.1. CTIP2 as a transcription factor

As state before, CTIP2 can regulate the gene expression by binding to the promoter through the interaction with Sp1 bound to the G/C regions.

While we will focus later in detail on the repressor effect of CTIP2 on the transcription, the cellular state, environment, and the post-translational modification on CTIP2 will make it act as an activator of the transcription. In the thymocytes activated by the T-cell receptor (TCR), the activation of the Mitogen-activated protein kinase (MAP) kinase will cause the dephosphorylation and SUMOylating of CTIP2 which will then become and activator of the Inter-leukin 2 gene (the IL2 activator effect of CTIP2 has been proven in the CD4 + lymphocytes as well). The post-translational modification of CTIP2 allows it to recruit a transcriptional coactivator, p300, which will act as an histone acetyltransferase and will remodel the chromatin environment form the condensed heterochromatin state to the relax euchromatin allowing the transcription (Cismasiu et al., 2006). Recently, a more in deep study has discovered the existence of a specific post-translational modification that act as a switch in this repressor/activator state of CTIP2 that is triggered with the TCR signalling cascade through the PKC, and lead to the phosphorylation of the Ser2 of CTIP2 switching it from a repressor to an activator in CD4+ Lymphocytes (Dubuissez et al., 2016). This modification on this Serine has also been documented to impair the recruitment by CTIP2 of the Nucleosome Remodelling and Deacetylation complex (NuRD) by dampening the interaction of CTIP2 with the MTA1 and MTA2 proteins (Cismasiu et al., 2005). Without the activation of the TCR, if CTIP2 maintains it repressor effect, the interaction with MTA1 and MTA2 will recruit the NuRD complex on the promoter to repress the transcription by remodelling the chromatin. While this interaction of CTIP2 with chromatin remodelling agents as NuRD it is specific for the T CD4+ lymphocytes, in microglial cells CTIP2 has a similar role by interacting with the Lysine-Specific Demethylase 1 (LSD1) and recruiting on the promoter (after binding Sp1) different chromatin remodelling proteins. In this case CTIP2 act as a scaffold or a dock site on which

these complexes bind and exert their function with LSD1 recruiting the human Complex proteins associated with Set1 (hCOMPASS) and CTIP2 recruiting the complex composed by HDAC1, HDAC2 and SUV39H1 (Marban et al., 2007; Schwartz et al., 2010; Le Douce et al., 2014).

1.7.2. CTIP2 seize the elongation factor P-TEFb

Another role in the repression of the transcription that has been linked to CTIP2 is its ability to seize in an inactive form the elongation factor P-TEFb. P-TEFb is an heterodimer composed by the cyclin T and the kinase CDK9; these two proteins have the function to phosphorylate (Marshall et al., 1996) the carboxy-terminal domain of the RNA polymerase II when it is in the early stages of transcription. This phosphorylation activates completely the polymerase allowing the elongation of the transcripts. The amount of P-TEFb available to reactivate the RNApolII is controlled by a balance between an active "free" complex and an inactive complex bound to the non-coding RNA 7sk (already described in the Tat-dependent phase of the HIV-1 gene transcription, Figure 8). This inactive complex includes CTIP2 and HEXIM1 and it is in a catalytically inactive state (Michel et al, 2004; Cherrier et al., 2013). The balance between the inactive and active form of the P-TEFb is under tight cellular control to adapt precisely to the need of the cell in the date moment. Perturbation of this equilibrium is often associated to the arise of a disease (Muniz et al., 2012). Our lab has shown the role of CTIP2 in the regulation of this complex. A global analysis of the gene expression upon CTIP2 overexpression showed that significant clusters of PTEF-B sensitive genes were coregulated by CTIP2 (Cherrier et al., 2013). In addition, we have demonstrated that CTIP2 has a direct negative effect on the catalytic subunit of PTEF-B : the CDK9 . In the context of viral latency, our lab described how the inactive P-TEFb complex is recruited on the HIV-1 promoter thanks to the mutual interaction of CTIP2 with the HMGA1 protein which binds the Transcription factors on the viral promoter. As I mentioned before, this complex is targeted by Tat, in the Tat-depend phase of the transcription. Tat is able to increase the active portion of the P-TEFb complex to allow the transcription of the viral genes. We can show then how CTIP2 is a major player in the establishment of the viral latency, by creating an heterochromatin environment but also by maintain the latency by inactivating the activity of the P-TEFb complex (Figure 18, Wallet et al., 2019)

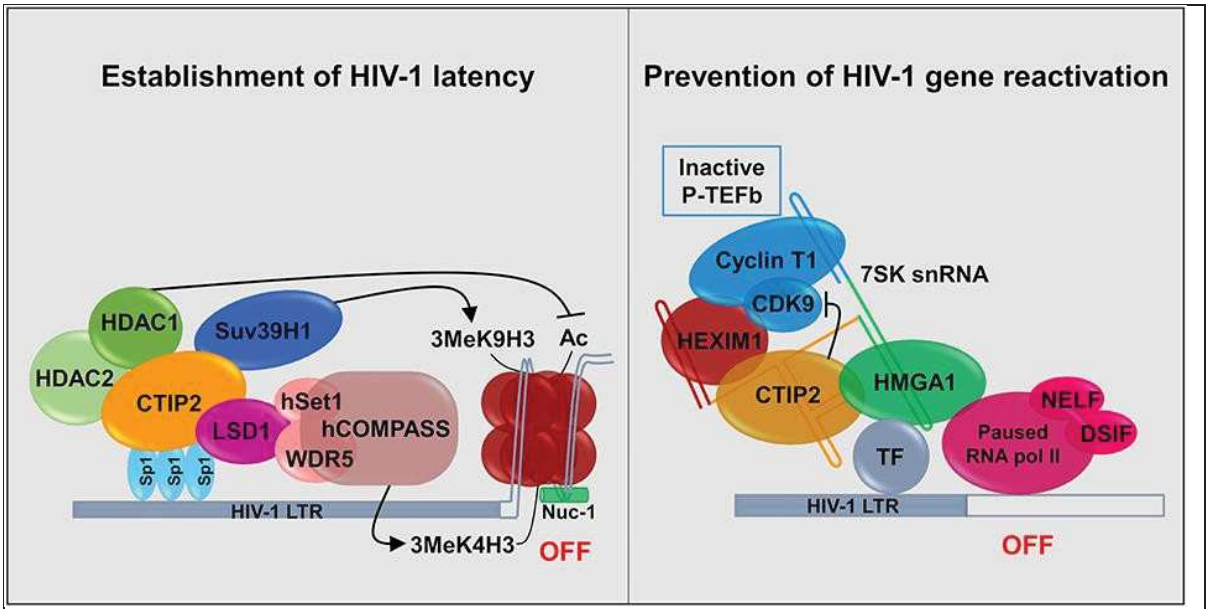


Figure 18: CTIP2 roles in the establishment and persistence of HIV-1 latency in microglial cells. On the left CTIP2 is involved in the establishment of HIV-1 latency by recruiting a chromatin modifying complex at the viral promoter. CTIP2 is anchored at Sp1 sites on the viral promoter and act as a scaffolding for the recruitment of several chromatin modifying proteins. On the right CTIP2 prevents transcription restart by sequestrating the elongation factor pTEF-B in an inactive complex where 7SK snRNA act as scaffold complex and is recruited at promoter via HMGA1 by interaction with transcription factors (TF). Transcription is inhibited at two levels. First, CTIP2 inhibits the kinase activity of CDK9, a pTEF-B subunit which prevent RNA pol II activation by inhibiting the phosphorylation of RNA pol II and the two subunits NELF and DSIF. Second, the persistence of inactive pTEF-B at the promoter prevents any recruitment of active pTEF-B. (Wallet et al., 2019)

1.7.3. CTIP2 and HIV-1

As we described in the previous chapter our group has explored the effect of CTIP2 on the transcription, especially of the HIV-1 proviral genes. In this chapter we will see more in detail the relationship of CTIP2 and HIV-1. To begin with patients infected with HIV-1 shows elevated amount of CTIP2 circulating in the cerebrospinal fluid (CSF) and post-mortem analysis showed this higher expression also within cells of the CNS such as astrocytes and microglials. This increase expression is directly correlated to the state of latency within the infected cells of the CNS when compared to asymptomatic patients or patients showing symptoms correlated with the inflammation of the brain (encephalitis) caused by the HIV-1 infection. This increase of CTIP2 levels is accompanied also by an increases in levels of HDAC1, HP1α and a deregulation of pro-inflammatory genes such as TNF- α and IL-6, which when are actually upregulated, offer an explanation to the brain damage in patients infected with HIV-1 (Desplats et al., 2013). As we stated before CTIP2 inhibits the transcription in the

microglial cells by recruiting the chromatin remodelling complex composed by HDAC1, HDAC2 and SUV39H1 to the viral promoter and in the lymphocytes T CD4+ by recruiting the NuRD complex but our group has shown other methods acted by CTIP2 to repress the HIV-1 transcription. One of these is by destabilizing the activity of Tat, more specifically CTIP2 delocalize Tat in specific sphere-like structures in the inactive part of the chromatin in complexes counting among them CTIP2-Tat and HP1α (Rohr et al., 2003a). The antagonistic relationship of Tat and CTIP2 is also seen, as explained in the previous chapter, on the role of Tat in recruiting the active P-TEFb complex, and CTIP2 and HGMA1 effect in reducing it to its inactive form; a depletion of both CTIP2 and HGMA1 as a matter of fact, causes an increase in both the initiation and the Tat-dependent phase of the transcription of the viral genes (Eilebrecht et al., 2014).

CTIP2 also exert a pleiotropic effect, indirectly promoting the latency of HIV-1 by downregulating genes favourable for the provirus transcription. One of these genes downregulated by CTIP2 is a similar fashion on how it repress the transcription of the HIV-1 genes, is the p21 which stimulates the viral gene expression in monocytes and macrophages thought the cyclin CDKN1A/$p21_{waf}$ (Vázquez et al., 2005; Cherrier et al., 2009).

1.8. PSF/SFPQ

PTB-associated splicing factor (PSF) also known as Splicing factor proline- and glutamine-rich (SFPQ) is an abundant and ubiquitous RBP that plays roles in many aspects of RNA biogenesis including transcriptional initiation and termination, transcriptional activation and repression, and splicing.

1.8.1. SFPQ is a member of the DBHS family

SFPQ is part of a family of multifunctional nuclear factors termed DBHS (Drosophila behavior human splicing) proteins. These proteins are highly conserved within three specific domain, a NonA/paraspeckle domain (NOPS), a N-terminal RNA recognition motif (RRMs) and a C terminal coiled-coil (Patton et al., 1993). The RRM2 and NOPS domain are necessary for the formation of homodimers and heterodimers among the proteins of these family (Fox et al., 2005; Kuwahara et al., 2006). Evolutionary this family of protein is present only in vertebrates and invertebrates and in human three protein are member of this family: SFPQ, the paraspeckle

protein component 1 (PSP1) and the Non-POU domain-containing octamer-binding protein (NONO also know with the name of p54nrb).

1.8.2. PSF/SFPQ localization

SFPQ, like the other proteins of the DHBS family, localize massively in the nucleus; in the nucleoplasm, the nucleolar caps and within specific nuclear structure named paraspeckles (Shav-Tal et al., 1992; Bond & Fox., 2009). As we will see in a later chapter, on all the different functions in which SFPQ has been connected to, the localization of SFPQ within all these different parts is suspected to be controlled by post-translational modifications (PTMs) of the protein as well as cellular environment. A specific mention is the localization within the paraspeckle structure, subnuclear bodies of roughly sphere like shape, localization shared among all the DHBS proteins, many of which, SFPQ included, are necessary for the formation of these structures (Fox & Lamond, 2010). Under specific circumstances, SFPQ is also found in the cytoplasm, but often this displacement it a sign of an undergoing disease progression such as Alzheimer's and Pick's disease (Ke et al., 2012).

1.8.3. PSF/SFPQ functions

SFPQ is the only member of the DHBS family to be essential for the cell. A study showed how even a reduction of SFPQ quantity of three-fold could induce apoptosis, in mice SFPQ depletion lead to an impairment of T-cell development and in zebrafish it is required for properly neuronal development. (Bond & Fox., 2009). Considering it is importance it does not come as a surprise to find SFPQ to be connected to many different and essential functions within the cell (Figure 19):

- Splicing: Splicing of the α-tropomyosin pre-mRNA was the first function to be connected to SFPQ. It was found for the first time connected to the RNA-binding protein PTB as, at the time, an unknown protein of roughly 100kDa MW (the protein itself is 72kDa but migrates in an abnormal fashion) (Patton et al., 1991). Nowadays is accepted the function of SFPQ is to be loosely associated with the splicing machinery

and its function is substrate-dependent especially in the second catalytic step of splicing, exon joining, of many, but not all, pre-mRNA.

- Nuclear retention: SFPQ, in conjunction with the other DHBS protein and MATR3 binds with high affinity a specific modification carried by some mRNA. A deamination event (catalyzed by the Adenosine De-Aminase RNA-Specific, ADAR) modifies the adenosine base into a rarer one called Inosine (I) which then act as a nuclear retention signal. This modification appears to be more common in mRNA that shows inverted repeats and lead to a double -stranded RNA structures (Chen & Carmichael, 2009). SFPQ is believed to directly recognize this Inosines and retain these mRNAs within the paraspeckle structure. (Nagawa & Hirose, 2012)
- Transcription: SFPQ can act as both a positive and negative regulator of transcription. SFPQ tighter with NONO are able to bind to the C-terminal domain of the RNA polymerase II and bridge the recruitment of polyadenylation and splicing factors to the site of transcription (Rosonina et al., 2005; Nagawa & Hirose, 2012). SFPQ also binds the exonuclease XRN2/Rat1 which is required for transcription termination and its knockdown prevent the function of this exonucleases causing an accumulation of 3' cleaved RNA in vitro (Kaneko et al., 2013). SFPQ is also known to negative regulate transcription by interaction with chromatin remodelling enzymes such as HDAC to repress transcription on circadian rhythm-controlling factors and DNA-bound nuclear hormone receptors. SFPQ can also directly repress transcription of genes close to location where it binds directly on DNA, as seen in the case of the IL-8 genes and RAB23 (Wang et al., 2009; Imamura et al., 2014)
- 3'-End processing. The first time SFPQ was linked to 3'-End processing was when it was found in HeLa cells tighter with NONO, to be interacting with the SF-A complex. Experiment with antibodies targeting SFPQ affected the polyadenylation (as well as the splicing) at suboptimal polyadenylation sites (Lian et al., 2006). To prove this correlation a Tethering experiment with SFPQ to the cyclooxygenase 2 (COX-2) suboptimal polyA signal was shown to activate the polyadenylation confirming the effect of SFPQ (Hall-Pogar et al., 2007). Even if the underlying mechanism has not yet been described, other studies have confirmed the role of SFPQ and other associated proteins, in the polyadenylation in where the signal is suboptimal.
- DNA-Repairing: It has been shown how many RNA-binding proteins are often connected to the DNA-repairing machinery and SFPQ is no exception (Ha et al., 2011; Morozumi

et al., 2009). Indeed SFPQ has a role in recognizing and repairing Double-Strand Breaks (DBS) through the interaction of its N-terminal domain with proteins such as the recombinase RAD51D (Morozumi et al., 2009) which plays a critical role in the main pathway of DBS repairing, Homologous recombination (HR). The binding of SFPQ cofactor NONO to DSB also recruit other proteins complexes responsible for the other main DSB repairing machinery, Non-homologous end-joining (NHEJ) (Salton et al., 2010; Rajesh et al., 2011)

- Viral infection: SFPQ has been reported to exert roles in the replication and infectivity of many viruses such as influenza (Landeras-Bueno et al., 2011), hepatitis (Greco-Stewart et al., 2006) and HIV (Zolotukhin et al., 2003; Kula et al., 2013). In influenza PSF seems to be required for RNA synthesis and viral replication since its depletion showed a reduced viral transcription and lower virus production. Interestingly, the splicing of viral RNA was not affected (Landeras-Bueno et al., 2011).

  SFPQ has been shown to interact directly with the Hepatitis Delta virus (HDV) RNA and it is believed to be involved in its life cycle (Greco-Stewart et al., 2006)

  Regarding HIV-1 infection, SFPQ act as a cofactor with MATR3 and Rev in the transport of unspliced viral RNA to the cytoplasm and it promotes the production of viral transcript as well through its interaction with Rev. SFPQ has been shown also to bind the cis-acting regulatory elements (INS) in the gag mRNA and to decrease the expression of gag-pol and env. The hypothesis of this reduction is supposed to be linked to a not yet described function of SFPQ in mRNA degradation (Zolotukhin et al., 2003)

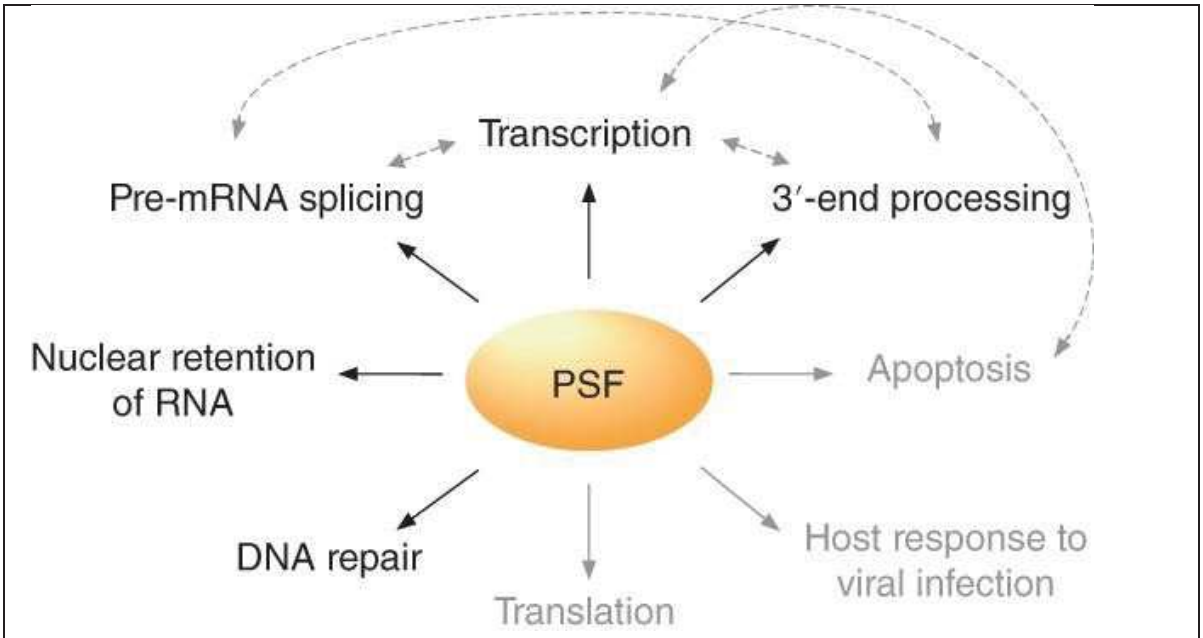


Figure 19: A summary of the cellular activities in which SFPQ is involved. With Black arrows are the well understood and proven function. In grey the function which to this day are not yet being well described and are based on hypothesis. Dotted lines represent coordination of activities of SFPQ. (Yarosh et al., 2015)

1.8.4. Post-Translational modification of PSF/SFPQ

The activity of SFPQ has been shown to be highly regulated by PTMs such as methylation, SUMOylation and the most well described in regard of SFPQ, phosphorylation. SFPQ has been reported to be methylated on the RGG box, and while the biological function of these specific modification has yet to be described, it is supposed that it may alter the interaction with other proteins (Ong et al., 2004; Snijders et al., 2013).

Four specific residues in the RRM1 domain are instead the target for SUMOylation (337–340, IKLE) and they have been shown to be required for the SFPQ interaction with HDAC1. SUMOylation of SFPQ is directly proportional to the decrease in histone acetylation and the activity on the tyrosine hydroxylase promoter in human (one of SFPQ targets) (Zhong et al.,2006).

Phosphorylation is the most abundant and well documented PTMs on SFPQ with more than 10 kinases have been shown to interact with it on many different phosphorylation sites.

Phosphorylation of SFPQ's serines 8 and 283 by MAP kinase interacting kinase (MNK) was shown to increase the binding of SFPQ to the 3′ UTR of TNFα mRNA (Buxade et al., 2008). The phosphorylation of SFPQ by the Glycogen synthase kinase 3 (GSK3) has been shown to regulate the splicing activity of SFPQ on the CD45 pre-mRNA (Heyd & Lynch, 2010).

Phosphorylation on the uncommon tyrosine residues of SFPQ have been discovered as marker in cancer cells where the phosphorylated Tyrosine 293 is suspected to be responsible of SFPQ misplacement in the cytoplasm (Galietta et al., 2007) by probably altering the interaction with carrier proteins.

This cytoplasmatic localization also inhibits cellular proliferation, probably due to a loss in essential functions of SFPQ within the nucleus (Galietta et al., 2007). Dephosphorylation is also regulated on SFPQ since different co-immunoprecipitation have shown its interaction with the protein phosphatase 1 (PP1) via the RRM1 region which contain a specific consensus binding sequence for PP1 (RVxF) (Hirano et al., 1996). Recent studies have suggested that PP1 was bridged to SFPQ by NONO (which also have an RVxF sequence). Since PP1 activity alters SFPQ transcriptional activity, it does not influence SFPQ binding to his transcriptional co-repressors (Liu et al., 2012).

1.9. Paraspeckle

Paraspeckles are small, subnuclear bodies that varies in number, depending on the cell type, between 5 and 20 foci per nucleus (Clemson et al. 2009). Paraspeckles have been found only in mammalian nuclei with the human embryonic stem cells (hESC) being the only mammalian cell type where paraspeckles were not found. (Clemson et al. 2009; Sasaki et al. 2009). While some of the proteins forming these structure are also found in other vertebrate and invertebrate species (see the SFPQ chapter on DHBS proteins) the fact that the NEAT1 RNA, which form the paraspeckle scaffold is only found in mammals may be the reason why these structures are limited to our Class. Paraspeckles are localized in the interchromatin space, close by to other nuclear speckles.

1.9.1. Paraspeckle components

The three most studied and well-known proteins of the paraspeckle are all member of the DBHS family and are PSPC1, NONO and SFPQ. The interplay between these proteins suggest that they exist in hetero and homo-dimers (Myojin et al., 2004; Fox et al., 2005) which is facilitated by the fact that they share more than 50% sequence identity within two N-terminal RNP-type RNA recognition motifs and a C-terminal coiled-coil domain. NONO and SFPQ are compulsory for the paraspeckle formation as it was discovered after knocking them down, but this does not seem to be the case for PSPC1 (Sasaki et al., 2009). Several other proteins have

been found to localize within the paraspeckle but only in experiments following overexpression and therefore they are not yet considered canonical member of this structure till localization of their endogenous form is examined (Table 1) (Fox et al. 2002; Dutton et al. 2006; Liu et al. 2006; Rivera et al. 2009).

Table 1: Summary of the main proteins and RNAs composing the paraspeckle structure (Adapted from Fox et al., 2018)

Gene Name	Importance in paraspeckle formation	Paraspeckle Zone
AHDC1	Dispensable	
AKAP8L	Dispensable	
CELF6	n.d.[(e)]	
CIRBP	Dispensable	
CPSF5	Dispensable	
CPSF6	Dispensable	
CPSF7	Important	
DAZAP1	Essential	
DLX3	n.d.	
EWSR1	Dispensable	
FAM98A	Important	
FIGN	Important	
FUS	Essential	Core
HNRNPA1	Important	
HNRNPA1L2	n.d.	
HNRNPF	n.d.	
HNRNPH1	n.d.	
HNRNPH3	Essential	
HNRNPK	Essential	
HNRNPR	Important	
HNRNPUL1	Important	
MEX3A	n.d.	
NONO	Essential	Core
PCED1A	Important	
PSPC1	Dispensable	Core
RBM3	Dispensable	
RBM4B	Dispensable	
RBM7	Dispensable	
RBM12	Important	
RBM14	Essential	Patch
RBMX	Dispensable	
RUNX3	Dispensable	
SFPQ	Essential	Core
SMARCA4 (BRG1)	Essential	Patch
SRSF10	Important	
SS18L1	n.d.	
TAF14	Important	
TDP43	n.d.	Shell
UBAP2L	Dispensable	
ZC3HG	Dispensable	
NEAT1	Essential	5' + 3' Shell, middle core
IR-containing RNAs	Dispensable	
AG-rich RNAs	Dispensable	Shell

While the paraspeckle were first identified as cluster of proteins, these proteins all shared the ability to bind RNAs, so a function within RNA processing or transport was suggested; moreover, treatment with RNase was able to disrupt the paraspeckle. The first RNA to ever be discovered within the paraspeckle was the Ctn RNA, a mouse-specific, nuclear-enriched, spliced poly(Aþ) transcript (Prasanth et al. 2005). While this mRNA contains all the exon of the CAT2 gene, so is ready for translation once exported, it was soon discovered it was produced under the control of a different promoter and had a distal poly(Aþ) site resulting in a much longer 30 untranslated region (UTR). This UTR was rich in edited adenine to inosine, and as I described before on SFPQ functions, it is believed that it may have the capability to recognize and bind this specific nucleotide. The retention of Ctn is linked to the control of the expression of its gene CAT2 which is a protein part of the nitric oxide response active in case of tissue damage and infections. Under stress stimuli it has been shown how the 30 UTR portion of the Ctn is cleaved off, reducing its nuclear levels with a concomitant increase of mCAT2 levels in the cytoplasm (and a followed-up increase in the CAT2 protein); this has been described as a system to rapidly produce a required protein when needed. While this RNA is mouse specific, analysis in human have found that more than 50% of all genes may produce transcripts with extended 30 UTRs that use a distal poly(Aþ) site (Iseli et al. 2002).

While Ctn was the first RNA discovered in the paraspeckle, as we could see, it was soon noticed how it was only one of the many RNA retained within the structure and it could not in fact be part of the paraspeckle formation: first, because it is mouse-specific and second, because knockdown of Ctn does not disrupt paraspeckles (Prasanth et al. 2005). In 2009, a lncRNA named NEAT1 was shown to be essential for paraspeckle structural integrity and formation (Clemson et al. 2009; Sasaki et al. 2009; Sunwoo et al. 2009).

Two variant transcripts of NEAT1 are transcribed, NEAT1_1 and NEAT1_2 which share approximately 3–4 kb of sequence at the 5' end that precisely delineates NEAT1_1, with the longer isoform, NEAT1_2, containing an additional approximately 20kb of RNA (Wilusz et al. 2008; Sunwoo et al. 2009). Knockdown and overexpression of NEAT1 showed a loss and increase of number in paraspeckle respectively suggesting the role of NEAT1 as the nucleating factor for paraspeckle (Clemson et al. 2009; Sasaki et al. 2009; Sunwoo et al. 2009). Another interesting fact is that NEAT1 is not A-to-I edited, a factor that also change how it interact with the DBHS proteins. An interesting finding was the report of increase levels of NEAT1 in the central nervous system of mice infected with Japanese encephalitis or rabies viruses (Saha et al. 2006) which led to stipulate the hypothesis that certain viruses may use paraspeckles for their processing, or that paraspeckles are part of a viral defence mechanism (Fox & Lamond,

2010). Electron microscopy experiments showed that both NEAT1 isoforms are found at the periphery of the paraspeckles, with the central sequence of the NEAT1_2 (the longest isoform) located at the core (Souquere et al., 2010). This different localization allowed to separate the paraspeckle in three different parts: a core, the patch and shell of paraspeckles (Figure 21) and to give a specific localization to the proteins originally discovered and also to some new members that were found to be indispensable for paraspeckle formation. SFPQ, NONO, PSPC1, and FUS localize at the core in tight association with the central part of NEAT1_2. The patch is composed by RBM14 and BRG1 while the protein TDP43 was found specifically at the periphery forming the shell (West et al.,2016). The localization of RNAs, which are not a constituent part of the paraspeckles has been addressed as well. As it was already being shown by Prasanth et al., RNAs with a longer untranslated region and inverted repeats where directly interacting with NONO and are found in close association with the core proteins of the paraspeckle but spliced mRNAs and spliced introns that both have AG-rich motifs partially colocalized with *Neat1_2* on the surface of paraspeckles (West et al.,2016). The retention of the core RNAs has already been shown to change in response at specific cellular stimuli, and the discovery of possible cellular stimuli to capture and release the AG rich RNA on the surface of the paraspeckle could uncover new interesting functions of these nuclear structures.

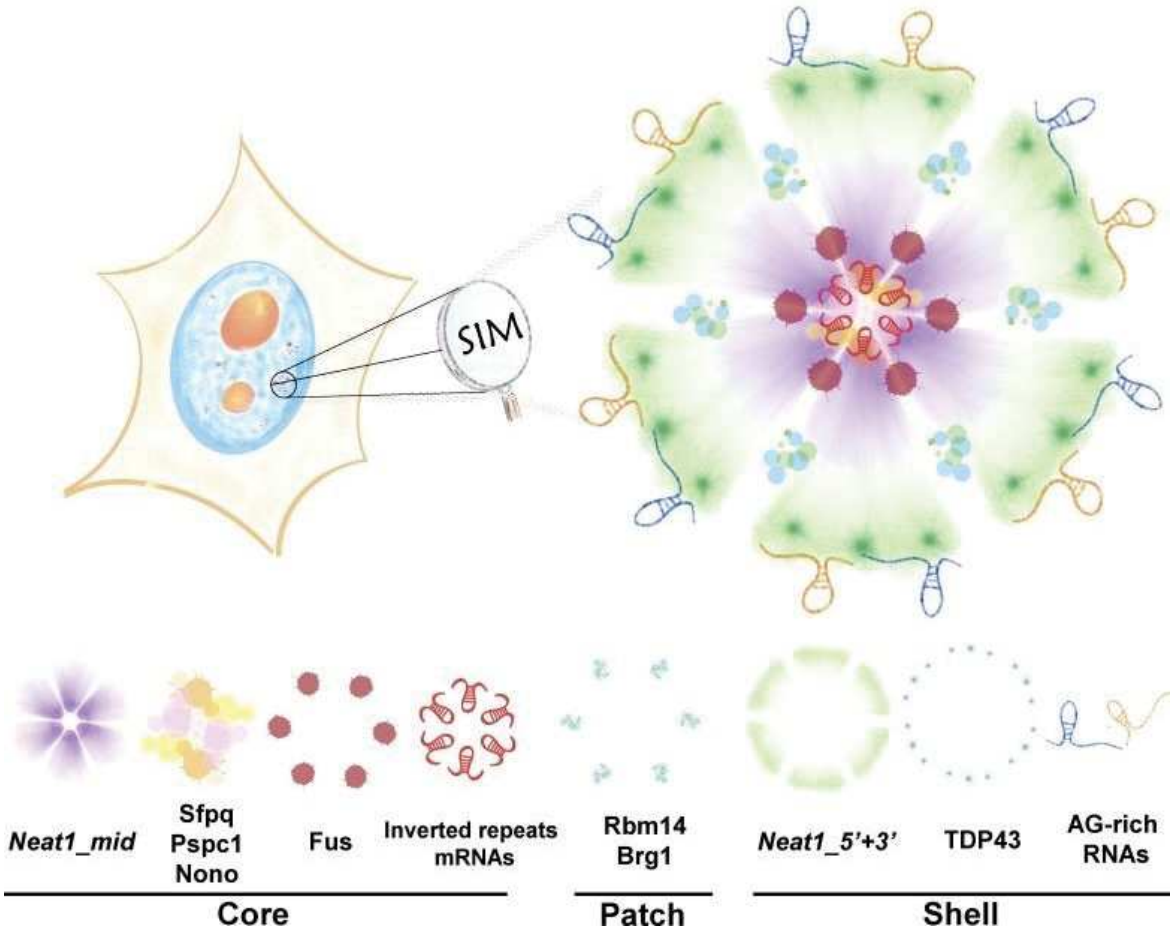


Figure 21: Paraspeckle structure under Structured Illumination Microscopy (SIM). Paraspeckles can be divided in a core part composed by the proteins SFPQ, FUS, PSPC1 and NONO and the central part of the NEAT1_2 lncRNA. The central part or patch of the paraspeckle is composed mostly by the proteins Rbm14 and Brg1 while the shells is occupied by the 5' and 3' end of NEAT1_2, the TDP43 protein. The localization of other RNAs within the paraspeckle are determined by their composition and sequence where AG rich RNAs are enriched on the surface and the RNAs carrying inverted repeats interact with the DHBS core proteins. (Hu et al., 2016)

2. Material and methods

2.1. Cell culture

Our experiments were performed in two main cells lines, HEK 293T and human microglial cells (obtained from M. Tardieu, France) (Janabi et al., 1995). Both cells lines were kept at 37° Celsius in the presence of 5% CO2. The culture medium provided was Dulbecco's Modified Eagle Medium (DMEM) High glucose, containing 10% fetal bovine serum and 1% (100U/mL) of penicillin-streptomycin.

2.2. Plasmids

The pcDNA3 plasmid and the full-length (1-813) human pFLAG-CTIP2 were previously described (Avram et al., 2000) as well the deletions mutants constructs of pFLAG-CTIP2 (Rohr et al., 2003); pRFP-CTIP2 (Marban et al., 2005); The following plasmids were gently provided by the following investigators: GFP-SFPQ by A. Marcello (Kula et al., 2013) plko1 SH SFPQ, plko1 SH CTRL, Laco SH SFPQ, laco SH Luc by A. Kula: Myc-SFPQ and Myc-SFPQ ΔRRM1+2 were purchased from (Addgene).

2.3. Immunoprecipitation assay (IP):

HEK 293T cells were cultured in 10cm dish plates and transfected at 40% confluency with PEI (PolyPlus) for 24h, with the plasmid of interest (pFLAG-CTIP2, Myc-SFPQ, GFP-SFPQ or the pCDNA3 control) to a maximum amount of 20µg of total plasmid used. After 48h the cells were harvested and washed with ice cold PBS and lysed in RIPA (50mM TRIS-HCL pH 7.4, 150mM NaCl, 1% NP40, 0,1% SDS 1,5 mM MgCl2) buffer in the presence of the protease inhibitor (Roche) for 30' on rotation at 4° Celsius. The lysate was separated from the cellular debris by centrifugation at 10000 rpm and 4° Celsius. 500 µg of the lysates were immunoprecipitated by Over Night (ON) incubation on rotation at 4° Celsius with 900µg M2 anti-FLAG beads (SIGMA) per samples or with the previously coated Magnetic beads (SIGMA) with the antibody of interest (5µg of antibody per sample) (GFP, Myc). The beads were then washed 6 times with ice cold RIPA buffer and the immune complexes obtained by this process were recovered from the beads by adding 40 µl of 2% Sample buffer (90% Lamelli

Blue and 10% β-mercaptoethanol), boiled at 95° Celsius for 15' and processed by SDS-Page and Western Blot analysis.

2.4. SDS-PAGE and Western Blot Analysis:

SDS page were performed as the standard technique, SDS PAGE run on acrylamide gels (8% to 12% depending on the target proteins) and transferred on nitrocellulose membrane for 70'. The membrane were blocked in 4% milk-PBS blocking buffer used as well for the preparation and dilution of the primary and secondary antibodies directed against the FLAG epitope (M2 mouse monoclonal, SIGMA), Myc epitope (SantaCruz Biotechnology), GFP (CloneTech), CTIP2 (Bethyl), SFPQ(Bethyl), FUS (SantaCruz Biotechnology), Actin (Sigma).

2.5. Mass spectrometry

HEK cells were transfected with the flag-CTIP2 expressing vector. Nuclear proteins were extracted and subjected to anti-flag immunoprecipitation. Immunoprecipitated proteins were eluted and then digested as described in this previous study (Turriziani et al., 2014). The peptides were separated and analysed by reverse phase liquid chromatography coupled to a high-resolution quantitative mass spectrometer (Q-Exactive connected to an Ultimate Ultra3000 chromatography system (Thermo Scientific, Germany)). Mass spectra were analysed using MaxQuant software. The results are representative of three independent experimental duplicates. The results were obtained in the form of protein intensities calculated by summarizing the ion intensities of the peptides which uniquely correspond to the protein sequence identified in all the samples (Turriziani et al., 2014). The proteins interacting with CTIP2 were determined after normalization with the trypsin used for peptide digestion. Normalization with a control condition represented by cells transfected with the empty vector pcDNA3 has also been carried out.

2.6. RNA-Immunoprecipitation (RIP):

Microglial cells were cultured in 10cm dish, at 80%-90% confluence the cell were lysate with the TNTN Buffer (Tris HCl pH 7.4 20mM, NaCl 20mM, Glycerol 10%, EDTA 0.5 mM, MgCl2 5mM, Triton X 0.05%, Tween 0.1%, DTT 1mM, PMSF added fresh 0.5mM) plus 400U

RNAase OUT (Invitrogen) for 30’ on rotation at 4° Celsius. The lysate was separated from the cellular debris by centrifugation at 10000 rpm at 4° Celsius. The lysates were pushed through a clearing step of 1h at 4° Celsius with 25ul magnetic beads coated with unspecific IgG and then immunoprecipitated by ON incubation on rotation at 4° Celsius with 30µl of magnetic beads (SIGMA) previously coated with 5µg of the antibody of interest or unspecific IgG as negative control for a minimum of 3h at 4° Celsius. The RNA-Protein-Antibody complexes were disrupted by performing an RNA extraction directly on the beads with 500µL TRIzol® (Ambien), incubating 5’ at 37 ° Celsius, follow up by adding 100 µl chloroform, centrifuging at 13000rpm for 15 minutes. After this step, the liquid phase containing the RNA were separated, mixed with isopropanol and Glycoblue (Invitrogen) to facilitate the formation of the RNA pellet. 10’ incubation in ice followed up by 10’ centrifugation at 13000rpm at 4° Celsius. Then Isopropanol was removed, the pellet washed with ethanol 80% and resuspended in 16 µl of DEPC H_20. The full amount is then used to retrotranscribe the RNA in cDNA with the iScriptTM Reverse Transcription Supermix for RT-qPCR (BIO-RAD) in a final volume of 20µl. cDNA is then diluted 5 times and 2uL are used for the qPCR analysis with SYBR Green (BioRad).

2.7. Cross-Linking Immunoprecipitation Sequencing (CLIP-Seq):

HEK 293 T cells were cultured in 10cm dish (2 plate per condition and transfected with the pFLAG-CTIP2 and pCDNA3 as control); microglial cells were cultured in 10cm dish (4 plate per condition). The day prior the cell lysis the magnetic beads were prepared, 3 for each condition, 50µl of protein G Dynabeads in each, one incubated only in lysis buffer, the second one with unspecific IgG and the third one with 5µg of the antibody of interest, αFLAG for the HEK cells, αCTIP2 for Microglia with unspecific IgG as control. Prior lysis the plate and cells were exposed to UV light without the plate lid for the crosslinking step by using the Stratalinker for 10’ for a total of 4000J while standing on a ice tray and covered in ice cold PBS 1X. The cells were then lysed with lysis buffer (Tris HCl 50mM and pH 7.4, CaCl 0.08mM, SDS 0.1%, KCl 100 mM, MgCl2 2mM, NP40 1%, Deoxycholate 5g/L. RNasin 40U (Promega) per ml and Proteinase Inhibitor tablet (ROCHE) added fresh (50x final dilution). 5µl TURBO-DNase 2u/uL (Ambion) was added to the tubes which were incubated 5’ at 37 ° Celsius, 400 rpm. The lysates were centrifugated for 15’ at 13000rpm 4° Celsius. In the meantime the tubes with the beads were washed with lysis buffer. The lysate was added to the tubes containing the beads

alone to eliminate any unspecific binding to the beads and putted in rotation at 4° Celsius for 1h. After this step the lysate was moved to the beads coated with unspecific IgG for another step to remove the unspecific binding following the same previous condition at which follow up the incubation with the beads coated with the antibody of interest O/N at 4° Celsius on rotation. The day after the beads were washed three times with an high salt washing buffer (Tris HCl pH 7.4 50mM, EDTA 1mM, SDS 0.1%, KCl 1M, NP40 1%, Deoxycholate 5g/L) followed up by a Proteinase K treatment, 200µL per tube (Tris HCl pH 7.4 100mM, EDTA 10mM,NaCl 40mM, Proteinase K 2mg/ml (ROCHE) for 20' at 37 ° Celsius. The proteinase K activity was then stopped by adding 200uL PK Urea buffer (Tris HCl pH 7.4 100mM, EDTA 10mM,NaCl 40mM, Urea 7M) at 37 ° Celsius. The RNA extraction step followed up by adding 500uL RNA phenol/chloroform incubated 5' at 37 ° Celsius, 1100 rpm, the tubes were then centrifuged at RT at 13000rpm and the aqueous phase transferred to a new tube. A solution of 3M sodium acetate pH 5.5 and 1µL Glycoblue was added tighter with 1ml of 1:1 Ethanol:Isopropanol, the tubes were mixed and the RNA was left to precipitate ON at -20° Celsius. The tubes were then centrifuged at 13000rpm at 4 ° Celsius for 10 minutes, the RNA pellet was washed twice with 500µL of 80% ethanol and left to dry for 2 minutes with the lid open. The RNA was then dissolved in 10µL DPEC water and sent to sequence.

2.8. Luciferase Assay:

Microglial cells were seeded in 48 well plate 24h before transfection to obtain the following day a confluence of roughly 70% per well. The cells were transfected with the PEI for 48h using the indicated vectors for each experiment and LTR-Luc, HIV-1 NL4-3 ΔEnv Luciferase Reporter Vector (pNL4-3.Luc.R-E-) and a pRenilla vector for normalization of the transfection. Cells were then harvested, and the luciferase activity was determined using the Dual-GloTM Luciferase Assay System (Promega)

2.9. CLIP sequencing and bioinformatics analysis

Analysis of CLIP-Seq data was described in (König et al., 2010). For short, the RNA library was sequenced on Illumina Hiseq 4000 sequencer as Single-Read 50 base reads following Illumina's instructions. FastQC was run to produce base quality. After adapters were cut, the

short reads less than 10 bases were removed, which leads to 1.5 to 3 million unique hits. Reads were mapped onto the Homo sapiens genome (hg19 assembly) using STAR aligner (version 2.5.3a). The genome assembly file were downloaded from GENCODE (https://www.gencodegenes.org/). The annotations for protein-coding genes and TSS (knownGene) were obtained from the University of California, Santa Cruz Genome Browser. To read number of reads per feature, feature Counts tool, Subread package (http://subread.sourceforge.net/) was used. The transcript annotation file was downloaded from BioMart, Ensembl database (https://grch37.ensembl.org/biomart/martview). RNA annotation file was downloaded from (https://rnacentral.org/). The uniquely mapped reads (MAPQ=255) were selected. Four conditions in two cell lines were used, which leads to a cluster of 740 genes. The cluster of reads should be over 7-folds the control. For over-expressed clusters, the significant cluster is two-fold the control.

Second method, we divided the genome into 10kb-long bins, and counted number of features, which should be 5-fold the control. The significant clusters were used for gene ontology and disease ontology. For downstream GO analysis, we used combination of methods, including the GO construction (http://geneontology.org/), DAVID database (https://david.ncifcrf.gov/), and GREAT tool (http://bejerano.stanford.edu/great/). Additionally, we downloaded OBO gene annotation files for proteins and ncRNA from (http://geneontology.org/docs/download-ontology/) in 2018.

2.10. Primers:

All primers have been acquired from Eurofins Genomics

NAME	SEQUENCE
GAPDH Fw	GGACCTGACCTGCCGTCTAGAA
GAPDH Rw	GGGTGTCGCTGTTGAAGTCAGAG
NEAT 1.1 Fw	TTGTTGCAGAGCCCATGAT
NEAT 1.1 Rw	TGAAAACCTTTACCCCAGGA
NEAT 1.2 Fw	GATCTTTTCCACCCCAAGAGTACATAA
NEAT 1.2 Rw	CTCACACAAACACAGATTCCACAAC
7SK Fw	GAGGGCGATCTGGCTGCGACAT

7SK Rw	ACATGGAGCGGTGAGGGAGGAA
MALAT1 Fw	GACGGAGGTTGAGATGAAGC
MALAT1 Rw	ATTCGGGGCTCTGTAGTCCT
MEG3 Fw	GCATTAAGCCCTGACCTTTG
MEG3 Rw	TCCAGTTTGCTAGCAGGTGA
SNORD133 Fw	AGCCCAGTGATGATCACTATTC
SNORD 133 RwS	TACAGCCCTCAGATCTCATAATC
U11 Fw	AAGCACAATGATGAACAATACTGC
U11 Rw	AAGCAATCAGATCTTATCAGTTTGAC

2.11. Statistical Analysis:

Statistical analysis was performed on a minimum of three independent experiments (or otherwise stated). The results were analysed with a T-test in Microsoft Excel regarding three different samples. An ANOVA test was performed regarding more than two data sets, and Grubbs test was utilized to remove aberrant data set. Values of $p < 0,05$, $p < 0,01$ et $p < 0,001$ were considered significant.

3. Objective of the thesis

After almost 40 years HIV-1 remain one of the biggest health problems in the world with almost 40 million people currently living with HIV-1. While in the past years the antiretroviral therapies have been perfectionated to a point which an infected person can look to an almost normal lifespan, we are still lacking a sterilizing cure, able to completely eradicate the virus from the organism. The strategies that are currently under study are tasked to face the formation and persistence of HIV latent reservoirs which are the main reason cART failed to be an effective cure. By studying, and well understanding, the mechanism behind the establishment and persistence of the viral latency, lies the objective to identify and discover new Latency Reversing Agent, which could hinder, amplify, modify and overall control the biological process responsible for the HIV-1 latency.

Our lab has focused much of its attention on the role of the transcription factor CTIP2 in the establishment and persistence of the post-integrative latency in microglial cells, the main viral reservoir of the CNS. Our lab previous works showed how CTIP2 is part of a biological machinery that create a heterochromatin environment on the viral promoter inhibiting the viral gene expression (Marban et al., 2007). We also identify a new association of CTIP2 with a ribonucleoprotein complex formed by the 7SK RNA which hijack the P-TEFb elongation factor in an inactive form, which is necessary for HIV-1 gene expression and the reactivation of the reservoir (Cherrier et al., 2013). In one of the publications here presented we also show how a specific viral protein Vpr is able to target CTIP2 for degradation, reactivating the silenced provirus (Forouzanfar et al., 2017)

Due to the central importance in the establishment and maintenance of the latency exerted by CTIP2 in microglial cells, the main objective of my thesis was an overall double project, with the focus of exploring, listing, and study all the partners of CTIP2. The project was divided in two, one focusing on protein partners of CTIP2 (Quantitative Mass Spectrometry screening), and study their function within the HIV-1 infection, and another focusing the identification of the RNAs interacting with CTIP2 (CLIP-Seq and RNA-IP).

4. Publication 1

HIV-1 Vpr mediates the depletion of the cellular repressor CTIP2 to counteract viral gene silencing

F. Forouzanfar, S. Ali, C. Wallet, M. De Rovere, C. Ducloy, H. El Mekdad, El Maassarani, A. Aït-Ammar, J. Van Assche, E. Boutant, F. Daouad, F. Margottin-Goguet, C. Moog, C. Van Lint, C. Schwartz & O. Rohr

Scientific Reports **volume 9**, Article number: 13154 (2019)

Constant evolution is the weapon that both host and viruses use to inhibit each other for their own ability to survive. As an organism develop a new system of antiviral factors, viruses respond even more rapidly with the evolution of new factors aimed at hindering these host responses. Our group has extensively showed how CTIP2 contribute to HIV-1 gene silencing and, in this paper, we demonstrated a viral response to block and revert the CTIP2 induced silencing. After a rising in CTIP2 quantity at the beginning of HIV-1 infection, a suddenly decrease of the protein levels, both in Jurkat cells and Microglial, made us suspect a viral ability to target this specific factor. While mRNA level of CTIP2 were not altered, the protein was targeted for degradation. Studies with ΔVpr mutant HIV-1 constructs showed a loss in this CTIP2 degradation, and over expression of Vpr in latently infected cells showed a reactivation similar to cell treated with TNFα. Moreover, cells treated with the proteasome inhibitor MG132 showed a less degradation of CTIP2 upon Vpr overexpression and immunoprecipitation assay showed an interaction of CTIP2 with the proteasome pathway upon Vpr overexpression.

In this paper we have proved how the viral protein Vpr was capable of targeting CTIP2 for degradation through the proteasome pathway.

My main contribution for this paper was to perform a major part of the experiments required during the reviewing process.

SCIENTIFIC REPORTS
nature research

OPEN

HIV-1 Vpr mediates the depletion of the cellular repressor CTIP2 to counteract viral gene silencing

Received: 11 February 2019
Accepted: 7 August 2019
Published online: 11 September 2019

F. Forouzanfar[1], S. Ali[1,2], C. Wallet[1], M. De Rovere[1], C. Ducloy[3], H. El Mekdad[1], M. El Maassarani[1], A. Aït-Ammar[1,8], J. Van Assche[1], E. Boutant[4], F. Daouad[1], F. Margottin-Goguet[5,6,7], C. Moog[3], C. Van Lint[8], C. Schwartz[1] & O. Rohr[1]

Mammals have evolved many antiviral factors impacting different steps of the viral life cycle. Associated with chromatin-modifying enzymes, the cellular cofactor CTIP2 contributes to HIV-1 gene silencing in latently infected reservoirs that constitute the major block toward an HIV cure. We report, for the first time, that the virus has developed a strategy to overcome this major transcriptional block. Productive HIV-1 infection results in a Vpr-mediated depletion of CTIP2 in microglial cells and CD4+ T cells, two of the major viral reservoirs. Associated to the Cul4A-DDB1-DCAF1 ubiquitin ligase complex, Vpr promotes CTIP2 degradation via the proteasome pathway in the nuclei of target cells and notably at the latent HIV-1 promoter. Importantly, Vpr targets CTIP2 associated with heterochromatin-promoting enzymes dedicated to HIV-1 gene silencing. Thereby, Vpr reactivates HIV-1 expression in a microglial model of HIV-1 latency. Altogether our results suggest that HIV-1 Vpr mediates the depletion of the cellular repressor CTIP2 to counteract viral gene silencing.

Jawed vertebrates including mammals have developed innate and acquired immune responses to fight infections. Mammals have evolved many antiviral factors such as SAMHD1 or APOBEC3G that limit virus expression[1–3]. We have reported that the cellular cofactor CTIP2 (Bcl11b) favors the establishment and the persistence of HIV-1 post-integration latency in microglial cells, the main target of the virus in the brain[4–6]. CTIP2 works as a scaffold protein to recruit at least two different complexes in microglial cells. As part of a chromatin remodelling complex CTIP2 is associated with the lysine demethylase LSD1, the histone deacetylases HDAC1 and HDAC2, and the histone methyltransferase SUV39H1[4,7–9]. In CD4+ T cells, CTIP2 associates with the NuRD chromatin remodelling complex[10]. In addition to this epigenetic silencing, CTIP2 contributes to control the elongation process of gene transcription. Indeed, we found CTIP2 associated with an inactive P-TEFb complex including HEXIM1, HMGA1 and the 7SK snRNA[11,12]. As part of this new complex, CTIP2 represses P-TEFb functions and thereby contributes to limit the reactivation of latently integrated HIV-1 proviruses. Altogether, CTIP2 favors the establishment and the persistence of HIV-1 gene latency in microglial cells, the major reservoirs of virus in the central nervous system.

Lentiviruses such as SIV and HIV have developed several mechanisms to evade innate and adaptive immune responses. "Accessory" proteins (Nef, Vif, Vpu, Vpx) are encoded by these viruses to allow immune evasion. For example, Vpx is involved in the degradation of SAMHD1 via the proteasome pathway (reviewed in[13]). Proteasome-mediated degradation is a privileged pathway subverted by HIV-1 to counteract cellular defenses. Associated with a Cul4A-based ubiquitin ligase complex through DCAF1 (VprBP), the HIV-1 accessory protein Vpr has long been hypothesized to target an as-yet undiscovered host restriction factor for degradation in order to induce G2 arrest[14,15]. Vpr has been proposed to promote the degradation of cellular factors including UNG2[16], Dicer[17], IRF3[18], ZIP[19], the HTLF DNA translocase[20] and the post-replication DNA repair factor Exo1[21]. By

[1]Université de Strasbourg, EA7292, FMTS, IUT Louis Pasteur, Schiltigheim, France. [2]Institute of Microbiology, University of Agriculture, Faisalabad, Pakistan. [3]INSERM U1109, Fédération de Médecine Translationnelle (FMTS), Université de Strasbourg, Strasbourg, France. [4]Université de Strasbourg, UMR 7021 CNRS, Illkirch, France. [5]Inserm, U1016, Institut Cochin, 22 rue Méchain, 75014, Paris, France. [6]CNRS, UMR8104, Paris, France. [7]Université Paris Descartes, Sorbonne Paris Cité, Paris, France. [8]Université Libre de Bruxelles (ULB), Service of Molecular Virology, Department of Molecular Biology (DBM), Gosselies, Belgium. Correspondence and requests for materials should be addressed to C.V.L. (email: cvlint@ulb-ac.be) or C.S. (email: christian.schwartz@unistra.fr) or O.R. (email: olivier.rohr@unistra.fr)

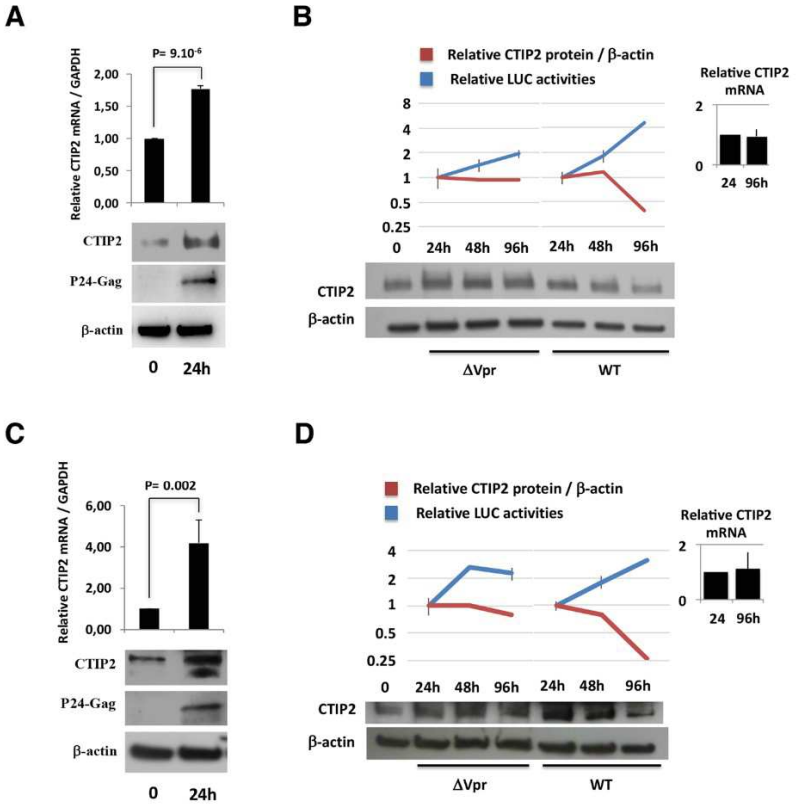


Figure 1. CTIP2-induced expression in infected cells is counteracted by HIV-1 in a Vpr dependent manner. (**A**,**C**) Jurkat T cells (**A**) and Microglial cells (**C**) were infected with 25×10^3 copies/ml of a VSV-G pseudotyped NL4.3ΔEnv-LUC provirus for 24 h. The presence of CTIP2 in infected cells was assessed by western blot and the CTIP2 mRNA quantified by RT-Q-PCR. (**B**,**D**) Jurkat T cells (**B**) and Microglial cells (**D**) were infected with 150×10^3 copies/ml of a VSV-G pseudotyped NL4.3ΔEnv-LUC (WT) and a NL4.3ΔEnvΔVpr -LUC (ΔVpr) provirus for 24, 48 and 96 h. The presence of CTIP2 in infected cells was assessed by western blot and the viral expression was quantified by luciferase assay. The quantifications are presented relative to the quantities obtained after 24 h of infection. The results are representative of at least three independent experiments.

activation of the SLX4 complex, Vpr prevents an appropriate innate immune response[22]. Recently, Vpr has been shown to repress HIV-1 expression by targeting APOBEC3G[23] and an as-yet unidentified macrophage restriction factor[24]. Finally, two recent papers described that the human silencing hub complex (HUSH) is also targeted to degradation by Vpr and HIV-2/SIV Vpx[25,26]. Interestingly, the HUSH complex is presented by the authors as a potent restriction factor regulating the transcription step of the viral life cycle.

Here we show that HIV-1 developed a strategy to counteract CTIP2-mediated repression. We report that the viral accessory protein Vpr promotes the degradation of CTIP2 through physical and functional interactions with the Cul4A-DDB1-DCAF1 complex. Vpr targets CTIP2 associated with the heterochromatin-modifying complex. Interestingly, Vpr expression also correlated with a reduced binding of CTIP2 at the HIV-1 promoter and a reactivation of the latent virus in the CHME5-HIV microglial model of latency.

Results

Hiv-1 Expression promotes a vpr-dependent degradation of ctip2. To investigate a possible modulation of CTIP2 expression in response to HIV-1 infection, we measured its protein and its RNA levels in time course experiments in Jurkat T-cells and microglial cells.

Following an initial induction of both CTIP2 mRNA and protein expression after the first 24 h post HIV-1-infection (Fig. 1, panels A and C), CTIP2 protein expression levels decreased in Jurkat T cells (Fig. 1B WT panel) and in microglial cells (Fig. 1D, WT panel) at later time points. Interestingly, while levels of CTIP2 mRNA were not affected along the time course (from 24 h to 96 h of infection), deletion of the Vpr gene (Vpr-deleted virus - ΔVpr) abrogated CTIP2 depletion upon viral expression (Fig. 1B,D, ΔVpr Panels). To confirm that the depletion of CTIP2 mediated by Vpr occurs at a post-transcriptional step, we next overexpressed a FLAG tagged CTIP2 in the presence of normalized quantities of WT and ΔVpr proviruses. The WT-, but not the ΔVpr provirus, promoted a depletion of CTIP2 expressed under the control of a CMV promoter (Figure S1). These results

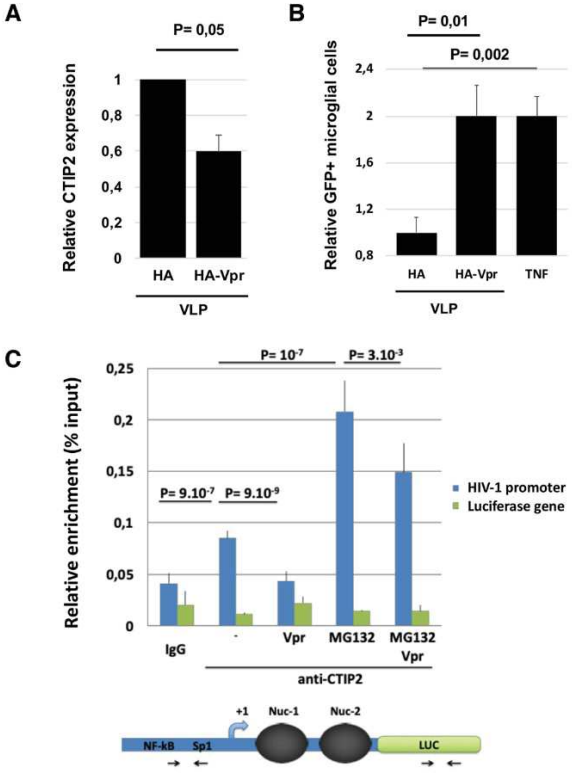


Figure 2. Vpr promotes depletion of CTIP2 in primary CD4+ T cells and at the HIV-1 promoter to induce viral reactivation in microglial cells. (**A**) Relative CTIP2 expression in CD4 T cells was quantified by flow cytometry. Expression levels in cells incubated with Virus Like-Particles VLP-HA-Vpr were compared to the control VLP-HA particles taken as 1. (**B**) CHME5-HIV latently infected microglial cell line was infected with VPLs or treated by TNFα. The number of GFP+ cells was quantified by flow cytometry and presented relative to GFP+ cells obtained after VLP-HA infections. (**C**) Chromatin Immuno-precipitation experiments targeting endogenous CTIP2 at the HIV-1 promoter were performed with chromatin from microglial cells expressing the chromatinized LTR-LUC episomal vector and Vpr as indicated. The cells were subjected or not to MG132 treatment. Enrichments are presented as percentages of the inputs. As a control, the presence of CTIP2 has been concomitantly quantified at the luciferase gene region. Results are representative of at least three independent experiments.

confirmed that CTIP2 expression was impacted post-transcriptionally. Vpr molecules are known to be incorporated into the virions[27]. To determine if incoming Vpr proteins are sufficient to induce a depletion of CTIP2 in primary cells, we next generated virus-like particles (VLPs) to deliver Vpr in purified CD4+ T cells. As shown in Fig. 2A, incoming Vpr induced a significant depletion of CTIP2 in primary CD4+ T cells, suggesting that Vpr produced in cells upon HIV-1 replication and Vpr delivered by the virions upon infection reduce CTIP2 expression. Since CTIP2 has been described to limit the reactivation of latently integrated viruses in microglial cells[9], we next investigated the impact of incoming Vpr on the CHME5-HIV microglial model of HIV-1 latency. Interestingly, Vpr delivery increased by two-fold the number of HIV-1 expressing cells (Fig. 2B). This magnitude of reactivation was similar to that obtained with TNFα treatment (Fig. 2B column HA-Vpr vs TNFα). As we previously reported that CTIP2 associates with the latent HIV-1 promoter in microglial cells, we next performed chromatin IP experiments targeting CTIP2 in the presence or not of Vpr (Fig. 2C). As expected, CTIP2 was found associated with the viral promoter but not with the luciferase gene of the construct in microglial cells (Fig. 2C: IgG vs (-) mock CTIP2 IP column). However, Vpr expression significantly reduced the binding of CTIP2 (Fig. 2C: Vpr vs (-) mock), in line with the previously observed depletion of the protein in nuclear extracts. Vpr has been shown to interact directly with the 26 S proteasome and to degrade target cellular factors via the proteasome pathway[28]. We indeed found CTIP2 associated with the 19S subunit of the proteasome but not Vpr (Figure S5). To further determine the mechanism underlying CTIP2 depletion, we next treated the cells with the proteasome inhibitor MG132. As shown, MG132 treatment favored CTIP2 binding to the viral promoter (MG132 vs (-) mock) and strongly impaired the impact of Vpr (MG132+ Vpr vs MG132). Altogether, these results suggest that Vpr produced in cells upon HIV-1 replication and Vpr particles incorporated in virions both contribute to the depletion of CTIP2 in cell lines and primary cells. This depletion correlates with reduced binding of CTIP2

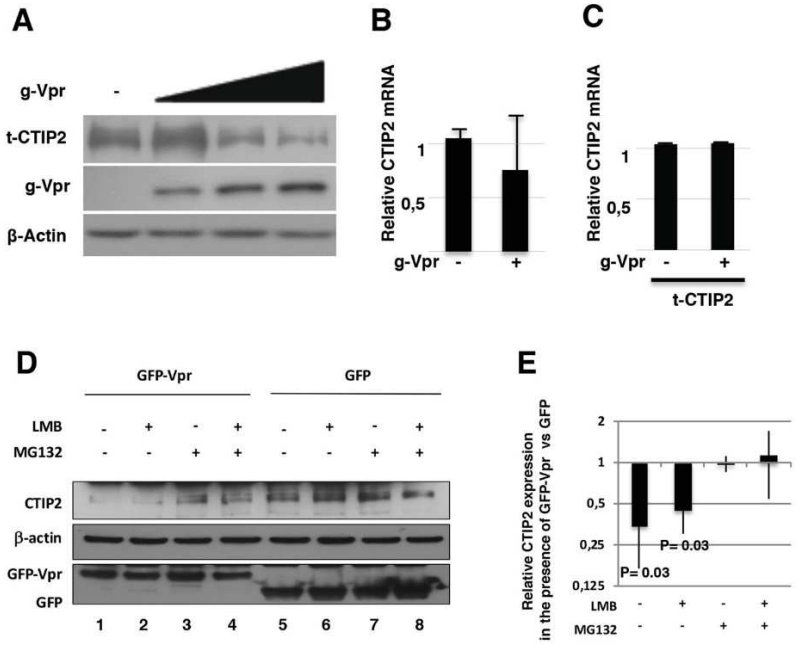


Figure 3. Vpr targets nuclear CTIP2 to degradation via the proteasome pathway. (**A**) Nuclear extracts from HEK cells expressing Tap-CTIP2 (t-CTIP2) and increasing amount of GFP-Vpr (g-Vpr) were analyzed by western blot for the presence of the proteins indicated. (**B**,**C**) CTIP2 mRNA were quantified by RT-Q-PCR in the presence or not of Vpr in the absence (**B**) and in the presence (**C**) of CTIP2 overexpression. The results are presented relative to the pCDNA3 control vector taken as one. (**D**) Endogenous CTIP2 expression was analyzed by western blot in extracts from cells expressing GFP-Vpr or the control GFP and subjected to the indicated treatments. Cells were treated with 10 ng/ml leptomycin B for 4 h and then with MG132 (5 μg/μl) for 6 h before lysis and nuclear protein extraction. Protein expression was assessed by western blot. (**E**) Expression of CTIP2 in the presence of GFP-Vpr was quantified relative to β-actin and presented relative to the quantities obtained in the presence of the GFP control protein[57]. Results are representative of at least three independent experiments.

at the viral promoter and reactivation of the latent HIV-1. In addition, the impact of MG132 on Vpr-mediated depletion of CTIP2 and the interaction of CTIP2 with the 19S subunit of the proteasome suggest an involvement of the proteasome pathway and a degradation of the protein rather than translational regulation. To further investigate Vpr-induced depletion of CTIP2 biochemically, we next expressed increasing amounts of Vpr in CTIP2 expressing cells. As expected, Vpr expression promoted a dose-dependent depletion of CTIP2 (Fig. 3A). Of note, levels of CTIP2 mRNA from control and from CTIP2-overexpressing cells were not significantly affected by Vpr (Fig. 3B,C, respectively) confirming the post-transcriptional impact of Vpr on CTIP2 expression. To further investigate the involvement of the proteasome pathway, microglial cells were treated with the proteasome inhibitor MG132. In accordance with the ChIP results presented in Fig. 2C, MG132 increased the level of endogenous CTIP2 expression in nuclear extracts (Fig. 3D, row 7 vs 5), suggesting that the physiological turnover of CTIP2 needs a functional proteasome pathway. Again, Vpr reduced CTIP2 expression (Fig. 3D, column 5 vs 1), and MG132 impaired Vpr-mediated depletion of CTIP2 (Fig. 3D, column 1 vs 3). These observations strongly advocate for a degradation of CTIP2 by the proteasome pathway. To determine if this degradation occurs in the nucleus, cells were treated with leptomycin B (LMB), an inhibitor of nuclear export. As shown, LMB treatment did not significantly affect the level of CTIP2 degradation by Vpr (Fig. 3D, column 2 vs 1). We next took advantage of the replicates of these experiments to quantify the impact of Vpr on CTIP2 expression levels in the presence of LMB and MG132. As shown in Fig. 3E, Vpr promoted a more than two-fold depletion of CTIP2 in microglial cells nuclei (first column). However, MG132 abolished the Vpr-mediated depletion of CTIP2 in the presence or absence of LMB (two last columns). Altogether, our results suggest that Vpr degrades CTIP2 via the proteasome pathway in the nuclei of HIV-1 target cells.

Ctip2 Interacts with the cull4/ddb1-dcaf1 ubiquitin ligase complex. Vpr has been reported to target cellular proteins for degradation through recruitment of the Cul4A-DDB1-DCAF1 complex. Since inhibiting the proteasome pathway increased CTIP2 expression levels in the absence of Vpr, we first investigated its possible interaction with the ubiquitin ligase complex. Confocal microscopy observations were performed on microglial cells expressing RFP-CTIP2, GFP-DCAF1, and HA-Vpr alone (Fig. 4A) or in combination (Fig. 4B). As previously described, CTIP2 is localized in dense- and structured-nuclear domains[29]. The expression of HA-Vpr and GFP-DCAF1 was mainly observed in the nuclear compartment. However, the combined expression of

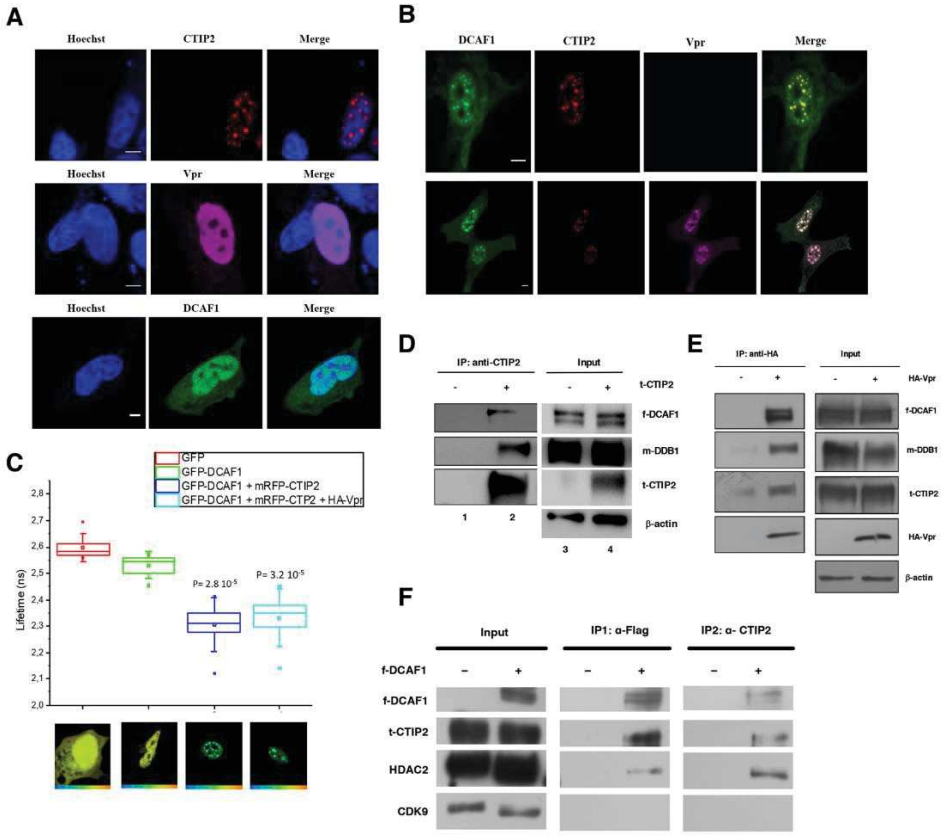


Figure 4. CTIP2 interacts with the Cull4-DDB1-DCAF1 ubiquitin ligase complex. (**A**,**B**) Microglial cells expressing RFP-CTIP2, HA-Vpr and GFP-DCAF1 alone (panel A) or in combination (panel B) were subjected to confocal microscopy. Scale bars are set to 5 µm. (**C**) Lifetimes from GFP proteins FRET experiments are presented for each condition. The results are representative of 5 independent experiments and from 6 to 10 quantifications for each condition per experiment. (**D**,**E**) Nuclear extracts from cells expressing the proteins indicated were subjected to immunoprecipitation experiments targeting Tap-CTIP2 (t-CTIP2) (**D**) and HA-Vpr (**E**). The inputs and the immunoprecipitated complexes were analyzed by western blot. (**F**) Nuclear extracts from cells expressing the proteins indicated were subjected to a first immunoprecipitation targeting FLAG-DCAF1 (f-DCAF) with anti-FLAG antibodies. The DCAF1 associated complexes were eluted by incubation with FLAG peptides and subjected to a second immuno-precipitation targeting CTIP2. The presence of the indicated proteins was analyzed at each step by western blot.

RFP-CTIP2 with GFP-DCAF1 promoted a relocation of DCAF1 proteins to CTIP2-induced structures (Fig. 4B, upper panel). The yellow staining of the CTIP2-induced structures results from this colocation. Interestingly, Vpr expression further turned the staining of CTIP2 structures white, suggesting a triple colocation of Vpr, DCAF1 and CTIP2 (Fig. 4B, lower panel). We have previously reported that CTIP2 associates with Vpr in CTIP2-induced nuclear structures[7]. To determine whether CTIP2 interacts directly with DCAF1, we next performed fluorescence resonance energy transfer (FRET) experiments. As shown in Fig. 4C, the presence of RFP-CTIP2 reduced significantly the lifetime of GFP-DCAF1 in the nuclear structures, indicating a direct interaction between these two proteins. Surprisingly, the presence of Vpr did not significantly impact on this effect. To confirm the physical interaction between DCAF1 and CTIP2 biochemically, we performed co-immuno-precipitation experiments targeting first CTIP2. DCAF1 and DDB1 were found in association with CTIP2 (Fig. 4D). These associations are in accordance with the increased CTIP2 expression observed following MG132 treatment. In addition, we found Vpr associated with CTIP2 and the DDB1-DCAF1 complex in nuclear extracts from MG132-treated cells (Fig. 4E). These results suggest that Vpr takes advantage of the cellular degradation machinery to deplete CTIP2 from HIV-1 infected cells and favor HIV-1 expression.

We have reported that CTIP2 associates with two different and exclusive complexes: A complex dedicated to epigenetic silencing of gene expression (chromatin modifying enzymes)[9] and an inactive form of the P-TEFb elongation complex including the 7SK snRNA[11]. To characterize CTIP2-associated DCAF1 containing complexes, we performed sequential immunoprecipitation (IP) experiments (Fig. 4F). In the first IP that targeted DCAF1-associated complexes, we found CTIP2 and HDAC2 (chromatin-modifying enzyme) but not CDK9,

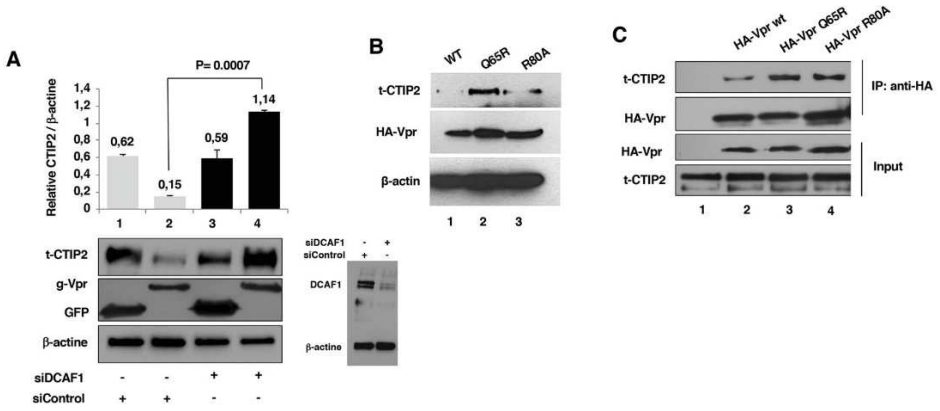


Figure 5. DCAF1 binding is necessary for Vpr-mediated degradation of CTIP2. (**A**) Nuclear extracts from control or DCAF1 knock-down cells were analyzed by western blot for the presence of the indicated proteins in the presence of GFP-Vpr and the control GFP. CTIP2 expression was quantified relative to the loading control using ImageJ software. DCAF1 knock-down efficiency in our experiments is presented. (**B**) Nuclear extracts from cells expressing the indicated proteins were analyzed by western blot. β-actin is presented as a loading control. (**C**) Nuclear extracts from cells expressing large amount of Tap-CTIP2 (t-CTIP2) and the indicated HA-Vpr proteins were subjected to immunoprecipitation experiments targeting the HA tag. The input and the Vpr-associated complexes were analyzed for the presence of CTIP2 by western blot.

the kinase subunit of the P-TEFb complex. This result suggests that DCAF1 may interact with CTIP2 and its chromatin-modifying complex. To confirm that CTIP2, HDAC2 and DCAF1 are part of the same complex, proteins were eluted from the first IP and submitted to a second IP targeting CTIP2. DCAF1 and HDAC2 were found associated with CTIP2 in the eluted DCAF1-containing complex, suggesting that CTIP2, DCAF1 and HDAC2 are members of the same complex (Fig. 4F). Altogether, these results strongly suggest that DCAF1 interacts with and degrades CTIP2 bound to the chromatin modifying enzyme complex (HDAC1/2, SUV39H1) but not with CTIP2 bound to the inactive P-TEFb complex. This conclusion is in accordance with the degradation of HDAC1 by Vpr upon HIV-1 infection (Figure S3 and[30]).

Association with dcaf1 is crucial for vpr-mediated degradation of ctip2. To further characterize mechanistically the degradation of CTIP2 by Vpr, we explored whether DCAF1 is needed for the Vpr-mediated degradation process by using a Si RNA-based knock-down approach. As expected, depletion of DCAF1 abrogated the degradation of CTIP2 (Fig. 5A column 4 vs 2). We next quantified CTIP2 expression levels in the presence of WT or mutated Vpr proteins (Fig. 5B). The mutations Q65R (which abrogates DCAF1 binding and G2 arrest) and R80A (supports efficient DCAF1 binding but abolishes G2 arrest) were tested[14,31,32]. As expected, CTIP2 was barely detectable in the presence of WT and R80A Vpr proteins (Fig. 5B columns 1 and 3). However, the Q65R mutation, abolished Vpr-mediated degradation of CTIP2 (Fig. 5B column 2). These results demonstrated that Vpr-mediated G2 arrest does not impact CTIP2 expression and confirmed the requirement for DCAF1 for the depletion of CTIP2 by Vpr. Interestingly, upon a strong overexpression of CTIP2, allowing its detection in the presence of Vpr, we were able to find WT, R80A and Q65R Vpr proteins associated with CTIP2 (Fig. 5C). Thereby, our results suggest that recruitment of DCAF1 is crucial to ensure the Vpr-mediated degradation of CTIP2 but dispensable for the physical interactions between the proteins.

Discussion

CTIP2/Bcl11B is involved both in the establishment and the persistence of HIV-1 latency in CD4+ T cells and microglial cells[8,9,11,33]. By silencing integrated HIV-1 proviruses and disfavoring its reactivation, this cellular cofactor constitutes a major block to viral expression. To prevent abortive infections, HIV-1 has evolved sophisticated mechanisms to evade innate and adaptive immune responses. Accessory proteins such as Nef, Vif, Vpu, Vpx and Vpr are dedicated to some aspects of this immune evasion. Indeed, several restriction factors are targeted to degradation by these accessory proteins. Vif targets APOBEC3G[23], Vpu targets tetherin by sequestration from the site of budding and proteasomal degradation[34,35] and Vpx targets SAMHD1[2]. Among the so-called accessory proteins which have to be rather considered as key virulence factors, Vpr remained enigmatic. Vpr is a small protein (around 14 kDa) which is involved in several biological functions including nuclear import of the pre-integration complex[36]; envelope protein expression and virion production[24]; HIV-1 transcription[37]; post-replication DNA repair[21], HIV-1 gene splicing[38]; induction of apoptosis[39] and cell cycle arrest[22] (reviewed in[40]). Vpr is produced upon HIV-1 replication and a significant amount (275 molecules) is incorporated in the new virion[27]. We report here that Vpr expressed in infected cells (Fig. 1) and incoming-Vpr molecules (Fig. 2) favor the degradation of CTIP2 as previously described for the translocase HTLF[20]. Interestingly, incoming Vpr is necessary to overcome the HIV-1 transcriptional block observed in dendritic cells[41]. The influence of CTIP2 in dendritic cells has not yet been studied. However, our results suggest that Vpr contributes to unlock CTIP2-mediated blocks of HIV-1 gene expression. Vpr has been described associated with the E3 ubiquitin ligase complex to promote a

proteasome-mediated degradation of cellular targets (for review[15]). Here we report that CTIP2 interacts with this ubiquitin ligase complex via DCAF1/VprBP. Of note, DCAF1 was reported to interact and to protect Vpr from proteasomal degradation[42]. We found Vpr, CTIP2 and DCAF1 co-localized in dense subnuclear structures, and FRET studies demonstrated a direct association between CTIP2 and DCAF1. These observations suggested to us that Vpr may subvert the proteasome pathway to promote CTIP2 degradation. Interestingly, we showed that CTIP2 protein expression is higher in cells treated with MG132 suggesting a physiological regulation of CTIP2 through the proteasome pathway. Such a CTIP2 protein degradation has been described in a physiological condition such as TCR activation[43,44]. Therefore, Vpr promotes a constitutive DCAF1-dependent CTIP2 turnover. Such a mechanism of action for Vpr has already been described for UNG2[16] and discussed elsewhere[15]. In agreement with this, we show that Vpr enhances the ubiquitination of CTIP2 (Figure S2). An increase of the basal ubiquitination has already been described for HDAC1 and HDAC3, proteins associated with the formation of heterochromatin, the compact-inactive form of the promoters[30,45]. We further report that Vpr promotes CTIP2 degradation in the nucleus, at the HIV-1 promoter. This observation strongly argues for an indirect impact of Vpr on HIV-1 gene transcription. As a consequence of CTIP2 degradation, Vpr contributes to overcome the transcriptional silencing induced by CTIP2-associated complexes, and to favors Tat function (Figure S4). Interestingly, we found CTIP2, HDAC2 and DCAF1 in the same complex. Since P-TEFb is excluded from this complex, our results suggest that Vpr promotes the degradation of CTIP2 associated with HDAC1/2 and SUV39H1. The concomitant degradation of CTIP2 and HDAC1 (Figure S3) and the recent demonstration that Vpr targets HDAC1 to a proteasome-mediated degradation further support these conclusions[30]. Moreover, results of a recent screen for factors targeted by the virulence factors Vpr and HIV-2/SIV Vpx strengthen our results[25,26]. This screen identified the human silencing hub (HUSH) complex as a new target of Vpr[46,47]. By the degradation of CTIP2 and its associated heterochromatin-promoting enzymes, Vpr unlocks an epigenetic block and activates HIV-1 gene transcription. As commented in[48], our results suggest a link between the intrinsic immunity and the epigenetic control of HIV-1 gene transcription, which might lead to new ways to combat HIV-1.

There are now several pieces of evidence that Vpr is a key virulence factor. It has been shown to disturb several cellular pathways including cell cycle arrest (reviewed in[40]) and DNA repair (reviewed in[49]). The involvement of Vpr in the latter pathway has been associated with the viral escape from innate immune sensing[22]. We and others reported that Vpr is involved in the epigenetic regulation of HIV-1 gene transcription. Vpr promotes the degradation of members of chromatin remodeling complexes such as CTIP2, HDAC1[30], ZIP and sZIP, the adaptors of the NuRD complex[19]. Interestingly CTIP2 recruits NuRD to silence HIV-1 gene transcription in CD4+ T cells[10]. Recently, we have shown that PKC-mediated phosphorylation on Ser2 of CTIP2 abrogates the recruitment of NuRD upon activation via the TCR[50]. Post-translational modifications drive the association of CTIP2 with different complexes and thereby determine CTIP2 functions. Whether or not Vpr is able to discriminate between the diverse forms of CTIP2 will require further investigation.

The recognition of Vpr functions in the viral escape from the immune responses in general and in preventing HIV-1 latency, as described in the present article, may have profound therapeutic implications[51]. We have recently reviewed the possibilities of targeting post-replication DNA repair complexes to prevent HIV-1 latency[52]. Taking advantage of Vpr-induced depletions of key cellular factors may be an interesting option, as recently described for the post replication DNA repair protein Exo1[21].

Materials and Methods

Plasmids. Most of the constructs used in our assays have been described previously: pcDNA3, pNTAP, pFLAG-CTIP2 (f-CTIP2), pRFP-CTIP2, pNTAP-CTIP2 (t-CTIP2) were used[8,9,11]. pGFP-Vpr WT (g-Vpr), pNL-4.3 ΔEnv-luc wt, pNL-4.3 ΔEnv-luc ΔVpr[19] and pMyc-DDB1 (m-DDB1), pFLAG-DCAF1 (f-DCAF1), pHA- Vpr wt, pHA-Vpr Q65R were kindly provided by Dr. F. Margottin-Goguet.

Cell culture. The human microglial (provided by M. Tardieu, Paris, France)[53], HEK293T and CHME5-HIV cell lines were maintained in Dulbecco's modified Eagle medium (DMEM) and Jurkat cell line in RPMI 1640. The culture mediums were completed with 10% (vol/vol) heat-inactivated fetal bovine serum (FBS) and 100 U/mL penicillin/streptomycin. Where indicated, the cells were treated with MG132 (5 μg/μl) for 6 h, 10.0 ng/ml leptomycin B (LMB) for 5 h before harvesting the cells.

Antibodies and reagents. CTIP2 (ab18465) antibodies were purchased from Abcam, BCL11b (A300–384A) from Bethyl Laboratories, CDK9 (sc-8338), Myc epitope (9 E10), mouse and rat HRP-conjugated from Santa Cruz Biotechnology, anti-FLAG M2 affinity gel and β-actin (A5441) from Sigma-Aldrich, and anti-HA.11 Tag Antibody (AFC-101P) was from Eurogentec. Rabbit HRP-conjugated antibodies were purchased from Thermo Fisher Scientific, anti-GFP (632592) from Clontech, and Protein G Magnetic beads from Millipore. MG132 and leptomycin B molecules were provided by Sigma-Aldrich, and the protease inhibitor mixture by Roche. VprBP (DCAF1) and non-targeting siRNAs were from Dharmacon. SYBR Select Master Mix used for Q-PCR was from Thermo Fisher. CTIP2-FITC (25B6) ab123449 and IgG control were purchased from Abcam. Anti-HA.11-AF647 was used for the confocal microscopy observations.

HIV-1 infections and viral production in cell lines. The VSV-G pseudotyped ΔEnv-LUC HIV-1 viruses were produced by transfecting the pNL4.3 ΔEnv-LUC HIV-1 (WT) or pNL4.3 ΔEnvΔVpr-LUC HIV-1 (ΔVpr) constructs together with the VSV-G expression plasmid in 293 T cells. 48 h after transfection, the culture supernatants were collected and viral production was measured by clinical HIV-1 viral load test (quantification of the genomic RNAs).

CTIP2 expression in primary CD4+ T cells infected with VLP. Primary CD4+ T cells were isolated and purified from human peripheral blood mononuclear cells (PBMC) of healthy blood donors by positive

selection using CD4 MicroBeads selection kits and autoMACS Pro Separator (Miltenyi Biotec) as previously described[54,55]. $CD4^+$ T lymphocytes were activated with phytohemaglutinin A (PHA) (2 µg/ml) (Sigma-Aldrich) for 3 days in RPMI 1640 plus Gibco GlutaMAX medium (Thermo Fisher) supplemented with 10% fetal calf serum (FCS). Cells were stored frozen. They were thawed the night before use. CD4 T cells were incubated with Viral-like particles VLP-HA or VLP-HA-Vpr (obtained from FMG) at 0.1 µg for 0.1 million cells for 1 days. For the quantification of intracellular CTIP2 expression, cells were incubated for 10 min at 4 °C with CD3 cell surface molecules before fixation with the Cytofix solution (BD Biosciences) and permeabilized with a solution of 0.1% Triton X-100. CTIP2 or IgG control isotype were incubated for 30 min at 4 °C. After extensive washing, the percentage of CD3-PEC T cells positive for CTIP2 expression was defined by flow cytometry (MACSQuant, Miltenyi) and analyzed with Kaluza software.

Co-immunoprecipitation assays and western blotting. Cells cultured in 150-mm diameter dishes were transfected using a calcium phosphate co-precipitation method with the plasmids indicated. Lysates were prepared two days post-transfection as previously described[11]. Immunoprecipitations were performed using the standard technique, as in[11]. Finally, the immunoprecipitated complexes were processed for SDS-PAGE and immunoblot analysis. Proteins were visualized by chemiluminescence using the SuperSignal Chemiluminescence Detection System (Thermo Fisher).

Chromatin immunoprecipitation (ChIP). Cells expressing the chromatinized HIV-1 episomal LTR-luciferase construct were subjected to ChIP experiments with IgG control or anti-CTIP2 IgG antibodies, as in[8,9]. The presence of CTIP2 at the promoter- or the luciferase gene- regions was quantified by qPCR. Results are presented as percentages of inputs. Values are representative of three independent experiments.

Confocal microscopy. Microglial cells were transfected in 24-well glass-bottomed plates (Greiner Bio-One, Ref. 662892) with a DNA mixture containing eGFP-DCAF1, mRFP-CTIP2 or HA-Vpr. At 24 h post-transfection, cells were extensively washed, fixed with 4% PFA/PBS solution and kept in 1x PBS at 4 °C until observation with a Leica SPE confocal microscope equipped with a Leica 63x1.4NA oil immersion objective (HXC PL APO 63x/1.40 OIL CS).

Immunofluorescence staining. Microglial cells expressing HA-tagged Vpr were immunodetected as described before[56]. Anti-HA (BioLegend, 901501) was diluted 1/1000 and Alexa Fluor goat anti-mouse (A-21237, Thermo Fisher) 1/250. Cells were analyzed by confocal microscopy.

Fluorescence lifetime imaging microscopy (FLIM). Time-correlated single-photon counting FLIM was performed on a home-made laser scanning set-up based on an Olympus IX70 inverted microscope with an Olympus 60× 1.2NA water immersion objective, as described[56]. Two-photon excitation at 900 nm was provided by an InSight DeepSee laser (Spectra Physics). For FLIM, the laser power was adjusted to give count rates with peaks up to as few as 10^6 photons.s^{-1}, so that the pile-up effect can be neglected. Photons were collected using a short pass filter with a cut-off wavelength of 680 nm (F75-680, AHF, Germany), and a band-pass filter of 520 ± 17 nm (F37-520, AHF, Germany). The fluorescence was directed to a fiber-coupled APD (SPCM-AQR-14-FC, Perkin Elmer), which was connected to a time-correlated single photon counting module (SPC83, Becker and Hickl, Germany). Typically, the samples were scanned continuously for about 60 s to achieve appropriate photon statistics to analyze the fluorescence decays. The FRET phenomenon causes a decrease in the fluorescent lifetime (τ) of the donor, which can be measured by the FLIM technique at each pixel or group of pixels. Experiments were performed with eGFP as a FRET donor and mRFP as acceptor. The FRET efficiency (E) was calculated according to: $E = 1 - \tau_{DA}/\tau_D = 1/1 + (R/R_0)^6$ where τ_{DA} is the lifetime of the donor in the presence of the acceptor and τ_D is the lifetime of the donor in the absence of the acceptor. A Förster R_0 distance (where the transfer efficiency is 50%) of 56 Ä was calculated between eGFP used as a donor and mRFP used as acceptor.

Luciferase assays. Microglial cells cultured in 48-well plates were transfected with the indicated vectors and the CMV-Renilla luciferase control vector using the calcium phosphate co-precipitation method. Two days later, cells were collected and firefly luciferase activity was determined using the Dual-Glo Luciferase Assay System (Promega, Madison, WI, USA) and normalized to the Renilla luciferase activity. Values correspond to an average of at least three independent experiments performed in triplicates.

mRNA quantification. The total RNA was extracted using an RNeasy Plus Mini Kit (Qiagen, Germantown, MD, USA) according to the manufacturer's instructions. Using SuperScript III Reverse Transcriptase (Thermo Fisher) and oligo (dT)12–18 Primers (Thermo Fisher), RNA was reverse transcribed into cDNA. The cDNA was diluted 10-fold with DNase-free water. Quantitative PCR was performed on the diluted cDNA using SYBR Select Master Mix. Primers used in qPCR include: CTIP2-F 5′-ATTGGAACCTGCCACTTG-3′, CTIP2-R 5′-TTTGCCTGTGTTCCACGA-3′ GAPDH-F5′-GAGAAGGCTGGGGCTCATTT-3′, GAPDH-R 5′-GCAGTGATGGCATGGACTGT-3′. Data were normalized to GAPDH and calculated using the $2^{-(\Delta\Delta CT)}$ method.

Statistical analysis. Data were statistically analyzed by using *t* tests with Excel (Microsoft) or StatView (SAS Institute Inc.). A value of $p < 0.05$ was considered statistically significant. Results were expressed as S.D. and represent data from a minimum of three independent experiments.

References

1. Descours, B. *et al.* SAMHD1 restricts HIV-1 reverse transcription in quiescent CD4(+) T-cells. *Retrovirology* **9**, 87 (2012).
2. Laguette, N. *et al.* SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* **474**, 654–7 (2011).
3. Lahouassa, H. *et al.* SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. *Nat. Immunol.* **13**, 223–8 (2012).
4. Le Douce, V., Cherrier, T., Riclet, R., Rohr, O. & Schwartz, C. The Many Lives of CTIP2: From AIDS to Cancer and Cardiac Hypertrophy. *J. Cell. Physiol.* **229**, 533–537 (2014).
5. Le Douce, V. *et al.* HIC1 controls cellular- and HIV-1- gene transcription via interactions with CTIP2 and HMGA1. *Sci. Rep.* **6**, 34920 (2016).
6. Marban, C. *et al.* Targeting the Brain Reservoirs: Toward an HIV Cure. *Front. Immunol.* **7**, 397 (2016).
7. Cherrier, T. *et al.* p21(WAF1) gene promoter is epigenetically silenced by CTIP2 and SUV39H1. *Oncogene* **28**, 3380–9 (2009).
8. Le Douce, V. *et al.* LSD1 cooperates with CTIP2 to promote HIV-1 transcriptional silencing. *Nucleic Acids Res.* **40**, 1904–15 (2012).
9. Marban, C. *et al.* Recruitment of chromatin-modifying enzymes by CTIP2 promotes HIV-1 transcriptional silencing. *Embo J* **26**, 412–423 (2007).
10. Cismasiu, V. B. *et al.* BCL11B functionally associates with the NuRD complex in T lymphocytes to repress targeted promoter. *Oncogene* **24**, 6753–6764 (2005).
11. Cherrier, T. *et al.* CTIP2 is a negative regulator of P-TEFb. *Proc Natl Acad Sci USA* **110**, 12655–60 (2013).
12. Eilebrecht, S. *et al.* HMGA1 recruits CTIP2-repressed P-TEFb to the HIV-1 and cellular target promoters. *Nucleic Acids Res.* **42**, 4962–71 (2014).
13. Mashiba, M. & Collins, K. L. Molecular mechanisms of HIV immune evasion of the innate immune response in myeloid cells. *Viruses* **5**, 1–14 (2013).
14. Le Rouzic, E. *et al.* HIV1 Vpr arrests the cell cycle by recruiting DCAF1/VprBP, a receptor of the Cul4-DDB1 ubiquitin ligase. *Cell Cycle* **6**, 182–8 (2007).
15. Romani, B. & Cohen, E. A. Lentivirus Vpr and Vpx accessory proteins usurp the cullin4-DDB1 (DCAF1) E3 ubiquitin ligase. *Curr. Opin. Virol.* **2**, 755–63 (2012).
16. Wen, X., Casey Klockow, L., Nekorchuk, M., Sharifi, H. J. & de Noronha, C. M. C. The HIV1 protein Vpr acts to enhance constitutive DCAF1-dependent UNG2 turnover. *PLoS One* **7**, e30939 (2012).
17. Casey Klockow, L. *et al.* The HIV-1 protein Vpr targets the endoribonuclease Dicer for proteasomal degradation to boost macrophage infection. *Virology* **444**, 191–202 (2013).
18. Okumura, A. *et al.* HIV-1 accessory proteins VPR and Vif modulate antiviral response by targeting IRF-3 for degradation. *Virology* **373**, 85–97 (2008).
19. Maudet, C. *et al.* HIV-1 Vpr induces the degradation of ZIP and sZIP, adaptors of the NuRD chromatin remodeling complex, by hijacking DCAF1/VprBP. *PLoS One* **8**, e77320 (2013).
20. Lahouassa, H. *et al.* HIV-1 Vpr degrades the HLTF DNA translocase in T cells and macrophages. *Proc. Natl. Acad. Sci. USA* **113**, 5311–6 (2016).
21. Yan, J. *et al.* HIV-1 Vpr Reprograms CLR4 ^{DCAF1} E3 Ubiquitin Ligase to Antagonize Exonuclease 1-Mediated Restriction of HIV-1 Infection. *MBio* **9** (2018).
22. Laguette, N. *et al.* Premature activation of the SLX4 complex by Vpr promotes G2/M arrest and escape from innate immune sensing. *Cell* **156**, 134–45 (2014).
23. Zhou, D. *et al.* The HIV-1 accessory protein Vpr induces the degradation of the anti-HIV-1 agent APOBEC3G through a VprBP-mediated proteasomal pathway. *Virus Res.*, https://doi.org/10.1016/j.virusres.2014.08.021 (2014).
24. Mashiba, M., Collins, D. R., Terry, V. H. & Collins, K. L. Vpr overcomes macrophage-specific restriction of HIV-1 Env expression and virion production. *Cell Host Microbe* **16**, 722–35 (2014).
25. Yurkovetskiy, L. *et al.* Primate immunodeficiency virus proteins Vpx and Vpr counteract transcriptional repression of proviruses by the HUSH complex. *Nat. Microbiol.*, https://doi.org/10.1038/s41564-018-0256-x (2018).
26. Chougui, G. *et al.* HIV-2/SIV viral protein X counteracts HUSH repressor complex. *Nat. Microbiol.* **3**, 891–897 (2018).
27. Müller, B., Tessmer, U., Schubert, U. & Kräusslich, H. G. Human immunodeficiency virus type 1 Vpr protein is incorporated into the virion in significantly smaller amounts than gag and is phosphorylated in infected cells. *J. Virol.* **74**, 9727–9731 (2000).
28. Li, G. *et al.* HIV-1 replication through hHR23A-mediated interaction of Vpr with 26S proteasome. *PLoS One* **5**, e11371 (2010).
29. Rohr, O. *et al.* Recruitment of Tat to heterochromatin protein HP1 via interaction with CTIP2 inhibits human immunodeficiency virus type 1 replication in microglial cells. *J Virol* **77**, 5415–5427(2003).
30. 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.* **291**, 2696–711 (2016).
31. DeHart, J. L. *et al.* HIV-1 Vpr activates the G2 checkpoint through manipulation of the ubiquitin proteasome system. *Virol. J.* **4**, 57 (2007).
32. Lv, L. *et al.* Vpr Targets TET2 for Degradation by CRL4 VprBP E3 Ligase to Sustain IL-6 Expression and Enhance HIV-1 Replication. *Mol. Cell* **70**, 961–970.e5 (2018).
33. Cismasiu, V. B. *et al.* BCL11B is a general transcriptional repressor of the HIV-1 long terminal repeat in T lymphocytes through recruitment of the NuRD complex. *Virology* **380**, 173–181 (2008).
34. Mitchell, R. S. *et al.* Vpu antagonizes BST-2-mediated restriction of HIV-1 release via beta-TrCP and endo-lysosomal trafficking. *PLoS Pathog.* **5**, e1000450 (2009).
35. Douglas, J. L. *et al.* Vpu directs the degradation of the human immunodeficiency virus restriction factor BST-2/Tetherin via a {beta} TrCP-dependent mechanism. *J. Virol.* **83**, 7931–47 (2009).
36. Popov, S. *et al.* Viral protein R regulates nuclear import of the HIV-1 pre-integration complex. *EMBO J.* **17**, 909–17 (1998).
37. Kino, T. *et al.* Human immunodeficiency virus type 1 (HIV-1) accessory protein Vpr induces transcription of the HIV-1 and glucocorticoid-responsive promoters by binding directly to p300/CBP coactivators. *J. Virol.* **76**, 9724–34 (2002).
38. Kuramitsu, M. *et al.* A novel role for Vpr of human immunodeficiency virus type 1 as a regulator of the splicing of cellular pre-mRNA. *Microbes Infect.* **7**, 1150–60 (2005).
39. Nishizawa, M., Kamata, M., Mojin, T., Nakai, Y. & Aida, Y. Induction of apoptosis by the Vpr protein of human immunodeficiency virus type 1 occurs independently of G(2) arrest of the cell cycle. *Virology* **276**, 16–26 (2000).
40. Guenzel, C. A., Hérate, C. & Benichou, S. HIV-1 Vpr-a still "enigmatic multitasker&quot. *Front. Microbiol.* **5**, 127 (2014).
41. Miller, C. M. *et al.* Virion-Associated Vpr Alleviates a Postintegration Block to HIV-1 Infection of Dendritic Cells. *J. Virol.* **91**, e00051–17 (2017).
42. Le Rouzic, E. *et al.* Assembly with the Cul4A-DDB1 ^{DCAF1} Ubiquitin Ligase Protects HIV-1 Vpr from Proteasomal Degradation. *J. Biol. Chem.* **283**, 21686–21692 (2008).
43. Zhang, L. *et al.* Coordinated Regulation of Transcription Factor Bcl11b Activity in Thymocytes by the Mitogen-activated Protein Kinase (MAPK) Pathways and Protein Sumoylation. *J. Biol. Chem.* **287**, 26971–26988 (2012).
44. Selman, W. H., Esfandiari, E. & Filtz, T. M. Alteration of Bcl11b upon stimulation of both the MAP kinase- and Gsk3-dependent signaling pathways in double-negative thymocytes. *Biochem. Cell Biol.* **97**, 201–213 (2019).

45. Romani, B. *et al.* HIV-1 Vpr reactivates latent HIV-1 provirus by inducing depletion of class I HDACs on chromatin. *Sci. Rep.* **6**, 31924 (2016).
46. Robbez-Masson, L. *et al.* The HUSH complex cooperates with TRIM28 to repress young retrotransposons and new genes. *Genome Res.* **28**, 836–845 (2018).
47. Tchasovnikarova, I. A. *et al.* Epigenetic silencing by the HUSH complex mediates position-effect variegation in human cells. *Science (80-.).* **348**, 1481–1485 (2015).
48. Van Lint, C. Stop HUSHing on SIV/HIV. *Nat. Microbiol.* **3**, 1336–1338 (2018).
49. Brégnard, C., Benkirane, M. & Laguette, N. DNA damage repair machinery and HIV escape from innate immune sensing. *Front. Microbiol.* **5**, 176 (2014).
50. Dubuissez, M. *et al.* Protein Kinase C-Mediated Phosphorylation of BCL11B at Serine 2 Negatively Regulates Its Interaction with NuRD Complexes during CD4+ T-Cell Activation. *Mol. Cell. Biol.* **36**, 1881–98 (2016).
51. Schwartz, C. *et al.* On the way to find a cure: Purging latent HIV-1 reservoirs. *Biochem. Pharmacol.*, https://doi.org/10.1016/j.bcp.2017.07.001 (2017).
52. Schwartz, C., Rohr, O. & Wallet, C. Targeting the DNA-PK complex: its rationale use in cancer and HIV-1 infection. *Biochem. Pharmacol.* https://doi.org/10.1016/J.BCP.2018.12.002 (2018).
53. Janabi, N., Peudenier, S., Héron, B., Ng, K. H. & Tardieu, M. Establishment of human microglial cell lines after transfection of primary cultures of embryonic microglial cells with the SV40 large T antigen. *Neurosci. Lett.* **195**, 105–108 (1995).
54. Su, B. *et al.* Neutralizing antibodies inhibit HIV-1 transfer from primary dendritic cells to autologous CD4 T lymphocytes. *Blood* **120**, 3708–3717 (2012).
55. Holl, V. *et al.* Stimulation of HIV-1 Replication in Immature Dendritic Cells in Contact with Primary CD4 T or B Lymphocytes. *J. Virol.* **84**, 4172–4182 (2010).
56. Fritz, J. V. *et al.* Direct Vpr-Vpr interaction in cells monitored by two photon fluorescence correlation spectroscopy and fluorescence lifetime imaging. *Retrovirology* **5**, 87 (2008).
57. McDonald, D., Carrero, G., Andrin, C., de Vries, G. & Hendzel, M. J. Nucleoplasmic beta-actin exists in a dynamic equilibrium between low-mobility polymeric species and rapidly diffusing populations. *J. Cell Biol.* **172**, 541–52 (2006).

Acknowledgements

We are grateful to Andrea Janossy for careful and critical reading of the manuscript, and to James Hutchins for correcting the scientific English. Work in the OR laboratory was supported by grants from the French Agency for Research on AIDS and Viral Hepatitis (ARNS), Sidaction, the European Union's Horizon 2020 research and innovation program under grant agreement No. 691119 — EU4HIVCURE — H2020-MSCA-RISE-2015, Alsace Contre le Cancer, Institut Universitaire de France and university of Strasbourg.

Author Contributions

F.F., A.S., D.R.M., D.C., E.H., E.M.M., D.F., V.A.J. and W.C. performed experiments. A.A.A., B.E., M.C. and M.G.F. provided technical resources, scientific and technical expertise. V.L.C., S.C. and R.O. designed and supervised the experiments, wrote the draft and provided funds.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-48689-x.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

5. Publication 2

Microglial Cells: The Main HIV-1 Reservoir in the Brain

Clementine Wallet, Marco De Rovere, Jeanne Van Assche, Fadoua Daouad, Stéphane De Wit, Virginie Gautier, Patrick W. G. Mallon, Alessandro Marcello, Carine Van Lint, Olivier Rohr and Christian Schwartz

Front. Cell. Infect. Microbiol., 24 October 2019

While the main reservoirs for HIV-1 are the resting CD4+ T cells in the past years many other cell reservoirs were described such as hematopoietic stem cells, dendritic cells, microglial cells and cells from the monocyte/macrophage lineage. In this review we described all the currently available knowledge on how microglial cells become, and should be considered, the main HIV-1 cell reservoirs in the CNS. This review also suggests how these cells should be the main target in the CNS in order to achieve cure. A description of molecular mechanisms of the establishment and the maintenance of HIV-1 latency is also discussed with and a further analysis on the therapeutic solutions to target these mechanisms. We also analyzed the difficulties of working with these cells, like their presence within an anatomical sanctuary protected by the Blood Brain Barrier and how to circumvent these limitations.

My main contribution for this review was to share my expertise maturated in these years in our lab working extensively with microglial cells in regards of HIV-1 infection.

frontiers in Cellular and Infection Microbiology

REVIEW
published: 24 October 2019
doi: 10.3389/fcimb.2019.00362

Microglial Cells: The Main HIV-1 Reservoir in the Brain

Clementine Wallet[1], Marco De Rovere[1], Jeanne Van Assche[1], Fadoua Daouad[1], Stéphane De Wit[2], Virginie Gautier[3], Patrick W. G. Mallon[3], Alessandro Marcello[4], Carine Van Lint[5], Olivier Rohr[1] and Christian Schwartz[1]*†*

[1] Université de Strasbourg, EA7292, FMTS, IUT Louis Pasteur, Schiltigheim, France, [2] Division of Infectious Diseases, Saint-Pierre University Hospital, Université Libre de Bruxelles (ULB), Brussels, Belgium, [3] UCD Centre for Experimental Pathogen Host Research (CEPHR), School of Medicine, University College Dublin, Dublin, Ireland, [4] Laboratory of Molecular Virology, International Centre for Genetic Engineering and Biotechnology (ICGEB), Trieste, Italy, [5] Service of Molecular Virology, Department of Molecular Biology (DBM), Université Libre de Bruxelles (ULB), Gosselies, Belgium

OPEN ACCESS

Edited by:
Alexander O. Pasternak,
Academic Medical Center (AMC), Netherlands

Reviewed by:
Larisa Y. Poluektova,
University of Nebraska Medical Center, United States
Eliseo A. Eugenin,
University of Texas Medical Branch at Galveston, United States

***Correspondence:**
Olivier Rohr
olivier.rohr@unistra.fr
Christian Schwartz
schwartz.christian@unistra.fr

†Lead contact

Specialty section:
This article was submitted to Virus and Host, a section of the journal Frontiers in Cellular and Infection Microbiology

Received: 10 July 2019
Accepted: 07 October 2019
Published: 24 October 2019

Citation:
Wallet C, De Rovere M, Van Assche J, Daouad F, De Wit S, Gautier V, Mallon PWG, Marcello A, Van Lint C, Rohr O and Schwartz C (2019) Microglial Cells: The Main HIV-1 Reservoir in the Brain. Front. Cell. Infect. Microbiol. 9:362. doi: 10.3389/fcimb.2019.00362

Despite efficient combination of the antiretroviral therapy (cART), which significantly decreased mortality and morbidity of HIV-1 infection, a definitive HIV cure has not been achieved. Hidden HIV-1 in cellular and anatomic reservoirs is the major hurdle toward a functional cure. Microglial cells, the Central Nervous system (CNS) resident macrophages, are one of the major cellular reservoirs of latent HIV-1. These cells are believed to be involved in the emergence of drugs resistance and reseeding peripheral tissues. Moreover, these long-life reservoirs are also involved in the development of HIV-1-associated neurocognitive diseases (HAND). Clearing these infected cells from the brain is therefore crucial to achieve a cure. However, many characteristics of microglial cells and the CNS hinder the eradication of these brain reservoirs. Better understandings of the specific molecular mechanisms of HIV-1 latency in microglial cells should help to design new molecules and new strategies preventing HAND and achieving HIV cure. Moreover, new strategies are needed to circumvent the limitations associated to anatomical sanctuaries with barriers such as the blood brain barrier (BBB) that reduce the access of drugs.

Keywords: HIV-1, Ctip2, microglial cells, reservoirs, latency, brain

INTRODUCTION

Since the introduction of the combination antiretroviral therapy (cART) in 1996 the fatal evolution of the HIV-1 infection has stopped and became a chronic disease. However, despite continual efforts definitive cure has not been achieved. This is mainly due to the lack of efficient vaccination and to the existence of viral reservoirs. Indeed, as soon as the treatment is interrupted, a viral rebound is observed within a couple of weeks and viremia is comparable to the one encountered before the introduction of cART (Davey et al., 1999; García et al., 1999; Mata et al., 2005). Moreover, with ultrasensitive methods, we can still notice a residual viremia in patients under cART (Dornadula et al., 1999; Bouchat et al., 2016). The main reservoirs are the resting CD4+ T cells but we have now enough evidence which argues for the existence of many other cell reservoirs such as hematopoietic stem cells, dendritic cells, microglial cells and cells from the monocyte-macrophage lineage (reviewed in Marcello, 2006; García et al., 2018; Sung and Margolis, 2018).

In addition, some of these reservoirs are found in sanctuaries such as the genital tract, the adipose tissue, the bone marrow and the brain. Targeting the infected cells in these reservoirs is even more difficult (Sadowski and Hashemi, 2019). The brain for instance constitutes a typical anatomical site, which has poor drug penetration and reduced immune surveillance due to the blood brain barrier (BBB). As a consequence, the genetical information exchange of the virus confined in the brain and other compartments such as the blood is poor (Canestri et al., 2010; Dahl et al., 2014). The measureable pool of latently infected, resting CD4+ T cells does not account fully for the early rebounding plasma HIV-1. It has been shown that the rebounding HIV-1 was genetically different suggesting that these viruses came from other cell reservoirs (Chun et al., 2000). In addition, the result of recent phylogenetic studies is in favor of other physiological observations suggesting that HIV-1 production in the brain is associated with the emergence of virus resistance (Smit et al., 2004; Strain et al., 2005) and with HIV-1 associated neurocognitive disease (HAND) (Salemi et al., 2005; Lamers et al., 2010). Up to 50% of patients under cART with undetectable virus in the blood suffer of less severe forms of HAND (Eggers et al., 2017). The pathogenesis of HAND is not fully understood but appears to be multifactorial: drug neurotoxicity of the cART, HIV replication in the central nervous system (CNS) from quiescent/latent infected cells, CNS inflammation and co-infection with other viruses such as hepatitis C virus (reviewed in Sutherland and Brew, 2018). Indeed production of newly synthesized viral protein such as Tat or gp120 from infected cells might lead to direct neural injury (reviewed in Rao et al., 2014). Recent reports described new mechanisms for HIV-1 Tat-mediated microglial inflammation which involves a novel miRNA-34a- NOD-like receptor C5 (NLRC5)-NFκB signaling axis or activation of the methyl CpG binding protein 2 (MEPC2)-STAT3 axis (Periyasamy et al., 2018, 2019). Activation by Tat of the microglial nucleotide-binding domain leucin-rich repeat and pyrin-containing receptor 3 (NLRP3) inflammasome also lead to the occurrence of neuroinflammation (Chivero et al., 2017). Indirect impact on the integrity of neurons is also observed during sustained chronic inflammation induced by secreting perivascular macrophages or microglial cells (reviewed in Eggers et al., 2017). Following activation these immune cells release neurotoxic host factors such as pro-inflammatory cytokines (TNFα, IFNα, IL6, IL8, IL1β) (Nolting et al., 2012; Zhou et al., 2016; Abassi et al., 2017) and chemokines (CCL2 and CCL5) (Woods et al., 2006; Zhou et al., 2016). All together these indirect evidences strongly support the existence of cellular reservoirs in the CNS. At least, three different types of infected cells have been described in the CNS referred as a compartment (Bednar et al., 2015): astrocytes, cells from the monocyte lineage (Wong et al., 2019) and microglial cells (Joseph et al., 2015). These putative true reservoirs are established very early during HIV-1 infection (within 3 to 5 days following infection) (Koenig et al., 1986; Whitney et al., 2014). It is well-documented that infection of the CNS results from transmigration of infected CD4+ cells (reviewed in Williams et al., 2012) and of infected monocytes across the BBB (Williams et al., 2012). However, these infected cells entering the CNS do not constitute an HIV reservoir since they have a very short half-life. Instead they will be the source of the infection of three long lived cell types: astrocytes, perivascular macrophages and microglial cells. Whether or not these cells constitute true cell reservoirs is still debated (Gray et al., 2014; Al-Harti et al., 2018; Vanhamel et al., 2019). Criteria for a cell reservoir are the followings: (i) presence of HIV-1 DNA integrated in the host genome of long lived cells, (ii) existence of mechanisms which allow the virus to persist for long period in latent cells which include mechanisms allowing establishment and maintenance of a latent infection and (iii) formation of replication-competent particles following activation of the reservoirs (Eisele and Siliciano, 2012). Two criteria of a true latency reservoir have been described in these cells: the presence of HIV-1 integrated DNA in long lived cells and the existence of mechanisms which allow the virus to persist for long period in latent cells (Blankson et al., 2002). On the other hand, it cannot be explored whether these cells can produce replication-competent viruses in human due to ethical and technical problems.

The astrocyte is the most abundant infected cell type in the CNS with up to 10–20% having HIV-1 DNA in the cell genome (Churchill et al., 2009). However a recent study using sensitive methods detected HIV-1 in brain macrophages and microglial cells but not in astrocytes (Ko et al., 2019). Moreover, the infection appears to be non-productive. Thus they might not constitute a true viral reservoir for HIV-1 (Gorry et al., 2003). On the contrary there is evidence that both macrophages and microglial cells are susceptible to HIV-1 infection and support productive infection (reviewed in Joseph et al., 2015). As already mentioned, the productive infection in these later cells in the CNS has been associated with HAND in humans and in animal models such as the macaques. Therefore, both macrophages and microglial cells can be considered as true reservoirs in the brain. However, their relative importance is not the same due to their differences in size and constitution in the brain. The macrophages are bone marrow-derived with a half-life of months and they do not undergo cell division (Koppensteiner et al., 2012). They are repopulated from monocytes crossing the BBB. Under a very efficient cART or by using therapeutic strategies preventing the specific subset of infected monocytes to cross the BBB we might expect in the future to drastically decrease this reservoir in the brain. The microglial cells originate from erythromyeloid progenitors in the yolk sac during embryogenesis and colonize the forming CNS (Kierdorf et al., 2013). They constitute therefore the main resident cells of the brain which can serve as brain macrophages. Due to their long half-life of years they constitute a very stable population (Réu et al., 2017). In addition unlike macrophages they can undergo cell division enabling HIV-1 to persist in the brain (Lawson et al., 1992). Moreover, a recent report suggests that microglial cells are highly susceptible to HIV-1 infection (Cenker et al., 2017). Altogether, the microglial cells might constitute one of the main HIV-1 reservoir in the brain.

In the review we will first present microglial cells and discuss why they might be the main HIV-1 cell reservoirs in the cART era. We suggest that microglial cells should be the main target in the brain in order to achieve cure. As a

prerequisite it is needed to decipher the molecular mechanisms of the establishment and the maintenance of HIV-1 latency in these cells. In further section, we will discuss the therapeutic implications of these mechanisms followed by the discussion of the challenges facing when targeting microglial cells and how to circumvent these limitations.

MICROGLIAL CELLS, A KEY PLAYER DURING THE COURSE OF HIV-1 INFECTION

The Biology of Microglial Cells

Microglial cells constitute the CNS resident innate immune cells (Bilimoria and Stevens, 2015). They are the most abundant mono nuclear macrophages which are found in the totality of the brain parenchyma. These cells appear very early during embryogenesis and derive from early myeloid precursors in the embryogenic yolk sac (Gomez Perdiguero et al., 2015; Sheng et al., 2015). Microglial cells colonize the entire brain parenchyma before the formation of the BBB and subsist by slow cell division during life (Gomez Perdiguero et al., 2015). The median rate of human microglial cell renewal is around 30% per year. In a recent report, it was estimated that microglial cells are on average 4.2 years old and most of the cells in the population regenerates throughout life (Réu et al., 2017). Contrary to the other potentially infected cells in the brain this slow cell division allows the persistence of HIV-1 in the brain throughout the life of the patient. Both resting and activated microglial cells are ramified cells with multiple branches and processes which contact neurons, astrocytes and blood vessels (reviewed in Sominsky et al., 2018). Resting cells are transformed to activated macrophage like state when sensing a change in their microenvironment (reviewed in Colonna and Butovsky, 2017). They are also very mobile and rapidly reach the site of brain damage. When activated, these cells release many cytokines, chemokines and other neurotoxic proteins. These secretory factors might contribute to the pathological neuroinflammation observed in HAND (Hong and Banks, 2015). Many other diseases have been associated with pathological neuroinflammation and are described elsewhere (Salter and Stevens, 2017). Microglial cells express numerous pattern-recognition receptors (PRRs) such as TLR1/2, TLR3, and TLR4 (reviewed in Colonna and Butovsky, 2017). Many other immune receptors are also expressed in microglial cells including CD4+, chemokine receptors, receptors which regulate the activation of microglial cells such as Triggering Receptor Expressed on Myeloid cell 2 (TREM2). These receptors might serve as a marker of inflammation during the course of HIV-1 infection (Gisslén et al., 2019). Many receptors for pro-inflammatory factors (TNFα, IL1β, IL10) and anti-inflammatory factors (IFNα/β) are also expressed (Lee et al., 2002). Interestingly, released anti-inflammatory factors transfer signals to the cell interior via a specific receptor complex to generate an antiviral response (Taniguchi and Takaoka, 2002). An exhaustive review of the wide range of immune receptors is described elsewhere (Colonna and Butovsky, 2017). Apart its function in immune surveillance to eliminate the microbes, microglial cells can also eliminate dead cells, redundant synapses, protein aggregates such as amyloid plaques (reviewed in Colonna and Butovsky, 2017; Salter and Stevens, 2017; Sominsky et al., 2018). These cells are involved in several functions of the CNS, including neurogenesis, the control of synaptic density and connectivity and the regulation of synaptic plasticity through the release of many cytokines such as Brain-Derived Neurotrophic Factor (BNDF). Of note HIV-1 Tat protein can inhibit BNDF by up regulating the microRNA mir-34a (Santerre et al., 2019). Microglial cells can also engulf neurotransmitters released in excess and therefore prevent excitotoxicity and participate in the cross talk between microglial cells and astrocytes. Due to their multifunction, any dysfunction in microglia cells will lead to many diverse diseases including brain aging or neurodegenerative diseases such as Alzheimer disease, Parkinson disease and amyotrophic lateral sclerosis. They are also central players in the occurrence of pathological neuroinflammation observed in patients under cART. HAND is thought to be related to excessive secretion of cytokines and chemokines (Yadav and Collman, 2009; Burdo et al., 2013; Eggers et al., 2017). In a recent study the gene expression profile in microglial cells from infected patients under cART with or without HIV encephalitis (HIVE) has been investigated which showed differential expression of several classes of genes compared to non-infected patients. 64% of the genes were effected in HIVE positive patients and 24% in HIVE negative patients (Ginsberg et al., 2018). This report strongly suggests that neuroinflammation, which leads to neurodegenerative diseases such as HAND, is associated with microglial impairment. Neuroimaging such as positron emission tomography (PET) has shown that neuroinflammation develops even under effective cART (Tavazzi et al., 2014; Vera et al., 2016). This approach has highlighted the importance of microglial activation in HIV+ patients under efficient cART with or without any cognitive impairments. PET also allows to follow the level of microglial activation by targeting translocator protein (TSPO) (Sinharay and Hammoud, 2019). Indeed, upregulation of TSPO during microglial activation is related to the degree of neuroinflammation (Venneti et al., 2004). Activation of microglial cells will further lead to the progression of neurodegenerative diseases by impacting the BBB (reviewed in da Fonseca et al., 2014; Osipova et al., 2018). Indeed microglial cells activation modulates tight junctions of endothelial cells of the BBB (Sumi et al., 2010). Of note, inhibition of microglial activation by minocycline is associated with preservation of BBB integrity *in vitro* and with reduced disruption of the BBB *in vivo* (Yenari et al., 2006; Sheng et al., 2018).

In conclusion, microglial cells fulfill several criteria of a brain reservoir. Most importantly they can subsist for a very long time in the brain and they can colonize the brain parenchyma. Contrary to other potential reservoirs in the brain, these cells divide slowly expanding the viral reservoirs in the brain and thus allowing virus persistence and reseeding of the blood. They are also involved in many functions including immune surveillance. As a consequence, any dysfunction of these cells might explain the occurrence of HAND.

Evidence Supporting That Microglial Cells Are Susceptible to HIV-1 Infection and They Contribute to the Formation of a Cell Reservoir in the Brain

It is believed that microglial cell infection arises from transmigration of infected monocytes occurring very early in the course of infection. Recently, a specific subset of infected monocytes which preferentially cross the BBB, the HIV CD14 + CD16+ monocytes, has been characterized (Veenstra et al., 2017). These cells express abundantly junctional proteins such as Junctional Adhesion Molecule-A (JAM-A) and Activated Leukocyte Cell Adhesion Molecule (ALCAM) and chemokine receptors CCR2 which help these cells to cross the BBB. In turn these infected monocytes may infect microglial cells. Alternatively, but still debated, infected CD4+ T cells migrating into the brain might be ingested by microglial cells (Murooka et al., 2012). Although it has not been clearly demonstrated, this later mechanism could be more efficient to spread the virus than exposure to the free virus (Baxter et al., 2014). Whatever the mechanism of infection, it appears that brain microglial cells are permissive to HIV-1 infection. This is despite high level of the restriction factor SAM domain and HD domain 1 (SAMHD1) (Rodrigues et al., 2017). The absence of restriction by SAMDH1 is due to its phosphorylation by the cyclin kinase 1 (CDK1) which is induced in microglial cells that cycle between G0 to G1 state (Cribier et al., 2013; Micochova et al., 2017). There is now evidence supporting that microglial cells are infected by HIV-1 both *in vitro* and *in vivo* (reviewed in Joseph et al., 2015). Previous studies from autopsy have identified HIV-1 DNA, RNA and protein in microglial cells (Cosenza et al., 2002; Churchill et al., 2006). However, it was pointed out that these patients died from severe form of HAND. A recent study confirmed that microglial cells are infected in patients whose viral level is suppressed but died from an HIV-1 unrelated outcome (Ko et al., 2019). In this study the authors used a unique cohort from the National Neuro AIDS Tissue Consortium (NTTC) which comprised 16 patients on cART with well-documented, sustained control of HIV-1. They used highly specific technology to detect and quantify both HIV-1 DNA and RNAs at the cellular level. Very interestingly they showed that both perivascular macrophages and microglial cells but not astrocytes harbored HIV-1 DNA. In 6 out 16 cases they also found HIV 1 RNA in these cells when HIV 1 RNA was undetectable in the cerebro Spinal Fluid (CSF) and in the blood. This result strongly argues in favor that virus production can take place in the CNS. Other studies have also shown that microglial cells are susceptible to infection *in vitro*. Several *in vitro* models of human microglial cells susceptible of infection have been developed (Janabi et al., 1995; Nagai et al., 2001; Garcia-Mesa et al., 2017; Rawat and Spector, 2017; Dello Russo et al., 2018). Some models for latency have been derived from these previous models and constitute valuable tools to study the mechanism of infection and the molecular mechanisms underlying the establishment and maintenance of HIV-1 latency in microglial cells (Garcia-Mesa et al., 2017; Alvarez-Carbonell et al., 2019). Evidence that the virus was detected in the CSF in patients under effective cART who had otherwise undetectable plasma HIV 1 also argues in favor for the production of HIV 1 in the brain (Edén et al., 2010, 2016; Ferretti et al., 2015). In addition, phylogenetic analyses also suggest an important compartmentalization of HIV-1 in the CSF even early in infection (Salemi and Rife, 2016; Bavaro et al., 2019).

Finally, infection of microglial cells is clearly demonstrated in animal models such as the macaque and the humanized mouse (reviewed in Kumar et al., 2016; Marsden and Zack, 2017; Gama et al., 2018; Abreu et al., 2019). In these models HIV-1 DNA, RNA and protein were detected in the brain. Importantly, studies on these models showed that the viral reservoir is established very early at 3 days post infection (Whitney et al., 2014). More interestingly, there is evidence that microglial cells can be latently infected (Avalos et al., 2017; Gama et al., 2017). A macaque model has recently been designed and a mechanism for the establishment of HIV-1 transcription suggested (Barber et al., 2006). In this model, microglial cells can be reactivated in response to cytokine stimulation (Qin et al., 2015; Xu et al., 2017). Interestingly, in a cell model developed in Karn's laboratory, the glucocorticoid receptor and the toll-like receptor 3 appeared to be crucial receptors for HIV 1 activation (Alvarez Carbonell et al., 2017, 2019). In brain autopsies from patients whose infection was controlled HIV-1 DNAs in microglial cells and macrophages were also detected (Thompson et al., 2011; Desplats et al., 2013). Humanized mouse models were generated in which microglial cells were infected by HIV-1 *in vivo* (Asahchop et al., 2017; Llewellyn et al., 2018; Mathews et al., 2019; Su et al., 2019). These models will allow us to study the pathophysiology of microglial cell activation and to develop strategies aiming to reduce the pool of these reservoirs.

Altogether, there are now numerous evidence supporting that microglial cells constitute a major cellular reservoir in the brain. Perivascular macrophages might also be considered as a cellular reservoir but not a true lifelong existing one. Moreover, we might expect to reduce drastically this pool by intensifying cART or by preventing the transmigration of infected monocytes into the brain. Concerning astrocytes it is still debated whether they constitute a true reservoir (Al-Harti et al., 2018; Ko et al., 2019).

Thus, a complete understanding of the molecular mechanisms underlying establishment and persistence of HIV-1 latency in microglial cells is needed in order to design original strategies aiming to target these reservoirs.

ESTABLISHMENT AND PERSISTENCE OF HIV-1 LATENCY IN MICROGLIAL CELLS

Infection and the constitution of a microglial cell reservoir might occur very early in the evolution of HIV-1 infection (Whitney et al., 2014). Some features of the microglial cells allow for the persistence of HIV-1 in the brain. When infected they are far more resistant to cytopathic effects and on the contrary of CD4+ T cells they are non-lytic. In addition, these cells are resistant to apoptosis. The mechanism of apoptosis resistance to HIV-1 deserves considerable attention since it helps to design alternative strategies based on the increase of apoptosis susceptibility

(Kumar et al., 2014). Indeed, strategies aiming to reactivate HIV-1 from infected cells should consider the problem of resistance to apoptosis since the reinforced therapy might not be efficient against the newly productive but non-apoptotic cells. Finally, the antiretroviral drug efficiency is drastically reduced in anatomical and pharmacological sanctuaries such as the brain where microglial cells are mostly found and thus contributing also to viral persistence (Asahchop et al., 2017; Honeycutt et al., 2017).

Two forms of latency have been described which explain HIV-1 persistence in microglial cells. Latency which occurs before the integration of the HIV-1 genome into the host is called pre-integration latency. This form of latency has been frequently observed in CD4+ T cells and in cells of the monocyte-macrophage lineage (discussed in Le Douce et al., 2010). In the latter cells, pre-integration latency might constitute a viral defense mechanism which could lead to the establishment and persistence of HIV-1 infection (Cara and Klotman, 2006; Hansen et al., 2016). This type of latency has not been described in microglia yet. We call post integration latency when HIV-1 genome has been incorporated into the host DNA genome. The integrated HIV-1 provirus can then be in an active-replication state enabling strong transcription of the HIV-1 promoter or in an inactive-state with a marked repression of HIV-1 transcription (Pai and Weinberger, 2017; Cao et al., 2018). In CD4 T cells, latency is a very rare event with 1 to 10^6-10^7 cells being latently infected (Weinberger and Weinberger, 2013). This rare event has not yet been described but might occur in the other reservoirs infected with HIV 1. Mechanisms involved in CD4+ T cell latency are well described (Khoury et al., 2018). The strength of the HIV Tat positive feedback loop regulates principally the viral fate selection (**Figure 1**) (Pai and Weinberger, 2017). In other word, there is stochasticity in HIV-1 transcription which explains that each single cell differs in many parameters (stage of cell cycle, level of expression of transcriptional activators/repressors, metabolic state…). Any mechanism reducing the level of HIV 1 Tat expression under a threshold level might therefore facilitate the establishment of latency (Karn, 2011). Integration of the viral genome into repressive heterochromatin environment (Jordan et al., 2001), transcriptional interference (Han et al., 2008), the lack of host transcriptional activators (Le Douce et al., 2010; Churchill et al., 2015), the presence of host transcriptional repressors (Le Douce et al., 2010; Churchill et al., 2015) or the influence of viral miRNA on the chromatin environment (Wang et al., 2009) are thought to favor the establishment of HIV-1 latency since all these mechanisms significantly decrease the strength of the Tat positive feedback. Of note, the mechanism involving viral miRNAs is however controversial and deserves further studies (Balasubramaniam et al., 2018). Additionally, post-transcriptional mechanisms regulating reactivation from latency have also been proposed (Sarracino et al., 2018). On the other hand, in microglial cells host transcription factors that facilitate the establishment of latency have been described. Intriguingly, it was observed that a new CNS strain of the virus has evolved with specific promoters (the 5' long terminal repeat or 5' LTR) probably due to the compartmentalization of the virus in the brain (Churchill et al., 2015; Gray et al., 2016). Moreover, in contrast to other reservoirs, the NF-κB and the core region including Sp1 sites are sufficient for HIV-1 transcription in microglial cells (Rohr et al., 2003b). In particular, Sp1 proteins allow the anchorage of many transcriptional activators such as NF-IL6 also referred as liver-enriched transcriptional activator protein (LAP), CREB and Coup-Tf (**Figure 1**). Unlike CD4− T cells which express only Sp1, microglial cells express both Sp1 and Sp3 which operate as transcriptional repressors. In addition, we and others have shown that in macrophages and microglial cells, a truncated form of NF-IL6, liver-enriched transcriptional inhibitory protein (LIP) and/or C-EBPγ is expressed and acts as repressor by competing with the transcriptional activator NF-IL6 (Schwartz et al., 2000) and (Tesmer and Bina, 1996). Noteworthy, this latter mechanism explains HIV-1 latency in microglial cells of a macaque model (Barber et al., 2006). Epigenetic regulation is also involved in the establishment of latency. Coup-TF interacting protein 2 (CTIP2 also known as BCL11b) is a key factor in promoting the formation of heterochromatin, the compact-inactive form of the promoter, in microglial cells. Indeed, CTIP2, apart its crucial role in establishing latency, is also involved in the persistence of latency by preventing HIV-1 reactivation, and in the generation of a cell microenvironment favorable to the establishment of HIV-1 latency. In addition, CTIP2 is described as an anti-apoptotic protein (Kamimura et al., 2007) and thus might also be involved in the apoptosis resistance in microglial cells (Le Douce et al., 2010; Kumar et al., 2014). Interestingly a new mechanism to prevent apoptosis in infected macrophage/microglia cells which survived the infection and became latently infected has been described (Castellano et al., 2017). They notably showed that apoptosis is blocked at a very early step and that the pro-apoptotic protein BIM which accumulate into mitochondria might serve as a new bio-marker of latently-infected macrophage/microglia cells *in vivo*.

CTIP2 is a Zinc finger transcription regulator which has been implicated in many functions required during development and in adult animal (reviewed in Le Douce et al., 2014). Germline disruption of CTIP2 is mortal stressing its importance during development. In fact, the essential role of CTIP2 in T cell development, odontogenesis, skin development and neuronal development are now well-documented (Le Douce et al., 2014). It was shown that CTIP2 is a transcriptional regulator binding directly on a promoter specific sequence (Avram et al., 2002) or binding indirectly to other promoter-bound transcription factors such as Sp1 (Avram et al., 2000).

CTIP2 is involved in the establishment of HIV-1 latency by favoring the formation of heterochromatin in the vicinity of the viral promoter. It serves as a platform anchoring a chromatin modifying complex (Marban et al., 2007). Indeed, our work showed that in the presence of CTIP2 histone deacetylases HDAC1, HDAC2, and histone methyltransferase (HMT) SUV39H1are simultaneously recruited on the viral promoter which generates the H3K9me3 epigenetic mark (trimetylation of H3K9) (**Figure 2**). This work was the first to demonstrate the recruitment of HMT at the latent HIV-1 promoter. It paved the way for reactivation strategies that use a combination of molecules targeting multiple enzymatic activities. Furthermore, we characterized a new factor working in synergy with CTIP2 i.e., the lysine specific demethylase 1

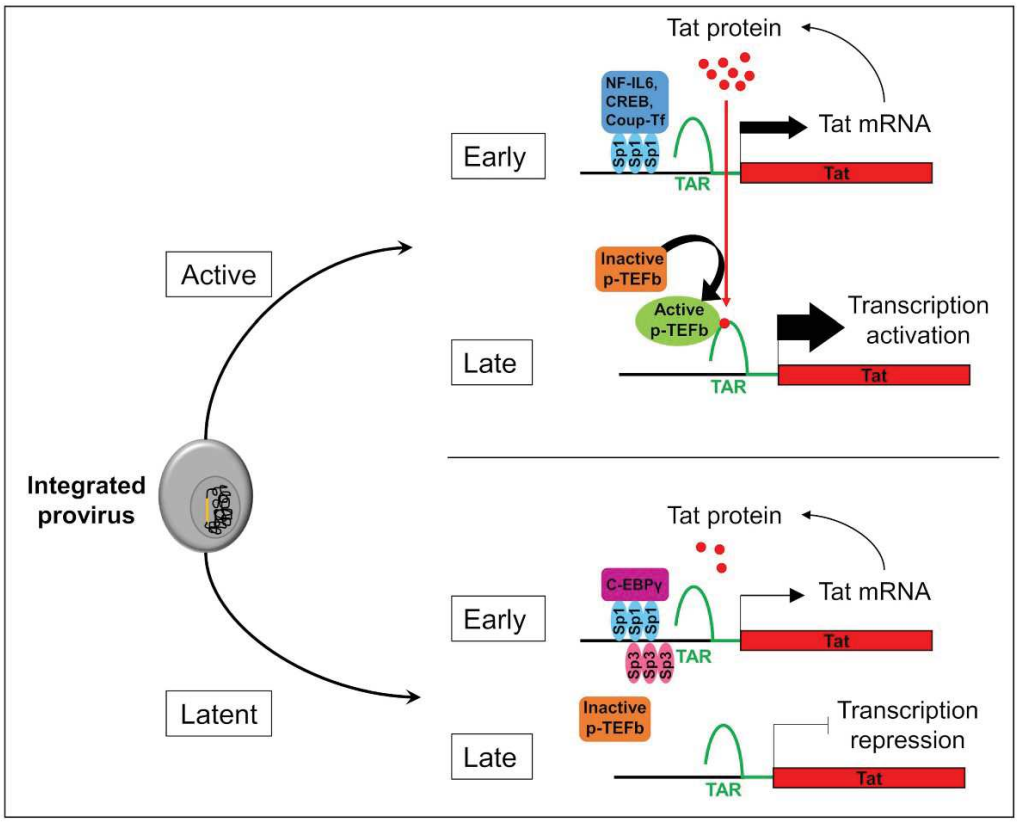


FIGURE 1 | Stochasticity of transcription in microglial cells. The provirus (in yellow) integration in the genome of the microglial host cell (in gray) leads to two possible fates of the cell: active or latent. In active cells, Sp1 allows the anchorage of transcriptional activators like NF-IL6, CREB, and Coup-Tf. It enhances Nf-κB activity and allows the transcription of the viral transactivator Tat in the early phase of transcription. Then, Tat promotes the recruitment of active p-TEFb elongation complex at TAR regions. This recruitment allows viral RNA elongation and transcription activation during the late phase. It results a Tat production which promotes a Tat positive feedback loop and enables a strong transcription of the HIV-1 promoter. In the early phase in latently infected microglial cells, Sp3 act as a transcriptional repressor. In addition, C-EBPγ competes with the transcriptional activator NF-IL6. These repressors reduce Tat expression, p-TEFb recruitment, and viral RNA elongation during the late phase of transcription. If Tat production is under a threshold level, transcription is inhibited and the virus is inactive.

(LSD1), which we showed to be a key factor in the molecular mechanisms of establishment of HIV-1 latency (Le Douce et al., 2012). In particular we established that LSD1 inhibits HIV-1 transcription and viral expression by recruiting hSet1 and WDR5 on the HIV-1 promoter, two members of the hCOMPASS complex, which induce H3K4me3 epigenetic mark (trimetylation of H3K4) (**Figure 2**). The association of H3K4me3 and H3K9me3 epigenetic marks might represent a new level of eukaryotic gene regulation. It was suggested that gene repression linked to H3K4me3 prevents the expression of cryptic promoters (Pinskaya and Morillon, 2009). This is supported by the findings that HIV-1 better integrates in active genes and thus can be considered as a cryptic gene (Wang et al., 2007).

Recently, we showed that CTIP2 is in another complex, which is able to counteract HIV-1 reactivation. In the initial report it was suggested that CTIP2 represses, the Tat-dependent late phase of HIV-1 transcription (Rohr et al., 2003a). Later experiments showed that an inactive form of the elongation factor pTEFb is recovered in a complex containing the non-coding 7SK snRNA and the proteins CTIP2 and HEXIM1 are attached to viral and cellular gene promoters in the absence of the HIV-1 factor Tat (Cherrier et al., 2013). pTEFb is a dimer comprising a regulatory subunit CyclinT1 and a catalytic subunit CDK9. CDK9 is a kinase which phosphorylates the Ser2 residue of the carboxyl terminal end of the RNA polymerase II and the negative transcriptional elongation factors NELF and DSIF. We showed that in this complex CTIP2 significantly inhibited CDK9 kinase activity thus inhibiting pTEFb function (Cherrier et al., 2013). In a following paper, we showed that the cellular protein High Mobility Group AT-hook 1 (HMGA1), which also belongs to the 7SK snRNA complex, recruited the inactive CTIP2/pTEFb complex onto the HIV-1 and cellular target promoters thus regulating their gene expressions (Eilebrecht et al., 2014). HMGA1 has been previously shown to interact with

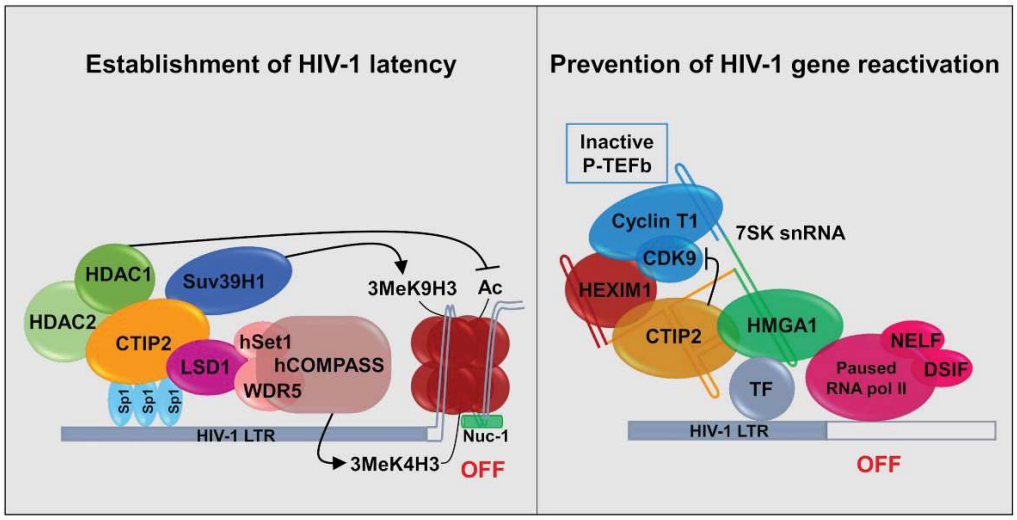


FIGURE 2 | Establishment and persistence of HIV-1 latency in microglial cells: CTIP2 as a central actor. **(Left)** CTIP2 is involved in the establishment of HIV-1 latency by recruiting a chromatin modifying complex at the viral promoter (HIV-1 LTR). CTIP2 is anchored at Sp1 sites on the viral promoter and act as a scaffolding for the recruitment of several chromatin modifying proteins. They perform epigenetic modifications of histones of the nucleosome located at the promoter. CTIP2 recruits simultaneously HDAC2, HDAC1 and the HMT Suv39H1. HDAC1 deacetylate histones while Suv39H1 generates the trimethylation of H3K9 (H3K9me3). In addition, LSD1 promotes the trimethylation of H3K4 by the hCOMPASS methyltransferase complex through hSet1 and WDR5 recruitment. These two epigenetic marks are sufficient to promote chromatin compaction (heterochromatin) and inhibit transcription. **(Right)** CTIP2 prevents transcription restart by sequestrating the elongation factor pTEF-B in an inactive complex. 7SK snRNA complex is recruited at promoter via HMGA1 by interaction with transcription factors (TF). HMGA1 then recruits CTIP2 associated with the inactive form of PTEF-B (Cyclin T1 and CDK9) and HEXIM1, on the 7SK snRNA. Transcription is inhibited at two levels. First, CTIP2 inhibits the kinase activity of CDK9, a pTEF-B subunit which prevent RNA pol II activation by inhibiting the phosphorylation of RNA pol II and the two subunits NELF and DSIF. Second, the persistence of inactive pTEF-B at the promoter prevents any recruitment of active pTEF-B.

TAR to modulate Tat-dependent HIV transcription (Eilebrecht et al., 2013). It seems that HMGA1 is critical for the repressive function of CTIP2. We propose a model in which HMGA1 facilitates the recruitment of CTIP2-inactivated P-TEFb onto cellular and viral gene promoters (**Figure 2**).

As suggested above, CTIP2 has an important role in the regulation of the expression of many infected cells genes. Microarray analysis of a microglial cell line knocked down for CTIP2 showed up or down regulation of several genes (Cherrier et al., 2013). Among them the cellular cyclin-dependent kinase inhibitor CDKN1A/p21^{waf} was up regulated. When recruited to the p21 promoter, CTIP2 represses p21 gene transcription with the same mechanism described for the HIV-1 promoter (Cherrier et al., 2009). The p21 gene repression may favor HIV-1 latency since activation of p21 gene stimulates viral expression in macrophages (Vazquez et al., 2005). In addition, CTIP2 counteracts HIV-1 Vpr which is required for p21 expression. Interestingly we recently showed that HIV-1 Vpr mediates the depletion of the cellular repressor CTIP2 to counteract viral gene silencing (Forouzanfar et al., 2019). Altogether our findings strongly support the findings that CTIP2 generates a cellular environment which is favorable for HIV-1 latency. Furthermore, our results support that CTIP2 is a major actor involved in HIV-1 latency in microglial cells. In agreement with our findings, CTIP2 colocalized with the microglial marker Ionized calcium binding adaptor molecule 1 (Iba1) in latently infected patients but not in HIV encephalitis (HIVE) patients (Desplats et al., 2013). Moreover, the levels of CTIP2 and markers of heterochromatin such as HDACs and heterochromatin 1 (HP1) were increased in microglial cells from HIV-1 positive latent cases. This work and others suggested that CTIP2 can be used as an HIV brain latency biomarker (Cysique et al., 2019).

Recently we have identified and characterized a new cellular factor i.e., Hypermethylated in Cancer 1 (HIC1). This factor regulates together with CTIP2 and HMGA1 cellular and HIV-1 gene transcriptions (Le Douce et al., 2016). However, the repressive activity of the complex requires the deacetylase activity of Sirtuin 1 (SIRT1) in microglial cells. We showed that HIC1 interacted and cooperated with HMGA1 and modulated Tat dependent HIV-1 transcription. However, intriguingly it occurs in the presence of Tat but it is independent of the elongation factor PTEFb. The need for HMGA1 to interact with HIC1 to repress the Tat dependent HIV-1 transcription might suggest a new role for HMGA1 in this process (discussed in Le Douce et al., 2016). Our findings support the idea that the TAR element serves as a HIC-1 reservoir to promote HIC-1/Tat interaction.

We are far from elucidating all the molecular mechanisms which underlie HIV-1 latency in microglial cells. Further investigation in this field is needed in order to identify new potential targets in HIV-1 therapy.

TARGETING THE MICROGLIAL CELLS

Targeting all the HIV-1 reservoirs including the microglial cells in the brain is important in order to achieve either a sterilizing or a functional cure. Indeed, these cells are potential sources of HIV-1 reseeding in the blood. In addition, production of the virus in these cells has been associated with HIV-1 resistance and the development of HAND. However, targeting these cells which are located in anatomic and pharmacologic sanctuaries might be very challenging (Marban et al., 2016). Most importantly the access of drugs used in cART is limited by the blood brain barrier. Poor access of the drugs contributes to the persistence of HIV-1 in microglial cells (Cosenza et al., 2002). Another limitation is the existence of a residual neuroinflammation which is responsible for the occurrence of HAND in up to 50% of HIV-1 infected patients. A main concern is to prevent deleterious neuroinflammation associated with infected microglial cells (Rock and Peterson, 2006). To date, three strategies are used to target infected microglial cells (Veenhuis et al., 2019): (i) the Shock and Kill strategy, (ii) the Block and Lock strategy, and (iii) gene therapy. We will discuss in the following section the principles of these three approaches and outline their limitations. Finally, we briefly discuss how these limitations can be circumvented and how to rationalize treatments aiming the eradication or reduction of the pool of latently infected microglial cells.

Strategies to Struggle Latently-Infected Microglial Cells

The Shock and Kill Strategy

The Shock and Kill strategy is based on the reactivation of the latent virus. In this strategy the intensification of cART aims to clear the reservoirs either by the cytopathic effect of the reactivated virus or by the immune system via the actions of cytotoxic T cells (CTLs) (Schwartz et al., 2017). Latency reversing agents (LRAs) used in the reactivation target identified and characterized cellular factors involved in HIV-1 latency (**Figure 3**). Many LRAs are now used *ex vivo* and in clinical trials. However, the main targets are circulating CD4+ T cells and not the microglial cells (Kumar et al., 2015; Spivak and Planelles, 2018). There are currently 160 compounds used as LRAs which belong to two main families or to a third one which includes uncommon drugs with unique or unknown mechanisms (e.g., disulfiram and ixazomid) (Abner and Jordan, 2019). Screening of new LRAs is still a field of intensive research (Richard et al., 2018).

One family of LRAs targets epigenetic mechanisms that occur in latency (Darcis et al., 2016). This class comprises HDAC inhibitors (HDACi) such as Valproic acid, Vorinostat, panobinostat, and romidepsin; Histone methyl transferases inhibitors (HMTi) such as chaetocin and BIX 01294; and DNA methylation inhibitors such as 5-AzadC.

In the second family T cell activating agents are listed. These drugs induce the expression of positive cellular factors and/or their release from inactive complex (Jiang and Dandekar, 2015; Rice, 2016). Prostratin, bryostatin, and ingenol B by activating the PKC pathway release NF-KB and pTEFb from inactive complexes and increase pTEFb expression (Sung and Rice, 2006; Fujinaga et al., 2012; Pandeló José et al., 2014) which leads to HIV-1 reactivation (Darcis et al., 2015). Bromodomain inhibitors such as JQ1 and others are able to release pTEFb from the BRD4-pTEFb complex (Bartholomeeusen et al., 2012; Li et al., 2013; Darcis et al., 2015). Many other compounds including TLR agonists, TNFα, cytokines, and antibodies are also used as LRAs. A comprehensive review of these compounds are described elsewhere (Abner and Jordan, 2019).

However, these drugs used alone failed to fully reactivate HIV-1 expression *ex vivo* (Cary et al., 2016; Darcis et al., 2016). This might reflect the multi factorial mechanisms involved in promoting latency and the stochastic nature of latency (Sengupta and Siliciano, 2018). Another important inability of LRAs relies in their lack of specificity and to the heterogeneity of the reservoirs (Ait-Ammar et al., submitted). Indeed, the LRAs are mainly used in clinical assays to target circulating CD4+ cells but not the other reservoirs such as microglial cells. Therefore, combination of LRAs is now used to circumvent these problems. We expect synergistic effect which improves the efficiency of reactivation and reduces toxicity because of the lower doses used. Moreover, a recent report demonstrated the importance of the timing of the LRA combination treatment in the reactivation of HIV (Bouchat et al., 2016).

Reactivation of reservoirs should be followed by the elimination of the virus (Kim et al., 2018). There is now evidence that eradication of latently-infected reservoirs in patients involves humoral and cell mediated immune responses (Schwartz et al., 2017). Humoral immune response plays an important role in controlling HIV infection (reviewed in Ferrari et al., 2017). Over the past years, a new class of antibodies, called broadly neutralizing antibodies, have been shown to neutralize a wide range of HIV strains (Halper-Stromberg and Nussenzweig, 2016). Preclinical studies in which SHIVs infected macaque monkeys and HIV-1 infected humanized mice were treated by antibodies gave promising results (Halper-Stromberg and Nussenzweig, 2016). The efficacy of the treatment was even enhanced when various broadly neutralizing antibodies in combination was applied (Bruel et al., 2016) or when using multi-specific engineered antibodies like bi and tri-specific antibodies (Ferrari et al., 2016; Sun et al., 2016). Numerous strategies were proposed to enhance cell-mediated immune responses, e.g., CD8+ T cell or natural killer cell activities (Scully and Alter, 2016; Trautmann, 2016). Some methods aim to redirect HIV-specific cell mediated immune responses (Patel et al., 2016). In one strategy, redirected T cells recognizing a range of HIV antigens are expanded *ex vivo*. In another strategy genetically modified lentiviruses expressing artificial T cell receptors (TCRs) or chimeric antigen receptors (CARs) are used in the treatment (Patel et al., 2016). Increasing the specificity and the affinity of the epitopes of the receptors to achieve wider HIV epitope recognition is also realistic (Patel et al., 2016). The possibility to increase HIV-specific CD8+ T cell responses with heterocyclic peptides is currently tested. It is assumed that these peptides stimulate deeper the cell-mediated immune responses than native epitopes (Buhrman et al., 2013).

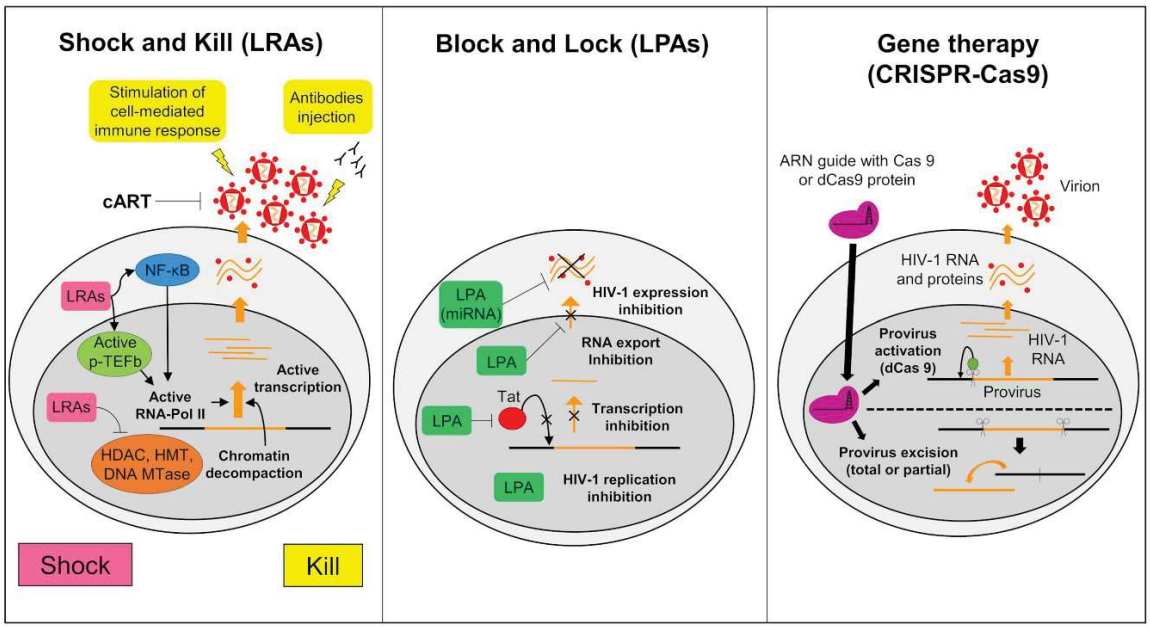


FIGURE 3 | Strategies to struggle latently-infected microglial cells. The shock and kill: the strategy is based on the reactivation of virus expression, in order to eliminate reservoirs and to target latently-infected cells. Latency reversing agents (LRAs) are investigated to reactivate transcription for the "shock." A first family of LRAs are HDAC, HMT, or DNA Mtases inhibitors and target epigenetic mechanisms. These LRAs induce chromatin decompaction and promote virus transcription. Another family of LRAs induce the expression of positive cellular factors (e.g., NF-κB and pTEFB) and/or their recruitment under their active form. These LRAs promote ARN pol II recruitment and favor transcription. The cART is maintained during this phase to clear reservoir without virus propagation in other cells. The "kill" can be enhanced by stimulation of the cell-mediated immune response or by using neutralizing antibodies and/or engineered antibodies. The block and lock: this strategy relates on the induction of a state of deep-latency which prevents any HIV-1 transcription. Several latency-promoting agents (LPAs) are investigated. They inhibit various step of virus expression like replication, transcription by Tat inhibition and RNA export. One promising LPA is a miRNA which inhibit virus expression. The gene therapy (CRISPR-Cas9): this strategy is based on the use of molecular scissors to excise total, or more reasonably, partial sequence of the provirus. Another strategy would be to use deficient Cas9 (dCas9) to carry reactivators or repressors of HIV-1 expression.

An original approach took advantage of the importance of co-stimulatory and co-inhibitory molecules involved in the regulation of T-cell responses (Sharpe and Abbas, 2006). Indeed, treatment by specific antibodies targeting the co-inhibitory molecule PD-1 is thought to reduce the size of the latently-infected reservoir and help to recover CD8+ T cell function from collapse. Finally, therapeutic vaccines might help CTLs to target HIV-1 infected cells derived from latent reservoirs (Deng et al., 2015).

The Block and Lock Strategy

An alternative approach to achieve long term control over HIV-1 infection without the need of cART is to induce a long lasting inhibition of HIV-1 gene expression (**Figure 3**) (Darcis et al., 2017). The molecules used to induce this process are called latency-promoting agents (LPAs). LPAs are expected to suppress the HIV-1 gene expression by inducing a deep latency state (the block) and thus preventing HIV-1 gene transcription (the lock) (Elsheikh et al., 2019). The first LPA identified from a marine sponge is a chemical derivative of corticostatin i.e., Didehydro-corticostatin (dCA). It inhibits Tat activity by preventing Tat-TAR interaction (Mousseau et al., 2015). To date many other LPAs have also been described (reviewed in Castro-Gonzalez et al., 2018). One of them, the ABX464, an inhibitor of Rev which is involved in the RNA export blocks HIV-1 replication *in vitro* as well as in animal models (Campos et al., 2015). Cellular mi RNAs are also potential molecules to be used in the Block and Lock strategy. It has been known for a long time that miRNAs are important modulator of HIV-1 latency (Huang et al., 2007). Screening and computational analysis have allowed the characterization of many miRNAs which impact HIV-1 transcription. One such promising miRNA, i.e., miRNA29a, which inhibits the virus transcription in Jurkat cells should be investigated for its effects in microglial cells (Hariharan et al., 2005). It is suggested that inducing miRNA29a expression in these cells could lead to suppression of HIV-1 gene expression.

The Gene Therapy

Gene therapy has been widely considered as a tool to combat HIV-1. It relies on nuclease-mediated gene editing tools such as the zinc fingers nuclease (ZFN1), the transcription activator-like effector nucleases (TALEN) and the CRISPR/Cas9 technologies (Panfil et al., 2018; Wang et al., 2018). Various gene therapies have been launched since the sterilizing cure of the Berlin patient

(Allers et al., 2011). This patient received transplantation of hematopoietic stem cells from a donor who was homozygous for the CCR5-delta32 mutation (ΔCCR5) conferring resistance to HIV infection. Recently a second patient who shared similar medical contexts, the London patient, was also cured from HIV-1 infection (Gupta R. K. et al., 2019). From this example, the idea to build HIV-1 resistant cells has arisen (Barmania and Pepper, 2013). Clinical trials using ZFN targeting CCR5 are ongoing and deserve attention (Mehta et al., 2017). However, The CRISPR/Cas9 gene editing has become the tool of choice, since ZFN and TALEN are costly and time consuming. This tool uses a guided RNA and a Cas9 nuclease that excise DNA sequences of target specific cellular factors such as CCR5 or HIV-1 DNA thus eliminating the HIV-1 provirus (**Figure 3**). The first targeted DNA sequence by the CRISPR/Cas9 technology was the NF-KB binding site located in the HIV-1 LTR (Ebina et al., 2013). Since then, many other HIV-1 sequences have been targeted (Xiao et al., 2019). This technology may serve to redesign the gene expression of cells such as CTLs so that they target more specifically HIV-1 infected cells. It is believed that by this way the antiviral immunity system could be more efficiently boosted against HIV-1 infected cells and activated reservoirs producing new virions (Mehta et al., 2017). We might even use a variation of the CRISPR/Cas9 technology to reactivate latently infected cells or to induce deep latency (Wang et al., 2018). For this purpose, a defective Cas9 (dCas9) protein is fused to activators to reactivate latently-infected cells (**Figure 3**) or to repressors to suppress HIV-1 expression (Wang et al., 2018). *In vitro* trials suggested a potential application of CRISPR/dCas9 in the reactivation of latent HIV (Zhang et al., 2015). These trials showed reactivation of HIV expression in CD4+ T cells and in microglial cell lines (Cary and Peterlin, 2016). Interestingly, CRISPR/dCas9 when associated with HDAC inhibitors and PKC activators reactivated HIV in a synergistic manner (Limsirichai et al., 2016).

Limitations and Challenges When Targeting Brain Microglial Cells

The current therapeutic strategies have limited efficacy to achieve cure. Their limitations comes from (i) the intrinsic nature of the strategy, (ii) the nature of the cells targeted, or (iii) the attributes of the virus to escape antiviral responses. Regarding microglial cells their specific cell properties and their location in anatomical and pharmacological sanctuaries make them difficult to target. In consequence, the presence of the BBB drastically reduces drug access to currently used drugs in cART (Asahchop et al., 2017). New antiviral drugs with improved CNS penetrance have been introduced in the last years (Veenhuis et al., 2019). However, some showed important side effects. For example, a new integrase inhibitor which showed important CNS uptake (Letendre et al., 2014) was neurotoxic showing severe neuropsychiatric effects (Scheper et al., 2018). Various technics improving delivery of drugs in the CNS and in the CNS-cell types are currently explored. Mechanisms underlying weak drug penetrance are mainly related to differential expression of efflux transporter and to multidrug resistance proteins (Valcour et al., 2011; da Fonseca et al., 2014). Several invasive and non-invasive methods for drug delivery are currently explored including modulation of brain barrier, ultrasound based BBB opening, endogenous transporter, nanoparticles, liposomes, dendrimers… (reviewed in Barnabas, 2019). Original route of drug delivery to the CNS are also explored such as drug delivery from the nose to the brain (Gupta S. et al., 2019).

We should also consider immune-based therapeutics to clear reservoirs activated by the LRAs. Animal models have shown the importance of the CTLs and neutralizing antibodies raised against HIV-1 in immune activation and clearing infected reservoirs in the CNS. Therefore it deserves considerable attention to ensure the feasibility of the Shock and Kill strategy (Brockman et al., 2015; Trautmann, 2016; Subra and Trautmann, 2019).

An important limitation of the use of the Shock and Kill strategy is that reactivation of microglial cells during the Shock therapy leads to neurcinflammation, since the release of pro-inflammatory factors and some viral proteins such as Tat and gp120 during the shock therapy are neurotoxic. Therefore, LRAs induced reactivation of reservoirs has to be associated with drugs preventing the adverse effects of neuroinflammation. One way to prevent inflammation is to improve cART via targeting HIV-1 transcription and/or by inhibiting RNA export in order to counteract the effects of the pro-inflammatory cytokines and to prevent synthesis of viral proteins (Schwartz et al., 2017). Molecules which have anti-inflammatory effects such as curcurmin are also investigated in microglial inflammatory responses (Chen et al., 2018). A new drug formulation using lipid nano-carrier and administration by intra nasal route are currently investigated (Vaz et al., 2017). Remarkably, many drugs targeting the NF-κB pathway (G1V001), the protein Tat (dCA, triptolides…) and rev (ABX4641) induce a deep latency and are thought to be used in the Block and Lock strategy as an alternative to the Shock and Kill strategy. We might use both strategies in a sequential time schedule. First, the Shock and Kill strategy which reactivates and decreases the pool of reservoirs. Followed by application of drugs exploiting the Block and Lock strategy, which suppresses transcription in persisting reservoirs and induces deep latency. Although, it is still an emerging field of research the Block and Lock strategy appears to be more than just an alternative to other strategies. Interestingly, less adverse effects with drugs targeting viral proteins are awaited. We believe that combination of the Block and Lock strategy with the Shock and Kill might allow the reduction of all HIV-1 reservoirs, including the microglial cell reservoir to a level, which does not need any further treatment to achieve a functional cure.

Gene therapy compared to a pharmacological approach such as the Shock and Kill strategy is still in its infancy but appears to be very promising. A recent review compares the advantages and weaknesses of these two approaches i.e., the pharmacological vs. CRISPR/Cas9-based shock strategies (Darcis et al., 2018). Gene therapy, in the same way as pharmacological approaches, does not resolve the problem of drug penetrance in the CNS where microglial cells are located. Indeed, currently used adeno-viruses in *in vitro* essays induce severe inflammation

(Chew et al., 2016). Moreover, another important limitation of gene editing was discovered; HIV-1 can escape CRISPR/cas9 mediated suppression. Indeed, following an initial attempt to establish proof of concept of the use of gene editing targeting specifically HIV-1 (Kaminski et al., 2016), it soon appeared that the virus was able to subvert the DNA repair machinery to evolve quickly into CRISPR/Cas9 resistant strains (Liang et al., 2016; Wang G. et al., 2016; Wang Z. et al., 2016; Yoder and Bundschuh, 2016). On the other hand combination of guided RNAs targeting different HIV-1 specific sequences may drastically reduce the occurrence of HIV-1 escape (Wang et al., 2018). A major off-target of this gene editing tool is the increase of undesirable gene mutations and chromosomal translocations (Yee, 2016). Efforts in reducing these off-target effects are required in order to improve this elegant and attractive approach. For example, bioinformatics could help to select the correct guided RNAs that avoid unexpected DNA cleavage (Zhu et al., 2017). Another limitation of the use of gene editing technology is that it has been mainly tested *in vitro* and mainly in CD4+ T cells. Experiments conducted *ex vivo* showed that redirected immune cells reached only a tiny fraction of the HIV-1 reservoirs. Therefore, this technology does not make possible to drastically reduce the number of infected cells, which is required to achieve functional cure (Wang et al., 2018).

Rationalization of these strategies is required to optimize either of these strategies to target microglial cells. Since we cannot make brain biopsies due to ethical concerns, it is critical to design these strategies and test their feasibility in adequate animal models. Indeed, the efficacy and the pharmacokinetic and pharmacodynamic characteristics of drugs are tested only in preclinical studies. Non-human primates (NHP) and humanized mice are currently the best-suited animal models to study either strategies. However, there are many ways to evaluate the efficiency of a therapy and to evaluate HIV infection in the brain. The search of specific biomarkers for latently infected microglial cells may lead to the development of original strategies, which specifically target brain reservoirs. One such specific marker of the CD4+ T cell reservoir, CD32a, has been described (Descours et al., 2017), although its specificity is debated (Badia et al., 2018). Much research has been focused on finding new specific biomarkers for brain reservoirs, nonetheless their use will be very challenging. Indeed, it will be difficult to target brain reservoirs *ex vivo*, even though it is technically possible for blood cells, and very difficult to target them *in vivo* due to their paucity and localization in the entire brain parenchyma. For example, the invasive techniques used to evaluate the evolution of biomarkers in the CSF such as neopterin, CTIP2, TREM2, and NFL after a treatment cannot be performed too frequently (Jessen Krut et al., 2014). Non-invasive methods such as neuroimaging are therefore needed for brain evaluation (discussed in Garrido and Margolis, 2015). Comprehensive neuropsychological testing is also required to detect subclinical deficits (Haddow et al., 2013). It will be essential to evaluate the efficiency of each strategy that target microglial cells. For this purpose, highly sensitive methods have been developed such as the single copy assay (SCA) which allows the detection of HIV RNA in the CSF from infected patients on cART or from elite controllers whose HIV RNA level was initially undetectable in the plasma and the CSF (Dahl et al., 2013).

CONCLUSION

HIV cure is theoretically possible but not yet feasible with current approaches. Complete cure demands targeting not only the CD4+ T cells reservoirs but also all the other potential reservoirs located in sanctuaries with low drug penetrance. This is particularly true for the microglial cells in the CNS which is considered as the main reservoir for HIV. Indeed, due to their properties (slow turn over, long half-life…), these cells are thought to persist lifelong. Other true reservoirs such as macrophages may be eliminated by intensifying cART and/or by targeting HIV infected monocytes which is the source of infection. Other reservoirs such as astrocytes or the newly identified pericytes (Bertrand et al., 2019) may constitute true reservoirs, but these are still debated and deserve far more investigations. Currently, three strategies are considered, which allow functional cure: two pharmacological approaches including the shock and kill strategy and the alternative block and lock strategy and the CRISPR/cas9 based technology, which is a gene editing tool based strategy. However, the Shock and Kill strategy is not suitable to eliminate reservoirs like microglial cells, since the reactivation of microglial cell reservoirs leads to neuroinflammation, which is the origin of HAND. The search for potential epigenetic regulators that control microglial cells are therefore needed. The most investigated epigenetic regulators are the long non-coding RNAs (Qu et al., 2019) and the mi RNAs (reviewed in Cheray and Joseph, 2018). However, due to the stochasticity of HIV-1 transcription the Shock and Kill strategy does not allow reactivation of all latently infected microglial cells. Alternative strategies are therefore considered. The Block and Lock strategy has the advantage that the risk of inflammation in the brain is drastically reduced. More importantly, this strategy since it targets transcription and/or RNA export might be used following the Shock and Kill strategy. The combination of strategies should lead to the decrease of the pool of the microglial reservoir. It may also induce deep latency in reservoirs that are not reactivated by LRAs. The gene editing tool, such as the CRISPR/Cas9 technology, because it operates on specific sequences is also very promising. Still in its infancy, this elegant approach can target either cellular factors, which are involved in HIV-1 resistance, or viral factors. This technology can be used either to reactivate the virus or to excise and eliminate the provirus from its host genome. However, several challenges have to be overcome such as *in vivo* delivery in the brain and long-term toxicity before its use in the clinic and thus needs far more investigations. An original way to deliver CRISPR/cas9 guided RNA across the BBB has been developed recently and hold promises (Kaushik et al., 2019). Notably, it was shown that this non-invasive method which uses magnetically guided delivery of RNAs inhibits latent infection of microglial cells and cross the BBB in *in vitro* models.

In conclusion, targeting microglial cells is crucial but might be very challenging. We have to keep in mind that these cells are resistant to the cytopathic effects of the virus and resistant to apoptosis. Thus strategies aiming to induce apoptosis in infected microglial cells have to be developed (Le Douce et al., 2010). It appears that the cellular factor CTIP2 (Bcl11b) has a pivotal role in the establishment and persistence of HIV latency, in the genesis of a favorable environment for HIV latency and in the resistance to apoptosis (Le Douce et al., 2014). CTIP2 silencing by RNA interference has been proposed in cancer since downregulation of CTIP2 leads to apoptosis in malignant cells but not in normal mature cells (Grabarczyk et al., 2007; Huang et al., 2007). We might also design drugs targeting CTIP2 interacting proteins specifically involved in HIV-1 latency and apoptosis resistance. Moreover, establishing a protein interaction network focused on CTIP2 interactions in microglial cells will allow the identification and characterization of new actors involved in HIV-1 latency and apoptosis resistance.

AUTHOR CONTRIBUTIONS

CS and OR wrote the review. CW designed the figures and wrote the legends. CW, MD, JV, FD, SD, VG, PM, AM, and CV participated in the writing of the review.

FUNDING

This work in OR and CS laboratory was supported by grants from the Agence Nationale de Recherches sur le SIDA (ANRS) to OR and CS, from Sidaction to OR and CS. Work in CV lab was funded by the Belgian Fund for Scientific Research (FRS-FNRS, Belgium), the Fondation Roi Baudouin, the NEAT program (Networking to Enhance the Use of Economics in Animal Health Education, Research and Policy Making), the Walloon Region (Fonds de Maturation), Les Amis des Instituts Pasteur à Bruxelles, asbl., and the University of Brussels [Action de Recherche Concertée (ARC) grant]. The laboratory of CV is part of the ULB-Cancer Research Centre (U-CRC). The work in VG laboratory was supported by University College Dublin. SD, VG, PM, AM, OR, and CV have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 691119-EU4HIVCURE-H2020-MSCA-RISE-2015.

ACKNOWLEDGMENTS

We are grateful to Andrea Janossy for a careful and critical reading of the manuscript.

REFERENCES

Abassi, M., Morawski, B. M., Nakigozi, G., Nakasujja, N., Kong, X., Meya, D. B., et al. (2017). Cerebrospinal fluid biomarkers and HIV-associated neurocognitive disorders in HIV-infected individuals in Rakai, Uganda. *J. Neurovirol.* 23, 369–375. doi: 10.1007/s13365-016-0505-9

Abner, E., and Jordan, A. (2019). HIV "shock and kill" therapy: in need of revision. *Antiviral Res.* 166, 19–34. doi: 10.1016/j.antiviral.2019.03.008

Abreu, C., Shirk, E. N., Queen, S. E., Mankowski, J. L., Gama, L., and Clements, J. E. (2019). A quantitative approach to SIV functional latency in brain macrophages. *J. Neuroimmune Pharmacol.* 14, 23–32. doi: 10.1007/s11481-018-9803-8

Al-Harti, L., Joseph, J., and Nath, A. (2018). Astrocytes as an HIV CNS reservoir: highlights and reflections of an NIMH-sponsored symposium. *J. Neurovirol.* 24, 665–669. doi: 10.1007/s13365-018-0691-8

Allers, K., Hutter, G., Hofmann, J., Loddenkemper, C., Rieger, K., Thiel, E., et al. (2011). Evidence for the cure of HIV infection by CCR5 32/ 32 stem cell transplantation. *Blood* 117, 2791–2799. doi: 10.1182/blood-2010-09-309591

Alvarez-Carbonell, D., Garcia-Mesa, Y., Milne, S., Das, B., Dobrowolski, C., Rojas, R., et al. (2017). Toll-like receptor 3 activation selectively reverses HIV latency in microglial cells. *Retrovirology* 14:9. doi: 10.1186/s12977-017-0335-8

Alvarez-Carbonell, D., Ye, F., Ramanath,N., Dobrowolski, C., and Karn, J. (2019). The glucocorticoid receptor is a critical regulator of HIV latency in human microglial cells. *J. Neuroimmune Pharmacol.* 14, 94–109. doi: 10.1007/s11481-018-9798-1

Asahchop, E. L., Meziane, O., Mamik, M. K., Chan, W. F., Branton, W. G., Resch, L., et al. (2017). Reduced antiretroviral drug efficacy and concentration in HIV-infected microglia contributes to viral persistence in brain. *Retrovirology* 14:47. doi: 10.1186/s12977-017-0370-5

Avalos, C. R., Abreu, C. M., Queen, S. E., Li, M., Price, S., Shirk, E. N., et al. (2017). Brain macrophages in simian immunodeficiency virus-infected, antiretroviral-suppressed macaques: a functional latent reservoir. *MBio* 8:e01186-17. doi: 10.1128/mBio.01186-17

Avram, D., Fields, A., Pretty On Top, K., Nevrivy, D. J., Ishmael, J. E., and Leid, M. (2000). Isolation of a novel family of C(2)H(2) zinc finger proteins implicated in transcriptional repression mediated by chicken ovalbumin upstream promoter transcription factor (COUP-TF) orphan nuclear receptors. *J. Biol. Chem.* 275, 10315–10322. doi: 10.1074/jbc.275.14.10315

Avram, D., Fields, A., Senawong, T., Topark-Ngarm, A., and Leid, M. (2002). COUP-TF (chicken ovalbumin upstream promoter transcription factor)-interacting protein 1 (CTIP1) is a sequence-specific DNA binding protein. *Biochem. J.* 368, 555–563. doi: 10.1042/bj20020496

Badia, R., Ballana, E., Castellví, M., García-Vidal, E., Pujantell, M., Clotet, B., et al. (2018). CD32 expression is associated to T-cell activation and is not a marker of the HIV-1 reservoir. *Nat. Commun.* 9:2739. doi: 10.1038/s41467-018-05157-w

Balasubramaniam, M., Pandhare, J., and Dash, C. (2018). Are microRNAs important players in HIV-1 infection? An update. *Viruses* 10:E110. doi: 10.3390/v10030110

Barber, S. A., Gama, L., Dudaronek, J. M., Voelker, T., Tarwater, P. M., and Clements, J. E. (2006). Mechanism for the establishment of transcriptional HIV latency in the brain in a simian immunodeficiency virus–macaque model. *J. Infect. Dis.* 193, 963–970. doi: 10.1086/500983

Barmania, F., and Pepper, M. S. (2013). C-C chemokine receptor type five (CCR5): an emerging target for the control of HIV infection. *Appl. Transl. Genomics* 2, 3–16. doi: 10.1016/j.atg.2013.05.004

Barnabas, W. (2019). Drug targeting strategies into the brain for treating neurological diseases. *J. Neurosci. Methods* 311, 133–146. doi: 10.1016/j.jneumeth.2018.10.015

Bartholomeeusen, K., Xiang, Y., Fujinaga, K., and Peterlin, B. M. (2012). Bromodomain and Extra-terminal (BET) bromodomain inhibition activate transcription via transient release of positive transcription elongation factor b (P-TEFb) from 7SK small nuclear ribonucleoprotein. *J. Biol. Chem.* 287, 36609–36616. doi: 10.1074/jbc.M112.410746

Bavaro, D. F., Calamo, A., Lepore, L., Fabrizio, C., Saracino, A., Angarano, G., et al. (2019). Cerebrospinal fluid compartmentalization of HIV-1 and correlation with plasma viral load and blood–brain barrier damage. *Infection* 47, 441–446. doi: 10.1007/s15010-019-01268-8

Baxter, A. E., Russell, R. A., Duncan, C. J. A., Moore, M. D., Willberg, C. B., Pablos, J. L., et al. (2014). Macrophage infection via selective capture of HIV-1-infected $CD4^+$ T cells. *Cell Host Microbe* 16, 711–721. doi: 10.1016/j.chom.2014.10.010

Bednar, M. M., Sturdevant, C. B., Tompkins, L. A., Arrildt, K. T., Dukhovlinova, E., Kincer, L. P., et al. (2015). Compartmentalization, viral evolution, and viral latency of HIV in the CNS. *Curr. HIV/AIDS Rep.* 12, 262–271. doi: 10.1007/s11904-015-0265-9

Bertrand, L., Cho, H. J., and Toborek, M. (2019). Blood–brain barrier pericytes as a target for HIV-1 infection. *Brain* 142, 502–511. doi: 10.1093/brain/awy339

Bilimoria, P. M., and Stevens, B. (2015). Microglia function during brain development: new insights from animal models. *Brain Res.* 1617, 7–17. doi: 10.1016/j.brainres.2014.11.032

Blankson, J. N., Persaud, D., and Siliciano, R. F. (2002). The challenge of viral reservoirs in HIV-1 infection. *Annu. Rev. Med.* 53, 557–593. doi: 10.1146/annurev.med.53.082901.104024

Bouchat, S., Delacourt,N., Kula, A., Darcis, G., Van Driessche, B., Corazza, F., et al. (2016). Sequential treatment with 5-aza-2'-deoxycytidine and deacetylase inhibitors reactivates HIV-1. *EMBO Mol. Med.* 8, 117–138. doi: 10.15252/emmm.201505557

Brockman, M. A., Jones, R. B., and Brumme, Z. L. (2015). Challenges and opportunities for T-cell-mediated strategies to eliminate HIV reservoirs. *Front. Immunol.* 6:506. doi: 10.3389/fimmu.2015.00506

Bruel, T., Guivel-Benhassine, F., Amraoui, S., Malbec, M., Richard, L., Bourdic, K., et al. (2016). Elimination of HIV-1-infected cells by broadly neutralizing antibodies. *Nat. Commun.* 7:10844. doi: 10.1038/ncomms10844

Buhrman, J. D., Jordan, K. R., Munson, D. J., Moore, B. L., Kappler, J. W., and Slansky, J. E. (2013). Improving antigenic peptide vaccines for cancer immunotherapy using a dominant tumor-specific T cell receptor. *J. Biol. Chem.* 288, 33213–33225. doi: 10.1074/jbc.M113.509554

Burdo, T. H., Lackner, A., and Williams, K. C. (2013). Monocyte/macrophages and their role in HIV neuropathogenesis. *Immunol. Rev.* 254, 102–113. doi: 10.1111/imr.12068

Campos, N., Myburgh, R., Garcel, A., Vautrin, A., Lapasset, L., Nadal, E. S., et al. (2015). Long lasting control of viral rebound with a new drug ABX464 targeting Rev – mediated viral RNA biogenesis. *Retrovirology* 12:30. doi: 10.1186/s12977-015-0159-3

Canestri, A., Lescure, F., Jaureguiberry, S., Moulignier, A., Amiel, C., Marcelin, A. G., et al. (2010). Discordance between cerebral spinal fluid and plasma HIV replication in patients with neurological symptoms who are receiving suppressive antiretroviral therapy. *Clin. Infect. Dis.* 50, 773–778. doi: 10.1086/650538

Cao, Y., Lei, X., Ribeiro, R. M., Perelson, A. S., and Liang, J. (2018). Probabilistic control of HIV latency and transactivation by the Tat gene circuit. *Proc. Natl. Acad. Sci. U S A.* 115, 12453–12458. doi: 10.1073/pnas.1811195115

Cara, A., and Klotman, M. E. (2006). Retroviral E-DNA: persistence and gene expression in nondividing immune cells. *J. Leukoc. Biol.* 80, 1013–1017. doi: 10.1189/jlb.0306151

Cary, D. C., Fujinaga, K., and Peterlin, B. M. (2016). Molecular mechanisms of HIV latency. *J. Clin. Invest.* 126, 448–454. doi: 10.1172/JCI80565

Cary, D. C., and Peterlin, B. M. (2016). Targeting the latent reservoir to achieve functional HIV cure. *F1000Res* 5:F1000. doi: 10.12688/f1000research.8109.1

Castellano, P., Prevedel, L., and Eugenin, E. A. (2017). HIV-infected macrophages and microglia that survive acute infection become viral reservoirs by a mechanism involving Bim. *Sci. Rep.* 7:12866. doi: 10.1038/s41598-017-12758-w

Castro-Gonzalez, S., Colomer-Lluch, M., and Serra-Moreno, R. (2018). Barriers for HIV cure: the latent reservoir. *AIDS Res. Hum. Retroviruses* 34, 739–759. doi: 10.1089/aid.2018.0118

Cenker, J. J., Stultz, R. D., and McDonald, D. (2017). Brain Microglial cells are highly susceptible to HIV-1 infection and spread. *AIDS Res. Hum. Retroviruses* 33, 1155–1165. doi: 10.1089/aid.2017.0004

Chen, G., Liu, S., Pan, R., Li, G., Tang, H., Jiang, M., et al. (2018). Curcumin attenuates gp120-induced microglial inflammation by inhibiting autophagy via the PI3K pathway. *Cell. Mol. Neurobiol.* 38, 1465–1477. doi: 10.1007/s10571-018-0616-3

Cheray, M., and Joseph, B. (2018). Epigenetics control microglia plasticity. *Front. Cell. Neurosci.* 12:243. doi: 10.3389/fncel.2018.00243

Cherrier, T., Le Douce, V., Eilebrecht, S., Riclet, R., Marban, C., Dequiedt, F., et al. (2013). CTIP2 is a negative regulator of P-TEFb. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12655–12660. doi: 10.1073/pnas.1220136110

Cherrier, T., Suzanne, S., Redel, L., Calao, M., Marban, C., Samah, B., et al. (2009). p21WAF1 gene promoter is epigenetically silenced by CTIP2 and SUV39H1. *Oncogene* 28, 3380–3389. doi: 10.1038/onc.2009.193

Chew, W. L., Tabebordbar, M., Cheng, J. K. W., Mali, P., Wu, E. Y., Ng, A. H. M., et al. (2016). A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat. Methods* 13, 868–874. doi: 10.1038/nmeth.3993

Chivero, E. T., Guo, M.-L., Periyasamy, P., Liao, K., Callen, S. E., and Buch, S. (2017). HIV-1 Tat primes and activates microglial NLRP3 inflammasome-mediated neuroinflammation. *J. Neurosci.* 37, 3599–3609. doi: 10.1523/JNEUROSCI.3045-16.2017

Chun, T. W., Davey, R. T., Ostrowski, M., Shawn Justement, J., Engel, D., Mullins, J. I., et al. (2000). Relationship between pre-existing viral reservoirs and the re-emergence of plasma viremia after discontinuation of highly active anti-retroviral therapy. *Nat. Med.* 6, 757–761. doi: 10.1038/77481

Churchill, M. J., Cowley, D. J., Wesselingh, S. L., Gorry, P. R., and Gray, L. R. (2015). HIV-1 transcriptional regulation in the central nervous system and implications for HIV cure research. *J. Neurovirol.* 21, 290–300. doi: 10.1007/s13365-014-0271-5

Churchill, M. J., Gorry, P. R., Cowley, D., Lal, L., Sonza, S., Purcell, D. F., et al. (2006). Use of laser capture microdissection to detect integrated HIV-1 DNA in macrophages and astrocytes from autopsy brain tissues. *J. Neurovirol.* 12, 146–152. doi: 10.1080/13550280600748946

Churchill, M. J., Wesselingh, S. L., Cowley, D., Pardo, C. A., McArthur, J. C., Brew, B. J., et al. (2009). Extensive astrocyte infection is prominent in human immunodeficiency virus-associated dementia. *Ann. Neurol.* 66, 253–258. doi: 10.1002/ana.21697

Colonna, M., and Butovsky, O. (2017). Microglia function in the central nervous system during health and neurodegeneration. *Annu. Rev. Immunol.* 35, 441–468. doi: 10.1146/annurev-immunol-051116-052358

Cosenza, M. A., Zhao, M.-L., Si, Q., and Lee, S. C. (2002). Human brain parenchymal microglia express CD14 and CD45 and are productively infected by HIV-1 in HIV-1 encephalitis. *Brain Pathol.* 12, 442–455. doi: 10.1111/j.1750-3639.2002.tb00461.x

Cribier, A., Descours, B., Valadão, A. L. C., Laguette,N., and Benkirane, M. (2013). Phosphorylation of SAMHD1 by Cyclin A2/CDK1 regulates its restriction activity toward HIV-1. *Cell Rep.* 3, 1036–1043. doi: 10.1016/j.celrep.2013.03.017

Cysique, L. A., Jugé, L., Lennon, M. J., Gates, T. M., Jones, S. P., Lovelace, M. D., et al. (2019). HIV brain latency as measured by CSF BcL11b relates to disrupted brain cellular energy in virally suppressed HIV infection. *AIDS* 33, 433–441. doi: 10.1097/QAD.0000000000002076

da Fonseca, A. C. C., Matias, D., Garcia, C., Amaral, R., Geraldo, L. H., Freitas, C., et al. (2014). The impact of microglial activation on blood-brain barrier in brain diseases. *Front. Cell. Neurosci.* 8:362. doi: 10.3389/fncel.2014.00362

Dahl, V., Gisslen, M., Hagberg, L., Peterson, J., Shao, W., Spudich, S., et al. (2014). An example of genetically distinct HIV type 1 variants in cerebrospinal fluid and plasma during suppressive therapy. *J. Infect. Dis.* 209, 1618–1622. doi: 10.1093/infdis/jit805

Dahl, V., Peterson, J., Spudich, S., Lee, E., Shacklett, B. L., Price, R. W., et al. (2013). Single-copy assay quantification of HIV-1 RNA in paired cerebrospinal fluid and plasma samples from elite controllers. *AIDS* 27, 1145–1149. doi: 10.1097/QAD.0b013e32835cf235

Darcis, G., Das, A., and Berkhout, B. (2018). Tackling HIV persistence: pharmacological versus CRISPR-based shock strategies. *Viruses* 10:157. doi: 10.3390/v10040157

Darcis, G., Kula, A., Bouchat, S., Fujinaga, K., Corazza, F., Ait-Ammar, A., et al. (2015). An in-depth comparison of latency-reversing agent combinations in various *in vitro* and *ex vivo* HIV-1 latency models identified bryostatin-1+JQ1 and ingenol-B+JQ1 to potently reactivate viral gene expression. *PLoS Pathog.* 11:e1005063. doi: 10.1371/journal.ppat.1005063

Darcis, G., Van Driessche, B., and Van Lint, C. (2016). Preclinical shock strategies to reactivate latent HIV-1: an update. *Curr. Opin. HIV AIDS* 11, 388–393. doi: 10.1097/COH.0000000000000288

Darcis, G., Van Driessche, B., and Van Lint, C. (2017). HIV latency: should we shock or lock? *Trends Immunol.* 38, 217–228. doi: 10.1016/j.it.2016.12.003

Davey, R. T., Bhat, N., Yoder, C., Chun, T. W., Metcalf, J. A., Dewar, R., et al. (1999). HIV-1 and T cell dynamics after interruption of highly active antiretroviral therapy (HAART) in patients with a history of sustained viral suppression. *Proc. Natl. Acad. Sci. U.S.A.* 96, 15109–15114. doi: 10.1073/pnas.96.26.15109

Dello Russo, C., Cappoli, N., Coletta, I., Mezzogori, D., Paciello, F., Pozzoli, G., et al. (2018). The human microglial HMC3 cell line: where do we stand? A systematic literature review. *J. Neuroinflammation* 15:259. doi: 10.1186/s12974-018-1288-0

Deng, K., Pertea, M., Rongvaux, A., Wang, L., Durand, C. M., Ghiaur, G., et al. (2015). Broad CTL response is required to clear latent HIV-1 due to dominance of escape mutations. *Nature* 517, 381–385. doi: 10.1038/nature14053

Descours, B., Petitjean, G., López-Zaragoza, J.-L., Bruel, T., Raffel, R., Psomas, C., et al. (2017). CD32a is a marker of a CD4 T-cell HIV reservoir harbouring replication-competent proviruses. *Nature* 543, 564–567. doi: 10.1038/nature21710

Desplats, P., Dumaop, W., Smith, D., Adame, A., Everall, I., Letendre, S., et al. (2013). Molecular and pathologic insights from latent HIV-1 infection in the human brain. *Neurology* 80, 1415–1423. doi: 10.1212/WNL.0b013e31828c2e9e

Dornadula, G., Zhang, H., VanUitert, B., Stern, J., Livornese, L., Ingerman, M. J., et al. (1999). Residual HIV-1 RNA in blood plasma of patients taking suppressive highly active antiretroviral therapy. *JAMA* 282, 1627–1632. doi: 10.1001/jama.282.17.1627

Ebina, H., Misawa,N., Kanemura, Y., and Koyanagi, Y. (2013). Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. *Sci. Rep.* 3:2510. doi: 10.1038/srep02510

Edén, A., Fuchs, D., Hagberg, L., Nilsson, S., Spudich, S., Svennerholm, B., et al. (2010). HIV-1 viral escape in cerebrospinal fluid of subjects on suppressive antiretroviral treatment. *J. Infect. Dis.* 202, 1819–1825. doi: 10.1086/657342

Edén, A., Nilsson, S., Hagberg, L., Fuchs, D., Zetterberg, H., Svennerholm, B., et al. (2016). Asymptomatic cerebrospinal fluid HIV-1 viral blips and viral escape during antiretroviral therapy: a longitudinal study. *J. Infect. Dis.* 214, 1822–1825. doi: 10.1093/infdis/jiw454

Eggers, C., Arendt, G., Hahn, K., Husstedt, I. W., Maschke, M., Neuen-Jacob, E., et al. (2017). HIV-1-associated neurocognitive disorder: epidemiology, pathogenesis, diagnosis, and treatment. *J. Neurol.* 264, 1715–1727. doi: 10.1007/s00415-017-8503-2

Eilebrecht, S., Le Douce, V., Riclet, R., Targat, B., Hallay, H., Van Driessche, B., et al. (2014). HMGA1 recruits CTIP2-repressed P-TEFb to the HIV-1 and cellular target promoters. *Nucleic Acids Res.* 42, 4962–4971. doi: 10.1093/nar/gku168

Eilebrecht, S., Wilhelm, E., Benecke, B.-J., Bell, B., and Benecke, A. G. (2013). HMGA1 directly interacts with TAR to modulate basal and Tat-dependent HIV transcription. *RNA Biol.* 10, 436–444. doi: 10.4161/rna.23686

Eisele, E., and Siliciano, R. F. (2012). Redefining the viral reservoirs that prevent HIV-1 eradication. *Immunity* 37, 377–388. doi: 10.1016/j.immuni.2012.08.010

Elsheikh, M. M., Tang, Y., Li, D., and Jiang, G. (2019). Deep latency: a new insight into a functional HIV cure. *EBioMedicine* 45, 624–629. doi: 10.1016/j.ebiom.2019.06.020

Ferrari, G., Haynes, B. F., Koenig, S., Nordstrom, J. L., Margolis, D. M., and Tomaras, G. D. (2016). Envelope-specific antibodies and antibody-derived molecules for treating and curing HIV infection. *Nat. Rev. Drug Discov.* 15, 823–834. doi: 10.1038/nrd.2016.173

Ferrari, G., Pollara, J., Tomaras, G. D., and Haynes, B. F. (2017). Humoral and innate antiviral immunity as tools to clear persistent HIV infection. *J. Infect. Dis.* 215, S152–S159. doi: 10.1093/infdis/jiw555

Ferretti, F., Gisslen, M., Cinque, P., and Price, R. W. (2015). Cerebrospinal fluid HIV escape from antiretroviral therapy. *Curr. HIV/AIDS Rep.* 12, 280–288. doi: 10.1007/s11904-015-0267-7

Forouzanfar, F., Ali, S., Wallet, C., Ducloy, C., El Mekdad, H., and El Maassarani, M., et al. (2019). HIV-1 Vpr mediates the depletion of the cellular repressor CTIP2 to counteract viral gene silencing. *Sci. Rep.* 9:13154. doi: 10.1038/s41598-019-48689-x

Fujinaga, K., Barboric, M., Li, Q., Luo, Z., Price, D. H., and Peterlin, B. M. (2012). PKC phosphorylates HEXIM1 and regulates P-TEFb activity. *Nucleic Acids Res.* 40, 9160–9170. doi: 10.1093/nar/gks682

Gama, L., Abreu, C., Shirk, E. N., Queen, S. E., Beck, S. E., Metcalf Pate, K. A., et al. (2018). SIV latency in macrophages in the CNS. *Curr. Top. Microbiol. Immunol.* 417, 111–130. doi: 10.1007/82_2018_89

Gama, L., Abreu, C. M., Shirk, E. N., Price, S. L., Li, M., Laird, G. M., et al. (2017). Reactivation of simian immunodeficiency virus reservoirs in the brain of virally suppressed macaques. *AIDS* 31, 5–14. doi: 10.1097/QAD.0000000000001267

García, F., Plana, M., Vidal, C., Cruceta, A., O'Brien, W. A., Pantaleo, G., et al. (1999). Dynamics of viral load rebound and immunological changes after stopping effective antiretroviral therapy. *AIDS* 13, F79–F86. doi: 10.1097/00002030-199907300-00002

García, M., Buzón, M. J., Benito, J. M., and Rallón,N. (2018). Peering into the HIV reservoir. *Rev. Med. Virol.* 28:e1981. doi: 10.1002/rmv.1981

Garcia-Mesa, Y., Jay, T. R., Checkley, M. A., Luttge, B., Dobrowolski, C., Valadkhan, S., et al. (2017). Immortalization of primary microglia: a new platform to study HIV regulation in the central nervous system. *J. Neurovirol.* 23, 47–66. doi: 10.1007/s13365-016-0499-3

Garrido, C., and Margolis, D. M. (2015). Translational challenges in targeting latent HIV infection and the CNS reservoir problem. *J. Neurovirol.* 21, 222–226. doi: 10.1007/s13365-014-0269-z

Ginsberg, S. D., Alldred, M. J., Gunnam, S. M., Schiroli, C., Lee, S. H., Morgello, S., et al. (2018). Expression profiling suggests microglial impairment in human immunodeficiency virus neuropathogenesis. *Ann. Neurol.* 83, 406–417. doi: 10.1002/ana.25160

Gisslén, M., Heslegrave, A., Veleva, E., Yilmaz, A., Andersson, L.-M., Hagberg, L., et al. (2019). CSF concentrations of soluble TREM2 as a marker of microglial activation in HIV-1 infection. *Neurol. Neuroimmunol. Neuroinflamm.* 6:e512. doi: 10.1212/NXI.0000000000000512

Gomez Perdiguero, E., Klapproth, K., Schulz, C., Busch, K., Azzoni, E., Crozet, L., et al. (2015). Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. *Nature* 518, 547–551. doi: 10.1038/nature13989

Gorry, P. R., Ong, C., Thorpe, J., Bannwarth, S., Thompson, K. A., Gatignol, A., et al. (2003). Astrocyte infection by HIV-1: mechanisms of restricted virus replication, and role in the pathogenesis of HIV-1-associated dementia. *Curr. HIV Res.* 1, 463–473. doi: 10.2174/1570162033485122

Grabarczyk, P., Przybylski, G. K., Depke, M., Völker, U., Bahr, J., Assmus, K., et al. (2007). Inhibition of BCL11B expression leads to apoptosis of malignant but not normal mature T cells. *Oncogene* 26, 3797–3810. doi: 10.1038/sj.onc.1210152

Gray, L. R., Cowley, D., Welsh, C., Lu, H. K., Brew, B. J., Lewin, S. R., et al. (2016). CNS-specific regulatory elements in brain-derived HIV-1 strains affect responses to latency-reversing agents with implications for cure strategies. *Mol. Psychiatry* 21, 574–584. doi: 10.1038/mp.2015.111

Gray, L. R., Roche, M., Flynn, J. K., Wesselingh, S. L., Gorry, P. R., and Churchill, M. J. (2014). Is the central nervous system a reservoir of HIV-1? *Curr. Opin. HIV AIDS* 9, 552–558. doi: 10.1097/COH.0000000000000108

Gupta, R. K., Abdul-Jawad, S., McCoy, L. E., Mok, H. P., Peppa, D., Salgado, M., et al. (2019). HIV-1 remission following CCR5Δ32/Δ32 haematopoietic stem-cell transplantation. *Nature* 568, 244–248. doi: 10.1038/s41586-019-1027-4

Gupta, S., Kesarla, R., and Omri, A. (2019). Approaches for CNS delivery of drugs – nose to brain targeting of antiretroviral agents as a potential attempt for complete elimination of major reservoir site of HIV to aid AIDS treatment. *Expert Opin. Drug Deliv.* 16, 287–300. doi: 10.1080/17425247.2019.1583206

Haddow, L. J., Floyd, S., Copas, A., and Gilson, R. J. C. (2013). A systematic review of the screening accuracy of the HIV dementia scale and international HIV dementia scale. *PLoS ONE* 8:e61826. doi: 10.1371/journal.pone.0061826

Halper-Stromberg, A., and Nussenzweig, M. C. (2016). Towards HIV-1 remission: potential roles for broadly neutralizing antibodies. *J. Clin. Invest.* 126, 415–423. doi: 10.1172/JCI80561

Han, Y., Lin, Y. B., An, W., Xu, J., Yang, H.-C., O'Connell, K., et al. (2008). Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. *Cell Host Microbe* 4, 134–146. doi: 10.1016/j.chom.2008.06.008

Hansen, E. C., Ransom, M., Hesselberth, J. R., Hosmane, N. N., Capoferri, A. A., Bruner, K. M., et al. (2016). Diverse fates of uracilated HIV-1 DNA during infection of myeloid lineage cells. *Elife* 5:e18447. doi: 10.7554/eLife.18447

Hariharan, M., Scaria, V., Pillai, B., and Brahmachari, S. K. (2005). Targets for human encoded microRNAs in HIV genes. *Biochem. Biophys. Res. Commun.* 337, 1214–1218. doi: 10.1016/j.bbrc.2005.09.183

Honeycutt, J. B., Thayer, W. O., Baker, C. E., Ribeiro, R. M., Lada, S. M., Cao, Y., et al. (2017). HIV persistence in tissue macrophages of humanized

Murooka, T. T., Deruaz, M., Marangoni, F., Vrbanac, V. D., Seung, E., von Andrian, U. H., et al. (2012). HIV-infected T cells are migratory vehicles for viral dissemination. *Nature* 490, 283–287. doi: 10.1038/nature11398

Nagai, A., Nakagawa, E., Hatori, K., Choi, H. B., McLarnon, J. G., Lee, M. A., et al. (2001). Generation and characterization of immortalized human microglial cell lines: expression of cytokines and chemokines. *Neurobiol. Dis.* 8, 1057–1068. doi: 10.1006/nbdi.2001.0437

Nolting, T., Lindecke, A., Hartung, H.-P., Koutsilieri, E., Maschke, M., Husstedt, I.-W., et al. (2012). Cytokine levels in CSF and neuropsychological performance in HIV patients. *J. Neurovirol.* 18, 157–161. doi: 10.1007/s13365-012-0091-4

Osipova, E. D., Semyachkina-Glushkovskaya, O. V., Morgun, A. V., Pisareva, N. V., Malinovskaya, N. A., Boitsova, E. B., et al. (2018). Gliotransmitters and cytokines in the control of blood-brain barrier permeability. *Rev. Neurosci.* 29, 567–591. doi: 10.1515/revneuro-2017-0092

Pai, A., and Weinberger, L. S. (2017). Fate-regulating circuits in viruses: from discovery to new therapy targets. *Annu. Rev. Virol.* 4, 469–490. doi: 10.1146/annurev-virology-110615-035606

Pandeló José, D., Bartholomeeusen, K., da Cunha, R. D., Abreu, C. M., Glinski, J., and da Costa, T. B. F., et al. (2014). Reactivation of latent HIV-1 by new semi-synthetic ingenol esters. *Virology* 462–463, 328–339. doi: 10.1016/j.virol.2014.05.033

Panfil, A. R., London, J. A., Green, P. L., and Yoder, K. E. (2018). CRISPR/Cas9 genome editing to disable the latent HIV-1 provirus. *Front. Microbiol.* 9:3107. doi: 10.3389/fmicb.2018.03107

Patel, S., Jones, R. B., Nixon, D. F., and Bollard, C. M. (2016). T-cell therapies for HIV: preclinical successes and current clinical strategies. *Cytotherapy* 18, 931–942. doi: 10.1016/j.jcyt.2016.04.007

Periyasamy, P., Thangaraj, A., Bendi, V. S., and Buch, S. (2019). HIV-1 Tat-mediated microglial inflammation involves a novel miRNA-34a-NLRC5-NFκB signaling axis. *Brain. Behav. Immun.* 80, 227–237. doi: 10.1016/j.bbi.2019.03.011

Periyasamy, P., Thangaraj, A., Guo, M.-L., Hu, G., Callen, S., and Buch, S. (2018). Epigenetic promoter DNA methylation of miR-124 promotes HIV-1 Tat-mediated microglial activation via MECP2-STAT3 Axis. *J. Neurosci.* 38, 5367–5383. doi: 10.1523/JNEUROSCI.3474-17.2018

Pinskaya, M., and Morillon, A. (2009). Histone H3 lysine 4 di-methylation: a novel mark for transcriptional fidelity? *Epigenetics* 4, 302–306. doi: 10.4161/epi.4.5.9369

Qin, Y., Sun, X., Shao, X., Cheng, C., Feng, J., Sun, W., et al. (2015). Macrophage-microglia networks drive M1 microglia polarization after mycobacterium infection. *Inflammation* 38, 1609–1616. doi: 10.1007/s10753-015-0136-y

Qu, D., Sun, W.-W., Li, L., Ma, L., Sun, L., Jin, X., et al. (2019). Long noncoding RNA MALAT1 releases epigenetic silencing of HIV-1 replication by displacing the polycomb repressive complex 2 from binding to the LTR promoter. *Nucleic Acids Res.* 47, 3013–3027. doi: 10.1093/nar/gkz117

Rao, V. R., Ruiz, A. P., and Prasad, V. R. (2014). Viral and cellular factors underlying neuropathogenesis in HIV associated neurocognitive disorders (HAND). *AIDS Res. Ther.* 11:13. doi: 10.1186/1742-6405-11-13

Rawat, P., and Spector, S. A. (2017). Development and characterization of a human microglia cell model of HIV-1 infection. *J. Neurovirol.* 23, 33–46. doi: 10.1007/s13365-016-0472-1

Réu, P., Khosravi, A., Bernard, S., Mold, J. E., Salehpour, M., Alkass, K., et al. (2017). The lifespan and turnover of microglia in the human brain. *Cell Rep.* 20, 779–784. doi: 10.1016/j.celrep.2017.07.004

Rice, A. P. (2016). Cyclin-dependent kinases as therapeutic targets for HIV-1 infection. *Expert Opin. Ther. Targets* 20, 1453–1461. doi: 10.1080/14728222.2016.1254619

Richard, K., Williams, D. E., de Silva, E. D., Brockman, M. A., Brumme, Z. L., Andersen, R. J., et al. (2018). Identification of novel HIV-1 latency-reversing agents from a library of marine natural products. *Viruses* 10:348. doi: 10.3390/v10070348

Rock, R. B., and Peterson, P. K. (2006). Microglia as a pharmacological target in infectious and inflammatory diseases of the brain. *J. Neuroimmune Pharmacol.* 1, 117–126. doi: 10.1007/s11481-006-9012-8

Rodrigues, V., Ruffin,N., San-Roman, M., and Benaroch, P. (2017). Myeloid cell interaction with HIV: a complex relationship. *Front. Immunol.* 8:1698. doi: 10.3389/fimmu.2017.01698

Rohr, O., Lecestre, D., Chasserot-Golaz, S., Marban, C., Avram, D., Aunis, D., et al. (2003a). Recruitment of Tat to heterochromatin protein HP1 via interaction with CTIP2 inhibits human immunodeficiency virus type 1 replication in microglial cells. *J. Virol.* 77, 5415–5427. doi: 10.1128/JVI.77.9.5415-5427.2003

Rohr, O., Marban, C., Aunis, D., and Schaeffer, E. (2003b). Regulation of HIV-1 gene transcription: from lymphocytes to microglial cells. *J. Leukoc. Biol.* 74, 736–749. doi: 10.1189/jlb.0403180

Sadowski, I., and Hashemi, F. B. (2019). Strategies to eradicate HIV from infected patients: elimination of latent provirus reservoirs. *Cell. Mol. Life Sci.* 76, 3583–3600. doi: 10.1007/s00018-019-03156-8

Salemi, M., Lamers, S. L., Yu, S., de Oliveira, T., Fitch, W. M., and McGrath, M. S. (2005). Phylodynamic analysis of human immunodeficiency virus type 1 in distinct brain compartments provides a model for the neuropathogenesis of AIDS. *J. Virol.* 79, 11343–11352. doi: 10.1128/JVI.79.17.11343-11352.2005

Salemi, M., and Rife, B. (2016). Phylogenetics and phyloanatomy of HIV/SIV intra-host compartments and reservoirs: the key role of the central nervous system. *Curr. HIV Res.* 14, 110–120. doi: 10.2174/1570162X13666151029102413

Salter, M. W., and Stevens, B. (2017). Microglia emerge as central players in brain disease. *Nat. Med.* 23, 1018–1027. doi: 10.1038/nm.4397

Santerre, M., Bagashev, A., Gorecki, L., Lysek, K. Z., Wang, Y., Shrestha, J., et al. (2019). HIV-1 Tat protein promotes neuronal dysregulation by inhibiting E2F transcription factor 3 (E2F3). *J. Biol. Chem.* 294, 3618–3633. doi: 10.1074/jbc.RA118.003744

Sarracino, A., Gharu, L., Kula, A., Pasternak, A. O., Avettand-Fenoel, V., Rouzioux, C., et al. (2018). Posttranscriptional regulation of HIV-1 gene expression during replication and reactivation from latency by nuclear matrix protein MATR3. *MBio* 9:e02158-18. doi: 10.1128/mBio.02158-18

Scheper, H., van Holten,N., Hovens, J., and de Boer, M. (2018). Severe depression as a neuropsychiatric side effect induced by dolutegravir. *HIV Med.* 19, e58–e59. doi: 10.1111/hiv.12538

Schwartz, C., Bouchat, S., Marban, C., Gautier, V., Van Lint, C., Rohr, O., et al. (2017). On the way to find a cure: purging latent HIV-1 reservoirs. *Biochem. Pharmacol.* 146, 10–22. doi: 10.1016/j.bcp.2017.07.001

Schwartz, C., Catez, P., Rohr, O., Lecestre, D., Aunis, D., and Schaeffer, E. (2000). Functional interactions between C/EBP, Sp1, and COUP-TF regulate human immunodeficiency virus type 1 gene transcription in human brain cells. *J. Virol.* 74, 65–73. doi: 10.1128/JVI.74.1.65-73.2000

Scully, E., and Alter, G. (2016). NK cells in HIV disease. *Curr. HIV/AIDS Rep.* 13, 85–94. doi: 10.1007/s11904-016-0310-3

Sengupta, S., and Siliciano, R. F. (2018). Targeting the Latent Reservoir for HIV-1. *Immunity* 48, 872–895. doi: 10.1016/j.immuni.2018.04.030

Sharpe, A. H., and Abbas, A. K. (2006). T-cell costimulation–biology, therapeutic potential, and challenges. *N. Engl. J. Med.* 355, 973–975. doi: 10.1056/NEJMp068087

Sheng, J., Ruedl, C., and Karjalainen, K. (2015). Most tissue-resident macrophages except microglia are derived from fetal hematopoietic stem cells. *Immunity* 43, 382–393. doi: 10.1016/j.immuni.2015.07.016

Sheng, Z., Liu, Y., Li, H., Zheng, W., Xia, B., Zhang, X., et al. (2018). Efficacy of minocycline in acute ischemic stroke: a systematic review and meta-analysis of rodent and clinical studies. *Front. Neurol.* 9:1103. doi: 10.3389/fneur.2018.01103

Sinharay, S., and Hammoud, D. A. (2019). Brain PET imaging: value for understanding the pathophysiology of HIV-Associated Neurocognitive Disorder (HAND). *Curr. HIV/AIDS Rep.* 16, 66–75. doi: 10.1007/s11904-019-00419-8

Smit, T. K., Brew, B. J., Tourtellotte, W., Morgello, S., Gelman, B. B., and Saksena,N.K. (2004). Independent evolution of Human Immunodeficiency Virus (HIV) drug resistance mutations in diverse areas of the brain in HIV-infected patients, with and without dementia, on antiretroviral treatment. *J. Virol.* 78, 10133–10148. doi: 10.1128/JVI.78.18.10133-10148.2004

Sominsky, L., De Luca, S., and Spencer, S. J. (2018). Microglia: key players in neurodevelopment and neuronal plasticity. *Int. J. Biochem. Cell Biol.* 94, 56–60. doi: 10.1016/j.biocel.2017.11.012

Spivak, A. M., and Planelles, V. (2018). Novel latency reversal agents for HIV-1 cure. *Annu. Rev. Med.* 69, 421–436. doi: 10.1146/annurev-med-052716-031710

Strain, M. C., Letendre, S., Pillai, S. K., Russell, T., Ignacio, C. C., Gunthard, H. F., et al. (2005). Genetic composition of human immunodeficiency virus type 1 in cerebrospinal fluid and blood without treatment and during failing antiretroviral therapy. *J. Virol.* 79, 1772–1788. doi: 10.1128/JVI.79.3.1772-1788.2005

Su, H., Cheng, Y., Sravanam, S., Mathews, S., Gorantla, S., Poluektova, L. Y., et al. (2019). Immune activations and viral tissue compartmentalization during progressive HIV-1 infection of humanized mice. *Front. Immunol.* 10:340. doi: 10.3389/fimmu.2019.00340

Subra, C., and Trautmann, L. (2019). Role of T lymphocytes in HIV neuropathogenesis. *Curr. HIV/AIDS Rep.* 16, 236–243. doi: 10.1007/s11904-019-00445-6

Sumi, N., Nishioku, T., Takata, F., Matsumoto, J., Watanabe, T., Shuto, H., et al. (2010). Lipopolysaccharide-activated microglia induce dysfunction of the blood–brain barrier in rat microvascular endothelial cells co-cultured with microglia. *Cell. Mol. Neurobiol.* 30, 247–253. doi: 10.1007/s10571-009-9446-7

Sun, M., Li, Y., Zheng, H., and Shao, Y. (2016). Recent progress toward engineering HIV-1-specific neutralizing monoclonal antibodies. *Front. Immunol.* 7:391. doi: 10.3389/fimmu.2016.00391

Sung, J. M., and Margolis, D. M. (2018). HIV persistence on antiretroviral therapy and barriers to a cure. *Adv. Exp. Med. Biol.* 1075, 165–185. doi: 10.1007/978-981-13-0484-2_7

Sung, T.-L., and Rice, A. P. (2006). Effects of prostratin on Cyclin T1/P-TEFb function and the gene expression profile in primary resting $CD4^+$ T cells. *Retrovirology* 3:66. doi: 10.1186/1742-4690-3-66

Sutherland, E. J., and Brew, B. J. (2018). Human immunodeficiency virus and the nervous system. *Neurol. Clin.* 36, 751–765. doi: 10.1016/j.ncl.2018.07.002

Taniguchi, T., and Takaoka, A. (2002). The interferon-alpha/beta system in antiviral responses: a multimodal machinery of gene regulation by the IRF family of transcription factors. *Curr. Opin. Immunol.* 14, 111–116. doi: 10.1016/S0952-7915(01)00305-3

Tavazzi, E., Morrison, D., Sullivan, P., Morgello, S., and Fischer, T. (2014). Brain inflammation is a common feature of HIV-infected patients without HIV encephalitis or productive brain infection. *Curr. HIV Res.* 12, 97–110. doi: 10.2174/1570162X12666140526114956

Tesmer, V. M., and Bina, M. (1996). Regulation of HIV-1 gene expression by NF-IL6. *J. Mol. Biol.* 262, 327–335. doi: 10.1006/jmbi.1996.0516

Thompson, K. A., Cherry, C. L., Bell, J. E., and McLean, C. A. (2011). Brain cell reservoirs of latent virus in presymptomatic HIV-infected individuals. *Am. J. Pathol.* 179, 1623–1629. doi: 10.1016/j.ajpath.2011.06.039

Trautmann, L. (2016). Kill. *Curr. Opin. HIV AIDS* 11, 409–416. doi: 10.1097/COH.0000000000000286

Valcour, V., Paul, R., Chiao, S., Wendelken, L. A., and Miller, B. (2011). Screening for cognitive impairment in human immunodeficiency virus. *Clin. Infect. Dis.* 53, 836–842. doi: 10.1093/cid/cir524

Vanhamel, J., Bruggemans, A., and Debyser, Z. (2019). Establishment of latent HIV-1 reservoirs: what do we really know? *J. Virus Erad.* 5, 3–9.

Vaz, G. R., Hädrich, G., Bidone, J., Rodrigues, J. L., Falkembach, M. C., Putaux, J.-L., et al. (2017). Development of nasal lipid nanocarriers containing curcumin for brain targeting. *J. Alzheimers. Dis.* 59, 961–974. doi: 10.3233/JAD-160355

Vazquez, N., Greenwell-Wild, T., Marinos, N. J., Swaim, W. D., Nares, S., Ott, D. E., et al. (2005). Human immunodeficiency virus type 1-induced macrophage gene expression includes the p21 gene, a target for viral regulation. *J. Virol.* 79, 4479–4491. doi: 10.1128/JVI.79.7.4479-4491.2005

Veenhuis, R. T., Clements, J. E., and Gama, L. (2019). HIV eradication strategies: implications for the central nervous system. *Curr. HIV/AIDS Rep.* 16, 96–104. doi: 10.1007/s11904-019-00428-7

Veenstra, M., León-Rivera, R., Li, M., Gama, L., Clements, J. E., and Berman, J. W. (2017). Mechanisms of CNS viral seeding by HIV^+ $CD14^+$ $CD16^+$ monocytes: establishment and reseeding of viral reservoirs contributing to HIV-associated neurocognitive disorders. *MBio* 8:e01280-17. doi: 10.1128/mBio.01280-17

Venneti, S., Lopresti, B. J., Wang, G., Bissel, S. J., Mathis, C. A., Meltzer, C. C., et al. (2004). PET imaging of brain macrophages using the peripheral benzodiazepine receptor in a macaque model of neuroAIDS. *J. Clin. Invest.* 113, 981–989. doi: 10.1172/JCI20227

Vera, J. H., Guo, Q., Cole, J. H., Boasso, A., Greathead, L., Kelleher, P., et al. (2016). Neuroinflammation in treated HIV-positive individuals: a TSPO PET study. *Neurology* 86, 1425–1432. doi: 10.1212/WNL.0000000000002485

Wang, G., Zhao,N., Berkhout, B., and Das, A.T. (2016). CRISPR-Cas9 can inhibit HIV-1 replication but NHEJ repair facilitates virus escape. *Mol. Ther.* 24, 522–526. doi: 10.1038/mt.2016.24

Wang, G., Zhao,N., Berkhout, B., and Das, A.T. (2018). CRISPR-Cas based antiviral strategies against HIV-1. *Virus Res.* 244, 321–332. doi: 10.1016/j.virusres.2017.07.020

Wang, G. P., Ciuffi, A., Leipzig, J., Berry, C. C., and Bushman, F. D. (2007). HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome Res.* 17, 1186–1194. doi: 10.1101/gr.6286907

Wang, X., Ye, L., Hou, W., Zhou, Y., Wang, Y.-J., Metzger, D. S., et al. (2009). Cellular microRNA expression correlates with susceptibility of monocytes/macrophages to HIV-1 infection. *Blood* 113, 671–674. doi: 10.1182/blood-2008-09-175000

Wang, Z., Pan, Q., Gendron, P., Zhu, W., Guo, F., Cen, S., et al. (2016). CRISPR/Cas9-derived mutations both inhibit HIV-1 replication and accelerate viral escape. *Cell Rep.* 15, 481–489. doi: 10.1016/j.celrep.2016.03.042

Weinberger, A. D., and Weinberger, L. S. (2013). Stochastic fate selection in HIV-infected patients. *Cell* 155, 497–499. doi: 10.1016/j.cell.2013.09.039

Whitney, J. B., Hill, A. L., Sanisetty, S., Penaloza-MacMaster, P., Liu, J., Shetty, M., et al. (2014). Rapid seeding of the viral reservoir prior to SIV viraemia in rhesus monkeys. *Nature* 512, 74–77. doi: 10.1038/nature13594

Williams, D. W., Eugenin, E. A., Calderon, T. M., and Berman, J. W. (2012). Monocyte maturation, HIV susceptibility, and transmigration across the blood brain barrier are critical in HIV neuropathogenesis. *J. Leukoc. Biol.* 91, 401–415. doi: 10.1189/jlb.0811394

Wong, M. E., Jaworowski, A., and Hearps, A. C. (2019). The HIV reservoir in monocytes and macrophages. *Front. Immunol.* 10:1435. doi: 10.3389/fimmu.2019.01435

Woods, S. P., Morgan, E. E., Marquie-Beck, J., Carey, C. L., and Grant, I., Letendre, S. L., et al. (2006). Markers of macrophage activation and axonal injury are associated with prospective memory in HIV-1 disease. *Cogn. Behav. Neurol.* 19, 217–221. doi: 10.1097/01.wnn.0000213916.10514.57

Xiao, Q., Guo, D., and Chen, S. (2019). Application of CRISPR/Cas9-based gene editing in HIV-1/AIDS therapy. *Front. Cell. Infect. Microbiol.* 9:69. doi: 10.3389/fcimb.2019.00069

Xu, H., Wang, Z., Li, J., Wu, H., Peng, Y., Fan, L., et al. (2017). The polarization states of microglia in TBI: a new paradigm for pharmacological intervention. *Neural Plast.* 2017, 1–11. doi: 10.1155/2017/5405104

Yadav, A., and Collman, R. G. (2009). CNS inflammation and macrophage/microglial biology associated with HIV-1 infection. *J. Neuroimmune Pharmacol.* 4, 430–447. doi: 10.1007/s11481-009-9174-2

Yee, J.-K. (2016). Off-target effects of engineered nucleases. *FEBS J.* 283, 3239–3248. doi: 10.1111/febs.13760

Yenari, M. A., Xu, L., Tang, X. N., Qiao, Y., and Giffard, R. G. (2006). Microglia potentiate damage to blood-brain barrier constituents: improvement by minocycline *in vivo* and *in vitro*. *Stroke* 37, 1087–1093. doi: 10.1161/01.STR.0000206281.77178.ac

Yoder, K. E., and Bundschuh, R. (2016). Host double strand break repair generates HIV-1 strains resistant to CRISPR/Cas9. *Sci. Rep.* 6:29530. doi: 10.1038/srep29530

Zhang, Y., Yin, C., Zhang, T., Li, F., Yang, W., Kaminski, R., et al. (2015). CRISPR/gRNA-directed synergistic activation mediator (SAM) induces

specific, persistent and robust reactivation of the HIV-1 latent reservoirs. *Sci. Rep.* 5:16277. doi: 10.1038/srep16277

Zhou, Y., Tang, H., and Xiong, H. (2016). Chemokine CCL2 enhances NMDA receptor-mediated excitatory postsynaptic current in rat hippocampal slices-a potential mechanism for HIV-1-associated neuropathy? *J. Neuroimmune Pharmacol.* 11, 306–315. doi: 10.1007/s11481-016-9660-2

Zhu, L. J., Lawrence, M., Gupta, A., Pagès, H., Kucukural, A., Garber, M., et al. (2017). GUIDEseq: a bioconductor package to analyze GUIDE-Seq datasets for CRISPR-Cas nucleases. *BMC Genomics* 18:379. doi: 10.1186/s12864-017-3746-y

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

6. Publication 3

Inhibition of HIV-1 gene transcription by KAP1 in myeloid lineage

Amina AIT AMMAR, Maxime BELLEFROID, Fadoua DAOUAD, Valerie MARTINELLI, Jeanne VAN ASSCHE, Clementine WALLET, Marco DE ROVERE, Birthe FAHRENKROG, Christian SCHWARTZ, Carine VAN LINT, Virginie GAUTIER and Olivier ROHR

(To be published)

In this paper we described a new cellular factor, KAP1, acting as a repressor of HIV-1 in microglial cells. KAP1 is an already described strong gene transcription negative factor. In this paper, we demonstrate that KAP1 interacts with CTIP2, a protein that our group has focused its attention for a long time due to its role in repressing the transcription of HIV-1 gene and promoting viral latency. We show that KAP1 interacts and localizes with the viral protein Tat to target it for proteasomal degradation. In the line, knocking-down KAP1, promoted an increase of HIV-1 gene RNA elongation and viral gene expression. KAP1 was found bound to CTIP2 on the latent viral promoter. However, this interaction was abrogated by Tat showing a dynamic interplay between the two cellular repressors and the viral proteins upon productive HIV-1 infection. Altogether, our results suggest that KAP1 contributes to the establishment and the persistence of HIV-1 latency in myeloid cells.

I contributed to this article by scientific and technical supports to the main investigator during the lab meetings and the experiments.

Inhibition of HIV-1 gene transcription by KAP1 in myeloid lineage

Amina AIT AMMAR[1,2,3], Maxime BELLEFROID[3], Fadoua DAOUAD[1], Valérie MARTINELLI[4], Jeanne VAN ASSCHE[1], Clémentine WALLET[1], Marco DE ROVERE[1], Birthe FAHRENKROG[4], Christian SCHWARTZ[1], Carine VAN LINT[3*], Virginie GAUTIER[2*] and Olivier ROHR[1*]

1. Université de Strasbourg, UR 7292 DHPI, FMTS, IUT Louis Pasteur, 1 Allée d'Athènes 67300 Schiltigheim, France
2. Center for Research in Infectious Diseases (CRID), School of Medicine and Medical Science (SMMS), University College Dublin (UCD), Dublin, Ireland
3. Service of Molecular Virology, Institute for Molecular Biology and Medicine (IBMM), Université Libre de Bruxelles (ULB), Gosselies, Belgium
4. Laboratory Biology of the Nucleus, Institute for Molecular Biology and Medicine, Université Libre de Bruxelles, Charleroi 6041

* Correspondence should be addressed to Olivier ROHR: Tel: +33 (0)3 68 85 25 80; Email: olivier.rohr@unistra.fr. Correspondence may also be addressed to Virginie GAUTIER: Tel +353 (1) 716 1229; Email: virginie.gautier@ucd.ie and to Carine Van Lint: Tel : +32 2 650 98 00; Email: cvlint@ulb.ac.be. The authors wish it to be known that, in their opinion, the last three authors should be regarded as joint senior authors.

ABSTRACT

HIV-1 latency generates reservoirs that prevent viral eradication by the current therapies. To find strategies toward an HIV cure, detailed understandings of the molecular mechanisms underlying establishment and persistence of the reservoirs are needed. The cellular transcription factor KAP1 is known as a potent repressor of gene transcription. Here we report that KAP1 represses HIV-1 gene expression in myeloid cells including microglial cells, the major reservoir of the central nervous system. Mechanistically, KAP1 interacts and colocalizes with the viral transactivator Tat to promote its degradation via the proteasome pathway and repress HIV-1 gene expression. In myeloid models of latent HIV-1 infection, the depletion of KAP1 increased viral gene elongation and reactivated HIV-1 expression. Bound to the latent HIV-1 promoter, KAP1 associates and cooperates with CTIP2, a key epigenetic silencer of HIV-1 expression in microglial cells. In addition, Tat and CTIP2 compete for KAP1 binding suggesting a dynamic modulation of the KAP1 cellular partners upon HIV-1 infection. Altogether, our results suggest that KAP1 contributes to the establishment and the persistence of HIV-1 latency in myeloid cells.

INTRODUCTION

The pandemic of HIV-1 infections is a global health problem. Current cART (combination antiretroviral therapy) efficiently suppresses viral expression below clinical detection levels but fails to eradicate latently infected reservoirs, the main obstacles toward an HIV cure. Indeed, most efforts have focused on understanding latency molecular mechanisms in resting memory CD4+ T-cell reservoirs. However, they are not the only source of viral rebound. Myeloid cells such as monocytes, tissue-resident macrophages and follicular dendritic cells are part of the viral reservoir(1–3). Protected by the blood brain barrier, the central nervous system is a major anatomic reservoir and a sanctuary for the virus (4). Indeed, integrated SIV DNA have been reported in the CNS of infected macaques, with undetectable plasma viral load (5). In the brain, microglial cells are the major reservoirs of latently integrated HIV-1 (6). The specific molecular mechanisms controlling HIV-1 gene silencing in these CNS resident macrophages should be taken into consideration to design strategies toward HIV cure.

We have previously reported the importance of the cellular co-factor CTIP2 (Coup-TF Interacting Protein 2) in the establishment of HIV-1 post-integration latency in microglial cells (7–9). CTIP2 promotes and maintains HIV-1 gene silencing by recruiting chromatin modifying complexes including HDAC1/2 (Histone Deacetylase 1/2), SUV39h1 (Suppressor of Variegation 3-9 Homolog 1) and LSD1 (Lysine-Specific histone Demethylase 1A) to the viral promoter (9, 10). In addition, CTIP2 targets and represses HIV-1 Tat transactivation function by promoting its relocation to heterochromatin structure via the formation of a Tat-CTIP2-HP1α (Heterochromatin Protein 1α) complex (11) and by repressing the elongation factor P-TEFb (12). P-TEFb is the key cofactor of Tat. CTIP2 associates with an inactive form of P-TEFb to repress the CDK9 (Cyclin Dependent Kinase 9) catalytic subunit and inhibit P-TEFb sensitive genes including HIV-1 (12). Finally, we reported that HMGA1 recruits this CTIP2-associated inactive P-TEFb complex to the viral promoter (13). Importantly, Desplats *et al*, revealed that CTIP2 expression is increased in the CSF, astrocytes and microglial cells from PLWH (Patients Living With HIV) on suppressive cART (14). Collectively, our results demonstrate the importance of CTIP2 and its partners in the regulation of HIV-1 gene silencing. The HIV-1 latency is a multifactorial phenomenon involving factors dedicated to epigenetic gene silencing. KAP1 (KRAB (Krugel-Associated Box) domain-Associated Protein 1), also known as TRIM28 (Tripartite Motif-containing Protein 28) or TIF1β (Transcriptional Intermediary Factor 1 beta) is one of them. It was identified in 1996 as an interaction partner of the KRAB-ZFPs (Krüppel-Associated Box Zinc Finger Proteins) transcription factors family members (15). KAP1 is a protein with multiple functional domains that regulates the chromatin environment through interactions with different partners (16). KAP1 is recruited to DNA loci via interactions of its RBCC (RING finger, 2 B-box zinc fingers, Coiled-Coil region) domain with KRAB proteins (17). The RING domain of KAP1 has SUMO and Ubiquitin E3 ligase activities and induces heterochromatin structures (18–20), the PHD (Plant Homeo Domain) C-terminal domain has a SUMO E3 ligase activity needed to SUMOylate its own Bromo domain and the SUMOylated Bromo domain recruits the NURD and SETDB1 repressor complexes (21). KAP1 is known to repress endogenous and latent retrovirus (22). In HIV-1 infected cells, KAP1 associates with HDAC1 to deacetylate the integrase and inhibit proviral integration (23). At the transcriptional level, KAP1 has been reported to contribute to ZBRK1 (Zinc finger and BRCA1 (Breast cancer type 1 susceptibility protein)-interacting protein with A KRAB domain 1) and ZNF10 (Zinc Finger Protein 10)-mediated HIV-1 LTR repression (24, 25). However, these studies have not yet shown the direct involvement of KAP1 and its mechanism of action. Moreover, KAP1-mediated recruitment of an inactive form of P-TEFb to most genes containing a paused RNA polymerase, including the HIV-1 LTR promoter has been described to favor gene expression upon stimulation (26) (27). Despite these controversies, KAP1 is mainly shown as a viral gene repressor. KAP1 restricts the activation of MLV (Murine Leukemia Virus) and HTLV-1 (Human T-Lymphotropic Virus type 1) genes (28, 29). More recently, the restriction factor APOBEC3A (Apolipoprotein B mRNA-Editing enzyme Catalytic polypeptide-like 3G) has been described to recruit KAP1 to suppress HIV-1 transcription (30). In addition, a very recent study suggests that KAP1 represses HIV-1 gene expression by mediating CDK9 SUMOylation resulting in P-TEFb repression in T cells (31). Surprisingly, since KAP1 functions have been extensively studied in T lineage, nothing has been done to define its role and its mechanism of action on HIV-1 infection in myeloid cells.

Here, we studied the influence of KAP1 on HIV-1 expression and its specific mechanism of action in monocytic and microglial cells, the resident macrophages and the main viral reservoirs in the brain. In microglial cells KAP1 repressed HIV-1 gene expression. While KAP1 was found associated with the silenced HIV-1 promoter in the CHME5-HIV microglial model of latency, stimulations released KAP1 from the provirus. RNA interference-mediated knockdown of KAP1 in monocytic cell line latently infected with HIV-1 showed an increase in viral reactivation and transcription. Mechanistically, KAP1 repressed the HIV-1 promoter activity in the absence and in the presence of the viral transactivator Tat. Moreover, KAP1 had a major repressive impact on Tat function. We report that KAP1 interacts physically and colocalizes with Tat in the nucleus of microglial cells to promote its degradation via the proteasome pathway. Finally, KAP1 was found associated with the repressor CTIP2. They both cooperated to repress Tat function. Altogether our results highlight the contribution and the specific mechanism of action of KAP1 in the establishment and the persistence of HIV-1 post-integration latency in cells from myeloid origin.

MATERIAL AND METHODS

Plasmids

The following plasmids: pcDNA3, flag CTIP2, flag-CTIP2 350-813, 350-717, 1-354 and 145-434, Tap-CTIP2, flag-Tat, DsRed-Tat, ShCTIP2, LTR-Luc, pNL4-3Δ Env Luc have been described previously (9, 11, 12). The flag-KAP1 plasmids, flag-KAP1 S824A/D come from Dr. David K. Ann (32). The vectors shKAP1 and shNT (non-target) were gift from Dr. Ivan D'orso (26). The plasmids flag-KAP1WT, flag-KAP1 ΔRBCC, ΔPHD, ΔBromo, ΔPHD/ΔBromo were provided by Dr. Florence Cammas (IRCM Montpellier). GFP-KAP1 comes from Addgene (#65397).

Cell culture

The human microglial cell line (provided by Prof. M. Tardieu, Paris, France) (33), HEK293T cell lines and CHME5 cell lines latently infected with HIV-1 (34) were maintained in Dulbecco's modifed Eagle medium (DMEM) containing 10% decomplemented fetal calf serum and 100 U/ml penicillin-streptomycin. When indicated, the cells were treated with 50 µM MG132 for 6 hours. THP89 cell lines were grown in RPMI 1640 medium supplemented with 10% decomplemented fetal calf serum and 50 U/ml penicillin-streptomycin.

Antibodies and reagents

The list of antibodies and other chemical reagents used in this work are listed in the following table 1.

Reagent	Reference
anti-CTIP2 antibody	A300-383A Bethyl
anti-KAP1 antibody	Ab10483 Abcam
anti-Tat antibody	ab43014 Abcam
anti-β-actin antibody	A1978 Sigma-Aldrich
anti-Flag antibody	F3165 Sigma-Aldrich
anti-GFP antibody	632592 Clontech
anti-α-tubulin antibody	Ab4074 Abcam
anti-UBC9 antibody	610749 BD Biosciences
anti-HA antibody	901 501 Biolegend
anti-mouse IgG-HRP	Sc-358914 Santa Cruz
anti-rabbit IgG-HRP	Sc-2004 Santa Cruz
anti-KAP1 ChIP grade antibody	C15410236-100 Diagenode
anti-RNA Pol II ChIP grade antibody	14958 Cell signaling
Rabbit IgG	C15410206 Diagenode
Mouse IgG	C15400001-15 Diagenode

Reverse PCR	1708890 Bio Rad
Quantitative PCR	Ssoadvanced™ Universal SYBR Green Supermix Bio Rad
anti-proteases	000000011873580001 Sigma-Aldrich
MG132	BML-PI102-0025 ENZO Life Sciences
TNFα	300-01A PeproTech
HMBA	224235 Sigma-Aldrich
DMA	20660 Thermoscientific
anti-flag immunoprecipitation beads	A-2220 Sigma-Aldrich
Protein G magnetic beads	C03010021-220 Diagenode
Protein A magnetic beads	C03010020-220 Diagenode
Protein A Agarose/Salmon Sperm DNA	16-157 Millipore

Lentiviral production

HEK cells seeded at a density of $4X10^6$ and grown in a 10 cm dish were transfected with the different shRNAs (shNT or shKAP1)-containing plasmids (9 µg), the pVSV-G (2.25 µg) and the psPAX2 (6.75 µg) packaging vectors by the calcium phosphate transfection method according to the manufacturer's protocol (TaKaRa; Calphos). 72h post-transfection, the virus-containing supernatant (8ml) was filtered (Steriflip; Merck) before being concentrated 20x (14000 rpm, 1h30, 4°c) and stored at -80°c.

Generation of stable THP89 knock-down for KAP1

THP89 cells were transduced by spinoculation as previously described (26). Briefly, 100µl of 20X concentrated lentiviral stock was added to 10^6 cells previously washed and resuspended in 400µl of complete growth medium containing polybrene (8µg/ml). Cells were then centrifuged (3200 rpm 1h, 32°c) before being incubated for 2h at 37°c. After this incubation, cells were washed and transferred in 1 ml complete growth medium in a 12-well dish. 48h post-infection, infected cells were selected by puromycin selection at a final concentration of 1 ug/ml for 2 weeks. GFP positive cells were measured by Flow Cytometry. Total proteins and RNAs were then extracted as described above for RT-qPCR and western blotting.

Nuclear protein extract

Cells seeded at a density of $3x10^6$ and grown in a 10 cm dish were transfected with the indicated vectors (30µg) using the calcium phosphate coprecipitation method. 48 hours post-transfection, cells were lysed for 10 minutes on ice in a buffer containing: 10 mM HEPES, 1.5 mM $MgCl_2$, 10 mM KCl and 0.5 mM DTT. After one minute of centrifugation at 13000 g, the nuclear pellet was lysed for 30 minutes on ice in another buffer containing: 20 mM HEPES, 25% glycerol, 1.5 mM $MgCl_2$, 420 mM NaCl, 0.2 mM EDTA, 0.5 mM DTT. After 2 minutes of centrifugation at 13000 g, the supernatant was analyzed by immunoprecipitation and Western blotting. Anti-proteases were systematically added to the lysis buffers.

Total protein extract

Cells seeded at a density of 10^6 cells per well and grown in a 6-well dish were transfected using the calcium phosphate coprecipitation method with the indicated vectors (12 µg). After 24 hours, the cells were lysed on ice for 50 minutes using a buffer containing: 10 mM Tris HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP40, 0.5% Na deoxycholate, 0.1% SDS. After 10 minutes centrifugation at 13000 g, the supernatant was stored for Western blot analysis. Antiproteases were systematically added to the lysis buffers.

Co-immunoprecipitation and Western blot analysis

Immunoprecipitations were performed on 500 µg of nuclear protein extracts using anti-flag M2 gel (Sigma-Aldrich) or by using Diagenode A/G protein-coupled magnetic beads to target endogenous proteins. Briefly, the nuclear protein extract was cleared for 1 hour under rotation at 4°C with the A/G protein-coupled magnetic beads. Meanwhile, the beads were coated under rotation with 1 µg of

antibody for 30 minutes at room temperature. After the preclearing step, the beads were removed, replaced by the antibody coated beads and incubated overnight at 4°C under rotation. The day after, the incubated beads were washed 2 times with a low salt buffer (50mM TrisHCl pH 7.5, 120 mM NaCl, 0.5mMEDTA, 0.5% NP40, 10% glycerol), 2 times with a high salt buffer (50mM TrisHCl pH 7.5, 0.5M NaCl, 0.5mM EDTA, 0.5% NP40, 10% glycerol) and one time with a low salt buffer. The immunoprecipitated proteins were eluted by heating at 100°C during 15 minutes in 4x Laemmli Sample Buffer (BIORAD). Immunoprecipitated protein complexes were subjected to Western blot analysis. The proteins were detected using the antibodies as indicated and visualized by a Thermo Scientific Chemiluminescence Detection System (Pierce ™ ECL Western Blotting Substrate 32106).

Fluorescence Microscopy

The microglial cells were seeded at $0.8x10^6$ and grown on coverslip in a 24-well plate. The cells were transfected after 24 hours with the indicated vectors by the Jetprime (Polyplus transfection) method, according to the instructions of the supplier. After 48 hours of transfection, the cells were fixed with PBS containing 2% formaldehyde for 15 minutes. The cells were then washed 3 times with PBS. The coverslip was mounted on a slide using the Mowiol 4-88 (Sigma-Aldrich) containing 1µg/ml DAPI. Cells fluorescence was observed under a fluorescence microscope and imaged using a Zeiss Axio Observer inverted microscope (Zeiss, Oberkochen, Germany) with Zeiss Plan-Apochromat 100x/1.4 oil objective. Images were acquired using the microscope system software AxioVision V 4.8.2.0 and processed using Image J software.

Luciferase assay

Microglial cells, seeded at $0.4x10^6$ in a 48-well dish, were transfected in triplicate by the calcium phosphate coprecipitation method with the indicated vectors. 48 hours post-transfection, the cells were collected and the luciferase activity was measured using the Dual-Glo Luciferase Assay system (Promega Madisson, USA) and normalized to Renilla Luciferase activity.

Chromatin immunoprecipitation assay

Chromatin immunoprecipitation assay was performed following ChIP assay kit (EMD Millipore) protocol. Briefly, HIV-1 CHME5 cells were seeded at $3x10^6$ in T175 Corning flask. 24 hours later, the cells were mock treated or stimulated with 10 ng/ml of TNFα (Tumor Necrosis Factor α) and 5 mM of HMBA (Hexamethylene bisacetamide). After 24 hours the cells were trypsinated, counted and then washed with PBS. The cells were resuspended at ratio of $4x10^6$ cell/ml of PBS. The cells were cross-linked with a freshly made DMA (dimethyl adipimidate) at 5 mM for 25 minutes followed by a second 8 minutes cross-link with formaldehyde at 1%. The cross-linking reaction was stopped by 0.125 M of Tris-Glycine solution. Cells were washed twice with ice cold PBS then resuspended to a ratio of $10x10^6$/300µl of SDS Lysis buffer (Tris HCl 50 mM pH 8, EDTA 10 mM, 1% SDS) containing proteases inhibitors. The chromatin was sonicated (Bioruptor Plus, Diagenode) 40 minutes (30s on, 30s off) to obtain DNA fragments of 200–400 bp. Chromatin immunoprecipitations were performed with chromatin from $5x10^6$ cells and 5 µg of antibodies. IgG was used as a control for immunoprecipitation. The DNA was purified using the High Pure PCR Product Purification Kit (11732676001 Sigma-Aldrich). Quantitative Real Time PCR reactions were performed with the TB Green™ Premix Ex Taq™ II (TAKARA) as recommended by the supplier. Relative quantification using standard curve method was performed for the primer pair, and 96-well plates (MU38900 Westburg) were read in a StepOnePlus PCR instrument (Applied Biosystems). The results are represented as a fold enrichment relative to the IgG condition. The LTR5' forward and reverse Primer sequences used are respectively: 5'-GCTTAAGCCTCAATAAAGCTTGC-3', 5'-TGACTAAAAGGGTCTGAGGGAT-3'.

Flow cytometry

CHME5 HIV-1 cells mock treated or stimulated with 10 ng/ml of TNFα and 5 mM of HMBA for 24 hours or either THP89 HIV-1 cells transduced with lentiviral particles were washed twice in PBS, resuspended in PBS containing 4% paraformaldehyde and fixed for 1 hour at 4°C in the dark. Cells were then washed twice in PBS and resuspended in FACS buffer (PBS, 0.1% BSA, 0.1% NaN_3). The percentage of GFP-

positive cells was measured on a CXP cytometer (Cytomics FC 500, Beckman Coulter) using CXP Software version 1.0 according to the manufacturer's instructions.

mRNA quantification

Total RNA was extracted from cells expressing the indicated vectors according to the QIAGEN RNeasy Plus Kit protocol or with Tri-Reagent (TRC118, MRC) according to the manufacturer's instructions. Following DNAse treatment (AM1907, Invitrogen), the cDNAs were produced using BioRad Reverse Transcription Kit (iScriptTM Reverse Transcription Supermix) or the PrimeScript RT reagent kit (RR037A, TaKaRa). cDNAs from HEK and THP89 cells were quantified by quantitative PCR using Bio Rad's SsoAdvanced Universal SYBR Green Supermix. Data were calculated using the $2^{-(\Delta\Delta CT)}$ method and normalized to the amount of GAPDH.

Cellular model	Target	Forward primer	Reverse primer
HEK	*tat*	ACTCGACAGAGGAGAGCAAG	GAGATCTGACTGTTCTGATGA
	GAPDH	GGACCTGACCTGCCGTCTAGAA	GGGTGTCGCTGTTGAAGTCAGAG
THP89	TAR	GTTAGACCAGATCTGAGCCT	GTGGGTTCCCTAGTTAGCCA
	tat	TGTTGCTTTCATTGCCAAGCTTGTTT	GTCTTCGTCGCTGTCTCCGCT
	gag	AAAAGCATTGGGACCAGGAG	CTTGCTTTATGGCCGGGT
	Ms RNA	GGATCTGTCTCTGTCTCTCTCTCCACC	ACAGTCAGACTCATCAAGTTTCTCTATCAAAGCA
	GFP	GAGGGCGATGCCACCTAC	GGTGGTGCAGATGAACTTCAG

Statistical analysis

Data were analyzed by performing the student t-test using Microsoft Excel. Values of p <0.05 (*), p <0.01 (**) and p <0.001 (***) were considered significant. The results were represented as mean and standard deviation of at least 3 independent experiments.

RESULTS

KAP1 cooperates with CTIP2 to repress HIV-1 gene expression in microglial cells

The functional impact of KAP1 overexpression or depletion on HIV-1 gene expression was examined on microglial cells expressing the pNL4-3 Δ ENV Luc HIV molecular clone (Figure 1 A and 1 B) or the episomal vector LTR-Luc (Figure 1 C and 1 D). In both cases we observed that expression of increasing amounts of KAP1 inhibited the HIV-1 promoter activity in a dose-dependent manner, while KAP1 depletion stimulated its activity.

These results are consistent with a transcriptional control of the viral expression by KAP1 and support a repressive role of KAP1 on HIV-1 gene expression. Of note, KAP1 protein expression levels were assessed by Western blot (Figures 1 E and F). Given KAP1 capacity to silence HIV-1 gene expression, we next examined the presence of KAP1 at the silenced viral promoter in CHME5-HIV microglial model of HIV-1 latency (34). Chromatin IP experiments revealed the presence of KAP1 associated with the silent 5'LTR region of the provirus (Figure 2 A). However, treatments with TNFα and HMBA to induce viral reactivation significantly displaced KAP1 from the promoter (Figure 2 A). The activation of HIV-1 promoter was validated by increased binding of the RNA Pol II (Figure 2 A) and by quantification of the GFP positive cells (Figure 2 B). Our results suggest that KAP1 targets the latent viral promoter to promote HIV-1 post-integration latency in microglial cells. To confirm these results in other models of infected myeloid cells, we took advantage of THP89 cells latently infected with recombinant p89.6 HIV-1 reporter virus wherein the GFP gene is inserted in the viral genome(35). We generated THP89 monocytic cell lines expressing nontarget (shNT) or KAP1 (shKAP1) small hairpin (sh) RNAs. The efficiency of KAP1 knockdown in the selected puromycin resistant cells was validated by Western blot (Figure 3 A). Using these cells, we assessed the role of KAP1 in viral latency and reactivation by monitoring GFP expression. Flow cytometry analysis and GFP transcripts quantifications revealed an increase of GFP positive cells upon KAP1 knockdown (Figure 3 A and B). Activation of HIV-1 gene transcription in THP89 cells following KAP1 depletion was further studied by quantifying initiated transcripts (TAR region), elongated transcripts (*gag* and *tat* regions), and the multiply spliced transcripts (Ms RNA) (Figure 3 C). Significant higher levels of elongated and multiply-spliced transcripts were observed in KAP1 knocked-down cells. Altogether, our results indicate that KAP1 depletion enables the release of HIV-1 transcriptional blocks underlying latency.

We have previously reported the importance of the cellular cofactor CTIP2 in the establishment and the persistence of HIV-1 latency (9, 10, 13). Since KAP1 and CTIP2 have been both involved in the control of HIV-1 gene expression, we next investigated whether they interact physically and functionally. Looking at the viral expression level, we observed that the concomitant overexpression of KAP1 and CTIP2 had a significantly stronger repressive activity than when overexpressed individually, (Figure 4 A, column 4 versus 2 and 3). Similarly, we observed a modest but significant cooperation in the context of a simultaneous depletion of the two proteins (Figure 4 B, column 4 versus 2 and 3). As shown in figure 4 C and 3 D, similar results were obtained upon analysis of the activity of 5'LTR viral promoter. To ensure that the effects can be correlated to the expression levels of CTIP2 and KAP1, we performed Western blot analysis. The results demonstrated the efficiencies of the overexpression and the knockdown at the protein level (Figures 4 E and F). We have reported that CTIP2 represses Tat function by relocating the viral regulatory protein in heterochromatin structures (11). We therefore investigated the impact of KAP1 on Tat function. As expected, Tat expression stimulated the viral promoter activity (Figure 5 A, column 2). However, while the overexpression of KAP1 repressed this transactivation (Figure 5 A, columns 3 to 5 versus 2), the knockdown of KAP1 promoted Tat activity (Figure 5 B). These results suggested us that the repressive function of KAP1 on HIV-1 expression results from a double impact on the initiation (-Tat) and the elongation (+Tat) steps of viral gene transcription. We next modulated the expression of KAP1 together with CTIP2. As expected, overexpression of KAP1 and CTIP2 alone inhibited Tat function (Figure 6 A, column 3 and 4). Moreover, their concomitant overexpression further repressed Tat-mediated activity (Figure 6 B, column 4). As shown in figure 6 C and 6 D, the efficiencies of the overexpression and the depletions were controlled by Western blot. Taken together, our results demonstrate that KAP1 and CTIP2 cooperate functionally to repress HIV-1 gene expression in microglial cells.

KAP1 interacts and colocalizes with Tat

Since our results suggest a physical interaction between Tat and KAP1, we performed co-immunoprecipitation targeting ectopically expressed KAP1 (Figure 7 A) and endogenously expressed KAP1 (Figure 7 B). We show that Tat associated with KAP1 in both conditions. To next visualize the association of KAP1 and Tat in the nucleus, microglial cells expressing GFP-KAP1 and DsRed-Tat were analyzed by direct fluorescence microscopy. As previously described in the literature, KAP1 was found in the nucleoplasm and excluded from the nucleoli, (Figure 7 C upper panel) (36). Tat localized in both nucleoplasm and nucleoli, (Figure 7 C, middle panel) (37, 38). We observed that Tat colocalizes with KAP1 in the nucleoplasm but not in the nucleoli, (Figure 7 C, lower panel).

KAP1 promotes Tat degradation by the proteasome pathway

To further characterize mechanistically the repression of Tat by KAP1. we observed the pattern of Tat expression in the presence of increasing amount of KAP1 (Figure 8 A). We found that Tat expression was inversely correlated to the protein level of KAP1 upon KAP1 overexpression (Figure 8 A) and KAP1 depletion (Figure 8 B). To determine if these modulations of Tat protein expression are controlled post-transcriptionally, we quantified Tat mRNAs by RT-qPCR, in the presence or lack of KAP1 overexpression. Since the amount of Tat mRNA remained unaffected by KAP1 (Figure 8 C), we hypothesized that KAP1 overexpression leads to Tat depletion by regulating the stability of the protein. Among the mechanisms controlling protein half-life in eukaryotic cells, the degradation via the proteasome pathway has been long studied. Interestingly, Tat has been described to be directed towards proteasome degradation (39–44). To determine whether the reduced expression of Tat in the presence of KAP1 is related to proteasome degradation, we quantified Tat expression levels in the presence of KAP1 and the proteasome inhibitor MG132, (Figure 8 D). As expected, in the absence of MG132, KAP1 expression reduced significantly Tat expression (Figure 8 D, column 2 versus 1). However, MG132 treatment increased Tat protein levels by about 50% (Figure 8 D column 3 versus 1) and counteracted KAP1-mediated degradation of Tat (Figure 8 D, column 4 versus 2). Thus, our results suggest that KAP1 targets Tat to degradation via the proteasome pathway.

Tat degradation is mediated by the PHD and Bromo domains of KAP1

Using different KAP1 deletion mutants, we next identified the KAP1 domain required for Tat degradation (Figure 9 A). While deletion of the RBCC domain (ΔRBCC) did not affect KAP1-mediated degradation of Tat (Figure 9 A, columns 2 and 3), deletions of the PHD (ΔPHD) and the Bromo domain (Δ Bromo) impaired Tat degradation (Figure 9 A, columns 4, 5 and 6 versus 2). We next used two mutants of KAP1, one mimicking a constitutive phosphorylation of Serine 824 (S824D) and the second which results in a non-phosphorylated form of AKP1 (S824A), (Figure 9 B). Of note, the phosphorylated status of the Serine 824 determine the repressive activity of KAP1 and its ability to interact with the NuRD and SETDB1 repressor complexes via the Bromo domain (32). While expression of KAP1 S824A promoted the degradation of Tat (Figure 9 B, column 5 versus 2), the S824D mutant did not (Figure 9 B, column 4). Overall, these results suggest that KAP1-mediated degradation of Tat may be controlled by phosphorylation of the protein and thereby may depend on the global physiology of the infected cells.

CTIP2 is interacting with KAP1 and competes with Tat for KAP1 binding

We demonstrated that KAP1 functionally cooperates with CTIP2 to inhibit HIV-1 gene expression. Since both CTIP2 and KAP1 have been reported to be associated with an inactive form of P-TEFb, we hypothesized that their functional interactions may result from physical association (12, 26). Immunoprecipitations targeting CTIP2 demonstrate that KAP1 associates with CTIP2 in RNA independent manner (Figure 10 A, column 4 and 6). Similar results were obtained after KAP1 immunoprecipitation (Figure 10 B, column 8 and 10) in the context of ectopic protein expression. We next confirmed these observations by immunoprecipitations targeting the endogenous proteins (Figure 10 C). To delineate the interface between KAP1 and CTIP2, we used different truncations of KAP1 and CTIP2 (Figure 10 D and E). We found that CTIP2 interacts with ΔPHD, ΔBromo, ΔPHD/ΔBromo deleted

KAP1 (Figure 10 D, columns 10, 11, 12), but not with the ΔRBCC KAP1, (Figure 10 D, column 9). On the other hand, KAP1 binds only the 1-354 and 145-434 truncated forms of CTIP2 (Figure 10 E, column 11 and 12). Altogether, we found that the PHD/Bromo domain of KAP1 associated with the central region of CTIP2 previously shown to mediate Tat binding (11). It should be noted that Tat has an additional CTIP2 binding site located at its C-terminal domain, between residues 717 and 813 (11). Co-immunoprecipitation experiments carried out with nuclear extracts expressing CTIP2, KAP1 and increasing amount of Tat, showed that Tat expression displaced CTIP2 from KAP1 suggesting that Tat may have a better affinity for KAP1 than CTIP2, (Figure 10 F, column 9 and 10).

DISCUSSION

Antiretroviral therapy has significantly improved the management of HIV-1 infection. However, despite ART efficiency, it is still impossible to cure HIV-1 infected patients. Latently infected reservoirs are the major hurdles toward an HIV cure. Among the major reservoirs are the resting CD4+ T cells and monocyte-macrophage lineage, including microglial cells. These reservoirs escape immune surveillance and cannot be targeted by current therapies. HIV-1 gene transcription is silenced by epigenetic modifications. These epigenetic changes remain a major obstacle to the virus eradication and patients cure. We have previously shown that CTIP2 is a major player in the establishment and persistence of HIV-1 latency in microglial cells (9, 10, 12). In the present work we report the role and the mechanism of action of KAP1 as new partner of CTIP2 in the control of HIV-1 expression in myeloid cells. Functionally, KAP1 suppressed HIV-1 expression and contributed to viral latency in myeloid cells. In accordance, we found KAP1 associated with the latent HIV-1 promoter but not after reactivation. Interestingly, in CD4+ T cells, KAP1 binding to the HIV-1 promoter has been reported to be unsensitive to reactivations (31). These controversies suggest that the mechanisms underlying KAP1 functions may be cell type specific. It is now well established that the repressive form of KAP1 is SUMOylated on its Bromo domain (45) and associates with the NuRD and SETDB1 repressor complexes (reviewed in (46)). The repressors CTIP2 and KAP1 have multiple similarities. They both interact with the NuRD complex (47, 48), the DNA repair enzymes PARP1 and PRKDC (49, 50) and a repressive form of the P-TEFb complex including the 7SKsnRNP (26). CTIP2 and KAP1 have also common target genes like p21 (32, 51). Our results demonstrate a functional cooperation between the two repressors to repress HIV-1 gene transcription. Moreover, KAP1 and CTIP2 interact physically. We report that the RBCC domain of KAP1 associates with the central domain of CTIP2. CTIP2 association with LSD1 and HMGA1 has been described to inhibit Tat-dependent viral gene transcription and to promote its relocation in heterochromatin structures (10, 11, 13). Our results show that KAP1 cooperates with CTIP2 to repress Tat function. In addition, we found that KAP1 interacts and colocalizes with Tat in the nucleoplasm of microglial cells. The HIV-1 Tat transactivator is essential to the elongation step of the viral gene mainly because of its ability to recruit P-TEFb at the HIV-1 initiated transcripts. Thereby, inhibiting Tat activity results in hindering the positive feedback loop required for efficient HIV-1 transcription. Interestingly, Tat has been shown to be sensitive to proteasome degradation (39–44, 52). We observed that Tat expression was inversely proportional to KAP1 expression level suggesting that KAP1-mediated repression of Tat function results from Tat degradation. KAP1 has been shown to promote proteasomal degradation of targets proteins by its RBCC domain (53–55). Surprisingly, we observed that Tat degradation was not sensitive to RBCC depletion but mediated by the PHD and the Bromo domains of KAP1. In line with these results, the point mutation of the KAP1 Serine 824 located in the Bromo domain abrogated Tat degradation. The Bromo domain is known to be modified by SUMOylation giving to KAP1 its repressive form. The phosphorylation of Serine 824 causes a loss of this SUMOylation state and a loss KAP1 repressive activity (32). Thus, our results suggest that Tat degradation is promoted by repressive SUMOylated forms of KAP1. Finally, we show that KAP1 interacts with Tat with higher affinity than with CTIP2. Indeed, upon increasing Tat expression, CTIP2 is displaced from KAP1 associated complex suggesting that the dynamic of KAP1 association with CTIP2 depends on the virological status of the cells. Altogether, our results suggest that KAP1 is a major player in the control of HIV-1 gene expression in microglial cells, the main HIV-1 reservoir in the brain. Moreover, they further highlight the cell type specific mechanism of action of the repressors

involved in viral latency. CTIP2 and KAP1 are highly expressed in resting T cells and depleted upon activation (31, 56). In addition, increased levels of CTIP2 are associated with persistence of HIV-1 latency in the brain (14) and increased level of KAP1 in the peripheral blood of gastric cancer patients is a biomarker predicting cancer stage progression (57–60). By transposition to what has been observed in oncology and virology, these repressors and their associated enzymatic activities may constitute good targets in therapies aiming at curing the HIV-1 infected patients

ACKNOWLEDGEMENT

We thank Dr. David K. Ann, Dr. Ivan D'orso and Dr. Florence Cammas for providing vectors.
The authors wish it to be known that, in their opinion, the last three authors should be regarded as joint senior authors.

FUNDING

This work was supported by the French agency for research on AIDS and viral hepatitis (ANRS); SIDACTION; the European Union's Horizon 2020 research and innovation programme under grant agreement No 691119-EU4HIVCURE-H2020-MSCA-RISE-2015; The Belgian Fund for Scientific Research (FRS-FNRS, Belgium), the "Fondation Roi Baudouin", the Walloon Region (Fonds de Maturation) and the university of Brussels (Action de Recherche Concertée (ARC) grant). The laboratory of CVL is part of the ULB-Cancer Research Centre (U-CRC). AAA is a fellow of the Wallonie-Bruxelles International program and of the Marie Skłodowska Curie COFUND action. MB is a doctoral fellow from the Belgian Fonds pour la formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA). CVL is Directeur de Recherches of the F.R.S-FNRS, Belgium.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest

REFERENCES

1. Kumar,A., Abbas,W. and Herbein,G. (2014) HIV-1 latency in monocytes/macrophages. *Viruses*, **6**, 1837–1860.
2. Kandathil,A.J., Sugawara,S. and Balagopal,A. (2016) Are T cells the only HIV-1 reservoir? *Retrovirology*, **13**.
3. Honeycutt,J.B., Thayer,W.O., Baker,C.E., Ribeiro,R.M., Lada,S.M., Cao,Y., Cleary,R.A., Hudgens,M.G., Richman,D.D. and Garcia,J.V. (2017) HIV persistence in tissue macrophages of humanized myeloid-only mice during antiretroviral therapy. *Nat. Med.*, **23**, 638–643.
4. Gelman,B.B., Lisinicchia,J.G., Morgello,S., Masliah,E., Commins,D., Achim,C.L., Fox,H.S., Kolson,D.L., Grant,I., Singer,E., *et al.* (2013) Neurovirological correlation with HIV-associated neurocognitive disorders and encephalitis in a HAART-era cohort. *J. Acquir. Immune Defic. Syndr.*, **62**, 487–495.
5. Clements,J.E., Gama,L., Graham,D.R., Mankowski,J.L. and Zink,M.C. (2011) A simian immunodeficiency virus macaque model of highly active antiretroviral treatment: viral latency in the periphery and the central nervous system. *Curr Opin HIV AIDS*, **6**, 37–42.
6. Llewellyn,G.N., Alvarez-Carbonell,D., Chateau,M., Karn,J. and Cannon,P.M. (2018) HIV-1 infection of microglial cells in a reconstituted humanized mouse model and identification of compounds that selectively reverse HIV latency. *J Neurovirol*, **24**, 192–203.
7. Le Douce,V., Cherrier,T., Riclet,R., Rohr,O. and Schwartz,C. (2014) The Many Lives of CTIP2: From AIDS to Cancer and Cardiac Hypertrophy. *J. Cell. Physiol.*, **229**, 533–537.
8. Le Douce,V., Forouzanfar,F., Eilebrecht,S., Van Driessche,B., Ait-Ammar,A., Verdikt,R., Kurashige,Y., Marban,C., Gautier,V., Candolfi,E., *et al.* (2016) HIC1 controls cellular- and HIV-1-gene transcription via interactions with CTIP2 and HMGA1. *Sci Rep*, **6**, 34920.
9. Marban,C., Suzanne,S., Dequiedt,F., de Walque,S., Redel,L., Van Lint,C., Aunis,D. and Rohr,O. (2007) Recruitment of chromatin-modifying enzymes by CTIP2 promotes HIV-1 transcriptional silencing. *EMBO J.*, **26**, 412–423.

10. Le Douce,V., Colin,L., Redel,L., Cherrier,T., Herbein,G., Aunis,D., Rohr,O., Van Lint,C. and Schwartz,C. (2012) LSD1 cooperates with CTIP2 to promote HIV-1 transcriptional silencing. *Nucleic Acids Res*, **40**, 1904–1915.
11. Rohr,O., Lecestre,D., Chasserot-Golaz,S., Marban,C., Avram,D., Aunis,D., Leid,M. and Schaeffer,E. (2003) Recruitment of Tat to Heterochromatin Protein HP1 via Interaction with CTIP2 Inhibits Human Immunodeficiency Virus Type 1 Replication in Microglial Cells. *J Virol*, **77**, 5415–5427.
12. Cherrier,T., Le Douce,V., Eilebrecht,S., Riclet,R., Marban,C., Dequiedt,F., Goumon,Y., Paillart,J.-C., Mericskay,M., Parlakian,A., *et al.* (2013) CTIP2 is a negative regulator of P-TEFb. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 12655–12660.
13. Eilebrecht,S., Le Douce,V., Riclet,R., Targat,B., Hallay,H., Van Driessche,B., Schwartz,C., Robette,G., Van Lint,C., Rohr,O., *et al.* (2014) HMGA1 recruits CTIP2-repressed P-TEFb to the HIV-1 and cellular target promoters. *Nucleic Acids Res*, **42**, 4962–4971.
14. Desplats,P., Dumaop,W., Smith,D., Adame,A., Everall,I., Letendre,S., Ellis,R., Cherner,M., Grant,I. and Masliah,E. (2013) Molecular and pathologic insights from latent HIV-1 infection in the human brain. *Neurology*, **80**, 1415–1423.
15. Moosmann,P., Georgiev,O., Le Douarin,B., Bourquin,J.P. and Schaffner,W. (1996) Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1. *Nucleic Acids Res.*, **24**, 4859–4867.
16. Iyengar,S. and Farnham,P.J. (2011) KAP1 Protein: An Enigmatic Master Regulator of the Genome. *J Biol Chem*, **286**, 26267–26276.
17. Iyengar,S., Ivanov,A.V., Jin,V.X., Rauscher,F.J. and Farnham,P.J. (2011) Functional analysis of KAP1 genomic recruitment. *Mol. Cell. Biol.*, **31**, 1833–1847.
18. Liang,Q., Deng,H., Li,X., Wu,X., Tang,Q., Chang,T.-H., Peng,H., Rauscher,F.J., Ozato,K. and Zhu,F. (2011) Tripartite motif-containing protein 28 is a small ubiquitin-related modifier E3 ligase and negative regulator of IFN regulatory factor 7. *J. Immunol.*, **187**, 4754–4763.
19. Riclet,R., Chendeb,M., Vonesch,J.-L., Koczan,D., Thiesen,H.-J., Losson,R. and Cammas,F. (2009) Disruption of the Interaction between Transcriptional Intermediary Factor 1β and Heterochromatin Protein 1 Leads to a Switch from DNA Hyper- to Hypomethylation and H3K9 to H3K27 Trimethylation on the MEST Promoter Correlating with Gene Reactivation. *Mol Biol Cell*, **20**, 296–305.
20. Yang,Y., Fiskus,W., Yong,B., Atadja,P., Takahashi,Y., Pandita,T.K., Wang,H.-G. and Bhalla,K.N. (2013) Acetylated hsp70 and KAP1-mediated Vps34 SUMOylation is required for autophagosome creation in autophagy. *Proc Natl Acad Sci U S A*, **110**, 6841–6846.
21. Garvin,A.J., Densham,R.M., Blair-Reid,S.A., Pratt,K.M., Stone,H.R., Weekes,D., Lawrence,K.J. and Morris,J.R. (2013) The deSUMOylase SENP7 promotes chromatin relaxation for homologous recombination DNA repair. *EMBO Rep*, **14**, 975–983.
22. Groh,S. and Schotta,G. (2017) Silencing of endogenous retroviruses by heterochromatin. *Cell. Mol. Life Sci.*, **74**, 2055–2065.
23. Allouch,A., Di Primio,C., Alpi,E., Lusic,M., Arosio,D., Giacca,M. and Cereseto,A. (2011) The TRIM Family Protein KAP1 Inhibits HIV-1 Integration. *Cell Host & Microbe*, **9**, 484–495.
24. Nishitsuji,H., Abe,M., Sawada,R. and Takaku,H. (2012) ZBRK1 represses HIV-1 LTR-mediated transcription. *FEBS Letters*, **586**, 3562–3568.
25. Nishitsuji,H., Sawada,L., Sugiyama,R. and Takaku,H. (2015) ZNF10 inhibits HIV-1 LTR activity through interaction with NF-κB and Sp1 binding motifs. *FEBS Letters*, **589**, 2019–2025.
26. McNamara,R.P., Reeder,J.E., McMillan,E.A., Bacon,C.W., McCann,J.L. and D'Orso,I. (2016) KAP1 Recruitment of the 7SK snRNP Complex to Promoters Enables Transcription Elongation by RNA Polymerase II. *Mol. Cell*, **61**, 39–53.
27. Morton,E.L., Forst,C.V., Zheng,Y., DePaula-Silva,A.B., Ramirez,N.-G.P., Planelles,V. and D'Orso,I. (2019) Transcriptional Circuit Fragility Influences HIV Proviral Fate. *Cell Rep*, **27**, 154-171.e9.
28. Lee,A., CingÖz,O., Sabo,Y. and Goff,S.P. (2018) Characterization of interaction between Trim28 and YY1 in silencing proviral DNA of Moloney murine leukemia virus. *Virology*, **516**, 165–175.
29. Wolf,D. and Goff,S.P. (2009) Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature*, **458**, 1201–1204.
30. Taura,M., Song,E., Ho,Y.-C. and Iwasaki,A. (2019) Apobec3A maintains HIV-1 latency through recruitment of epigenetic silencing machinery to the long terminal repeat. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 2282–2289.
31. Ma,X., Yang,T., Luo,Y., Wu,L., Jiang,Y., Song,Z., Pan,T., Liu,B., Liu,G., Liu,J., *et al.* (2019) TRIM28 promotes HIV-1 latency by SUMOylating CDK9 and inhibiting P-TEFb. *Elife*, **8**.

32. Li,X., Lee,Y.-K., Jeng,J.-C., Yen,Y., Schultz,D.C., Shih,H.-M. and Ann,D.K. (2007) Role for KAP1 Serine 824 Phosphorylation and Sumoylation/Desumoylation Switch in Regulating KAP1-mediated Transcriptional Repression. *J. Biol. Chem.*, **282**, 36177–36189.
33. Janabi,N., Peudenier,S., Héron,B., Ng,K.H. and Tardieu,M. (1995) Establishment of human microglial cell lines after transfection of primary cultures of embryonic microglial cells with the SV40 large T antigen. *Neurosci. Lett.*, **195**, 105–108.
34. Wires,E.S., Alvarez,D., Dobrowolski,C., Wang,Y., Morales,M., Karn,J. and Harvey,B.K. (2012) Methamphetamine activates nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) and induces human immunodeficiency virus (HIV) transcription in human microglial cells. *J. Neurovirol.*, **18**, 400–410.
35. Kutsch,O., Benveniste,E.N., Shaw,G.M. and Levy,D.N. (2002) Direct and Quantitative Single-Cell Analysis of Human Immunodeficiency Virus Type 1 Reactivation from Latency. *J Virol*, **76**, 8776–8786.
36. Ryan,R.F., Schultz,D.C., Ayyanathan,K., Singh,P.B., Friedman,J.R., Fredericks,W.J. and Rauscher,F.J. (1999) KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Krüppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. *Mol. Cell. Biol.*, **19**, 4366–4378.
37. Coiras,M., Camafeita,E., Ureña,T., López,J.A., Caballero,F., Fernández,B., López-Huertas,M.R., Pérez-Olmeda,M. and Alcamí,J. (2006) Modifications in the human T cell proteome induced by intracellular HIV-1 Tat protein expression. *Proteomics*, **6**, S63–S73.
38. Ruben,S., Perkins,A., Purcell,R., Joung,K., Sia,R., Burghoff,R., Haseltine,W.A. and Rosen,C.A. (1989) Structural and functional characterization of human immunodeficiency virus tat protein. *J. Virol.*, **63**, 1–8.
39. Ali,A. and Banerjea,A.C. (2016) Curcumin inhibits HIV-1 by promoting Tat protein degradation. *Sci Rep*, **6**, 27539.
40. Balachandran,A., Wong,R., Stoilov,P., Pan,S., Blencowe,B., Cheung,P., Harrigan,P.R. and Cochrane,A. (2017) Identification of small molecule modulators of HIV-1 Tat and Rev protein accumulation. *Retrovirology*, **14**.
41. Gargano,B., Fiorillo,M., Amente,S., Majello,B. and Lania,L. (2008) p14ARF is capable of promoting HIV-1 tat degradation. *Cell Cycle*, **7**, 1433–1439.
42. Hong,H.-W., Lee,S.-W. and Myung,H. (2013) Induced degradation of Tat by nucleocapsid (NC) via the proteasome pathway and its effect on HIV transcription. *Viruses*, **5**, 1143–1152.
43. Li,J., Chen,C., Ma,X., Geng,G., Liu,B., Zhang,Y., Zhang,S., Zhong,F., Liu,C., Yin,Y., *et al.* (2016) Long noncoding RNA NRON contributes to HIV-1 latency by specifically inducing tat protein degradation. *Nat Commun*, **7**.
44. Wan,Z. and Chen,X. (2014) Triptolide inhibits human immunodeficiency virus type 1 replication by promoting proteasomal degradation of Tat protein. *Retrovirology*, **11**, 88.
45. Ivanov,A.V., Peng,H., Yurchenko,V., Yap,K.L., Negorev,D.G., Schultz,D.C., Psulkowski,E., Fredericks,W.J., White,D.E., Maul,G.G., *et al.* (2007) PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. *Mol. Cell*, **28**, 823–837.
46. Cheng,C.-T., Kuo,C.-Y. and Ann,D.K. (2014) KAPtain in charge of multiple missions: Emerging roles of KAP1. *World J Biol Chem*, **5**, 308–320.
47. Cismasiu,V.B., Paskaleva,E., Suman Daya,S., Canki,M., Duus,K. and Avram,D. (2008) BCL11B is a general transcriptional repressor of the HIV-1 long terminal repeat in T lymphocytes through recruitment of the NuRD complex. *Virology*, **380**, 173–181.
48. Schultz,D.C., Friedman,J.R. and Rauscher,F.J. (2001) Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. *Genes Dev.*, **15**, 428–443.
49. Quénet,D., Gasser,V., Fouillen,L., Cammas,F., Sanglier-Cianferani,S., Losson,R. and Dantzer,F. (2008) The histone subcode: poly(ADP-ribose) polymerase-1 (Parp-1) and Parp-2 control cell differentiation by regulating the transcriptional intermediary factor TIF1beta and the heterochromatin protein HP1alpha. *FASEB J.*, **22**, 3853–3865.
50. Tomimatsu,N., Mukherjee,B. and Burma,S. (2009) Distinct roles of ATR and DNA-PKcs in triggering DNA damage responses in ATM-deficient cells. *EMBO Rep.*, **10**, 629–635.
51. Cherrier,T., Suzanne,S., Redel,L., Calao,M., Marban,C., Samah,B., Mukerjee,R., Schwartz,C., Gras,G., Sawaya,B.E., *et al.* (2009) p21(WAF1) gene promoter is epigenetically silenced by CTIP2 and SUV39H1. *Oncogene*, **28**, 3380–3389.
52. Imai,K., Asamitsu,K., Victoriano,A.F.B., Cueno,M.E., Fujinaga,K. and Okamoto,T. (2009) Cyclin T1 stabilizes expression levels of HIV-1 Tat in cells. *FEBS J.*, **276**, 7124–7133.

53. Jin,X., Pan,Y., Wang,L., Zhang,L., Ravichandran,R., Potts,P.R., Jiang,J., Wu,H. and Huang,H. (2017) MAGE-TRIM28 complex promotes the Warburg effect and hepatocellular carcinoma progression by targeting FBP1 for degradation. *Oncogenesis*, **6**, e312.
54. Okamoto,K., Kitabayashi,I. and Taya,Y. (2006) KAP1 dictates p53 response induced by chemotherapeutic agents via Mdm2 interaction. *Biochem. Biophys. Res. Commun.*, **351**, 216–222.
55. Pineda,C.T., Ramanathan,S., Fon Tacer,K., Weon,J.L., Potts,M.B., Ou,Y.-H., White,M.A. and Potts,P.R. (2015) Degradation of AMPK by a cancer-specific ubiquitin ligase. *Cell*, **160**, 715–728.
56. Avram,D. and Califano,D. (2014) The multifaceted roles of Bcl11b in thymic and peripheral T cells - impact on immune diseases. *J Immunol*, **193**, 2059–2065.
57. Czerwińska,P., Mazurek,S. and Wiznerowicz,M. (2017) The complexity of TRIM28 contribution to cancer. *J Biomed Sci*, **24**.
58. Czerwińska,P., Shah,P.K., Tomczak,K., Klimczak,M., Mazurek,S., Sozańska,B., Biecek,P., Korski,K., Filas,V., Mackiewicz,A., *et al.* (2017) TRIM28 multi-domain protein regulates cancer stem cell population in breast tumor development. *Oncotarget*, **8**, 863–882.
59. Hu,M., Fu,X., Cui,Y., Xu,S., Xu,Y., Dong,Q. and Sun,L. (2015) Expression of KAP1 in epithelial ovarian cancer and its correlation with drug-resistance. *Int J Clin Exp Med*, **8**, 17308–17320.
60. Wang,Y.-Y., Li,L., Zhao,Z.-S. and Wang,H.-J. (2013) Clinical utility of measuring expression levels of KAP1, TIMP1 and STC2 in peripheral blood of patients with gastric cancer. *World J Surg Oncol*, **11**, 81.

TABLE AND FIGURES LEGENDS

Figure 1: KAP1 suppresses HIV-1 expression (A-D) Microglial cells were transfected with the pNL4-3 Δ ENV Luc vectors (A, B) or with episomal vector LTR-Luc (C, D) in the presence of an increasing dose of KAP1 (A, C), or with a KAP1 knockdown (shKAP1) (B, D). Luciferase activity was measured 48 hours post-transfection and after lysis of the cells. Luciferase values were normalized to those obtained with LTR-Luc or pNL4-3 Δ ENV Luc alone. A t-test was performed on 3 independent experiments *P* (* *P* <0.05, ** *P* <0.01, *** *P* <0.001). (E, F) Overexpressions of KAP1, as well as knockdown efficiencies, were validated by Western blotting.

Figure 2: KAP1 binds to the latent HIV-1 promoter (A) Microglial cells latently infected with HIV-1 containing the GFP reporter gene were treated for 24h with TNFα and HMBA. Cells were subjected to ChIP-qPCR experiments targeting the binding to the HIV-1 5' LTR promoter (B) The same microglial cells were also analyzed by flow cytometry. This figure is representative of 3 independent experiments.

Figure 3: Reactivation of HIV-1 gene transcription in myeloid models of latency upon KAP1 depletion (A-C) The HIV-1 infected monocytic (THP89) cells containing a GFP reporter gene were transduced either with non-targeting shRNA (indicated as shNT) or shKAP1. (A) The GFP expression was examined by flow cytometry and the KAP1 protein expression level was assessed on total cellular extract by western blot. (B, C) Total RNA products from shNT or shKAP1 THP89 transduced cells were retrotranscribed. GFP and HIV-1 gene transcripts as Initiated (TAR region), elongated (*gag*, *tat*), and multiple-spliced RNA (Ms RNA) were quantitated by real-time RT–PCR. The relative mRNA level was first normalized to GAPDH then to the shNT-transduced condition. Results are means from duplicate.

Figure 4: Kap1 and CTIP2 together contribute to the silencing of HIV-1 gene expression (A-D) Microglial cells were transfected with the pNL4-3 Δ ENV Luc vectors (A, B) or LTR-Luc (C, D) under the indicated conditions. The cells were lysed after 48h of transfection. The luciferase activities were measured and normalized to the conditions where the pNL4-3 Δ ENV Luc and LTR-Luc vectors were transfected alone. A t-test was performed on 3 independent experiments *P* (* *P* <0.05, ** *P* <0.01, *** *P* <0.001). (E, F) Over-expression and knockdown of the indicated proteins were validated by Western blot.

Figure 5: KAP1 inhibits Tat activity (A, B) Microglial cells were transfected with LTR-Luc vector under the indicated conditions. After 48h of transfection, the luciferase activities were measured after cell lysis

and normalized under conditions where LTR-Luc vector were transfected alone. A t-test was performed on 3 independent experiments *P* (* *P* <0.05, ** *P* <0.01, *** *P* <0.001).

Figure 6: KAP1 cooperates with CTIP2 in Tat repression (A, B) Microglial cells were transfected with LTR-Luc vector under the indicated conditions. After 48h of transfection, the luciferase activities were measured after cell lysis and normalized under conditions where LTR-Luc vector was transfected alone. A t-test was performed on 3 independent experiments *P* (* *P* <0.05, ** *P* <0.01, *** *P* <0.001). (C, D) Over-expression and knockdown of the indicated proteins were validated by Western blot.

Figure 7: KAP1 interacts and colocalize with Tat (A, B) HEK cells were transfected with the indicated vectors. After 48h of transfection, the nuclear protein extracts were immunoprecipitated with KAP1 (A) and flag (B) antibodies. The eluted protein complexes were analyzed by Western blotting for the presence of the indicated proteins. (C) Microglial cells were transfected for 48h with the indicated vectors and subjected to observation under a fluorescence microscope. Scale bars are set to 5µm.

Figure 8: KAP1 promotes Tat degradation via the proteasome pathway (A, B) HEK cells were transfected with the indicated vectors. 48h later, the nuclear proteins were analyzed by Western blot for the presence of KAP1 and Tat. (B) Tat expression level whose relative quantification to the α-tubulin was carried out using the image J software. (C) The RNA extracts from HEK cells transfected with the indicated plasmids were submitted to RT-qPCR experiments against Tat. The relative mRNA level of Tat was normalized to the GAPDH gene. (D) HEK cells were transfected with the indicated vectors. After 18h of transfection, the cells were treated or not with 50 µM of MG132 for 6h. 24h post-transfection, the total protein extracts were analyzed by Western blot for the presence of KAP1 and Tat. Tat expression level was quantified relatively to the β-actin expression, using image J software. A t-test was performed on 3 (B, C) and 5 (D) independent experiments *P* (* *P* <0.05, ** *P* <0.01, *** *P* <0.001).

Figure 9: Tat degradation is mediated by the PHD and Bromo domains of KAP1 (A, B) HEK cells were transfected as indicated. 48h later, the nuclear proteins were extracted and tested by Western blot for the expression of the indicated proteins.

Figure 10: CTIP2 interacts with KAP1 in an RNA independent manner and competes with Tat for KAP1 binding (A, B) HEK cells were transfected with flag-CTIP2 (A) or flag-KAP1 and Tap-CTIP2 (B). 48 h post-transfection, the nuclear protein extracts, treated or not with RNase, were subjected to immunoprecipitation with the flag antibody. The immunoprecipitated complexes were tested by Western blot for the presence of KAP1 (A) and CTIP2 (B). (C) The nuclear protein extracts of microglial cells were analyzed by Western blot for the expression of CTIP2 and KAP1 and submitted to immunoprecipitation experiments with CTIP2 and KAP1 antibodies. (D) The HEK cells were transfected with Tap-CTIP2 and one of the vectors encoding f-KAP1 WT: wild type, or KAP1 deleted from the RBCC domain (ΔRBCC), or PHD (ΔPHD), or Bromo (ΔBromo) or both PHD and Bromo domains (ΔPHD / ΔBromo), or with f-CTIP2 deletion mutants, as indicated (E). After 48h, the nuclear protein extracts were immunoprecipitated with flag antibody. The eluted protein complexes were analyzed by Western blot for the presence of CTIP2 and KAP1. (F) HEK cells were transfected with the indicated vectors in the presence of an increasing dose of Tat. After 48h, the nuclear protein extracts were treated as in D. The results are representative of at least two independent experiments.

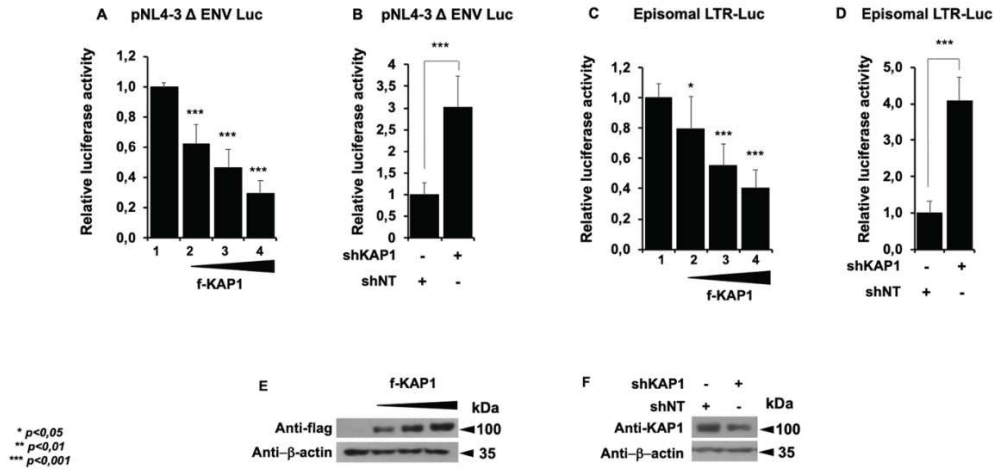


Figure 1: KAP1 suppresses HIV-1 expression (A-D) Microglial cells were transfected with the pNL4-3 Δ ENV Luc vectors (A, B) or with episomal vector LTR-Luc (C, D) in the presence of an increasing dose of KAP1 (A, C), or with a KAP1 knockdown (shKAP1) (B, D). Luciferase activity was measured 48 hours post-transfection and after lysis of the cells. Luciferase values were normalized to those obtained with LTR-Luc or pNL4-3 Δ ENV Luc alone. A t-test was performed on 3 independent experiments *P* (* *P* <0.05, ** *P* <0.01, *** *P* <0.001). (E, F) Overexpressions of KAP1, as well as knockdown efficiencies, were validated by Western blotting.

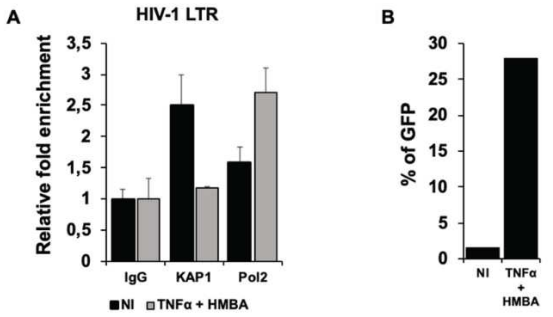


Figure 2: KAP1 binds to the latent HIV-1 promoter (A) The microglial cells latently infected with HIV-1 containing the GFP reporter gene were treated for 24h with TNFα and HMBA. The cells were subjected to ChIP-qPCR experiments targeting the binding to the HIV-1 5' LTR promoter (B) The same microglial cells were also analyzed by flow cytometry. This figure is representative of 3 independent experiments.

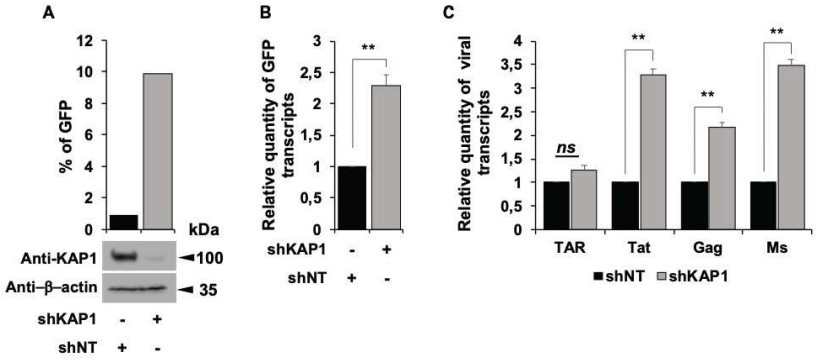


Figure 3: Reactivation of HIV-1 gene transcription in myeloid models of latency upon KAP1 depletion (A-C) The HIV-1 infected monocytic (THP89) cells containing a GFP reporter gene were transduced either with non-targeting shRNA (indicated as shNT) or shKAP1. (A) The GFP expression was examined by flow cytometry and the KAP1 protein expression level was assessed on total cellular extract by western blot. (B, C) Total RNA products from shNT or shKAP1 THP89 transduced cells were retrotranscribed. GFP and HIV-1 gene transcripts as Initiated (TAR region), elongated (*gag*, *tat*), and multiple-spliced RNA (Ms RNA) were quantitated by real-time RT–PCR. The relative mRNA level was first normalized to GAPDH then to the shNT-transduced condition. Results are means from duplicate.

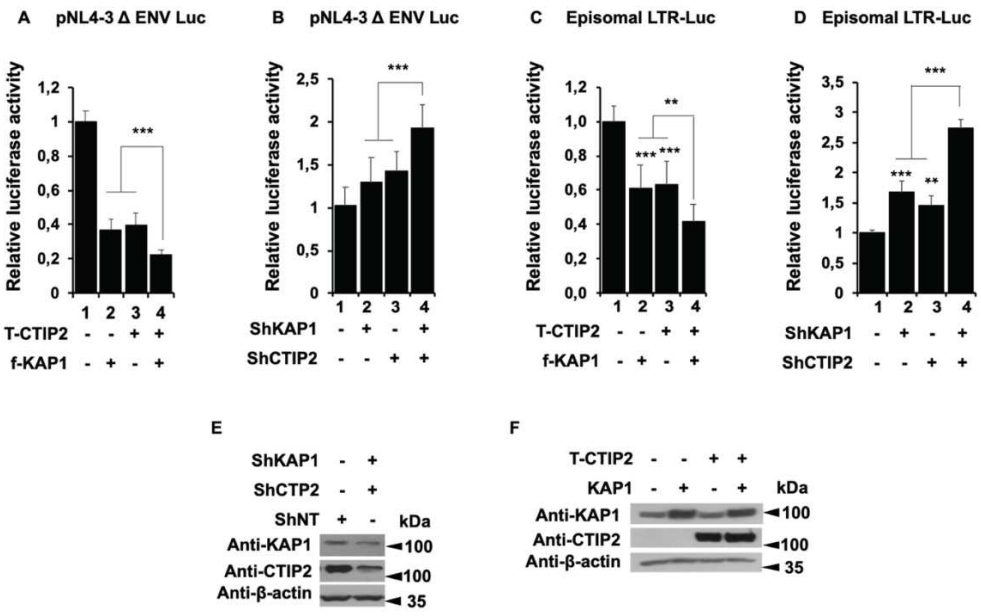


Figure 4: Kap1 and CTIP2 together contribute to the silencing of HIV-1 gene expression (A-D) Microglial cells were transfected with the pNL4-3 Δ ENV Luc vectors (A, B) or LTR-Luc (C, D) under the indicated conditions. The cells were lysed after 48h of transfection. The luciferase activities were measured and normalized to the conditions where the pNL4-3 Δ ENV Luc and LTR-Luc vectors were transfected alone. A t-test was performed on 3 independent experiments *P* (* *P* <0.05, ** *P* <0.01, *** *P* <0.001). (E, F) Over-expression and knockdown of the indicated proteins were validated by Western blot.

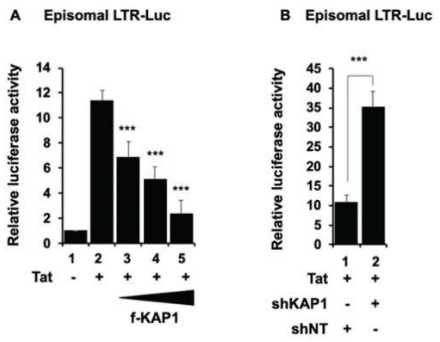


Figure 5: KAP1 inhibits Tat activity (A, B) Microglial cells were transfected with LTR-Luc vector under the indicated conditions. After 48h of transfection, the luciferase activities were measured after cell lysis and normalized under conditions where LTR-Luc vector were transfected alone. A t-test was performed on 3 independent experiments *P* (* *P* <0.05, ** *P* <0.01, *** *P* <0.001).

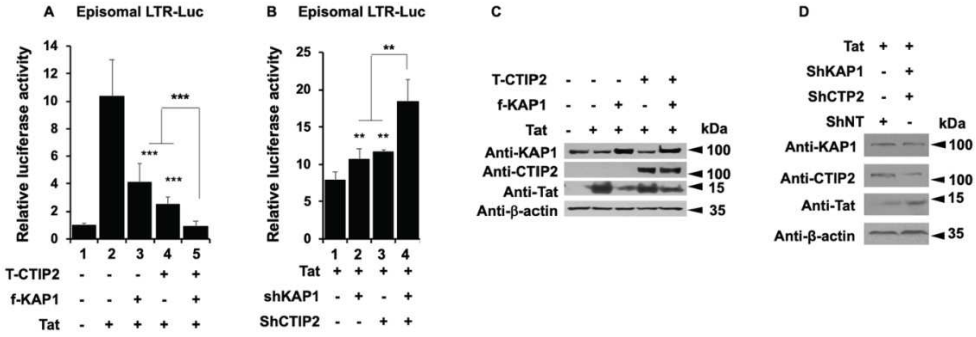


Figure 6: KAP1 cooperates with CTIP2 in Tat repression (A, B) Microglial cells were transfected with LTR-Luc vector under the indicated conditions. After 48h of transfection, the luciferase activities were measured after cell lysis and normalized under conditions where LTR-Luc vector was transfected alone. A t-test was performed on 3 independent experiments *P* (* *P* <0.05, ** *P* <0.01, *** *P* <0.001). (C, D) Over-expression and knockdown of the indicated proteins were validated by Western blot.

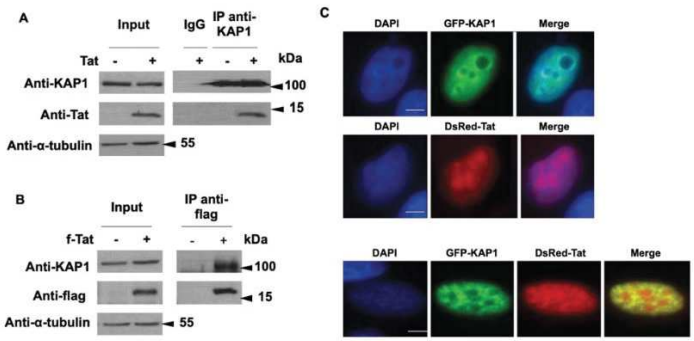


Figure 7: KAP1 interacts and colocalize with Tat (A, B) HEK cells were transfected with the indicated vectors. After 48h of transfection, the nuclear protein extracts were immunoprecipitated with KAP1 (A) and flag (B) antibodies. The eluted protein complexes were analyzed by Western blotting for the presence of the indicated proteins. (C) Microglial cells were transfected for 48h with the indicated vectors and subjected to observation under a fluorescence microscope. Scale bars are set to 5µm.

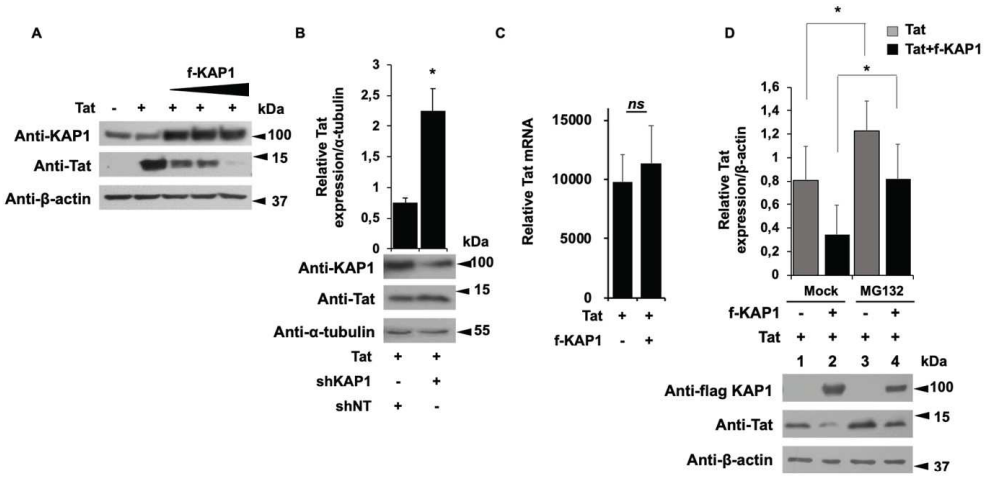


Figure 8: KAP1 promotes Tat degradation via the proteasome pathway (A, B) HEK cells were transfected with the indicated vectors. 48h later, the nuclear proteins were analyzed by Western blot for the presence of KAP1 and Tat. (B) Tat expression level whose relative quantification to the α-tubulin was carried out using the image J software. (C) The RNA extracts from HEK cells transfected with the indicated plasmids were submitted to RT-qPCR experiments against Tat. The relative mRNA level of Tat was normalized to the GAPDH gene. (D) HEK cells were transfected with the indicated vectors. After 18h of transfection, the cells were treated or not with 50 µM of MG132 for 6h. 24h post-transfection, the total protein extracts were analyzed by Western blot for the presence of KAP1 and Tat. Tat expression level was quantified relatively to the β-actin expression, using image J software. A t-test was performed on 3 (B, C) and 5 (D) independent experiments *P* (* *P* <0.05, ** *P* <0.01, *** *P* <0.001).

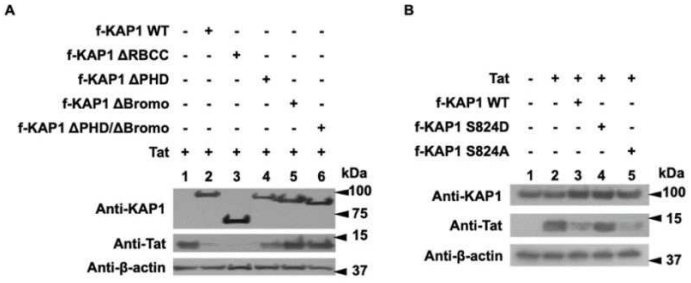


Figure 9: Tat degradation is mediated by the PHD and Bromo domains of KAP1 (A, B) HEK cells were transfected as indicated. 48h later, the nuclear proteins were extracted and tested by Western blot for the expression of the indicated proteins.

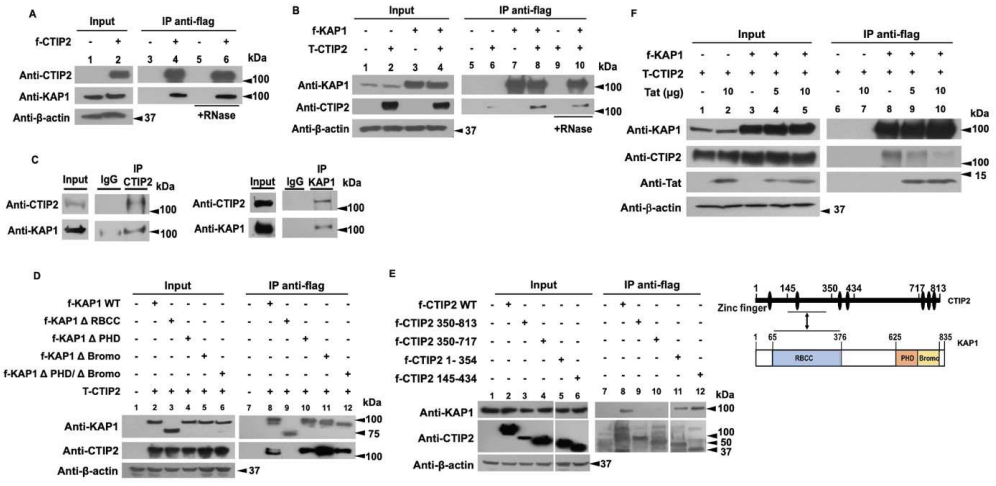


Figure 10: CTIP2 interacts with KAP1 in an RNA independent manner and competes with Tat for KAP1 binding (A, B) HEK cells were transfected with flag-CTIP2 (A) or flag-KAP1 and Tap-CTIP2 (B). 48 h post-transfection, the nuclear protein extracts, treated or not with RNase, were subjected to immunoprecipitation with the flag antibody. The immunoprecipitated complexes were tested by Western blot for the presence of KAP1 (A) and CTIP2 (B). (C) The nuclear protein extracts of microglial cells were analyzed by Western blot for the expression of CTIP2 and KAP1 and submitted to immunoprecipitation experiments with CTIP2 and KAP1 antibodies. (D) The HEK cells were transfected with Tap-CTIP2 and one of the vectors encoding f-KAP1 WT: wild type, or KAP1 deleted from the RBCC domain (ΔRBCC), or PHD (ΔPHD), or Bromo (ΔBromo) or both PHD and Bromo domains (ΔPHD / ΔBromo), or with f-CTIP2 deletion mutants, as indicated (E). After 48h, the nuclear protein extracts were immunoprecipitated with flag antibody. The eluted protein complexes were analyzed by Western blot for the presence of CTIP2 and KAP1. (F) HEK cells were transfected with the indicated vectors in the presence of an increasing dose of Tat. After 48h, the nuclear protein extracts were treated as in D. The results are representative of at least two independent experiments.

7. Publication 4

(In preparation)

Mass spectrometry and CLIP-seq of BCL11b reveal interactions with RNA processing pathways.

Haitham Sobhy, Marco De Rovere*, Amina Ait-Ammar*, Carine Van Lint, Christian Schwartz and Olivier Rohr*

University of Strasbourg, EA 7292, DHPI, IUT Louis Pasteur, France 2 Université Libre de Bruxelles, ULB, Gosselies, Belgium

* HS, MDR and AAA can be considered as equal contributors

Correspondence should be addressed to

Although BCL11b (B-cell lymphoma/leukemia 11B, CTIP2) is a well-known transcription repressor and tumor suppressor, its functions and cellular pathways are largely unknown. Here, we confirm that BCL11b interacts with RNA splicing/processing and nonsense-mediated decay (NMD) proteins, including FUS, SMN1, UPF1 and Drosha. Mass spectrometry analysis show that BCL11b interacts with histones, polymerases, and chromatin remodeling (CHD, SWI/SNF, and topoisomerase) proteins. BCL11b-bound RNAs were UV cross-linked and sequenced (CLIP-seq) showing that BCL11b binds to coding and noncoding RNAs (ncRNAs), which are involved in ribonucleoproteins or development. Surprisingly, BCL11b binds to RNA transcript and protein product of same genes (FUS, ESWR1, CHD and Tubulin). CLIP-seq analysis show BCL11b binds to NMD and retained introns transcripts. Our study is first genome-wide study of BCL11b-protein and BCL11b-RNA interactants, and first-time to show the direct interaction between BCL11b and gene agents of cancer and neurodegenerative diseases.

For this publication, I generated all the data concerning the RNA partners of CTIP2 including CLIP-Seq and RIP). I also contributed to validations of protein-protein bindings and to the discussions on the analysis of the data.

Introduction

The transcription regulator B-cell lymphoma/leukemia 11B (Bcl11b, also known as COUP-TF-interacting protein 2, CTIP2) plays crucial roles in the epigenetic regulation of gene transcription and specifically in the control of the elongation process by interacting with an inactive form of the p-TEFb complex (7SK RNA, CDK9, cyclin T1, and HEXIM1) [1] .

BCL11b, which harbors 6 zinc finger domains (ZnFs) of C2H2-type, shares 60% homology with the human BCL11a, as well as the mouse, chicken, and Xenopus homologues. BCL11 proteins were isolated from T-cell lines derived from patients with T-cell leukemia, suggesting the role of BCL11 proteins in blood cell development and lymphomagenesis, reviewed in [2-4] . Beside, BCL11b was deleted in γ-ray induced mouse lymphomas, suggesting its role as a tumor suppressor and modulator of radiation-induced DNA damages, reviewed in [5] . On the other hand, studying the function of BCL11b by knocking out the gene (loss- of-function) is challenging [6] . The mice that lack the two alleles of BCL11b die shortly after birth exhibiting defects in multiple tissues, including immune system, central nervous system (CNS), skin, hair cells in cochlea, teeth, and thymus among other organs [5] , which suggests the importance of BCL11b during the development. As consequences of absence of the BCL11b-deficient cells, studying the functions and pathways triggered by BCL11b is challenging.

In this study, we aimed to determine BCL11b-protein interactions and BCL11b-RNA interactions, using protein co-immunoprecipitation (co-IP) followed by mass spectrometry (MS) to identify the protein partners of BCL11b as well as UV cross-linking with immunoprecipitation (CLIP-seq), respectively, to define the pathways that could be triggered by BCL11b. Our results reveal interactions of BCL11b with cellular pathways dedicated to neuron development, the control of the cell cycle, RNA splicing and neurodegenerative diseases.

Results

Protein interactomics and pathways of BCL11b

To characterize the protein interactants and pathways triggered by BCL11b, we used HEK cells overexpressing Flag-BCL11b. After pulling-down proteins using anti-flag antibodies, we performed quantitative MS. After applying different statistical cut-offs, we identified 629

BCL11b protein partners, see method for details and supplementary data S1. Among the interacting proteins, we could confirm the previously described interactions between BCL11b and the P-TEFb complex (including CDK9 and CycT1) [1] , HDACs [7] , HMGA1 [8] , HP1 proteins [9] , and DCAF1 [10] as a validation of our experimental approach,
Figure 1A.

BCL11b interacts with proteins dedicated to genetic and epigenetic regulations

Performing gene ontology (GO) analysis (figure 1B and supplementary data S2), we observed that BCL11b interacts with proteins and complexes that are involved in cellular genetics including DNA replication and regulations of mitosis (Figure 1B). BCL11b is associated with proteins involved in DNA repair and cellular response to stimuli (such as BRCA2, MCMs, PARP1, and SMCs proteins). Indeed, BCL11b binds to more than 20% of the total protein components of the H2AX, the NuRD and the SNW1 complexes, among other proteins involved in chromatin organization, and DNA recombination (Figure 1C). As expected, BCL11b was found bound to protein complexes involved in the regulation of gene transcription (CCNT1, CDK9, SMARCAs and SMARCCs), and epigenetic regulators, including HDACs, CHDs, EHMTs, KMT2A (Figure 1B).

BCL11b interacts with structural proteins and the nuclear pore complex

Another significant GO term is the nucleocytoplasmic shuttling pathways, which include transport of the RNA or ribonucleoprotein complexes. Noteworthy, BCL11b-interacting nucleoporin (NUP133, NUP153, NUP93 and NUP214) have molecular roles beyond the nucleocytoplasmic shuttling, such as regulations of gene transcription and chromatin organization [11] . Additionally, BCL11b interacts with structural proteins involved in cytoskeleton, and motor activity, such as multiple isoforms of tubulin (supplementary data S1 and S2). BCL11b interacts with proteins involved in RNA translation and RNA splicing. The role of BCL11b during RNA processing have not been described by any previous research, interestingly, we found multiple BCL11b partner proteins are involved in RNA processing pathways such as translation of RNA and ribosome assembly, such as 40S and 60S ribosomal subunits (figure 1C). Moreover, BCL11b interacts with RNA-binding proteins, including RNA splicing and ribonucleoproteins (RNPs), such as spliceosomes. Spliceosomes are ribonucleoprotein complexes that regulate removing of the intronic regions from RNA.
Among these complexes, the survival motor neuron (SMN) protein complex, is made of SMN and GEMIN proteins [12] , which is assembled in nucleus and then translocalized to the cytoplasm. Other RNPs that are involved in post-transcriptional RNA regulation, include small

nuclear RNP (snRNPs), heterogeneous nuclear RNP (hnRNP), serine/arginine (SR)-rich proteins, and small nuclear RNA (snRNAs). Co-IP and immunoblot confirm that BCL11b interacts with SMN1 (Figure 1E and 1F). BCL11b interacts to 22 different hnRNPs or snRNPs, and about 17% of proteins that constitute the spliceosome (figure 1B-1D). Taken together, the GO analysis suggest that BCL11b may be involved in assembly and disassembly of the protein and ribonucleoprotein complexes.

BCL11b interacts with the Drosha and the DGCR8 protein complexes

Number of transcription factors can regulate RNA splicing, including FUS, which accumulates near the transcription start sites (TSS) and binds to transcription initiation factors to inhibit phosphorylations of the RNA polymerase 2 C-terminal domain (POLR2-CTD)
[13]. As a member of the Drosha and the DGCR8 complexes [14] , FUS is involved in DNA damage response, regulations of RNA splicing and transports of RNA [13] . By IP-MS, we found BCL11b associated with 15 of the 20 proteins of the Drosha complex and 10 proteins of the 11 proteins of the DGCR8 complex (Figure 1D, and 1E). Co-IP experiments confirmed the association of BCL11b with FUS and Drosha (Figure 1F, 1G and 1H) and further suggest that BCL11b may co-ordinate with FUS and Drosha to regulate gene expression at transcriptional and post-transcriptional levels. Additionally, Ago, FUS and Drosha have additional roles to regulate translation of mRNA, or degradation of mRNA of certain transcripts [14] .

BCL11b binds to protein from the nonsense-mediated mRNA decay (NMD) pathway NMD pathway is crucial to decay mRNA transcripts with premature stop codons [15] . These transcripts lead to gain- or loss-of-function of some genes, which could be deleterious during the development. UPF1 is the hallmark of NMD pathway. Among the IP-MS significant hits, we observed interactions between BCL11b and proteins involved in NMD pathway. Out of 120 proteins identified in this specific GO molecular pathway, 61 were found bound to BCL11b (figure 1C). Co-IP experiments confirmed that BCL11b associates with UPF1, which is the hallmark of NMD pathway, and thereby this confirms that BCL11b interacts with the protein complex dedicated to NMD (Figure 1I). Again, these results offer additional evidences that BCL11b is involved in the regulation of gene expression at the post-transcriptional level, and it can regulate mRNA translation (Figure 1F-1H).

BCL11b is associated with cellular RNA

Although BCL11b has six ZnF domains, to our knowledge the RNA binding activities has not been described for any of the BCL11 proteins, except for BCL11b that can form a complex with CDK9, CycT1, and 7SK ncRNA [1, 8] . Our MS proteomics data show that BCL11b binds to RNA-binding proteins in complexes dedicated to RNA processing and RNA splicing. To further study the RNA-binding activity and obtain genome-wide landscape of BCL11b-bound RNAs, we cross-linked the BCL11b protein to cellular RNAs using UV radiation before immunoprecipitation and sequencing (CLIP-seq). These experiments have been performed by targeting the endogenous BCL11b in microglial cells, and an overexpressed Flag-BCL11b in HEK293 cells. We analyzed the results by two approaches, as the following:

First, we applied statistical cut-offs (see method in Suppl file S1), which lead to a dataset of 740 different RNAs (figure 2, 3A, S1-S5 and Suppl data 3-5). We found that about 96% of the BCL11b-bound RNAs are mapped to protein-coding genes (Table S1). BCL11b tends to bind to transcripts of long genes. Among the genes that have high numbers of CLIP reads, we found that about 20% are over 100kb in length, whereas the average length of these genes is 61kb, which is two-fold the genomics length (~30kb). The longest gene is Arf-GAP with GTPase, ANK repeat and PH domain-containing protein 1 (AGAP1, NCBI GeneID: 116987), which is 637,753 bp in length. In agreement with previous studies focusing on BCL11b control of cellular gene expression [5, 16] , the GO terms of BCL11b-bound RNA include gene from neuronal development and neurogenesis of multiple parts of brain (e.g. midbrain, hindbrain and substantia nigra). BCL11b binds to RNA of proteins that have been described to regulate proliferation, regeneration and migration of neuronal and non-neuronal cells, including glial cell and Schwann cell (details are found in figure 3B, Suppl data 4, Figure S6- S8). However, BCL11b was also found associated with RNA from gene involved in the development of the immune system, hemopoiesis, blood vessels, and hair follicles. Finally, we found RNA coding for actin and tubulin associated with BCL11b as well as RNA coding for proteins of the major and the minor spliceosome (Figure 1E, Table S2, Suppl data 5). As a second analysis approach, we divide the whole genome into 10kb regions (bins) and we count the number of reads corresponding to each bin (see method in Suppl file S1). Only regions that pass 5 folds increase to the control were selected, resulting in 593 bins. The results show that over 60% of CLIP reads are mapped down-stream of the transcription start sites (TSS), and more than 30% are located within 5-50kb down-stream of the TSS (Figure S9, Suppl data 6). We concluded that about half of the CLIP reads are within the genes and at least one third could be located in the intergenic regions. Aligning the sequences, we found two conserved motifs (CUCRGCCU and

UCCCAGCW) in BCL11b-bound RNA regions (figure 3C). Gene ontology analysis for the genes localized in these regions suggested their involvement into structural process, extracellular matrix organization, and response to stress (Suppl data 6). The gene products are involved in angiogenesis, p53 and inflammation pathways, cytoskeletal regulation and multiple diseases such as cancers, Huntington's disease and Alzheimer (Figure 10, Suppl data 6).
Taken together, our results demonstrate that half of the BCL11b-binding RNAs are transcribed from long genes and mapped to the genic regions further suggesting an impact of BCL11b on protein expression or function at post-transcriptional levels.

BCL11b binds to RNA transcripts and proteins product of the same gene

We found BCL11b associated with 137 RNA transcripts and proteins encoded by the same gene (Figure 3A, and full list of genes can be found in Suppl data 7). Among them, we found EWSR1, FUS, SFPQ, UPF1, tubulins, CHDs and hnRNPs transcripts and proteins. Interestingly, GO analysis reveals that these proteins are involved in RNA processing, NMD pathway and translation (Figure 3D, Suppl data 8). Interestingly, tubulin, for example, may use a unique mechanism to regulate its expression level within the cell [17]. Over-expression of tubulin gene does not lead to increase its protein expression level. Tubulin auto-regulates its expression level to ensure a stable production of the protein in cells. The mechanism of tubulin auto-regulation is largely unknown [17] . However, tubulins α and β could polymerize through an unknown mediator, which then bind to the nascent tubulin peptide. By unknown mechanism, the mediator can bind to ribosomes and/or ribonucleases to terminate the translation. We found BCL11b interacts with multiple isoforms of tubulin proteins, tubulin RNAs, as well as ribonucleases, and ribosomes (Figure 3E). In addition, we have noted that BCL11b binds to multiple members of the same protein family, such as tubulins, actins, RNPs, or ribosomes. Together, these results may support the involvement of BCL11b in the regulation of protein expressions.

BCL11b binds differentially transcripts of the same genes

To further study the association of BCL11b with cellular RNA, we counted the number of reads per exon for CLIP reads that mapped to genes. We found that more than 75% of the CLIP reads were mapped to exons (Figure 3F). Having a deeper look to the CLIP-seq reads on genome browsers, we found that when reads were mapped to intronic regions of genes such as EWSR1, FUS, and SRRM2, the locations of the reads corresponded to the coordinates of RNA

isoforms that do not code for proteins (we refer to them here as INCP), such as retained introns (IR), nonsense mediated decay (NMD), non-stop decay or processed transcript, (Figure 2, Suppl Figure S2-5). This result is consistent with our results demonstrating that BCL11b interacts with RNA splicing and processing, and NMD pathway proteins, including, FUS, SMN1, RNPs, Drosha, and UPF1. These results provide additional evidences that BCL11b can regulate gene expression at post-transcriptional levels (Figure1F-1I).

To strengthen this finding, we collected the transcripts (or RNA isoforms) of 23 genes that encode for isoforms coding for proteins (ICPs), and isoforms that does not code for proteins (INCPs), such as NMD and IR. We then counted the number of reads that correspond to each isoform. We found that BCL11b binds to different isoforms of the same gene in differential manner (Figure 3G, Suppl data 9). For example, TOP3A gene encodes for 21 RNA isoforms, similarly, ESWR1, Drosha, SIRT7 and RAD52 genes encode for 20-21 RNA isoforms. The fold increase to control of CLIP-seq for TOP3A ranges from 1.1 to 2.7 fold (log2 fold change < 1.5), whereas, the binding to RNA of RAD52 is up to 7-fold increase (log2 fold change ranges from 0.8 to 2.8 fold) (Figure 3G). The same concept applies to SMN1 (10 isoforms), COL27A1 (8 isoforms) and RAD51 (11 isoforms). We counted the CLIP-seq reads per each transcripts (isoforms) and then normalized to the length of the transcripts. We found that isoforms coding for protein (ICPs) have 2 folds CLIP reads than INCPs (Figure 3H). The Spearman correlation rank (R) between the CLIP reads and transcripts length for ICPs is 0.51 and 0.44 for INCP, which suggest that BCL11b tends to bind to long isoforms of ICPs and less degree INCPs. Together these results suggests that BCL11b preferentially binds to long RNA isoforms that code for proteins.

We still noticed some CLIP reads that correspond to intronic regions within ICPs, but the CLIP reads are mapped to the coordinates of exons in another INCPs isoforms. Noteworthy, NMD or IR transcripts introduce additional exons instead of introns. To explain this concept, we highlighted the genomics introns in pink colors, see (Figure 2 and Figures S2-S5).

Although these intron regions are usually removed in case of ICP transcripts (thick dark bars in Figure 2, e panel), we can find that part of these regions correspond to exons in INCPs (light blue bars in Figure 2, e panel). Computationally, we counted number of reads mapped to exons and introns, we normalized them to number of exons or introns as shown in Figure 3I and Suppl data 10. If BCL11b binds to RNA that correspond to exons, we expect to find most of the CLIP reads are biased to exons coordinates. However, in Figure 3I, we observed that most of the reads are mapped to introns. In case of INCPs, the number of reads were almost equally distributed throughout the exons and introns. Statistically, we calculated the Spearman

correlation rank (R) between number of reads and number of exons, which show that number of exons correlate with number of CLIP reads, in case of ICPs, but not in case of INCPs (R of INCPs=0.2, and R of ICPs=0.46). This result suggests that INCPs with few numbers of exons have multiple CLIP reads. DNA (cytosine-5-)-methyltransferase 1 (DNMT1) transcripts are good examples that support this observation. Number of reads mapped to IR transcripts are 2-3 fold higher than the number of reads mapped to INCPs transcripts, such as IR transcripts (ENST00000591764), which has 5 exons, are 3-fold reads to those mapped to the protein coding transcript (ENST00000588952), which has 9 exons (Figure 3J, Suppl data 10).

Taken together, Our CLIP-seq results confirm that BCL11b can form complexes with RNA and ribonucleoproteins. The results support our proteomics findings that BCL11b interacts with RNA processing and RNA splicing proteins, such as FUS, SMN1, Drosha complex and NMD complex (Figure 1). Our results further suggest that BCL11b may be able to differentiate between the types of RNA isoforms to contribute to the selection of the one to be translated. Together, BCL11b could interact with other proteins, such as FUS or UPF1 to regulate mRNA translation.

Involvement of BCL11b- associated factors in diseases

Described as a tumor suppressor gene, BCL11b expression level was reported to be modulated in T-Cell Lymphoma [3, 5] . Previous reports studying Ewing Sarcoma [18, 19] , spinal muscular atrophy (SMA) [20, 21] , and Parkinson's disease (PD) [22] showed differential expression of BCL11b. In addition, since BCL11b was reported to be involved in development, it is speculated that BCL11b could have roles in neurodegenerative diseases [4] , particularly in motor neuron amyotrophic lateral sclerosis (ALS) disease [2] , and Huntington’s disease [23] . However, as consequence of the lack of studies related to protein and RNA interactions of BCL11b nor its functional complexes, it was challenging to understand the exact role of BCL11b in these diseases. As we discussed above, using IP-MS, and CLIP-seq experiments, we detected interactions between BCL11b and multiple proteins that are involved in diseases such as:

EWSR1, which is known to contribute to Ewing's sarcoma [19] ;

SMN1, which is involved in SMA [21] ;

FUS, TDP-43 and TAF15 proteins that are involved in ALS [24] ;

SRRM2, which is a key factor causing Parkinson's disease [25] and 7SK, MALAT1, NEAT1 and TUG1 ncRNAs, SFPQ protein, RNPs and spliceosomes proteins that are known to contribute to cancer, as well as the motor neuron SMA and ALS diseases [26-28] .

Noteworthy, FUS associates with SMN, RNPs and RNA-binding proteins to form large complex [24, 29] . One third of these proteins were found bound to BCL11 (table S2, S3). On the other hand, DNA damage pathway may regulate RNA splicing process by selecting specific RNA isoforms to be translated [30] . FUS, Drosha and Ago2 are examples of regulators that are recruited in response to DNA damage to regulate gene expression [13, 14] .

Our CLIP-seq and RNA IP results confirmed the binding of BCL11b to the RNAs coding for FUS, SFPQ and SRRM2 genes and specifically to exon and exon-intron junctions regions (Figure 2, 4B, S1-S3, S11 and S12, table S4, suppl data 10). Noteworthy, we could not observe significant increase of the gene expression of RNA encoded from the same regions (using RT-qPCR) after over-expression of BCL11b (Figure 4C). These results allow to speculate that BCL11b could be able to differentiate between different, which suggest the potential role of BCL11b in mRNA translation regulation.

Discussion and Conclusion

This study describes for the first time BCL11b-associated proteome and RNAome. Our results suggest that BCL11b is involved in multiple pathways and function. In addition to its function as transcription regulator, BCL11b binds to proteins involved in cytoskeleton and development, as well as ribonucleic complexes such as ribosomes and RNPs. We confirmed that BCL11b binds to proteins involved in RNA processing, such as RNA splicing and NMD pathway. Our results show that BCL11b binds to DNA damage repair proteins, RNA splicing proteins, in addition to the ncRNA that are involved in RNA splicing. Interestingly, BCL11b as a transcription regulator and repressor can bind to chromatin remodeling and histone-modifying proteins (epigenetic modifications). One major GO term found is the assembly and disassembly of the complexes. We speculate that the major function of BCL11b is to assemble various complexes to perform a function. Once the function finished it may dis-assemble the complex.

There are number of regulators found in close association to chromatin as checkpoint safeguard. As response to cell cycle checkpoints, stimuli, or DNA damage, the safeguard factors regulate the expression level of some specific genes. Here, it emerges the importance of a regulator that can coordinate between replication, transcription and translation [30] . First, we demonstrate that BCL11b binds to RNA processing proteins (such as FUS, Drosha, Ago2, UPF1, and RNA splicing factors, Figure 1F-1I) speculating a role in the regulation of gene expression in response to DNA damage. Then, we highlighted the potential role of BCL11b in

autoregulation of some specific gene translation such as tubulin (Figure 3E). Based on our results, we found GO hits related to cell cycle checkpoint, and DNA damage response. We found that BCL11b preferentially bind to RNA isoforms coding for proteins (Figure 3F). We also observed multiple MS hits that correspond to 40S and 60S ribosome complexes (Figure 1C). On the other hand, BCL11b was found associated with isoforms that do not code for proteins (INCPs), such as NMD and IR. Generally, translation is regulated by selecting the desired mRNA for being translated, whereas, the RNA decay pathways, such as Drosha, UPF1 (NMD pathway), and Ago (RNA interference) are able to decay or silencing the undesirable RNA isoforms [31, 32] . One possibility is that BCL11b is associated in selecting of RNA isoforms. Isoform selecting (or switching) is a crucial process through which cells select a specific RNA isoform during the development or immune response. Some isoforms might be needed at specific steps during development, cell cycle or response to stimuli. Therefore, malfunction of RNA splicing or processing has a great impact on cell proliferation and cancer prognosis [31] .

In conclusion, our analysis of BCL11b-protein interactions and BCL11b-RNA interactions highlights the function of BCL11b as developmental regulator and tumor suppressor. BCL11b may bind to RNA processing proteins to regulate translation and/or select the desirable RNA isoform during development. One can speculate that depletion of BCL11b could lead to isoform switching and translation of wrong isoform of the protein (misfolded or short). Although the exact role of BCL11b in diseases deserve to be studied by future research, our results offer potential genome-wide RNA and proteins datasets, which will help future efforts to understand the molecular basis of cancer and neurodegenerative
diseases.

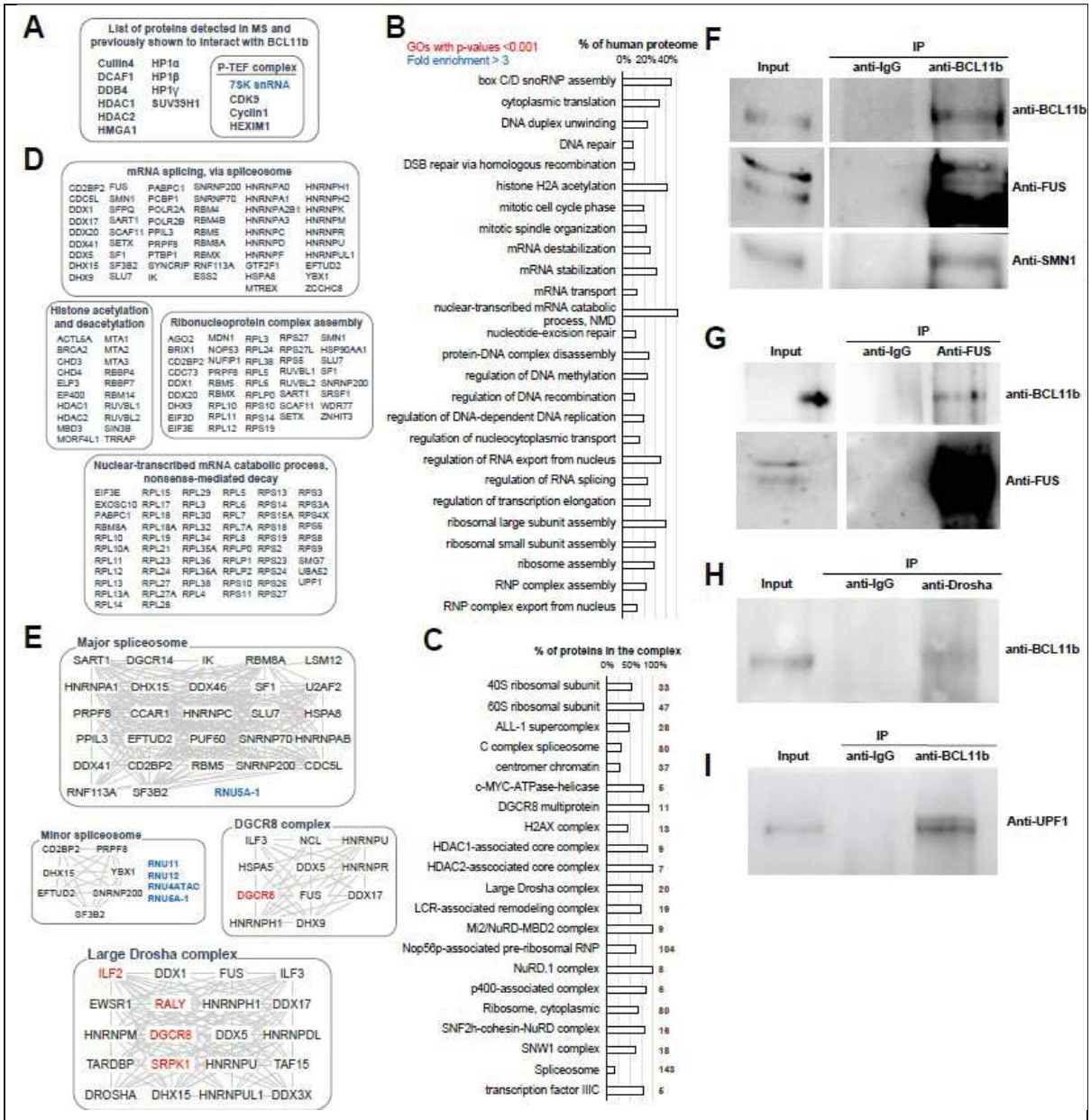


Figure 1. BCL11b binds to multiple complexes and pathways as deciphered by MS results. A) These proteins were previously shown to interact with BCL11b and we can detect them by IP-MS method. The gene names of BCL11b-bound ncRNAs are written in blue. B) GO terms of BCL11b-interacting proteins (BIPs), X-axis show percent of BIPs to the total number of proteins that identified in this category. C) Four representative GO terms with the list of proteins that detected to bind to BCL11b. D) The complexes that constitute of BIPs, X-axis show percent of BIPs to the total number of proteins in the complex; the total number of protein in the complex is written in dark blue outside the bars. E) Four representative complexes with the name of proteins. The BIPs are found in rectangle, BCL11b-interacting ncRNAs are in blue; whereas the non-BIPs are in hexagonal red shapes. The arrows refer to protein-protein interactions. F-I) western blot image show co-IP using antibodies directed to anti-BCL11b, anti-FUS, and anti-Drosha to detect BCL11b, FUS, SMN1, and UPF1. The antibodies (anti-) used to probe western blots are listed at left of the figure. The figures were constructed from Supplementary data files S1 and S2.

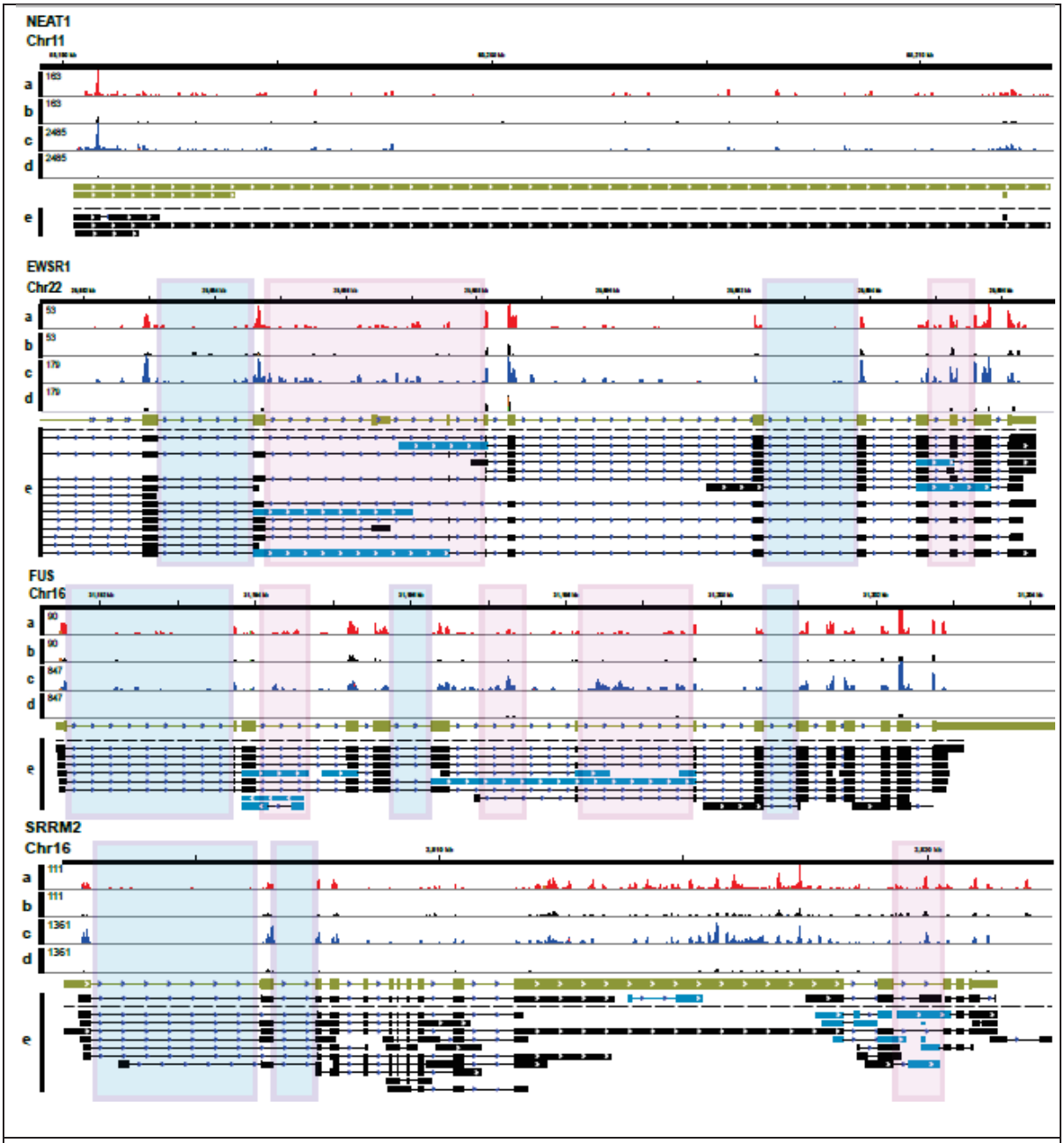


Figure 2. Representative CLIP-seq results of three protein-coding genes (EWSR1, FUS, and SRRM2), and non-protein coding gene (NEAT 1 ncRNA). The olive green bars refer to the NCBI RefSeq genes. X-axis represents the coordinates of the gene and the chromosome number. The Y-axis represents the number of CLIP reads in each condition. a) Over-expressed BCL11b in HEK cells, b) endogenous BCL11b in microglial cells, whereas, c and d are their control experiments, respectively. e) The black bars represent the GENCODE transcripts of the stated genes, the sky blue transcripts refer to the protein non-coding transcripts that harbor significant CLIP reads and their coordinates correspond to intron of RefSeq gene, these intronic regions are highlighted with light rose colored boxes. The light blue horizontal boxes refer to introns without significant BCL11b CLIP reads. For additional figures and CLIP-seq reads alignments, see Figure S1-S5.

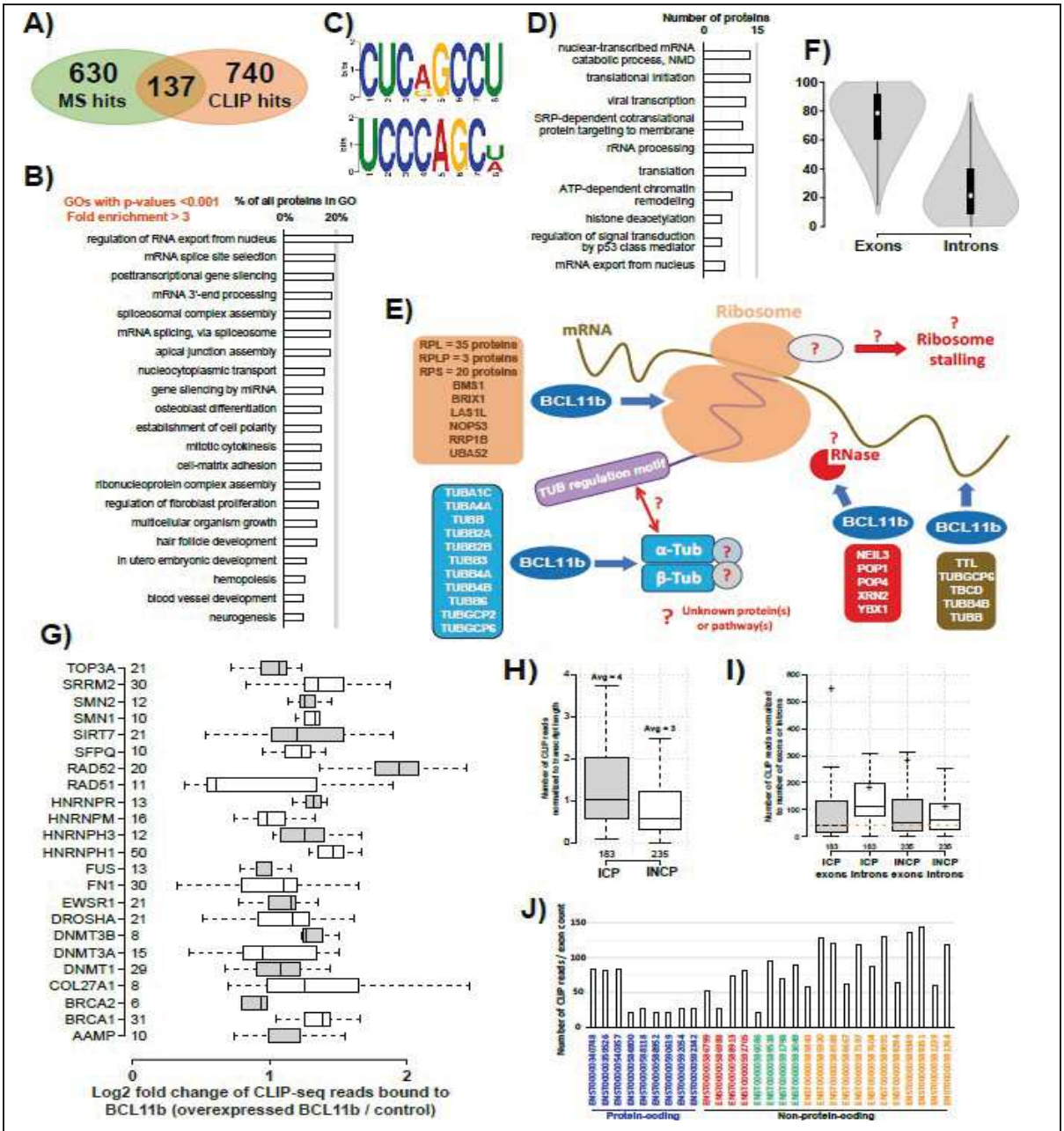
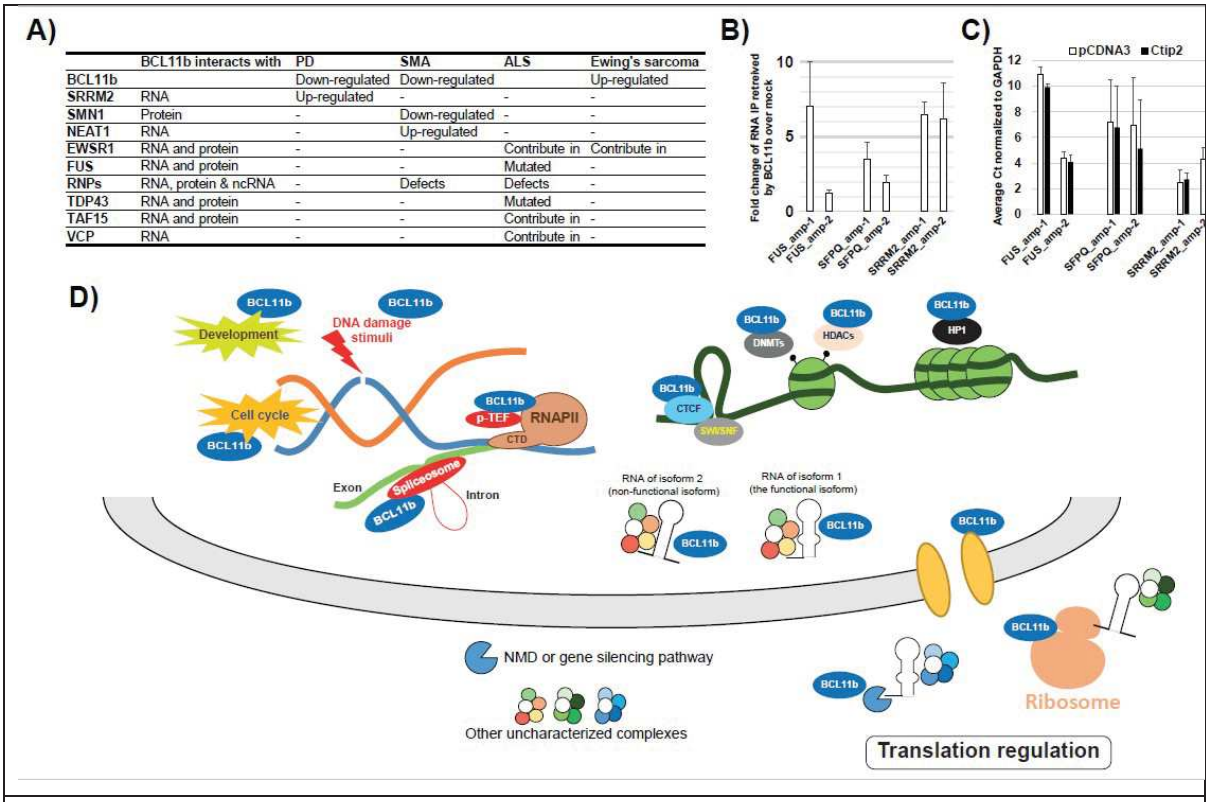


Figure 3. CLIP-seq results of BCL11b-bound RNA. A) Number of BCL11b-interacting proteins, from IP-MS experiment, that passed our cut-off (so-called MS hits); number of genes, which have CLIP reads passed our cut-off (so-called CLIP hits); 137 genes can bind to BCL11b in form of RNA or in form of protein. B) GO terms of the genes that passed CLIP- seq cut-offs. C) The predicted BCL11b-binding motif. D) GO terms of 137 shared genes (BCL11b binds to RNA and protein product). E) Schematic diagram of the hypothesis regarding tubulin autoregulation mechanism. BCL11b binds to multiple tubulin RNA, in brown box, it binds as well to ribosomes (oranges box), tubulins (blue box) and ribonucleases (red box). F) Violin shape represents the number of CLIP reads mapped to exons and introns of the significant genes. G) BCL11b binds differentially to the selected 23 genes. Number next ot the names are number of transcripts. H) Numbers of CLIP reads normalized to number of exons for isoforms coding for proteins (ICPs) or isoforms do not code for proteins (INCPs). I) Number of CLIP reads mapped to exons and those mapped to introns in case of ICPs or INCPs. J) Number of CLIP reads mapped to exons of DNMT1 gene are counted and normalized to number of exons. The figures 3G-3J were constructed from Supplementary data files S9.



	BCL11b interacts with	PD	SMA	ALS	Ewing's sarcoma
BCL11b		Down-regulated	Down-regulated		Up-regulated
SRRM2	RNA	Up-regulated	-	-	-
SMN1	Protein	-	Down-regulated	-	-
NEAT1	RNA	-	Up-regulated	-	-
EWSR1	RNA and protein	-	-	Contribute in	Contribute in
FUS	RNA and protein	-	-	Mutated	-
RNPs	RNA, protein & ncRNA	-	Defects	Defects	-
TDP43	RNA and protein	-	-	Mutated	-
TAF15	RNA and protein	-	-	Contribute in	-
VCP	RNA	-	-	Contribute in	-

Figure 4. A) The table shows the contribution of BCL11b and other genes in the Parkinson's disease (PD), SMA, ALS and Ewing sarcoma diseases. B) Results of RNA IP-qPCR (RIP-PCR) of three genes, the fold change of RNA bound to endogenous BCL11b to the input control. For the location of the amplicons on the gene, please refer to figure S3, and for PCR primers, Table S4. C) Gene expression of the amplicons used in G using overexpressed BCL11b plasmid and pCDNA3 vector. For raw Ct values presented in (B) and (C), see Supplementary data files S10. D) The graphical abstract shows the major proteins pathways, complexes and based on GO terms obtained after IP-MS and CLIP-seq experiments.

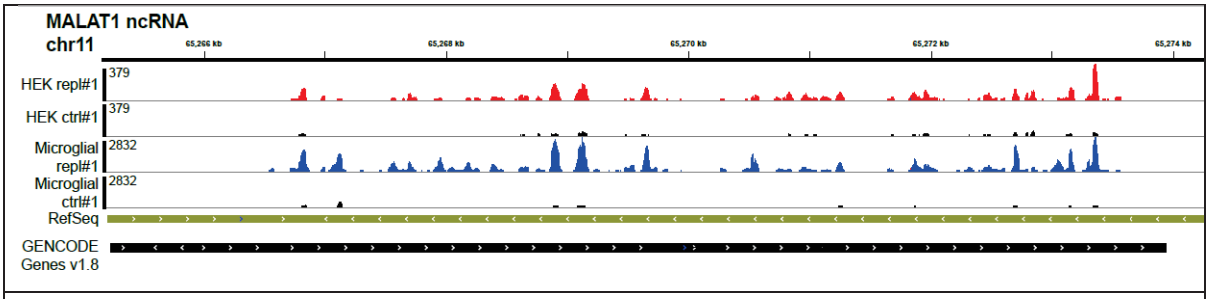


Figure S1

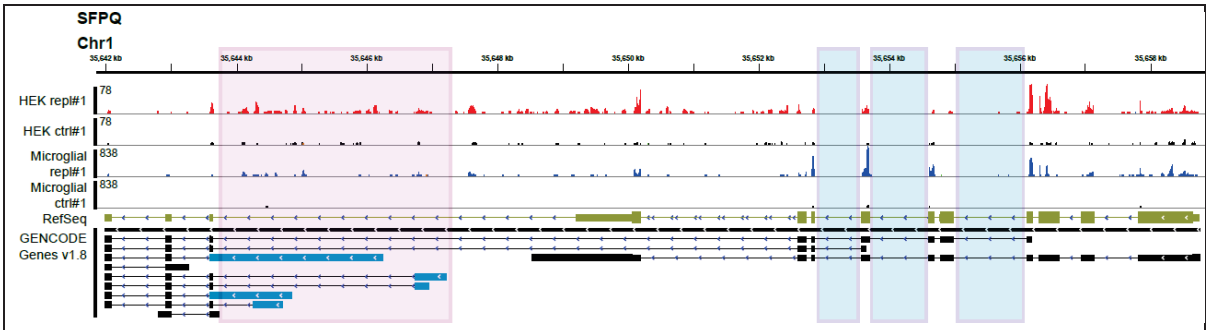


Figure S2

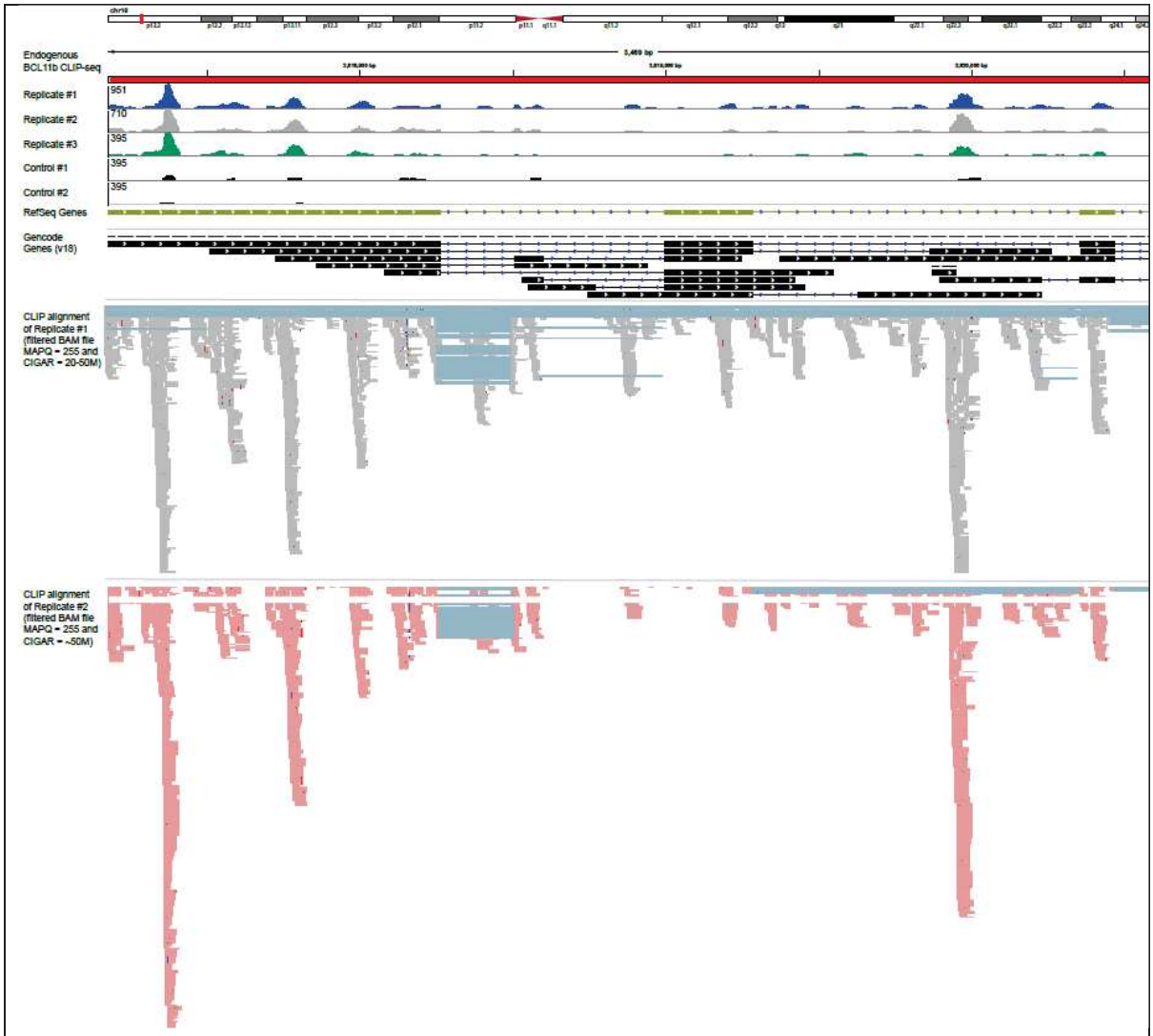

Endogenous
BCL11b CLIP-seq
Replicate #1
Replicate #2
Replicate #3
Control #1
Control #2
RefSeq Genes
Gencode
Genes (v18)
951
710
395
395
395
CLIP alignment
of Replicate #1
(filtered BAM file
MAPQ = 255 and
CIGAR = 20-50M)
CLIP alignment
of Replicate #2
(filtered BAM file
MAPQ = 255 and
CIGAR = ~50M)

Figure S3

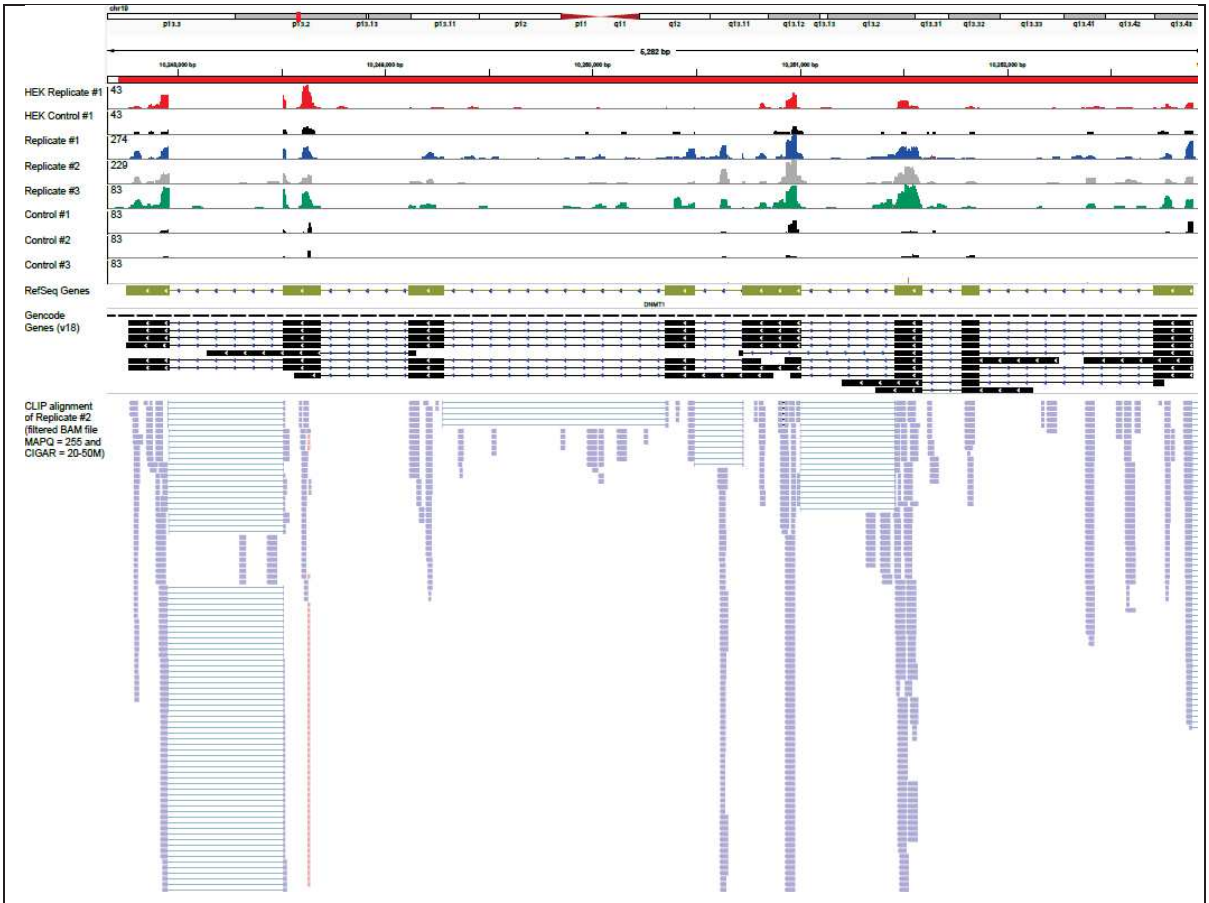

HEK Replicate #1
HEK Control #1
Replicate #1
Replicate #2
Replicate #3
Control #1
Control #2
Control #3
RefSeq Genes
Gencode Genes (v18)
CLIP alignment of Replicate #2 (filtered BAM file MAPQ = 255 and CIGAR = 20-50M)

Figure S4

Figure S5

References

1. Cherrier, T., et al., CTIP2 is a negative regulator of P-TEFb. Proc Natl Acad Sci U S A, 2013. 110(31): p. 12655-60.

2. Lennon, M.J., et al., Bcl11b: A New Piece to the Complex Puzzle of Amyotrophic Lateral Sclerosis Neuropathogenesis? Neurotoxicity Research, 2016. 29(2): p. 201-207.

3. Fu, W., et al., BCL11B-Mediated Epigenetic Repression Is a Crucial Target for Histone Deacetylase Inhibitors in Cutaneous T-Cell Lymphoma. J Invest Dermatol, 2017. 137(7): p. 1523-1532.

4. Lennon, M.J., et al., Bcl11b—A Critical Neurodevelopmental Transcription Factor—Roles in Health and Disease. Frontiers in Cellular Neuroscience, 2017. 11(89).

5. Kominami, R., Role of the transcription factor Bcl11b in development and lymphomagenesis. Proc Jpn Acad Ser B Phys Biol Sci, 2012. 88(3): p. 72-87.

6. Longabaugh, W.J.R., et al., Bcl11b and combinatorial resolution of cell fate in the T-cell gene regulatory network. Proceedings of the National Academy of Sciences, 2017. 114(23): p. 5800-5807.
7. Marban, C., et al., Recruitment of chromatin-modifying enzymes by CTIP2 promotes HIV-1 transcriptional silencing. EMBO J, 2007. 26(2): p. 412-23.
8. Eilebrecht, S., et al., HMGA1 recruits CTIP2-repressed P-TEFb to the HIV-1 and cellular target promoters. Nucleic Acids Res, 2014. 42(8): p. 4962-71.
9. Rohr, O., et al., Recruitment of Tat to heterochromatin protein HP1 via interaction with CTIP2 inhibits human immunodeficiency virus type 1 replication in microglial cells. J Virol, 2003. 77(9): p. 5415-27.
10. Forouzanfar, F., et al., HIV-1 Vpr mediates the depletion of the cellular repressor CTIP2 to counteract viral gene silencing. Sci Rep, 2019. 9(1): p. 13154.
11. Sun, J., Y. Shi, and E. Yildirim, The Nuclear Pore Complex in Cell Type-Specific Chromatin Structure and Gene Regulation. Trends in Genetics, 2019. 35(8): p. 579-588.
12. Matera, A.G. and Z. Wang, A day in the life of the spliceosome. Nature Reviews Molecular Cell Biology, 2014. 15(2): p. 108-121.
13. Chen, C., et al., Fused in Sarcoma: Properties, Self-Assembly and Correlation with Neurodegenerative Diseases. Molecules, 2019. 24(8): p. 1622.
14. Pong, S.K. and M. Gullerova, Noncanonical functions of microRNA pathway enzymes – Drosha, DGCR8, Dicer and Ago proteins. FEBS Letters, 2018. 592(17): p. 2973-2986.
15. Kurosaki, T., M.W. Popp, and L.E. Maquat, Quality and quantity control of gene expression by nonsense-mediated mRNA decay. Nature Reviews Molecular Cell Biology, 2019. 20(7): p. 406-420.
16. Wakabayashi, Y., et al., Bcl11b is required for differentiation and survival of alphabeta T lymphocytes. Nat Immunol, 2003. 4(6): p. 533-9.
17. Gasic, I. and T.J. Mitchison, Autoregulation and repair in microtubule homeostasis. Current Opinion in Cell Biology, 2019. 56: p. 80-87.
18. Orth, M.F., et al., High Specificity of BCL11B and GLG1 for EWSR1-FLI1 and EWSR1-ERG Positive Ewing Sarcoma. Cancers, 2020. 12(3): p. 644.
19. Wiles, E.T., et al., BCL11B is up-regulated by EWS/FLI and contributes to the transformed phenotype in Ewing sarcoma. PLoS One, 2013. 8(3): p. e59369.

20. Doktor, T.K., et al., RNA-sequencing of a mouse-model of spinal muscular atrophy reveals tissue-wide changes in splicing of U12-dependent introns. Nucleic Acids Research, 2016. 45(1): p. 395-416.
21. Maeda, M., et al., Transcriptome profiling of spinal muscular atrophy motor neurons derived from mouse embryonic stem cells. PLoS One, 2014. 9(9): p. e106818.
22. Scherzer, C.R., et al., Molecular markers of early Parkinson's disease based on gene expression in blood. Proceedings of the National Academy of Sciences, 2007. 104(3): p. 955-960.
23. Ahmed, I., et al., Huntington's disease: Neural dysfunction linked to inositol polyphosphate multikinase. Proceedings of the National Academy of Sciences, 2015. 112(31): p. 9751-9756.
24. Yamazaki, T., et al., FUS-SMN Protein Interactions Link the Motor Neuron Diseases ALS and SMA. Cell Reports, 2012. 2(4): p. 799-806.
25. Shehadeh, L.A., et al., SRRM2, a potential blood biomarker revealing high alternative splicing in Parkinson's disease. PLoS One, 2010. 5(2): p. e9104.
26. Briese, M., et al., hnRNP R and its main interactor, the noncoding RNA 7SK, coregulate the axonal transcriptome of motoneurons. Proc Natl Acad Sci U S A, 2018. 115(12): p. E2859-E2868.
27. Bhat, A.A., et al., Role of non-coding RNA networks in leukemia progression, metastasis and drug resistance. Molecular Cancer, 2020. 19(1): p. 57.
28. Wu, Y.-Y. and H.-C. Kuo, Functional roles and networks of non-coding RNAs in the pathogenesis of neurodegenerative diseases. Journal of Biomedical Science, 2020. 27(1): p. 49.
29. De Santis, R., et al., FUS Mutant Human Motoneurons Display Altered Transcriptome and microRNA Pathways with Implications for ALS Pathogenesis. Stem Cell Reports, 2017. 9(5): p. 1450-1462.
30. Shkreta, L. and B. Chabot, The RNA Splicing Response to DNA Damage. Biomolecules, 2015. 5(4): p. 2935-2977.
31. Wang, E. and I. Aifantis, RNA Splicing and Cancer. Trends in Cancer, 2020.
32. Tuck, A.C., et al., Mammalian RNA Decay Pathways Are Highly Specialized and Widely Linked to Translation. Molecular Cell, 2020. 77(6): p. 1222-1236.e13.

8. Results

Our lab has focused much of its attention on the role of the transcription factor CTIP2 in microglial cells in the establishment and persistence of the HIV-1 latency. In the last years we discovered and described two complexes responsible of establishing the viral latency and maintain it (Figure 22). Considering our results showing the interaction of CTIP2 with also a ncRNA such as 7SK we come to the conclusion, that CTIP2, while not having a canonical RNA Recognition Motif, could also interact with other RNAs. The aim of this project is to identify new partners interacting with CTIP2, both RNAs and proteins and to describe their function and localization, regarding the HIV-1 infection in Microglial cells. We split this main task in two distinctive projects, one focusing on the protein/protein interaction and another on the protein/RNA interaction. As we will see further in this chapters, results from the two projects started soon to intertwine, giving an early confirmation of the early treated results of the two projects.

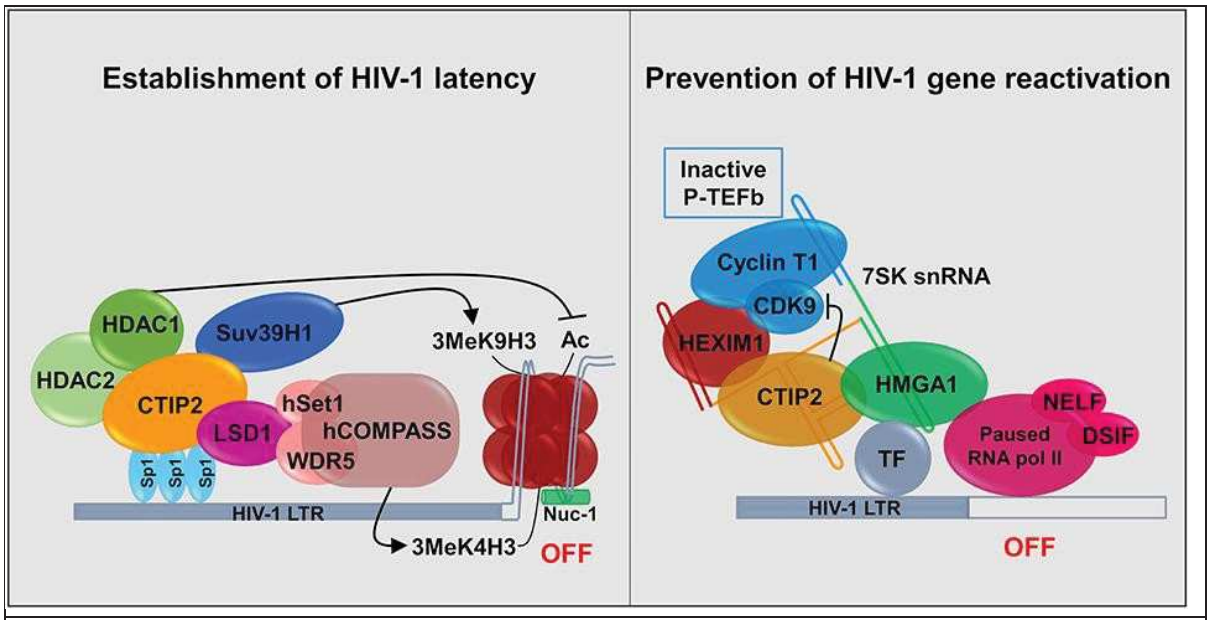


Figure 22: CTIP2 previously described complexes and partners responsible of HIV-1 latency in microglial cells. On the left CTIP2 recruit a complex of chromatin remodelling enzymes which create a heterochromatin environment on the viral promoter. On the right CTIP2 prevent the reactivation of the silenced provirus by hijacking the inactive for of the P-TEFb complex in a ribonucleoprotein complex. (Wallet et al., 2019)

8.1. Identification of new protein partners of CTIP2 by quantitative mass spectrometry

A series of experiments of immunoprecipitation of CTIP2 followed up by analysis through quantitative mass spectrometry found 888 proteins interacting with CTIP2 in a significant matter ($p < 0,05$) when compared to the control. A first analysis subdivided these interacting proteins by their function (if already known or described) and through this we discovered how CTIP2 was found within different functional groups of proteins with functions spanning from cellular division, replication, transcription and reparation of DNA and helicase activity, cellular adhesion, histone deacetylation, nucleotide binding, management of ribosomal RNA and mRNA splicing. Proteins already described to interact with CTIP2 were of course found, confirming previously obtained data such as P-TEFb, HDAC, HEXIM, LSD1 and Suv39H1 Previous student of our group took the responsibility to investigate a complex of proteins interacting with CTIP2 part of the PTMs machinery, specifically of the SUMOylation. My focus of attention was to investigate a series of proteins interacting with CTIP2, many of them described to interact among them, but with described different functions (Figure 23).

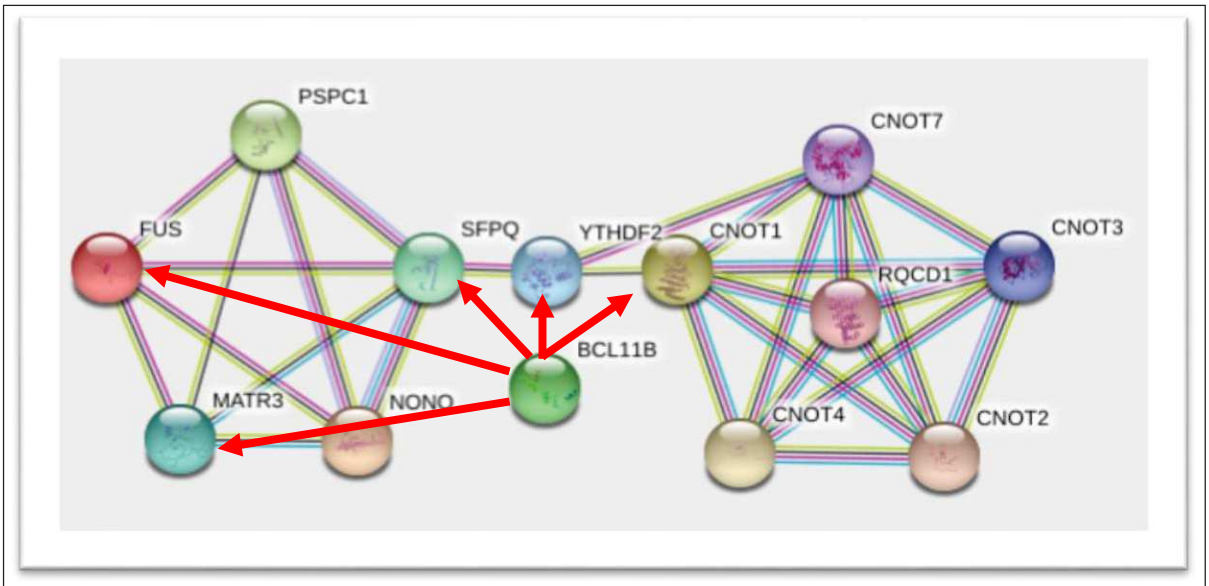


Figure 23: STRING Analysis of the newly found proteins interacting with CTIP2. A new set of protein were discovered interacting with CTIP2 after a Quantitative Mass Spectrometry analysis. The proteins indicated by a red arrows were selected for further investigation and confirmation of the Mass Spec data.

We decided to focus and try to confirm the interaction with the following proteins/complexes:

- MATR3 together with SFPQ have been shown to be required together with Rev to export partially unspliced and unspliced HIV1 RNA to the cytoplasm
- PSPC1, FUS and SFPQ together are part of the paraspeckle nuclear structure. SFPQ as we descried before is linked to many different cellular processes ranging from transcriptional initiation and termination, transcriptional activation and repression, and splicing.
- CNOT1 and the CCR4-NOT complex. CCR4-NOT is one of the major cellular mRNA deadenylases and is linked to various cellular processes including bulk mRNA degradation, miRNA- mediated repression, translational repression during translational initiation and general transcription regulation. The deadenylation involves the shortening of the polyA tail of mRNAs which represses mRNA translation, and it is the rate-limiting step that initiates mRNA degradation.
- YTHDF2: Specifically recognizes and binds N6-methyladenosine (m6A)-containing RNAs and regulates mRNA stability. M6A is a modification present at internal sites of mRNAs and some non-coding RNAs . YTHDF2 binding to m6A controls the stability of the methylated RNA by promoting localizations to mRNA decay sites, such as processing bodies (P-bodies), leading to mRNA degradation.

8.2. CTIP2/SFPQ interaction

The interaction between CTIP2 and SFPQ, seen for the first time in the mass spectrometry results, was confirmed by immunoprecipitation assay. In Figure 24 A we present both the endogenous SFPQ and overexpressed GFP-SFPQ being immunoprecipitated with and α-FLAG IP in the presence of Flag-CTIP2. These results immediately suggested us a strong interaction between the two proteins. Immunoprecipitation experiments performed with deleted mutants of CTIP2 (Figure 24B) show that SFPQ binds the N-terminal domain of CTIP2 (Figure 24C).

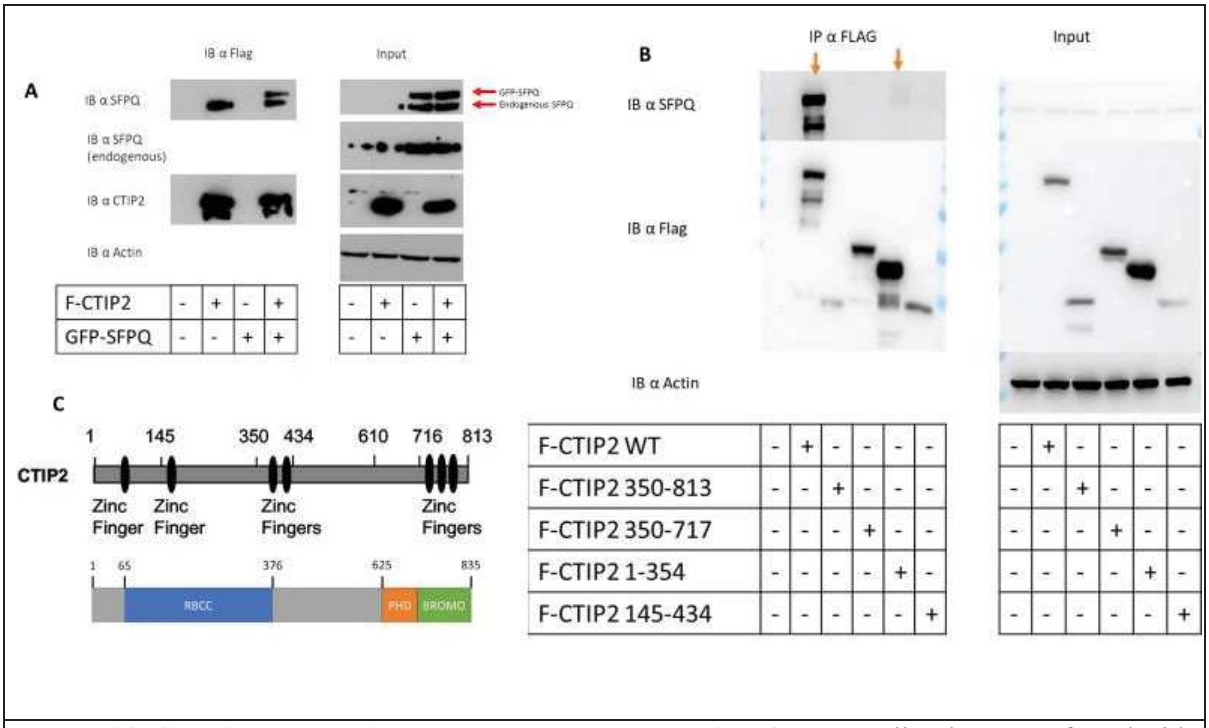


Figure 24: CTIP2 interact with paraspeckle component SFPQ. HEK cells where transfected with Flag-CTIP2 and/or GFP/SFPQ (A) and the cellular extract was immunoprecipitated with an α-FLAG antibody. The immunoprecipitated SFPQ appeared in two forms (higher and lower) showing CTIP2 interaction with both the endogenous SFPQ and the bigger GFP-SFPQ. A similar IP with delete CTIP2 mutants (B) showed how this interaction is mediated by the Zinc domain at the N-terminal, but the full protein still has a higher ability to bind SFPQ. (C) Domains composing the CTIP2 protein.

The next step was to confirm these interactions in microglial cells. To do so microglial were co-transfected with a GFP-SFPQ construct and an RFP-CTIP2, and the nucleus was stained with the Hoechst staining method; the transfection was followed by an analysis at the confocal microscopy.

As previously described by our group (Rohr et al., 2003) CTIP2 is found within sphere-like dense nuclear structures when overexpressed (also when the cell produce more of it naturally, like after Interferon gamma stimulation); when SFPQ is also overexpressed together with CTIP2 (Figure 25) is possible to notice an extremely strong co-localization, almost total, within these structures.

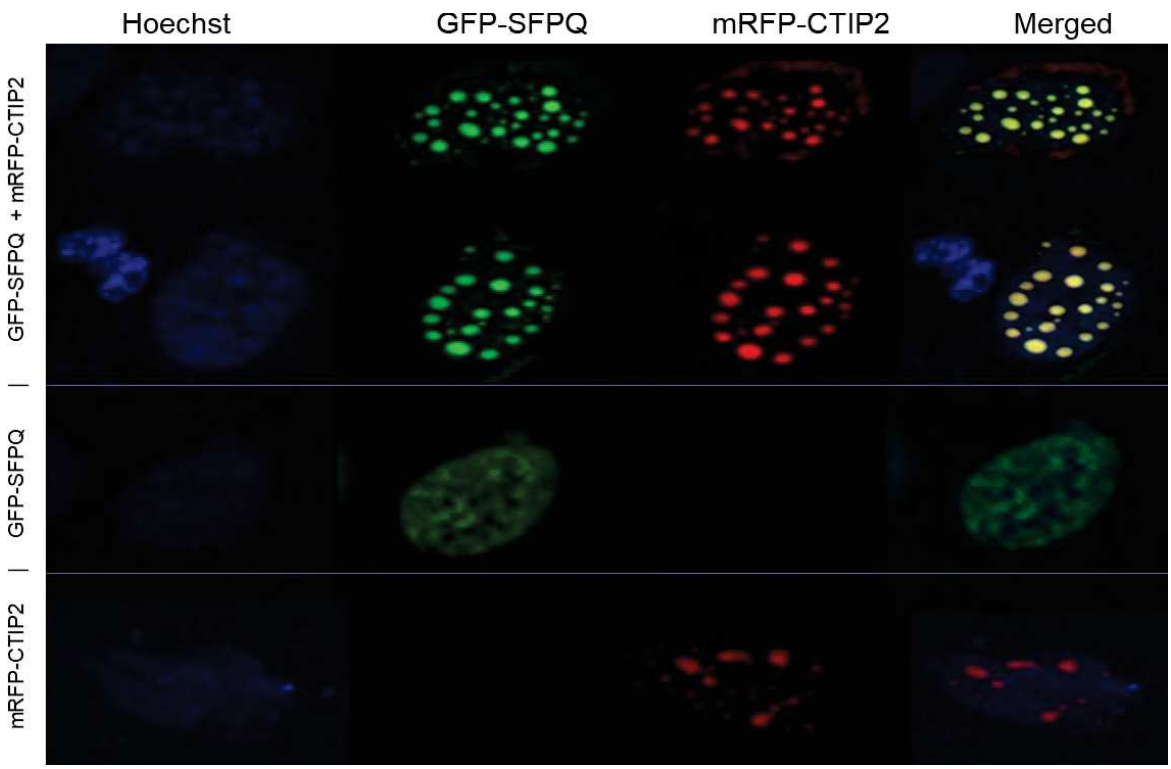


Figure 25: CTIP2 and SFPQ colocalization in microglial cells. RFP-CTIP2 and GFP-SFPQ vectors were transfected in microglial cells for 48h. Confocal microscopy analysis shows the high colocalization of the two proteins (first two rows). Endogenous SFPQ (third row) shows a general overall nuclear localization. RFP-CTIP2 alone (fourth row) localize in the sphere like dense nuclear structures.

To exclude the fact that this visible interaction could have been a results of an aberrant effect caused by the overexpression of the two proteins, confocal microscopy experiments were performed to investigate the localizations of CTIP2 and of the endogenous SFPQ in microglial cells. In Figure 26 A, one can observe that the nuclear localization of the endogenously expressed SFPQ is comparable to the one observed above with the GFP-tagged protein. The relocation of the endogenous SFPQ in CTIP2-induced nuclear structures confirmed our previous observations suggesting strong interactions between CTIP2 and SFPQ in the nucleus.

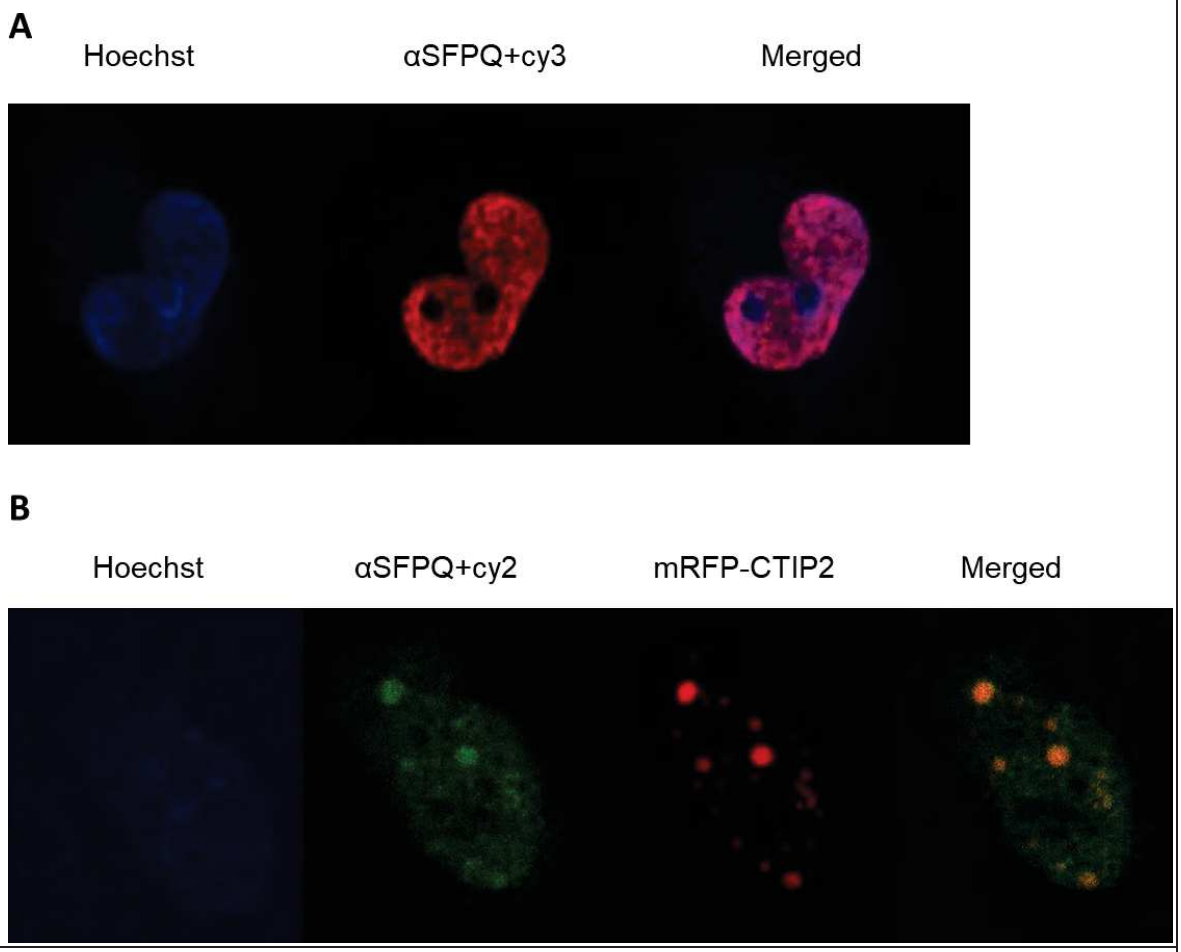


Figure 26: CTIP2 and endogenous SFPQ colocalization in microglial cells (A) Endogenous SFPQ immunostained with an α-SFPQ primary antibody and a CY3-secondary antibody. It is possible to see its localization all around the nucleus except for the nucleoli. (B) Complete colocalization of RFP-CTIP2 and the endogenous SFPQ in the sphere like structure formed by CTIP2.

8.3. CTIP2/SFPQ/Tat interaction

Our lab has already shown how CTIP2 interacts and inhibits the viral trans activator Tat by sequestration in specific ball-like structures. CTIP2 harbours two interaction interfaces, the 145-434 and the 717-813 domains which hijack Tat within CTIP2-induced structures and re-localize it within inactive regions of the chromatin (Rohr et al., 2003). After proving the interaction of CTIP2 with SFPQ, we logically wanted to explore if SFPQ was able to bind Tat. We performed an immunoprecipitation assay in HEK cells transfected with Myc-SFPQ and GFP-Tat (Figure 27). Two different Ips were performed: one targeting Myc-SFPQ (A), and one against GFP-Tat (B). While the α-Myc (SFPQ) immunoprecipitation (A) was unable to immune-precipitate Tat, α-GFP (Tat) immunoprecipitations (B) were able to pull down the

overexpressed Myc-SFPQ. In addition, while we demonstrated the interaction of endogenously expressed SFPQ with CTIP2 (Figure 24, A), we were unable to do it for Tat. As shown on the 4th line of Figure 25 B, the endogenous SFPQ was not found interacting with the overexpressed Tat, contrary to the overexpressed Myc-SFPQ.

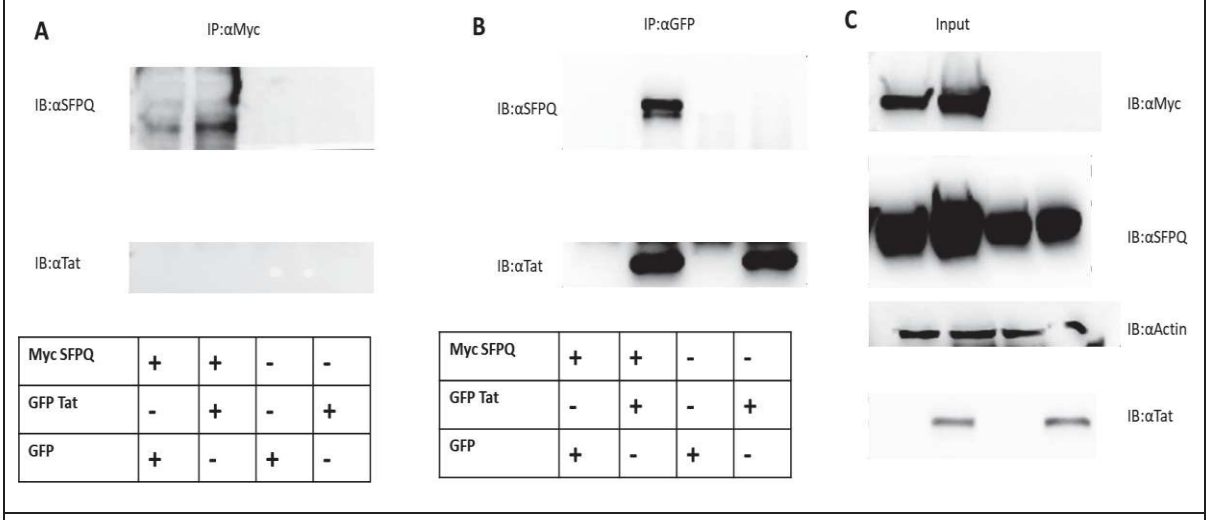


Figure 27: SFPQ interaction with Tat: A) α-Myc IP of HEK cells transfected with Myc-SFPQ and GFP-Tat and GFP as negative control. B) α-GFP IP of HEK cells transfected with Myc-SFPQ and GFP-Tat and GFP as negative control. C) Input

We performed another series of confocal microscopy experiment to try to confirm Tat interaction with SFPQ. Microglial cells were transfected with GFP-Tat and the endogenous SFPQ protein was stained with an α-SFPQ antibody and with a CY3 (Red) secondary antibody Figure 28 A). The results showed the classical localization of Tat within the nucleolus, regions where the endogenous SFPQ was not found (like already seen in previous slides); but in some small localization points, a small colocalization of the two proteins was observed (Figure 28A, 4th panel, light blue arrows). We next analysed how CTIP2 colocalized with SFPQ in Tat expressing cells. Microglial cells expressing RFP-CTIP2 together with GFP-Tat were immune-stained with anti-SFPQ and CY5 Ultrared fluorochrome-couple secondary antibodies (Figure 28 B). Our results show that Tat localized with the endogenous SFPQ protein CTIP2-induced structures. Interestingly, as SFPQ is considered as a key component of the paraspeckle nuclear structures, we started to believe that CTIP2-induced structures may be related to them.

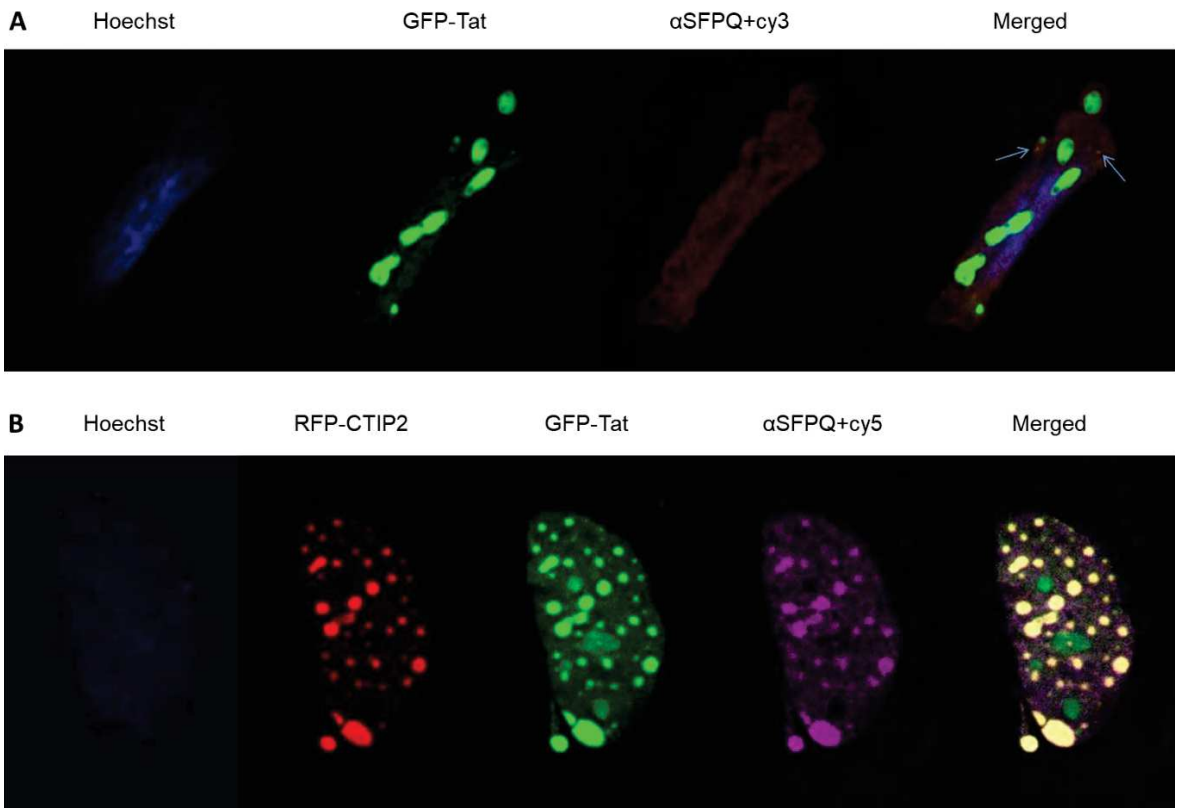


Figure 28: CTIP2, Tat and endogenous SFPQ colocalization in microglial cells (A) Microglial cells were transfected with a GFP-Tat vector and endogenous SFPQ was immunostained with an α-SFPQ primary antibody and a CY3-secondary antibody. Tat localize in the nucleolus while SFPQ localize everywhere in the nucleus except for the nucleolus. Two small point of colocalization were observed. (B) Microglial cells were transfected with vectors for GFP-Tat and RFP-CTIP2 and the endogenous SFPQ was immunostained with an α-SFPQ primary antibody and a CY5-secondary (UltraRed) antibody. Tat and SFPQ both localize together with CTIP2 in the sphere-like structures, while the nucleolus remains the sole localization region for Tat.

8.4. Role of SFPQ in HIV-1 gene transcription and viral replication

Our group have fully described the repressor effect of CTIP2 on the transcription of the HIV-1 genes. CTIP2 was shown to be able to reduce both the basal transcription of the proviral promoter, as well as the Tat-dependent transcription and the replication of the virus. After discovering and proving this tight association of CTIP2 with SFPQ we went on to explore if SFPQ had any functional activity on the transcription of the viral genes. Luciferase Assays performed in microglial cells expressing Myc-SFPQ shows a repressor effect on the basal level of transcription (Figure 29 A) but in the presence of Tat, SFPQ switch to a transcriptional activator phenotype (Figure 29 B); the graph showing the transcriptional activity in the

presence of Tat were normalized to the basal level of transcription in the presence of Tat alone, which is still around 30 to 40 fold higher than the basal level. Strangely enough, the replication of the pNL4-3 ΔEnv was repressed by SFPQ overexpression (Figure 29 C), which is surprising considering the importance of Tat function for the viral replication. This may be explained by some post-transcriptional regulation of SFPQ in HIV-1 replication. Like the Luciferase Assay in the presence of Tat, the pNL4-3 ΔEnv is also normalized to its own basal level, which still stand between 20 to 40 fold higher than the basal level of transcription of the LTR experiment without Tat.

Another battery of tests were performed using a deletion mutant of Myc-SFPQ missing the RNA Recognition Motif 1 and 2. Surprisingly, since the mutation had no impact on SFPQ-mediated repression of the LTR without Tat (Figure 29 D), it completely abrogated SFPQ stimulation of Tat function (Figure 29 E). Logically, expression of the SFPQ mutant repressed the expression of the pNL4-3 ΔEnv virus (Figure 29 F).

As last we compared the Myc-SFPQ overexpression with the ΔRRM1+2 mutant. The analysis confirmed the conclusions made above concerning the basal and the Tat-mediated regulation of HIV-1 gene transcription (Figure 29 G, H). In addition, they suggested us that the ability to bind RNA may be necessary for SFPQ stimulation of Tat. Of note, we observed that the mutation also partially impaired SFPQ-mediated repression of the provirus replication (Figure 29 H). HIV-1 pre-mRNA has been previously shown bound to SFPQ and diverted towards MATR3 (Kula et al., 2013) which, together with the viral protein Rev, contributes to the shuttling of the unspliced viral RNA to the cytoplasm. While we do not know if this is what we are hindering with this mutation, for sure the lost RNA binding capacity have a slight effect on the repressor effect of SFPQ on the replication of the pNL4-3 ΔEnv. Defining the relative importance of SFPQ functions at the transcriptional and the post-transcriptional steps of the viral life cycle remains to be done.

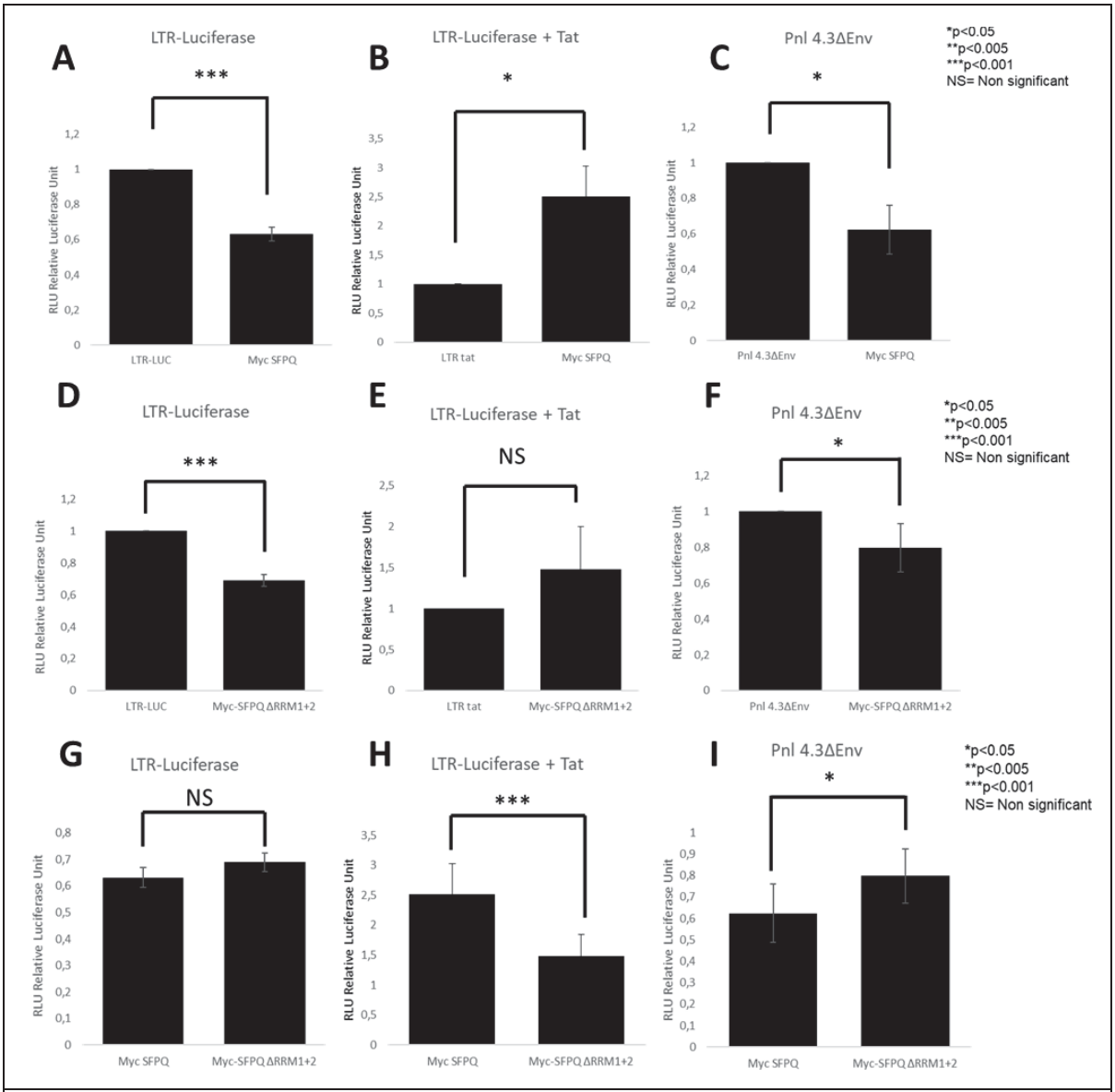


Figure 29: Functional effect of SFPQ on viral transcription and replication in microglial cells. Microglial cells were transfected with the LTR-Luciferase plasmid alone (A, D, G), with Tat (B, E, H), or with the luciferase reporter vector HIV-1 pNL4-3 ΔEnv (C, F, I) plus the plasmid encoding for Myc-SFPQ alone (A, B, C, G, H, I) or its mutant Myc-SFPQ ΔRRM1+2 (D, E, F, G, H, I). The cells were lysate 48h post-transfection and the luciferase activity analysed. The values were normalized against the sample non-transfected with the protein of interest LTR, LTR+Tat and the pNL4-3 ΔEnv without any other proteins (set at 1) and a T-test was performed on non-less than 3 independent experiments.

We then performed another series of Luciferase assay to compare CTIP2 and SFPQ impacts on HIV-1 gene transcription and the viral replication. We observed the typical repressor effect of CTIP2 on the basal transcription (Figure 30 A), in the presence of Tat (Figure 30 B) and on the pNL4-3 ΔEnv (Figure 30 C). Compared to the CTIP2-mediated repression of HIV-1 gene transcription, SFPQ has a weaker effect (Figure 30 D) but no significant difference was seen

on the viral replication (Figure 30 F). As a reminder that the results were normalized based on the basal level for each set of experiment (LTR, LTR+Tat and the pNL4-3 ΔEnv without any other proteins) which are set to 1.

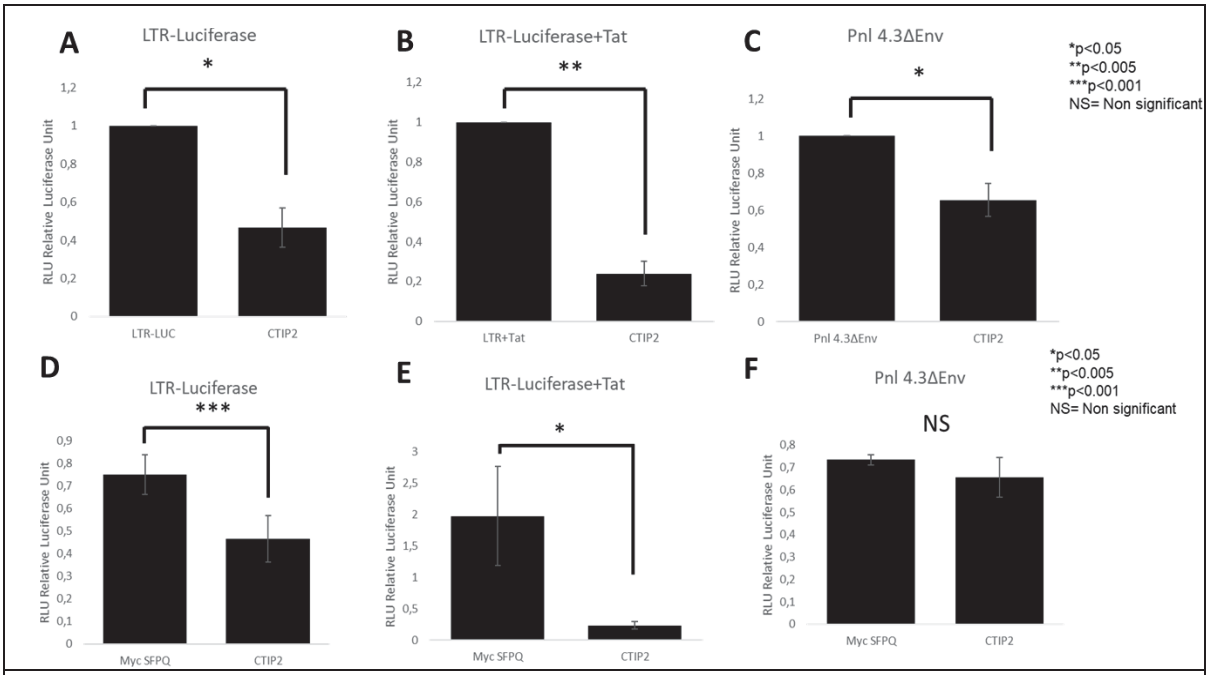


Figure 30: Functional effect of CTIP2 and SFPQ on viral transcription and replication in microglial cells: Microglial cells were transfected with the LTR-Luciferase plasmid alone (A, D), with Tat (B, E), or with the luciferase reporter vector HIV-1 pNL4-3 ΔEnv (C, F) plus the plasmid encoding for CTIP2 (A, B, C, D, E, F), Myc-SFPQ (D, E, F). The cells were lysate 48h post-transfection and the luciferase activity analysed. The values were normalized against the sample non-transfected with the protein of interest (set at 1) and a T-test was performed.

The next step was to analyse the effect on the transcription of the two proteins combined. First we compared the effect of the overexpression of the two proteins on the LTR and its basal level of transcription and we were not surprised to see a negative regulator effect since both of the proteins showed to have this phenotype when transfected alone (Figure 31 A). Same results were obtained on the virus replication (pNL4-3 ΔEnv) when the two proteins were expressed together (Figure 31 C). More surprising was the fact that whatever activator effect SFPQ showed on the viral transcription in the presence of Tat, was lost when co-transfected with CTIP2 (Figure 31 B). Taking in account all the now proven strong interaction between the two proteins, we can propose the hypothesis that the delocalization of SFPQ by CTIP2 in those described nuclear compartment, may affect its ability to boost Tat-mediated transcription. We then compared the single overexpression effect of one (and the other) proteins versus the two proteins combined. Compared to CTIP2 overexpression alone, the co-expression of both

proteins had no statistical effect on the transcription (Figure 31 D and E) but if compared to the SFPQ effect, the two proteins together had a stronger repressor effect on the LTR (Figure 31 G) and as already described before on the LTR+Tat the phenotype was shifted from an activator effect to a repressor one (Figure 31 H). CTIP2 and SFPQ also cooperated to repress the expression of the provirus (Figure 31 F and I) showing a stronger repression. Of note, due to the need to express both proteins in these experiments, the effects of SFPQ and CTIP2 expressed alone presented in the figure 31 have been obtained with half of the protein levels than those presented previously.

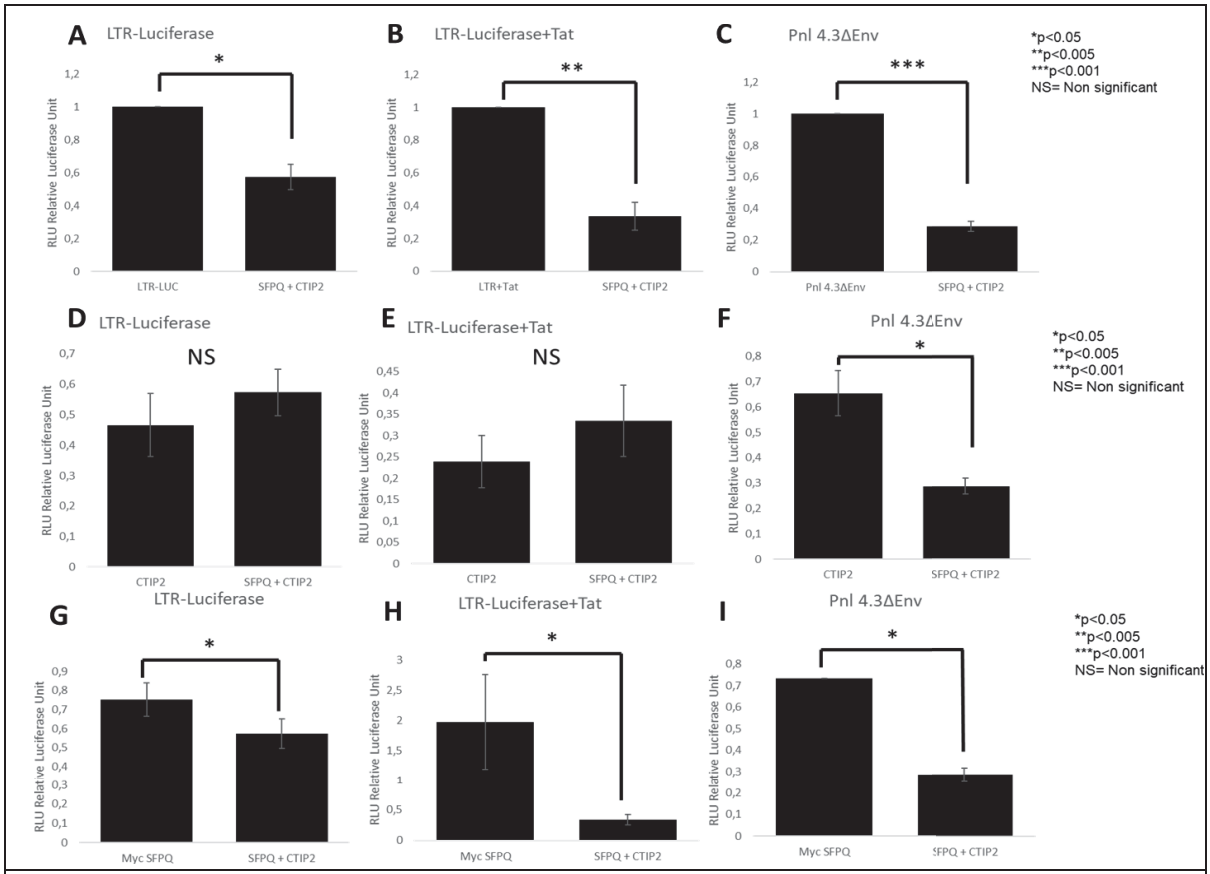


Figure 31: Functional effect of CTIP2 and SFPQ on viral transcription and replication in microglial cells together: Microglial cells were transfected with the LTR-Luciferase plasmid alone (A, D, G), with Tat (B, E, H), or with the luciferase reporter vector HIV-1 pNL4-3 ΔEnv (C, F, I) plus the plasmid encoding for CTIP2, Myc-SFPQ or the two combined. The cells were lysate 48h post-transfection and the luciferase activity analyzed. The values were normalized against the sample non-transfected with the protein of interest (set at 1) and a T-test was performed on non-less than 3 independent experiments.

We next performed new experiments with increasing amounts of SFPQ. No statistical difference on the basal transcription or the viral expression was noted (Figure 32 A and C). Interestingly, Increasing amounts of SFPQ stimulated Tat function in a dose dependant manner. (Figure 32 B).

We should notice that the global effect of increase doses of SFPQ on the provirus expression is clearly visible but still not statistically relevant at this step (Figure 32 C).

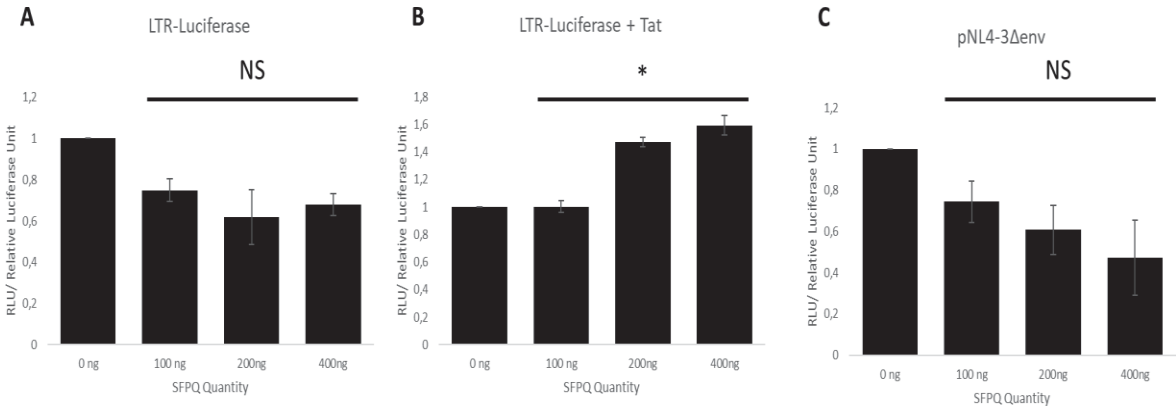


Figure 32: Functional dose dependent effect of SFPQ on viral transcription and replication in microglial cells. Microglial cells were transfected with the LTR-Luciferase plasmid alone (A), with Tat (B), or with the luciferase reporter vector HIV-1 pNL4-3 ΔEnv (C) plus the plasmid encoding for Myc-SFPQ at 3 different quantities. The cells were lysate 48h post-transfection and the luciferase activity analyzed. The values were normalized against the sample transfected with the control pCdna3 (set at 1) and an ANOVA test was performed on non-less than 3 independent experiments. (* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$)

8.5. CTIP2/FUS interaction

Another protein found in the Mass Spectrometry screening was FUS. Our attention on this protein was attracted by the fact that, together with SFPQ (and other proteins and RNAs) is a key component of the paraspeckle structures. FUS is an abundant nuclear protein that affects multiple levels of RNA biogenesis, including transcription, splicing, and mRNA transport (Yang et al., 2000, Wang et al., 2008, Polymenidou et al., 2012). FUS affects transcription by binding the RNA Pol II through its C-Terminal domain and locates at the start stie of transcription of almost 50% of expressed genes in HEK cells (Schwartz et al., 2012).

For these reasons, we started to perform some immunoprecipitation experiments to try to confirm CTIP2 interaction with FUS. An α-CTIP2 IP (vs a Mock IP using unspecific IgG) was performed in HEK cells, in the presence/absence of CTIP2 expression (using mock transfected cDNA3 as a secondary negative control) and with or without an RNase A treatment to determine if their interaction is RNA mediated. As shown in figure 33 A, we managed to co-immunoprecipitate FUS and CTIP2 without relevant impact of the RNase treatment (Figure 33 A). As described above, CTIP2 has been shown bound to SFPQ and Tat. We then performed a

second set of experiments in the presence of increasing quantities of Tat (Figure 33 B); we did not observe any relative change in the interaction between CTIP2 and FUS and SFPQ with increasing Tat quantities. We also performed an immunoprecipitation in the presence of Tat and the LTR (expressing TAR) (Figure 33 C). This time we used an α-FLAG antibody targeting FLAG-CTIP2 for the IP. while the IP for SFPQ was clear, with no relevant differences between the sample treated with RNase or also transfected with LTR-TAR and Tat, for FUS we obtained a more "dirty" immunoprecipitation. We can see how the sample lacking CTIP2 still had a strong unspecific binding of FUS to the beads, and how the sample treated with RNAase lost most of the binding (to level visually similar to the IP in picture 33 A). A theory behind this is that there must be a massive unspecific binding of FUS to the FLAG antibody, mediated by RNAs.

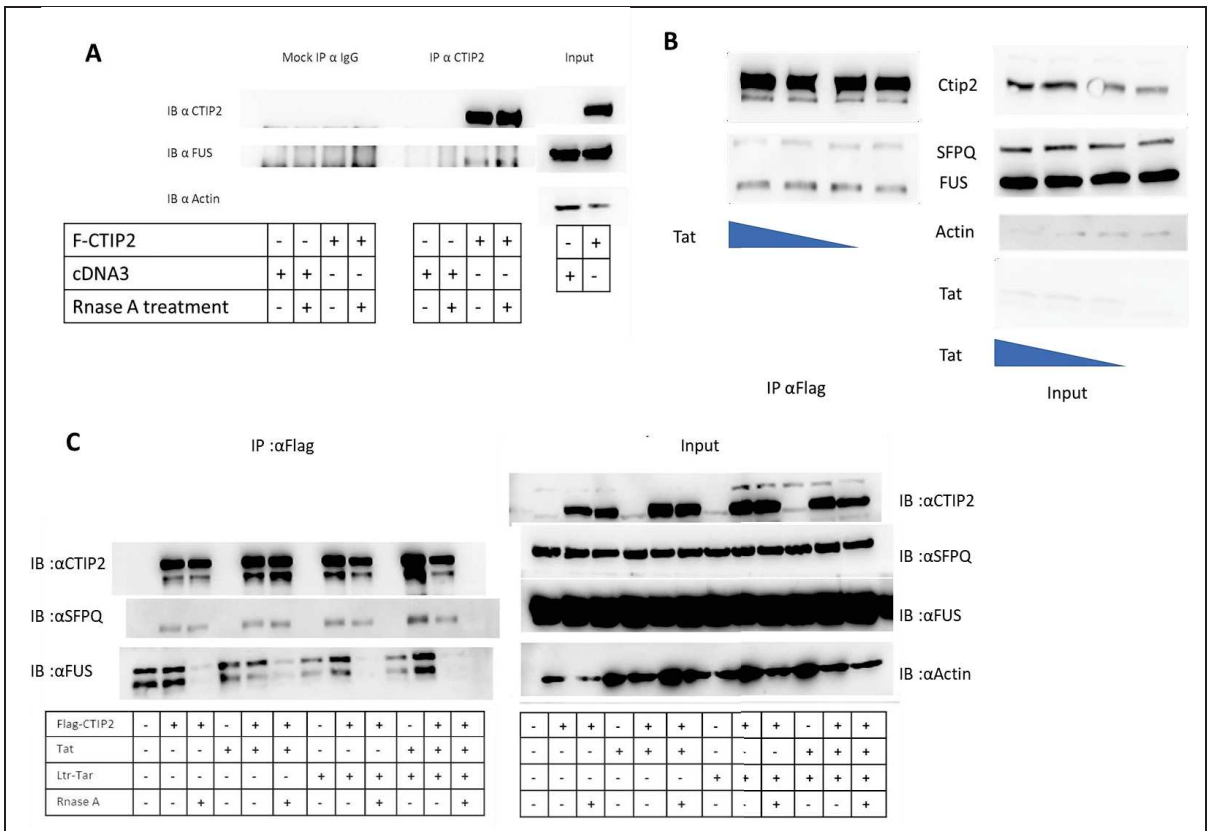


Figure 33: CTIP2-FUS interaction. HEK cells were transfected with the indicated plasmids for 48h, the whole cell lysate was then extracted and used to perform the immunoprecipitation with the indicated antibody. (A) An IP αCTIP2 vs a mock IP with unspecific IgG, and a Mock transfection with pcDNA3 in the presence/absence of RNase treatment proved an RNA independent interaction of CTIP2 with FUS. (B) HEK cells extract transfected with a FLAG-CTIP2 construct were immunoprecipitated with an αFLAg in the presence of increasing quantity of Tat showing how the interaction between CTIP2 and SFPQ and FUS is not altered by the presence of Tat. (C) HEK cells extracts transfected with a FLAG-CTIP2 were immunoprecipitated with an αFLAG antibody in the presence of the LTR-Tar and/or Tat.

To further study the interaction between CTIP2 and FUS, we perform another set of confocal analysis of CTIP2 and FUS localizations in microglial cells. Microglial cells expressing RFP-CTIP2 were immune-stained with an α-FUS antibody. This led to two different kind of phenotype regarding their colocalization. In cells expressing low levels of RFP-CTIP2, the colocalization within the sphere structures seems similar to SFPQ, with a strong association between the two proteins (Figure 34 A) while in cells showing an higher level of CTIP2, FUS looked to be completely excluded by these structures (Figure 34 B). In some of the early results obtained by our lab on CTIP2 and these structures, we noticed already in the past how the high level in density these structures were, and this is probably an explanation for these two type of phenotypes; when higher quantity of CTIP2 are packed within these structures, loosely associated protein may be misplaced on the outside rim or more simply the antibodies have an higher difficulty to access these structures.

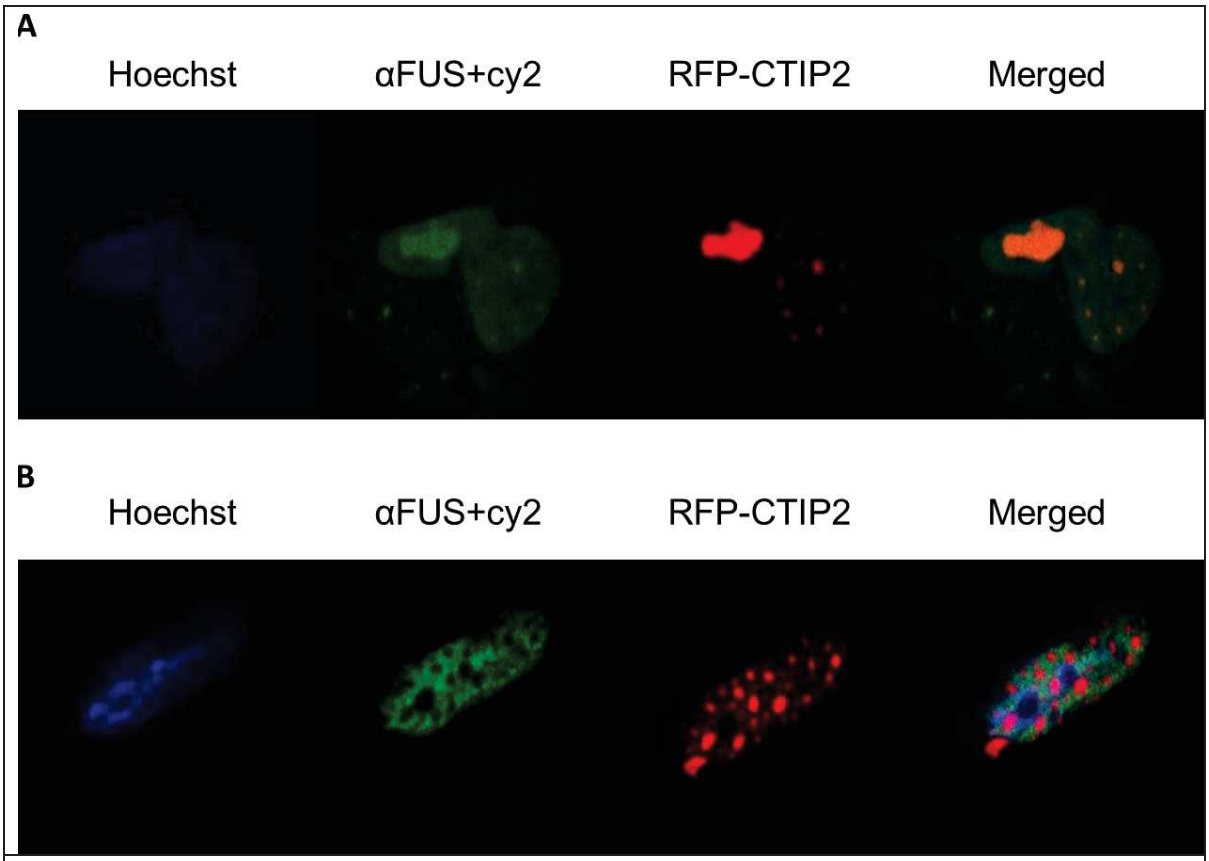


Figure 34: CTIP2 and endogenous SFPQ colocalization in microglial cells. Microglial cells were transfected with a RFP-CTIP2 vector and endogenous SFPQ was immunostained with an α-SFPQ primary antibody and a CY3-secondary antibody. Tat localize in the nucleolus while SFPQ localize everywhere in the nucleus except for the nucleolus. Two small point of colocalization were observed. (B) Microglial cells were transfected with vectors for GFP-Tat and RFP-CTIP2 and the endogenous SFPQ was immunostained with an α-SFPQ primary antibody and a CY5-secondary (UltraRed) antibody.

This “exclusion” effect is way more noticeable in the next set of slides showing the expression of overexpressed RFP-CTIP2, GFP-Tat and endogenously expressed FUS (CY5 staining). We can see again the colocalization of RFP-CTIP2 and GFP-Tat in these ball-like structures. FUS immunostaining is weak in the core of the structures and more concentrated at the rim and in dense speckles (Figure 35 A). These finding support our idea that FUS is part of these structures and it is pushed outward to the level of exclusion by CTIP2. In fact, we can see how other cells shows these structures to be empty of FUS while it still strongly localizes at its periphery (Figure 35 B). Surprisingly, no colocalization of tat and FUS has been observed in the absence of CTIP2. (Figure 35 C)

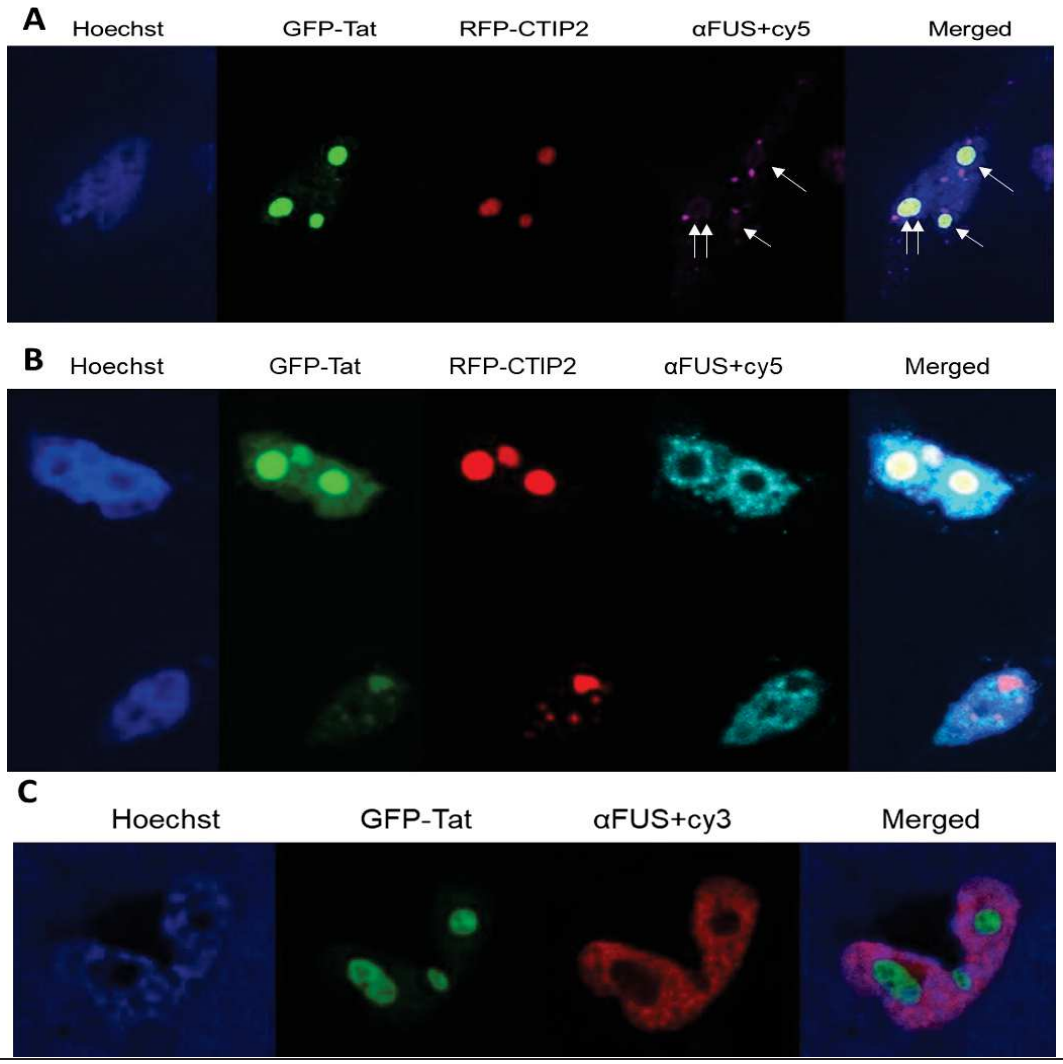


Figure 35: CTIP2, Tat and endogenous FUS colocalization in microglial cells. Microglial cells were transfected with a GFP-Tat vector (A, B, C), RFP-CTIP2 (A, B) for 48h and endogenous FUS was immunostained with an α-FUS primary antibody and a CY5 (UltraRed)(A, B) CY3 © secondary antibody.

8.6. CTIP2/YTHDF family interaction

The YTHDF family includes three paralogs (YTHDF1, YTHDF2, and YTHDF3; also called DF1, DF2, and DF3, respectively), each of which has different reported functions; DF1 enhances mRNA translation, DF2 promotes mRNA degradation, and DF3 enhances translation and degradation (Li et al., 2017, Shi et al., 2017, Wang et al., 2014a, Wang et al., 2015). These family of proteins exert their function by binding a specific modification harboured by many RNAs. The N^6-methyladenosine (m^6A) is the most abundant mRNA nucleotide modification and regulates critical aspects of cellular physiology and differentiation and the genomic deletion of the m^6A-forming methyltransferase caused defective seed development and sporulation, respectively in in *Arabidopsis* and yeast.

While these proteins are described to be located mostly in the cytoplasm, one of them specifically, YTHDF2, was found to interact with CTIP2 in the Quantitative Mass Spectrometry results. A recent paper showed how these three proteins appear to have similar binding preferences and affinities contrary at what was believed in the past and that all three together have a mRNA degradation role with YTHDF2 being the main one to be compulsory (Zaccara & Jaffrey 2020).

To confirm that YTHDF proteins associate with CTIP2, we performed an co-immunoprecipitation using an α-YTHDF1, α-YTHDF2, α-YTHDF3 IP, versus a Mock IP (unspecific IgG) as a negative control (Figure 36 A) in the presence or not of Flag-CTIP2. As shown in figure 36 A, we managed to specifically immunoprecipitate CTIP2 with the respective YTHDF proteins. Interestingly, YTHDF2 was found in the samples overexpressing or not CTIP2 suggesting that this interaction is strong enough to be detectable with the endogenous CTIP2 as well. While this happening with YTHDF2 specifically does not come as a surprise considering it was specifically YTHDF2 that was found in the Mass Spec results, it is also important to know that the basal level of expression of CTIP2 has always been extremely low to be easily detected by Western Blot. Further analysis will be required. Anyhow the interaction of CTIP2 with all three members of the YTHD family was confirmed.

Knowing the role of the YTHDF proteins in the binding of the specific N^6-methyladenosine RNA modification, we wanted to prove that CTIP2 was able to bind these specifically modified RNAs. We performed a Reverse RNA Immunoprecipitation, using an antibody targeting specifically this modification, an α-N^6-methyladenosine, versus a Mock IP using unspecific IgG (Figure 36 B). These experiments were also done in the presence/absence of the HIV-1

pNL4-3 ΔEnv due to the fact there are report showing the viral mRNA being methylated. We wanted to see if the presence of the HIV-1 RNA could affect the binding of CTIP2 to the modified RNA. We were able to immunoprecipitated CTIP2 bound to this modified RNA, confirming its interaction with this RNA processing machinery. While the interaction between CTIP2 and YTHDF proteins is bridged by RNA (or viceversa) still need to be explored.

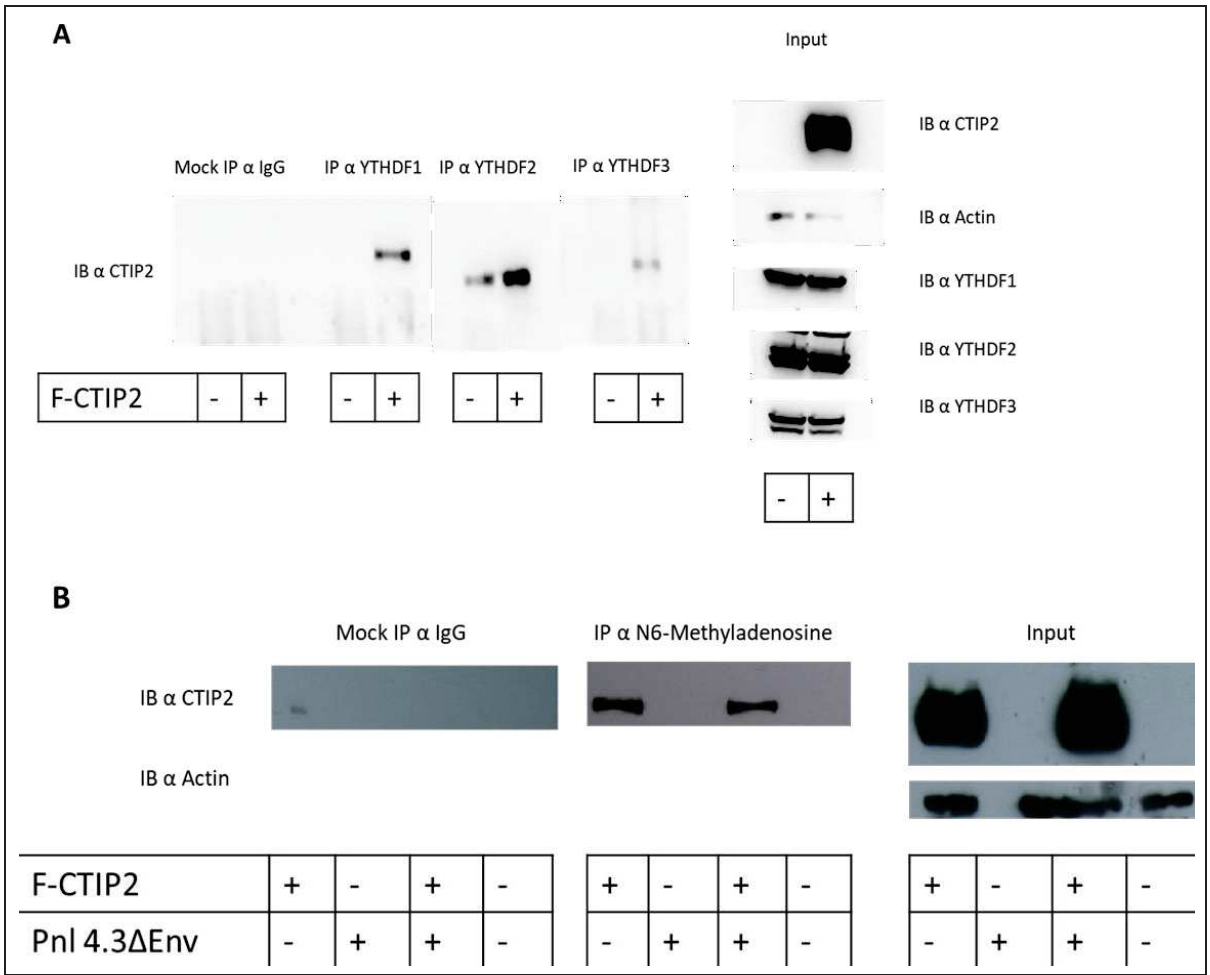


Figure 36: CTIP2 interacts with proteins of the YTHDF family and methylated RNAs : (A) HEK cells were transfected with Flag-CTIP2 for 48h. The whole cell lysate was then immunoprecipitated ON in the presence of the specific YTHDF antibody described, or an unspecific IgG as Mock control. A WB analysis was then performed. (B) HEK cells were transfected with Flag-CTIP2 and or pNL4-3 ΔEnv for 48h. The whole cell lysate was then immunoprecipitated ON with the specific α-N^6-methyladenosine antibody (or an unspecific IgG as Mock control) both in the presence of an RNAase Inhibitor. A WB analysis was then performed.

8.7. CTIP2/MATR3 interaction

Another protein found interacting with CTIP2 in the Mass Spec screening was MATR3. MATR3 is an abundant inner nuclear matrix protein widely expressed in various tissues. It may bind DNA and RNA and play a role in protein translation, DNA replication and repair, apoptosis (Zeitz et al., 2009, Zhang and Carmichael, 2001), mediating neuronal cell death in response to NMDA (N-methyl-D-aspartate) receptor (Giordano et al., 2005), nuclear retention of hyperedited RNA, and messenger RNA stabilization (Salton et al., 2011). Recently it has been described as positive regulator of HIV-1 acting at a posttranscriptional level in primary peripheral blood lymphocytes (PBLs). MATR3 forms a complex with SFPQ, which was already implicated in Rev-mediated export of HIV-1 RNAs. While SFPQ and Rev bind the viral pre-mRNA at the site of viral transcription, MATR3 interacts at a subsequent step involved in nuclear export (Sarracino et al., 2018).

To confirm the MS results, we performed an immunoprecipitation targeting CTIP2 in HEK cells expressing Flag-CTIP2. These experiments were performed in in the presence or not of RNase treatment (Figure 37 A). As shown, in figure 37 A, Matr3 was immunoprecipitated with CTIP2 in a RNA independent manner. We then overexpressed MATR3 in microglial to investigate the its function on HIV-1 gene transcription and the viral expression (Figure 37 B and C respectively). Our results suggest that MATR3 represses the basal transcription, Tat-mediated stimulation of transcription and the HIV-1 expression. These results highlight cell type specific regulations. Indeed, in the PBL cell line, MATR3 was described as a positive regulator of HIV-1 at the post-transcriptional level.

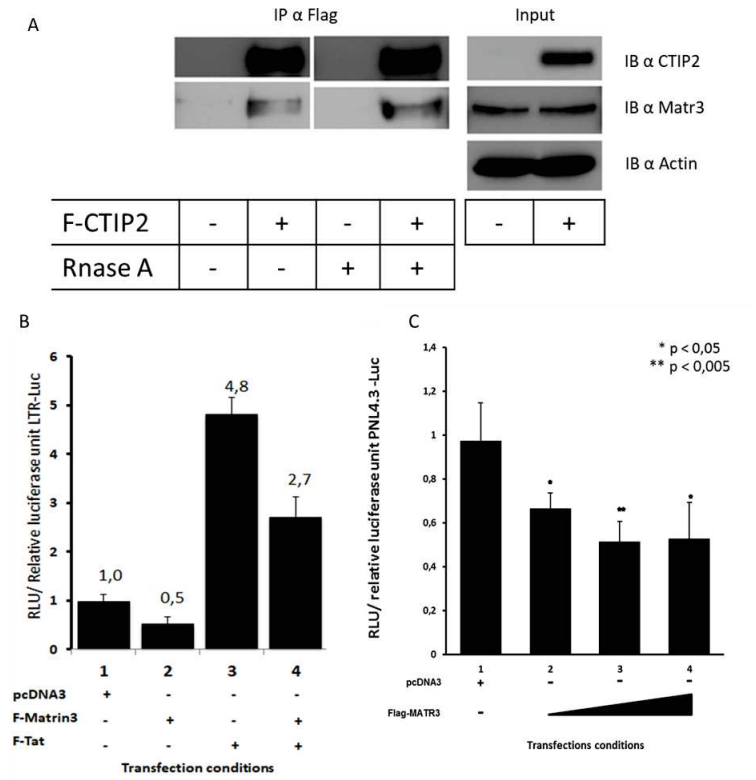


Figure 37: CTIP2 interaction with MATR3 and its effect on viral gene transcription. (A) HEK cells were transfected for 48h and the cell lysate was immunoprecipitated with an α-FLAG in the presence/absence of RNase A treatment. (B) Microglial cells were transfected with a LTR-Luciferase reporter in the presence or absence of Tat and Matr3 for 48h. After cell lysis the lysate where analysed to quantify the luciferase signal. (C) Luciferase Assay performed with increasing quantity of MATR3 on the pNL4-3 ΔEnv Luciferase.

8.8. Identification of new RNAs partners of CTIP2 by CLIP-Seq

Since our lab description of the RNP complex formed by 7SK, CTIP2, HEXIM etc. we knew the ability to of CTIP2 to interact, and bind, to RNAs. While we knew of the direct interaction of CTIP2 specifically with the 2nd stem loop of the sn7SK RNA, the Quantitative Mass Spec analysis results showed how CTIP2 was part of many different complexes, and proteins, with roles ranging from mRNA editing, degradation and splicing. This discovery shined a new light on the possibility of CTIP2 to be interacting with even more RNAs. From this the second part of the project was to identify and catalogue all the RNAs interacting with CTIP2. A battery of experiment called CLIP-Seq was performed, in triplicates in both HEK cells with overexpressed CTIP2 and Microglial cells against the endogenous CTIP2. The Cross-Linking Immunoprecipitation and Sequencing results have been used for the Publication 4 (currently in

preparation). From the first analysis we were able to predict the CTIP2-binding motif (Figure 38 A). Moreover, the analysis of the transcriptome found that more than 75% of the CLIP reads were mapped to exons (Figure 38 B) and when the reads were mapped to intronic regions, the location of the reads was on RNA isoforms that do not code for the protein such as retained introns (IR), nonsense mediated decay (NMD), non-stop decay or processed transcript.
Another interesting discovery was that by crosslinking the Mass Spec results with the CLIP-Seq ones, we found CTIP2 bound to 137 proteins and RNA encoded by the same gens. Of these proteins some of them were analysed and confirmed in the previous chapters such as FUS and SFPQ (Figure 38 C).

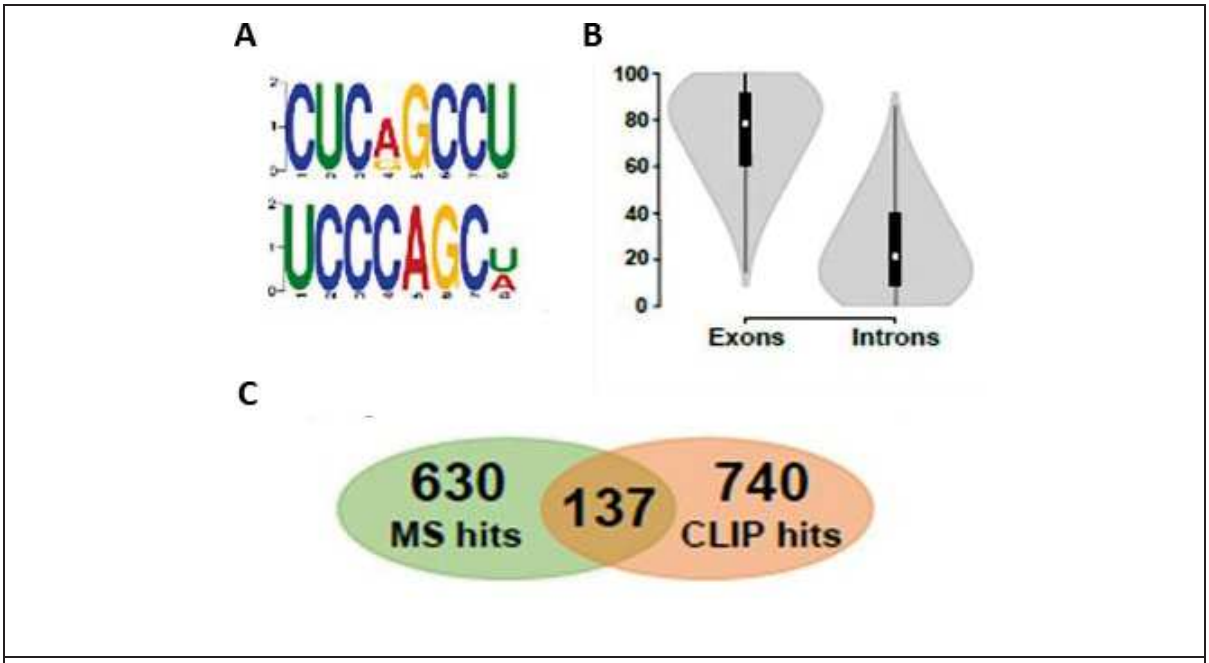


Figure 38: CTIP2-RNA interaction details. A) The two conserved motifs (CUCRGCCU and UCCCAGCW) in CTIP2-bound RNA regions. B) Number of CLIP reads mapped to exons and introns of the significant genes. C) CTIP2-interacting proteins, from IP-MS experiment, CTIP2-interacting RNAs from CLIP-Seq and the 137 genes that bind to CTIP2 in form of RNA or in form of protein

Of all the RNAs bound to CTIP2, around 97% were protein coding. While still a hypothesis, it is interesting to remember that many mRNAs are retained within the paraspeckle structures, ready to be translated if the necessity arises. A possible role in CTIP2 on acting on this mechanism could explain the interaction with so many mRNAs, while the fact that it specifically interacts with RNAs and proteins of the same gene has yet to be explained.

As said before, 97% of the RNAs were discovered to be protein coding. But our first glance on CTIP2 interaction with RNAs came from it role in the 7SK ribonucleoprotein complex. For this reason, we decided to take a look at the remaining 3% of the noncodingRNAs. Out of this group we decided to focus on five ncRNAs that showed a high hits count (high presence in the CTIP2 immunoprecipitation) and a relevant biological significance:

- Neat1: As described in the introduction chapter regarding the paraspeckle nuclear structure, NEAT1 (Nuclear paraspeckle assembly transcript 1) is a lncRNA forming the scaffold of these structures. It is present in two different forms NEAT1-1 (3.7 kb in length) and NEAT1-2 (23 kb in length) which are regulated by alternative 3′-end processing. Its expression has been shown to be induced by viral infection in mouse brains (Clemson et al., 2009; Naganuma et al., 2012)
- MALAT1: MALAT 1 (metastasis associated lung adenocarcinoma transcript 1) also known as NEAT2 (noncoding nuclear-enriched abundant transcript 2) is a large, infrequently spliced non-coding RNA, which is highly conserved amongst mammals and highly expressed in the nucleus. MALAT1 has been connected to alternative splicing, nuclear organization, epigenetic modulating of gene expression, and various pathological processes, ranging from diabetes complications to cancers (Hutchinson et al., 2007)
- MEG3: MEG3 (maternally expressed 3) is a maternally expressed, imprinted long non-coding RNA gene. MEG3 seems to have a negative role in tumorigenesis since its expression is lost in cancer cells. It acts as a growth suppressor in tumor cells, and activates p53 (Zhang et al., 2003)
- U11: The small nuclear ribonucleic acid U11 is an important non-coding RNA in the minor spliceosome protein complex, which activates the alternative splicing mechanism. U11 snRNA to recognize the 5' splice site region which possesses sequence complementarity with the 5' splice site of the eukaryotic U12 type pre-mRNA introns (Elliot and Ladomery, 2011)
- SNORD133 (SCARNA97): SNORD133 is a small RNA wich localize within the Cajal bodies. Together with SNORD97 (a small nucleolar RNA) they cooperatively 2′ -O-methylate tRNAMet(CAT), protecting it from stress-triggered angiogenin cleavage (Nostramo and Hopper, 2019).

We decide to perform RNA Immunoprecipitation in Microglial cells to confirm the interaction of CTIP2 with these RNAs. Considering how we already described CTIP2 interacting with SFPQ, and how SFPQ is a major component of the paraspeckle structure (with Neat1), we performed these series of RIP experiment with an αCTIP2, αSFPQ against a Mock IP using unspecific IgG to see the interaction of these two proteins with the lncRNA Neat1_1 and Neat1_2 (Figure 38). We can observe how we managed to confirm the interaction of CTIP2 with both the two variants of Neat, and of course also with SFPQ that it was acting here as our positive control. The fold increase was calculated against the Mock IP IgG (here set to 1) (Figure 39 A, B). While there was always an association of the RNAs with both proteins, a high grade of variance in the fold increase, among the triplicates, was observed. This has to do specifically with how much unspecific binding was eliminated (or not) from the unspecific IgG. We performed then a statistical analysis between the samples (Figure 39 C, D) to show that while there was a variance between the triplicates when compared to the IgG, the samples between each other maintained a stable ratio.

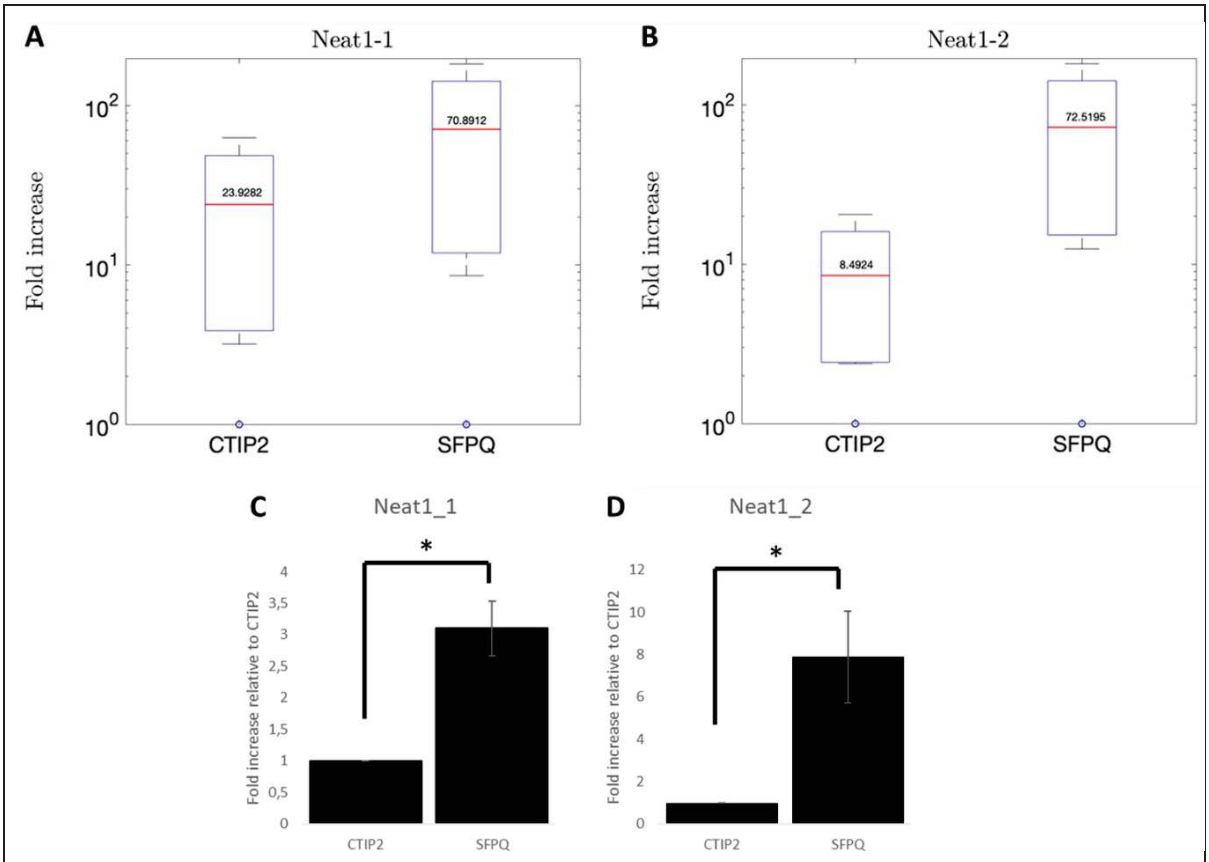


Figure 39: CTIP2 and SFPQ RIP of NEAT1. The whole cell lysate from microglial cell from 3 independent experiments was immunoprecipitated (in the presence of the RNase Inhibitor) with the chosen antibody (and IgG as mock) for 24h ON. The RNA was then extracted with the phenol-chloroform method, retrotranscribed in cDNA and analyzed through qPCR using primer for Neat1_1 (A) and Neat1_2 (B). The results from the CTIP2 and SFPQ RIP were then analyzed one against the other after normalizing them by CTIP2 (C,D). A t-test was done on 3 (C, D) independent experiments p (* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$).

Similar experiments were performed for Malat1 (Figure 40 A,C) and MEG3 (Figure 40 B,D). Again, the CLIP-Seq results were confirmed by our RNA IP experiments.

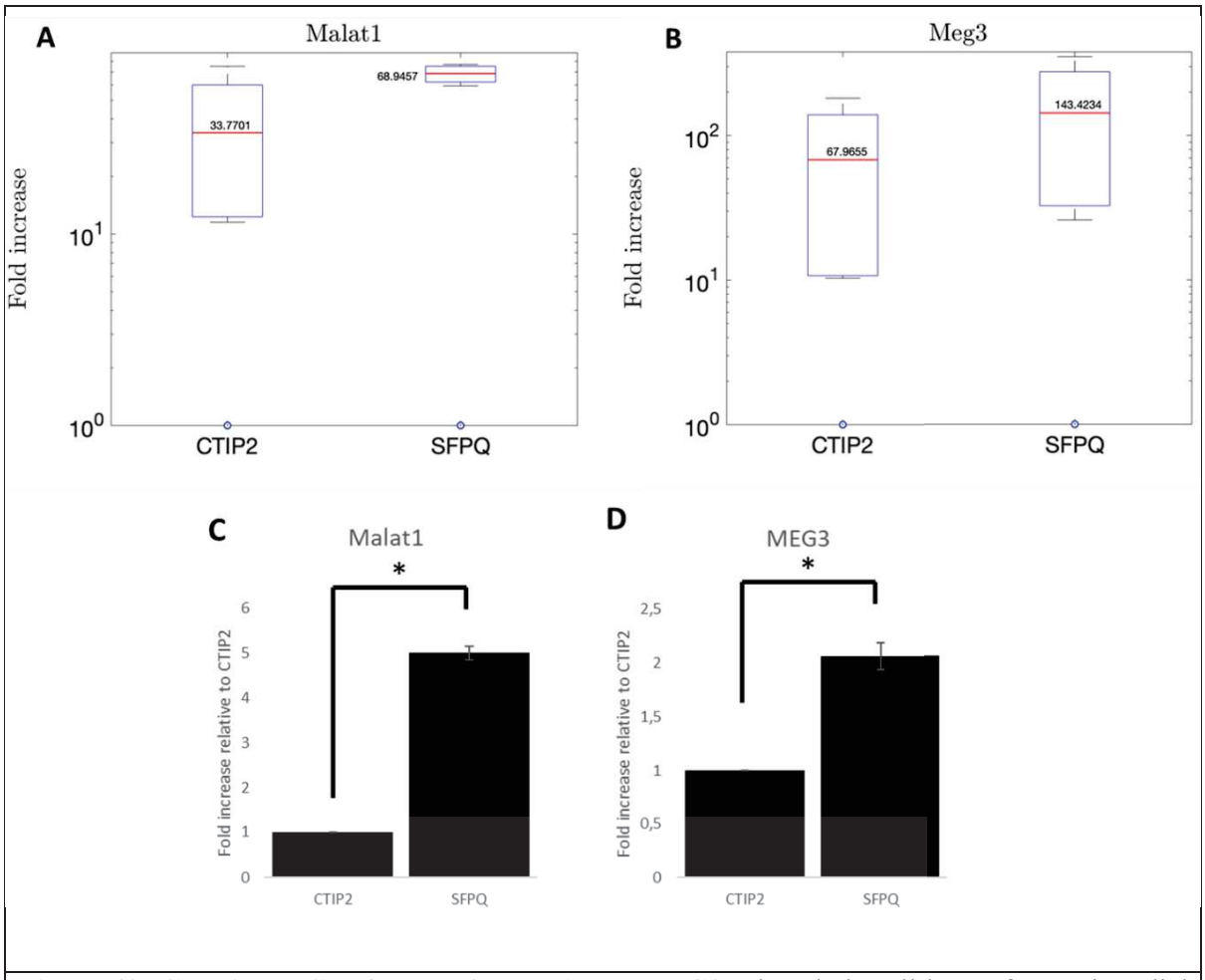


Figure 40: CTIP2 and SFPQ RIP of Malat1 and MEG3. The whole cell lysate form microglial cell from 3 independent experiments was immunoprecipitated (in the presence of the RNase Inhibitor) with the chosen antibody (and IgG as mock) for 24h ON. The RNA was then extracted with the phenol-chloroform method and retrotranscribed in cDNA and analysed through qPCR using primer for Malat1 (A) and MEG3 (B). The results from the CTIP2 and SFPQ RIP were then analysed one against the other after normalizing them by CTIP2 (C, D) (* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$).

Finally, we repeated the same experiment for SNORD133 (Figure 41 A,C) and U11 (Figure 41 B,D) and we again managed to confirm the interaction of these RNA with both CTIP2 and SFPQ. The ratio of the RNAs immunoprecipitated was again analysed, showing a rather stable ratio between the IPs for the sncRNA U11 (Figure 40 D) but a more variable one for SNORD133 (Figure 40 C)

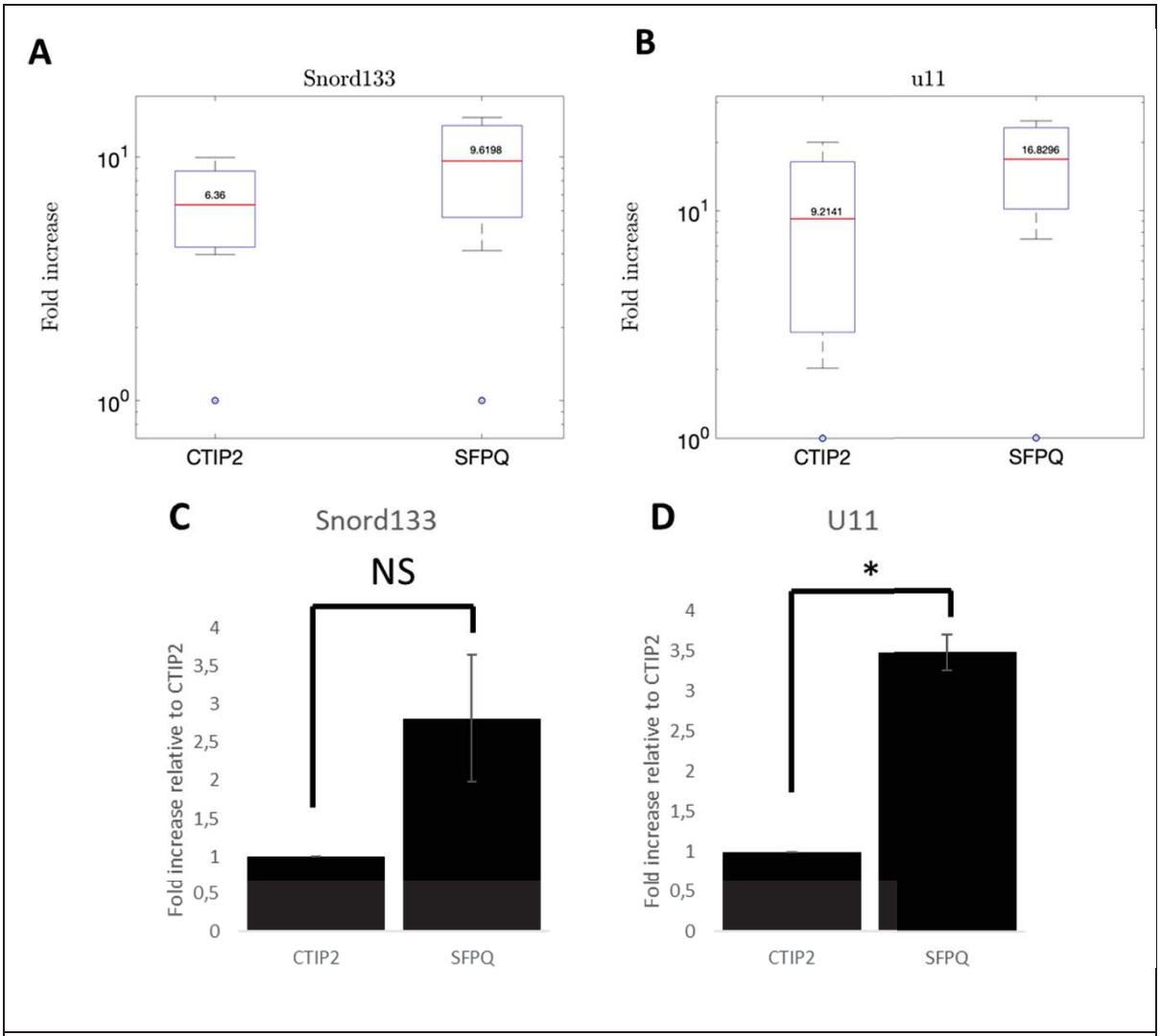


Figure 41: CTIP2 and SFPQ RIP of SNORD133 and U11. The whole cell lysate form microglial cell from 3 independent experiments was immunoprecipitated (in the presence of the RNase Inhibitor) with the chosen antibody (and IgG as mock) for 24h ON. The RNA was then extracted with the phenol-chloroform method and retrotranscribed in cDNA and analyzed through qPCR using primer for SNORD133 (A) and U11 (B). The results from the CTIP2 and SFPQ RIP were then analysed one against the other after normalizing them by CTIP2 (C, D) p (* $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$).

9. Discussion and future prospective

HIV has been one of the major challenges for the global public health for the past 40 years, with millions of deaths and millions of peoples currently infected. The introduction of cART and its evolution to the Highly active antiretroviral therapy (HAART) has saved millions of lives, and one of the main goals is still now to cover with the treatment as much as 90% of the infected population. The antiretroviral therapy while has saved lives has fallen short in being able to cure the infected people and a lingering infection still remain, ready to burst out once the regimen is interrupted. The discovery of population of latently infected cells acting as viral reservoir become then one of the main topics of research in the fight against HIV-1. While the CD4+ form most of the reservoir, the monocyte/macrophage line and microglial especially needed to be addressed. We have seen how the microglia, the resident macrophages of the CNS, are the main viral reservoir within the brain of infected patients, and in deep studies of this specific cell line is compulsory to achieve a sterilization cure. Microglial cells differ from the quiescent CD4+ reservoirs for many aspects, first of all their longevity, and it is important also to remember how protected are within the anatomically sanctuary of the brain, separated from the organism by the selective barrier which is the Blood-Brain-Barrier. Our lab has focused its attention on this cell line for years, analysing and studying the molecular mechanism behind the establishment of the latency of HIV-1. Among these molecular mechanisms, within the microglial nucleus, a protein become the protagonist of many researches we focused on. CTIP2 is a transcriptional factor that we described being at the centre of a chromatin remodelling complex able to silence the viral promoter by shifting the chromatin state from the euchromatin, easy to transcribe, to the heterochromatin, refractory to transcription and silenced. It was shown then how CTIP2 was also part of a ribonucleoprotein complex responsible for the maintenance of this latency, by sequestering in an inactive form the elongation factor PTEF-b necessary for viral gene transcription. With the years we started to analyse others proteins with roles in the suppression of HIV-1 transcription, proteins such KAP1 one that we described being able to interact with CTIP2 and cooperate in the silencing of HIV-1 genes; we also provided results showing how the virus itself has developed mechanisms to overcome these blocks like the ability to target for proteasomal degradation CTIP2 by the action of the viral protein Vpr.

By knowing how central CTIP2 is in these mechanisms, we started a wider project with the aim to identify, and possible describe all the interactants of CTIP2. At a protein level, a Quantitative Mass Spectrometry screening was performed, and many exciting new interactions

were discovered. Regarding its known interaction with the 7SK RNA, we knew that CTIP2 may had the ability to bind RNAs as well and a second project aimed at describing the RNAs interacting with CTIP2 was initiated as well. Interestingly the results obtained from the two projects confirmed each other's findings. Within the protein-protein interactions we discovered CTIP2 interacting with many different complexes exerting different functions ranging from RNA processing, splicing, the NMD pathway, to DNA damage, of course transcription regulation and chromatin structure modulation. Comparing also the two projects we saw how in many cases CTIP2 was found to be bound/interact with both the proteins and the RNA coding for the same proteins such as EWSR1, FUS, SFPQ, UPF1, tubulins, CHDs and hnRNPs; almost 97% of the RNAs bound to CTIP2 was protein coding RNA.

9.1. CTIP2 interaction with paraspeckle components

Among the protein-protein interaction we then moved to confirm some of these interactions that showed an interesting biological connection to HIV-1. We showed how CTIP2 interact in an RNA independent fashion and localize strongly with the protein SFPQ in sphere like structures; SFPQ is a main component, of the paraspeckle structures and we wanted to investigate if these were indeed the structures we were observing when we overexpressed CTIP2. To do so another paraspeckle protein, FUS was analysed and its interaction (and colocalization) was confirmed, its interaction with CTIP2 was again proved to be RNA independent. This localization was revealed to have a specific set of phenotypes, with proteins such FUS seemly pushed at the rim of the structures. A hypothesis maturated in the lab is that the excess CTIP2 relocated within these structures could push the paraspeckle structure to an almost collapse state. Interestingly enough, the viral protein Tat was shown to relocated in the same speckle in the presence of CTIP2 and we proved that also SFPQ may interact with Tat but not FUS. It is also worth to mention that increasing quantities of Tat did not affect the interaction of these two proteins with CTIP2.

We performed some functional studies to analyse the interplay between CTIP2 and SFPQ in the transcription of HIV-1 genes. What we observed was the known repressor effect of CTIP2, on the viral promoter, and the replication of the provirus. SFPQ showed a similar effect except a total reversed activator effect when the viral protein Tat was also expressed. This activator effect was lost in the presence of CTIP2 overexpression showing a definitely interplay between the two proteins, within the already described sphere-like structures. We also proved that the

two proteins cooperate to silence the HIV-1 expression, making us believe that their interplay may play a function also at post-transcriptional level. A confirmation of the hypothesis that CTIP2 may be part (or interact with) the paraspeckle structure came also from tests performed to confirm the CLIP-Seq data. In this case the NEAT1 lncRNA, major component of the paraspeckle structure, was selected for RIP and its interaction with CTIP2 (and of course SFPQ) was confirmed.

Putting these early data together different hypothesis can be made. Paraspeckle structures have been described to be temporary storage of sequestered proteins and mRNAs, that can be then promptly released and quickly translated when needed; these structures also seem to increase in number during some viral infections' events (Beeharry et al., 2018). Considering the interaction of CTIP2 with Tat and SFPQ, an idea could be that these structures, were these proteins colocalize, are indeed the paraspeckles. This would be confirmed by the interaction of CTIP2 with FUS and Neat1 as well, but the displacement of FUS make us believe that as CTIP2 is produced, under certain stimuli, it is relocated within the paraspeckle thanks to the interaction with SFPQ. This strong heap of proteins within this flexible structure may lead to a change in their basic structure conformation with FUS, a core protein, being pushed at the exterior, and probably breaking some of the scaffolding provided by NEAT1 (or the FUS antibody cannot access anymore these dense structures). Further analysis is required but paraspeckle or not, it is within these structures that SFPQ and CTIP2 exert together their repressor effect on HIV-1, and they need to be investigated further. It is interesting to remember that a study has shown SFPQ to be a negative regulator of the transcription of the IL8 gene, with is binding motif being located just 3' downstream of the TATA box of the human IL8 gene. Our lab has also found a similar repressor function for CTIP2 on the IL8 gene. What is interestingly is that expression of Neat1 RNA increases upon viral RNA detection such as influenza and Herpres simplex (Wang et al., 2017) but also HIV-1 (Zhang et al., 2013) and its overexpression was able to relocate SFPQ from the IL8 promoter to the paraspeckle, inducing IL8 transcription (Imamura et al., 2014). Considering our findings, it may have a similar effect on CTIP2 and the interplay between NEAT1, SFPQ and CTIP2 regarding IL8 and the immune response activation upon viral infection is worth being investigated further (Figure 42).

Another interestingly connection with our paper currently under writing, regarding CTIP2 Mass spectrometry and CLIP-seq results, showed how roughly 97% of the RNA bound to CTIP2 were protein coding. Of these, CTIP2 was bound to exons in the 75% of the cases; here

we present a possible hypothesis that may link these finding to some of the interaction described in this thesis. A role we described in the introduction regarding the paraspeckle function, is the retaining of edited mRNA (RNAs adenosine-to-inosine edited (A-to-I) in their 3′UTR) (Prasanth et al., 2005). These mRNA are "stored" in the paraspeckle and upon stress condition (as a viral infection) this UTR get cleaved and the RNA is ready to be exported and translated granting the cell a quick response in case of stressful situations. The connection of CTIP2 with the RNA modifying machinery, and the paraspeckle proteins, may be the results on CTIP2 actively binding these modified mRNAs. It is also interesting to note that , HIV-1 RNA has been reported to be subjected to A-to-I editing by the ADAR1 (adenosine deaminase acting on RNA 1) enzyme (Doria et al., 2009) providing an explanation and a confirmation of our hypothesis for the possibility to have HIV transcripts relocated within the paraspeckles and ADAR1 is a known to interact with SFPQ and we believe we could find an interaction with CTIP2 as well in future experiments.

Moreover, from the Mass spec results we linked CTIP2 to different proteins of the nuclear pore complex, and a role in the shuttling these mRNAs ready to be translated to the pore may be a yet not described function of CTIP2.

We do not know if the binding happens at the site of transcription, at the paraspeckle or if CTIP2 has an active role in shuttling these RNAs in the nucleus but our finding from the CLIP-Seq, and the proven interaction of CTIP2 with the paraspeckle proteins are now confirming each other in such way.

Regarding CTIP2 and its interaction with the paraspeckle proteins, a set of FISH experiments performed by Doctor Kula-Pacurar in her lab, showed an interesting effect of the overexpression of CTIP2 regarding its interaction with the NEAT1 lncRNA. It seems that the overexpression of CTIP2 causes the NEAT1 foci to disappear, and in some cells, where these CTIP2 structures are less numerous, a colocalization of CTIP2 and NEAT1 is noticeable. This early finding needs to be investigated further but could support the hypothesis here made that the increasing quantity of CTIP2 within the paraspeckle cause a shift in their conformational state with NEAT1 spreading in the nucleus, losing its specific concentration in the paraspeckle and FUS and other protein getting pushed on the outside of the structures.

Another interesting described association of the paraspeckle component with the HIV-1 replication was the report that NEAT1 knockdown was linked to enhanced virus production by increasing nucleus-to cytoplasm export of Rev-dependent INS-containing HIV-1 transcripts in Jurkat cells (Zhang et al., 2013). The knockdown of NEAT1 was of course followed by a decrease in number of paraspeckle but not in the quantity of the paraspeckle associated

proteins. Also it had no effect on the Tat activated LTR-transcription but results in increased cytoplasmic expression of HIV-1 INS-containing mRNAs suggests that paraspeckles may indeed represent the long-postulated nuclear compartment for storing HIV-1 Rev-dependent INS-containing RNAs (Berthold & Maldarelli, 1996) (Figure 42).

These previous finding, while on a different cell lines, should be tested on microglial cells as well, and it could explain the post-transcriptional repressor effect of CTIP2 and SFPQ, their common localization in the paraspeckle, and CTIP2 ability to relocate Tat within these structures.

Summarizing these results together and preparing a plan for future experiments we should further investigate the interaction of CTIP2 with other paraspeckle proteins. We should aim to link CTIP2 to the post-transcriptional machinery responsible for the Adenosine-to-inosine (A-to-I) RNA editing. ADAR, one of the protein responsible for this modification, has been shown to interact with proteins we already described interacting with CTIP2 such has SFPQ and MATR3. Moreover, we should try to identify this modification in the RNAs bound to CTIP2, proving its role in the retention of these RNAs within the paraspeckle structures. Also, a new set of confocal microscopy experiments and IP should be performed after Interferon stimulation. I believe that a more natural (and lower) overexpression of CTIP2 will keep the paraspeckle in a more steady state, and a more precise colocalization of CTIP2, especially with FUS, will be obtained; if my theory will be proved to be correct I expect to see similar results in the FISH experiment upon CTIP2 overexpression through Interferon stimulation regarding its interaction with Neat1. Finally, a Chromatin Immunoprecipitation assay should be performed upon Neat1 overexpression to confirm if it will have the same effect on CTIP2 binding on the IL8 gene promoter as it has on SFPQ.

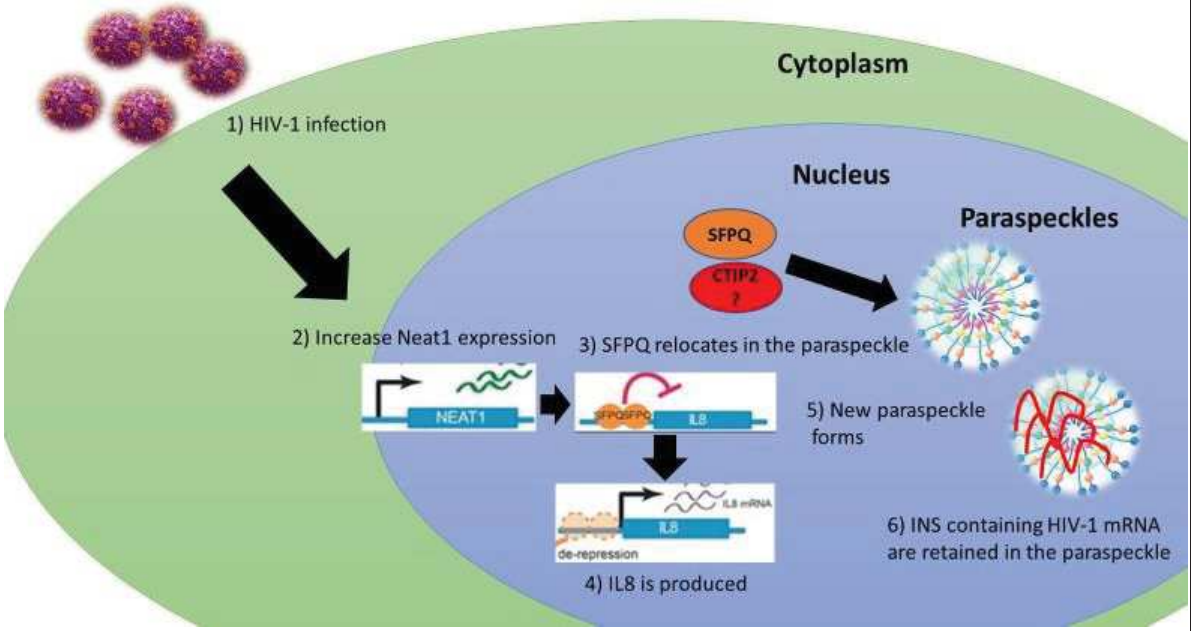


Figure 42: Transcriptional regulation of antiviral genes by action of Neat1 and paraspeckle proteins upon HIV-1 infection. HIV-1 infection causes an increase of NEAT1 expression which relocates SFPQ from the IL8 promoters where usually it exerts its repressor effect. CTIP2 has been shown to have the same effect and may be relocated as well. SFPQ is relocated and with the increased amount of NEAT1 from new paraspeckle which retain A-to-I modified viral RNAs repressing the viral replication.

9.2. CTIP2 interaction with the YTHDF family

While a role in this specific post-transcriptional modification is yet to be proven, another set of interesting proteins interaction were showed, with CTIP2 interacting with a family of proteins responsible for processing and modifying RNAs. The YTHDF family composed by three main proteins (DF1, DF2, DF3) were found to interact with CTIP2, especially the YTHDF2. These proteins are readers of a specific methylation of mRNAs the N^6-methyladenosine; this modification, and the activity of especially YTHDF2 results in mRNA degradation. YTHDF2 also regulates the innate immune response to infection by inhibiting the type I interferon response: acts by binding to m6A-containing IFNB transcripts and promoting their degradation (Winkler et al., 2019)

We were also able to perform a reverse RIP, finding CTIP2 associated with RNA bearing this modification, but to this day we are still not sure if this interaction is direct or bridged by CTIP2 interaction with the YTHDF proteins. Knowing that the HIV-1 RNA is methylated, a possible role in the viral mRNA degradation could be in the future elucidated.

Interestingly in T cells, Viral 3′ UTR m6A sites or analogous cellular m6A sites strongly enhanced mRNA expression in *cis* by recruiting the cellular YTHDF m6A "reader" proteins. Reducing YTHDF expression inhibited, while YTHDF overexpression enhanced, HIV-1 protein and RNA expression, and virus replication in CD4+ T cells. These data identify m6A editing and the resultant recruitment of YTHDF (especially YTHDF2) proteins as major positive regulators of HIV-1 mRNA expression (Kennedy et al., 2016). While this could be cell type specific, the strong interaction of CTIP2 with specifically YTHDF2 could exert a negative effect on the YHTDF2 enhancing role on HIV-1 expression. A set of Confocal microscopy experiments will be performed to confirm the localization of the interaction between CTIP2 and the proteins of the YTHDF family especially because, with some reserves, these proteins have been described to be predominantly cytoplasmic.

9.3. CTIP2 interaction with the MATR3

As last, we confirmed the interaction of CTIP2 with MATR3, a protein described to have a role in the shuttling of unspliced HIV-1 RNA (with Rev and SFPQ as cofactors) from the nucleus to the cytoplasm. Matr3 has been showed to strongly interact with SFPQ in exerting its function and also with ADAR1 in a recent interactome screening performed in HeLa cells (Iradi et al., 2018). We managed to confirm the interaction of CTIP2 with MATR3 in HEK cells, and to show this interaction to be RNA independent; moreover we described how Matr3 showed a repressor effect of the basal viral transcription and the Tat mediated one in microglial cells, an effect that could be cell-type specific considering opposing reports were showed by other groups.

9.4. CTIP2 newly discovered ncRNA partners

Moving on the CLIP-Seq data, we selected a small number of ncRNA based on their biological importance and the number of HITS from the Sequencing analysis and we performed a series of experiments with the aim to confirm our results. Regarding Neat1 we already described how our attention was drown to it by the fact that it is the scaffolding form the paraspeckle structures, but other RNAs were confirmed interacting with our protein.

MALAT1, a lncRNA connected to alternative splicing, nuclear organization and epigenetic modulating of gene expression, all cellular processes in which proteins and complexes we were

able to find interacting with CTIP2, was confirmed to be bound to CTIP2. It has also been reported to be upregulated upon HIV-1 infection in Jurkat and MT4 cell lines (Zhang et al., 2013).

MEG3 with its role as a tumor suppressor gene was also found bound to CTIP2, and again we see an RNA with proven functions mirroring CTIP2 described ones.

U11, an important component of the minor splicing complex, was another of the ncRNA found in the CLIP-Seq, confirmed in the RIP and double-confirming the results of CTIP2 being associated with the spliceosome machinery in the Quantitative Mass Spectrometry experiments.

SNORD133 an ncRNA found within the Cajal bodies, with function connected to the methylation of tRNAs.

10. Conclusion

In conclusion, these projects surrounding CTIP2 have opened an “opposite” Pandora vase; instead of all evil we found an exciting high number of interesting interactions that will spawn many future projects to be carried on. Many of these projects will shed light on the role of CTIP2 in many different cellular processes that were never linked to this protein before. Moreover, we shall not forget that the main objective is, at the end, to find the key stone in the latency mechanism of HIV-1. While we will continue in the perceivable future to analyse the SFPQ-CTIP2 interaction, seemingly the most promising, we should optimistically think that among all these CTIP2-protein interactions, and CTIP2-RNAs interaction we may have the weak spot that a targeted LRA, or genetic therapy could be aimed at to finally eradicate HIV-1, curing once and for all, one of the mains scourges of humanity.

11. References

Accola, Molly A., Åsa Öhagen, and Heinrich G. Göttlinger. "Isolation of human immunodeficiency virus type 1 cores: retention of Vpr in the absence of p6 gag." *Journal of virology* 74.13 (2000): 6198-6202.

Albright, A. V., Soldan, S. S., & González-Scarano, F. (2003). Pathogenesis of human immunodeficiency virus-induced neurological disease. *Journal of neurovirology*, *9*(2), 222-227.

Aloia, H Tian, F C Jensen "Lipid composition and fluidity of the human immunodeficiency virus envelope and host cell plasma membranes" Proceedings of the National Academy of Sciences Jun 1993, 90 (11) 5181-5185;

Archin NM, Liberty AL, Kashuba AD, et al. Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy [published correction appears in Nature. 2012 Sep 20;489(7416):460]. *Nature*. 2012;487(7408):482-485. Published 2012 Jul 25.

Archin, N. M., Bateson, R., Tripathy, M. K., Crooks, A. M., Yang, K. H., Dahl, N. P., Kearney, M. F., Anderson, E. M., Coffin, J. M., Strain, M. C., Richman, D. D., Robertson, K. R., Kashuba, A. D., Bosch, R. J., Hazuda, D. J., Kuruc, J. D., Eron, J. J., & Margolis, D. M. (2014). HIV-1 expression within resting CD4+ T cells after multiple doses of vorinostat. *The Journal of infectious diseases*, *210*(5), 728–735.

Arlotta, P., Molyneaux, B. J., Chen, J., Inoue, J., Kominami, R., & Macklis, J. D. (2005). Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. *Neuron*, *45*(2), 207-221.

Autran, B., Descours, B., Avettand-Fenoel, V., & Rouzioux, C. (2011). Elite controllers as a model of functional cure. *Current opinion in HIV and AIDS*, *6*(3), 181–187.

Avram, D., Fields, A., Pretty On Top, K., Nevrivy, D. J., Ishmael, J. E., & Leid, M. (2000). Isolation of a novel family of C(2)H(2) zinc finger proteins implicated in transcriptional repression mediated by chicken ovalbumin upstream promoter transcription factor (COUP-TF) orphan nuclear receptors. *The Journal of biological chemistry*, *275*(14), 10315–10322.

Bai, Y., Xue, H., Wang, K., Cai, L., Qiu, J., Bi, S., Lai, L., Cheng, M., Liu, S., & Liu, K. (2013). Covalent fusion inhibitors targeting HIV-1 gp41 deep pocket. *Amino acids*, *44*(2), 701–713.

Bajetto, A., Bonavia, R., Barbero, S., & Schettini, G. (2002). Characterization of chemokines and their receptors in the central nervous system: physiopathological implications. *Journal of neurochemistry*, *82*(6), 1311-1329.

Bannister, A., Zegerman, P., Partridge, J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410,** 120–124 (2001).

Barau J, Teissandier A, Zamudio N, Roy S, Nalesso V, Hérault Y, et al. (November 2016). "The DNA methyltransferase DNMT3C protects male germ cells from transposon activity". *Science*. **354** (6314): 909–912

Barre-Sinoussi, F, J. Chermann, F Rey, M. Nugeyre, S Chamaret, J Gruest, C Dauguet, et al. "Isolation of a T-Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS)." *Science* 220, no. 4599 (May 20, 1983): 868–71.

Baumgärtel, V., Müller, B., & Lamb, D. C. (2012). Quantitative live-cell imaging of human immunodeficiency virus (HIV-1) assembly. *Viruses*, *4*(5), 777-799.

Beeharry, Y., Goodrum, G., Imperiale, C. J., & Pelchat, M. (2018). The Hepatitis Delta Virus accumulation requires paraspeckle components and affects NEAT1 level and PSP1 localization. *Scientific reports*, *8*(1), 6031.

Bégin P, Nadeau KC. Epigenetic regulation of asthma and allergic disease. *Allergy Asthma Clin Immunol*. 2014;10(1):27. Published 2014 May 28. doi:10.1186/1710-1492-10-27

Berger, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature*, *447*(7143), 407-412.

Berthold, E., & Maldarelli, F. (1996). cis-acting elements in human immunodeficiency virus type 1 RNAs direct viral transcripts to distinct intranuclear locations. *Journal of virology*, *70*(7), 4667–4682.

Besnard, E.; Hakre, S.; Kampmann, M.; Lim, H.W.; Hosmane, N.N.; Martin, A.; Bassik, M.C.; Verschueren, E.; Battivelli, E.; Chan, J.; et al. The mTOR complex controls HIV Latency. Cell Host Microbe 2016, 20, 785–797.

Blazkova J, Trejbalova K, Gondois-Rey F, et al. CpG methylation controls reactivation of HIV from latency. *PLoS Pathog*. 2009;5(8)

Bond CS, Fox AH. Paraspeckles: nuclear bodies built on long noncoding RNA. J Cell Biol 2009, 186:637–644.

Bonsignori, M., Liao, H.-X., Gao, F., Williams, W.B., Alam, S.M., Montefiori, D.C. and Haynes, B.F. (2017), Antibody-virus co-evolution in HIV infection: paths for HIV vaccine development. Immunol Rev, 275: 145-160.

Bouchat, S., Gatot, J.-S., Kabeya, K., Cardona, C., Colin, L., Herbein, G., De Wit, S., Clumeck, N., Lambotte, O., Rouzioux, C., et al. (2012). Histone methyltransferase inhibitors induce HIV-1 recovery in resting CD4(+) T cells from HIV-1-infected HAART-treated patients. AIDS Lond. Engl. *26*, 1473–1482.

Bouchat, S., Delacourt, N., Kula, A., Darcis, G., Van Driessche, B., Corazza, F., Gatot, J.-S., Melard, A., Vanhulle, C., Kabeya, K., et al. (2015). Sequential treatment with 5-aza-2'-deoxycytidine and deacetylase inhibitors reactivates HIV-1. EMBO Mol. Med. *8*, 117–138.

Bouchat, S., Delacourt, N., Kula, A., Darcis, G., Van Driessche, B., Corazza, F., Gatot, J., Melard, A., Vanhulle, C., Kabeya, K., et al. (2016). Sequential treatment with 5-aza-2′-

deoxycytidine and deacetylase inhibitors reactivates HIV-1. EMBO Mol. Med. *8*, 117–138.

Bour S, Schubert U, Strebel K (March 1995). "The human immunodeficiency virus type 1 Vpu protein specifically binds to the cytoplasmic domain of CD4: implications for the mechanism of degradation". *Journal of Virology*. **69** (3): 1510–20.

Braaten, D., Franke, E. K., & Luban, J. (1996). Cyclophilin A is required for an early step in the life cycle of human immunodeficiency virus type 1 before the initiation of reverse transcription. *Journal of virology*, *70*(6), 3551-3560.

Bukrinsky, M., & Adzhubei, A. (1999). Viral protein R of HIV-1. *Reviews in medical virology*, *9*(1), 39-49.

Buxade M, Morrice N, Krebs DL, Proud CG. The PSF.p54nrb complex is a novel Mnk substrate that binds the mRNA for tumor necrosis factor α. J Biol Chem 2008, 283:57–65.

R C Aloia, H Tian, F C Jensen "Lipid composition and fluidity of the human immunodeficiency virus envelope and host cell plasma membranes" Proceedings of the National Academy of Sciences Jun 1993, 90 (11) 5181-5185; DOI: 10.1073/pnas.90.11.5181

Briggs JA, Wilk T, Welker R, Kräusslich HG, Fuller SD. Structural organization of authentic, mature HIV-1 virions and cores. *EMBO J*. 2003;22(7):1707-1715. doi:10.1093/emboj/cdg143

Cameron PU, Freudenthal PS, Barker JM, Gezelter S, Inaba K, et al. (1992) Dendritic cells exposed to human immunodeficiency virus type-1 transmit a vigorous cytopathic infection to CD4+ T cells. Science 257: 383–387.

Campbell-Yesufu, Rajesh T. Gandhi, Update on Human Immunodeficiency Virus (HIV)-2 Infection, *Clinical Infectious Diseases*, Volume 52, Issue 6, 15 March 2011, Pages 780–787,

Cane, P., Chrystie, I., Dunn, D., Evans, B., Geretti, A. M., Green, H., Phillips, A., Pillay, D., Porter, K., Pozniak, A., Sabin, C., Smit, E., Weber, J., Zuckerman, M., & UK Group on Transmitted HIV Drug Resistance (2005). Time trends in primary resistance to HIV drugs in the United Kingdom: multicentre observational study. *BMJ (Clinical research ed.)*, *331*(7529), 1368. https://doi.org/10.1136/bmj.38665.534595.55

Cihlar, T., & Fordyce, M. (2016). Current status and prospects of HIV treatment. *Current opinion in virology*, *18*, 50-56.

Cherrier, T., Suzanne, S., Redel, L., Calao, M., Marban, C., Samah, B., Mukerjee, R., Schwartz, C., Gras, G., Sawaya, B. E., Zeichner, S. L., Aunis, D., Van Lint, C., & Rohr, O. (2009). p21(WAF1) gene promoter is epigenetically silenced by CTIP2 and SUV39H1. *Oncogene*, *28*(38), 3380–3389.

Cherrier, Thomas, Valentin Le Douce, Sebastian Eilebrecht, Raphael Riclet, Céline Marban, Franck Dequiedt, Yannick Goumon, Jean-Christophe Paillart, Mathias Mericskay, Ara Parlakian, Pedro Bausero, Wasim Abbas, Georges Herbein, Siavash K. Kurdistani, Xavier Grana, Benoit Van Driessche, Christian Schwartz, Ermanno Candolfi, Arndt G.

Benecke, Carine Van Lint, and Olivier Rohr. “CTIP2 Is a Negative Regulator of P-TEFb.” *Proceedings of the National Academy of Sciences* 110, no. 31 (July 30, 2013): 12655.

Charneau P., M. Alizon, and F. Clavel. "A second origin of DNA plus-strand synthesis is required for optimal human immunodeficiency virus replication." *Journal of virology* 66.5 (1992): 2814-2820.

Charneau P, Borman AM, Quillent C, et al. Isolation and envelope sequence of a highly divergent HIV-1 isolate: definition of a new HIV-1 group. *Virology*. 1994;205(1):247-253.

Charneau, P., Mirambeau, G., Roux, P., Paulous, S., Buc, H., & Clavel, F. (1994). HIV-1 reverse transcription A termination step at the center of the genome. *Journal of molecular biology*, *241*(5), 651-662.

Chen, H., & Engelman, A. (1998). The barrier-to-autointegration protein is a host factor for HIV type 1 integration. *Proceedings of the National Academy of Sciences*, *95*(26), 15270-15274.

Chen, R., Le Rouzic, E., Kearney, J. A., Mansky, L. M., & Benichou, S. (2004). Vpr-mediated incorporation of UNG2 into HIV-1 particles is required to modulate the virus mutation rate and for replication in macrophages. *Journal of Biological Chemistry*, *279*(27), 28419-28425.

Chen LL, Carmichael GG. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. Mol Cell 2009, 35:467–478

Christ, F.; Voet, A.; Marchand, A.; Nicolet, S.; Desimmie, B.A.; Marchand, D.; Bardiot, D.; Van der Veken, N.J.; Van Remoortel, B.; De Maeyer, M.; et al. Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. Nat. Chem. Biol. 2010, 6, 442–448

Churchill, M. J., Deeks, S. G., Margolis, D. M., Siliciano, R. F., & Swanstrom, R. (2016). HIV reservoirs: what, where and how to target them. *Nature Reviews Microbiology*, *14*(1), 55.

Cismasiu, V. B., Adamo, K., Gecewicz, J., Duque, J., Lin, Q., & Avram, D. (2005). BCL11B functionally associates with the NuRD complex in T lymphocytes to repress targeted promoter. *Oncogene*, *24*(45), 6753-6764.

Cismasiu, V. B., Ghanta, S., Duque, J., Albu, D. I., Chen, H. M., Kasturi, R., & Avram, D. (2006). BCL11B participates in the activation of IL2 gene expression in CD4+ T lymphocytes. *Blood*, *108*(8), 2695-2702.

Clemson, C.M., J.N. Hutchinson, S.A. Sara, A.W. Ensminger, A.H. Fox, A. Chess, and J.B. Lawrence. 2009. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. Mol. Cell. 33:717–726.

Clutton, G. T., & Jones, R. B. (2018). Diverse Impacts of HIV Latency-Reversing Agents on CD8+ T-Cell Function: Implications for HIV Cure. *Frontiers in immunology*, *9*, 1452. https://doi.org/10.3389/fimmu.2018.01452

Coffin, John M.; Hughes, Stephen H.; Varmus, Harold E. (1997). *The Place of Retroviruses in Biology*. Cold Spring Harbor Laboratory Press

Colin L, Van Lint C. Molecular control of HIV-1 postintegration latency: implications for the development of new therapeutic strategies. *Retrovirology*. 2009;6:111. Published 2009 Dec 4.

Conticello, Silvestro G., Reuben S. Harris, and Michael S. Neuberger. "The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G." *Current Biology* 13.22 (2003): 2009-2013.

Contreras, X., Barboric, M., Lenasi, T., & Peterlin, B. M. (2007). HMBA releases P-TEFb from HEXIM1 and 7SK snRNA via PI3K/Akt and activates HIV transcription. *PLoS Pathog*, *3*(10), e146.

Cosson, P. (1996). Direct interaction between the envelope and matrix proteins of HIV-1. *The EMBO journal*, *15*(21), 5783-5788.

Cotter, A. G., & Mallon, P. W. (2014). The effects of untreated and treated HIV infection on bone disease. *Current opinion in HIV and AIDS*, *9*(1), 17-26.

Craigie, R., & Bushman, F. D. (2014). *Host factors in retroviral integration and the selection of integration target sites. Microbiol. Spectr. 2*. MDNA3-0026-2014.

Cullen, Bryan R (2003). "Nuclear mRNA export: Insights from virology". *Trends in Biochemical Sciences*. **28** (8): 419–424.

Darcis, G., Berkhout, B., & Pasternak, A. O. (2019). The Quest for Cellular Markers of HIV Reservoirs: Any Color You Like. *Frontiers in immunology*, *10*, 2251.

Darcis, G., Van Driessche, B., & Van Lint, C. (2017). HIV Latency: Should We Shock or Lock?. *Trends in immunology*, *38*(3), 217–228.

Darcis, G., Kula, A., Bouchat, S., Fujinaga, K., Corazza, F., Ait-Ammar, A., Delacourt, N., Melard, A., Kabeya, K., Vanhulle, C., Van Driessche, B., Gatot, J. S., Cherrier, T., Pianowski, L. F., Gama, L., Schwartz, C., Vila, J., Burny, A., Clumeck, N., Moutschen, M., … Van Lint, C. (2015). An In-Depth Comparison of Latency-Reversing Agent Combinations in Various In Vitro and Ex Vivo HIV-1 Latency Models Identified Bryostatin-1+JQ1 and Ingenol-B+JQ1 to Potently Reactivate Viral Gene Expression. *PLoS pathogens*, *11*(7), e1005063.

Das, K., & Arnold, E. (2013). HIV-1 reverse transcriptase and antiviral drug resistance. Part 1. *Current opinion in virology*, *3*(2), 111–118.

Debouck, C., Gorniak, J. G., Strickler, J. E., Meek, T. D., Metcalf, B. W., & Rosenberg, M. (1987). Human immunodeficiency virus protease expressed in Escherichia coli exhibits autoprocessing and specific maturation of the gag precursor. *Proceedings of the National Academy of Sciences*, *84*(24), 8903-8906.

Desplats, P., Dumaop, W., Smith, D., Adame, A., Everall, I., Letendre, S., Ellis, R., Cherner, M., Grant, I., & Masliah, E. (2013). Molecular and pathologic insights from latent HIV-1 infection in the human brain. *Neurology*, *80*(15), 1415–1423.

Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Heaphy, S., Karn, J., Valerio, R. (1989). Human immunodeficiency virus 1 tat protein binds trans-activation-responsive region (TAR) RNA in vitro. *Proceedings of the National Academy of Sciences*, *86*(18), 6925-6929.

Dalgleish, A., Beverley, P., Clapham, P. *et al.* The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* **312,** 763–767 (1984).

D'arc, Mirela, Ahidjo Ayouba, Amandine Esteban, Gerald H. Learn, Vanina Boué, Florian Liegeois, Lucie Etienne, et al. "Origin of the HIV-1 Group O Epidemic in Western Lowland Gorillas." *Proceedings of the National Academy of Sciences* 112, no. 11 (March 17, 2015): E1343–52.

Debouck, C., Gorniak, J. G., Strickler, J. E., Meek, T. D., Metcalf, B. W., & Rosenberg, M. (1987). Human immunodeficiency virus protease expressed in Escherichia coli exhibits autoprocessing and specific maturation of the gag precursor. *Proceedings of the National Academy of Sciences*, *84*(24), 8903-8906.

Deeks, S. G., Lewin, S. R., Ross, A. L., Ananworanich, J., Benkirane, M., Cannon, P., Chomont, N., Douek, D., Lifson, J. D., Lo, Y. R., Kuritzkes, D., Margolis, D., Mellors, J., Persaud, D., Tucker, J. D., Barre-Sinoussi, F., International AIDS Society Towards a Cure Working Group, Alter, G., Auerbach, J., Autran, B., … Zack, J. (2016). International AIDS Society global scientific strategy: towards an HIV cure 2016. *Nature medicine*, *22*(8), 839–850.

Descours, B., Petitjean, G., López-Zaragoza, J. L., Bruel, T., Raffel, R., Psomas, C., Reynes, J., Lacabaratz, C., Levy, Y., Schwartz, O., Lelievre, J. D., & Benkirane, M. (2017). CD32a is a marker of a CD4 T-cell HIV reservoir harbouring replication-competent proviruses. *Nature*, *543*(7646), 564–567.

Doms, R. W., & Moore, J. P. (2000). HIV-1 membrane fusion: targets of opportunity. *The Journal of cell biology*, *151*(2), F9-F14.

Doria, M., Neri, F., Gallo, A., Farace, M. G., & Michienzi, A. (2009). Editing of HIV-1 RNA by the double-stranded RNA deaminase ADAR1 stimulates viral infection. *Nucleic acids research*, *37*(17), 5848–5858.

Dubuissez, M., Loison, I., Paget, S., Vorng, H., Ait-Yahia, S., Rohr, O., Tsicopoulos, A., & Leprince, D. (2016). Protein Kinase C-Mediated Phosphorylation of BCL11B at Serine 2 Negatively Regulates Its Interaction with NuRD Complexes during CD4+ T-Cell Activation. *Molecular and cellular biology*, *36*(13), 1881–1898.

Eilebrecht, S., Le Douce, V., Riclet, R., Targat, B., Hallay, H., Van Driessche, B., Schwartz, C., Robette, G., Van Lint, C., Rohr, O., & Benecke, A. G. (2014). HMGA1 recruits CTIP2-repressed P-TEFb to the HIV-1 and cellular target promoters. *Nucleic acids research*, *42*(8), 4962–4971.

Elliott, J. H., McMahon, J. H., Chang, C. C., Lee, S. A., Hartogensis, W., Bumpus, N., Savic, R., Roney, J., Hoh, R., Solomon, A., Piatak, M., Gorelick, R. J., Lifson, J., Bacchetti, P., Deeks, S. G., & Lewin, S. R. (2015). Short-term administration of disulfiram for reversal of latent HIV infection: a phase 2 dose-escalation study. *The lancet. HIV*, *2*(12)

Elliot, David; Ladomery, Michael (2011). *Molecular Biology of RNA*. Oxford University Press. p. 124.

Endsley, M. A., Somasunderam, A. D., Li, G., Oezguen, N., Thiviyanathan, V., Murray, J. L., Rubin, D. H., Hodge, T. W., O'Brien, W. A., Lewis, B. & Ferguson, M. R. (2014). Nuclear trafficking of the HIV-1 pre-integration complex depends on the ADAM10 intracellular domain. *Virology*, *454*, 60-66.

Esposito, F., Corona, A., & Tramontano, E. (2012). HIV-1 reverse transcriptase still remains a new drug target: structure, function, classical inhibitors, and new inhibitors with innovative mechanisms of actions. *Molecular biology international*, *2012*.

Farnet, C. M., & Bushman, F. D. (1997). HIV-1 cDNA integration: requirement of HMG I (Y) protein for function of preintegration complexes in vitro. *Cell*, *88*(4), 483-492.

Figdor CG, van Kooyk Y, Adema GJ (2002) C-type lectin receptors on dendritic cells and Langerhans cells. Nat Rev Immunol 2: 77–84.

Forouzanfar, F., Ali, S., Wallet, C., De Rovere, M., Ducloy, C., El Mekdad, H., El Maassarani, M., Van Assche, J., Boutant, E., Daouad, F., Margottin-Goguet, F., Moog, C., Van Lint, C., Schwartz, C., Rohr, O., (2019). HIV-1 Vpr mediates the depletion of the cellular repressor CTIP2 to counteract viral gene silencing. *Scientific reports*, *9*(1), 1-10.

Fox AH, Lamond AI. Paraspeckles. Cold Spring Harb Perspect Biol 2010, 2:a000687.

Fox AH, Bond CS, Lamond AI. P54nrb forms a heterodimer with PSP1 that localizes to paraspeckles in an RNA-dependent manner. Mol Biol Cell 2005, 16:5304–5315.

Friedman, J., Cho, W. K., Chu, C. K., Keedy, K. S., Archin, N. M., Margolis, D. M., & Karn, J. (2011). Epigenetic silencing of HIV-1 by the histone H3 lysine 27 methyltransferase enhancer of Zeste 2. *Journal of virology*, *85*(17), 9078–9089.

Galietta A, Gunby RH, Redaelli S, Stano P, Carniti C, Bachi A, Tucker PW, Tartari CJ, Huang CJ, Colombo E, et al. NPM/ALK binds and phosphorylates the RNA/DNA-binding protein PSF in anaplastic large-cell lymphoma. Blood 2007, 110:2600–2609.

Gallay, P., Hope, T., Chin, D., & Trono, D. (1997). HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway. *Proceedings of the National Academy of Sciences*, *94*(18), 9825-9830.

Gallo RC, Sarin PS, Gelmann EP, Robert-Guroff M, Richardson E, Kalyanaraman VS, Mann D, Sidhu GD, Stahl RE, Zolla-Pazner S, Leibowitch J, Popovic M (May 1983). "Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS)". Science. 220 (4599): 865–67.

Ganguli-Indra, G., Wasylyk, C., Liang, X., Millon, R., Leid, M., Wasylyk, B., Abecassis, J., & Indra, A. K. (2009). CTIP2 expression in human head and neck squamous cell carcinoma is linked to poorly differentiated tumor status. *PloS one*, *4*(4), e5367.

Gao, Feng, Elizabeth Bailes, David L Robertson, Yalu Chen, Cynthia M Rodenburg, Scott F Michael, Larry B Cummins, George M Shaw, Paul M Sharp, and Beatrice H Hahn. "Origin of HIV-1 in the Chimpanzee" 397 (1999):

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., Myszka, D. G., & Sundquist, W. I. (2001). Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell*, *107*(1), 55–65.

Gasparian, A.V.; Burkhart, C.A.; Purmal, A.A.; Brodsky, L.; Pal, M.; Saranadasa, M.; Bosykh, D.A.; Commane, M.; Guryanova, O.A.; Pal, S.; et al. Curaxins: Anticancer compounds that simultaneously suppress NF-κB and activate p53 by targeting FACT. Sci. Transl. Med. 2011, 3, 95ra74

Gavegnano, C.; Detorio, M.; Montero, C.; Bosque, A.; Planelles, V.; Schinazi, R.F. Ruxolitinib and Tofacitinib Are Potent and Selective Inhibitors of HIV-1 Replication and Virus Reactivation in vitro. Antimicrob. Agents Chemother. 2014, 58, 1977–1986

Gelderblom HR, Hausmann EH, Ozel M, Pauli G, Koch MA. Fine structure of human immunodeficiency virus (HIV) and immunolocalization of structural proteins. *Virology*. 1987;156(1):171-176.

Geyer M, Fackler OT, Peterlin BM (July 2001). "Structure–function relationships in HIV-1 Nef". *EMBO Reports*. **2** (7): 580–5.

Geijtenbeek, T. B., Kwon, D. S., Torensma, R., van Vliet, S. J., van Duijnhoven, G. C., Middel, J., Cornelissen, I. L., Nottet, H. S., KewalRamani, V. N., Littman, D. R., Figdor, C. G., & van Kooyk, Y. (2000). DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell*, *100*(5), 587–597..

Glass, J. D., Fedor, H., Wesselingh, S. L., & McArthur, J. C. (1995). Immunocytochemical quantitation of human immunodeficiency virus in the brain: correlations with dementia. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *38*(5), 755-762.

Golonzhka, O., Liang, X., Messaddeq, N., Bornert, J. M., Campbell, A. L., Metzger, D., Chambon, P., Ganguli-Indra, G., Leid, M., & Indra, A. K. (2009). Dual role of COUP-TF-interacting protein 2 in epidermal homeostasis and permeability barrier formation. *The Journal of investigative dermatology*, *129*(6), 1459–1470.

Golonzhka, O., Metzger, D., Bornert, J. M., Bay, B. K., Gross, M. K., Kioussi, C., & Leid, M. (2009). Ctip2/Bcl11b controls ameloblast formation during mammalian odontogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(11), 4278–4283.

González-Scarano, Francisco, and Julio Martín-García. "The neuropathogenesis of AIDS." *Nature reviews immunology* 5.1 (2005): 69-81.

Gras, G., & Kaul, M. (2010). Molecular mechanisms of neuroinvasion by monocytes-macrophages in HIV-1 infection. *Retrovirology*, *7*(1), 1-11.

Greco-Stewart VS, Thibault CS, Pelchat M. Binding of the polypyrimidine tract-binding protein-associated splicing factor (PSF) to the hepatitis *δ* virus RNA. Virology 2006, 356:35–44.

Gupta, R. K., Peppa, D., Hill, A. L., Gálvez, C., Salgado, M., Pace, M., McCoy, L. E., Griffith, S. A., Thornhill, J., Alrubayyi, A., Huyveneers, L., Nastouli, E., Grant, P., Edwards, S. G., Innes, A. J., Frater, J., Nijhuis, M., Wensing, A., Martinez-Picado, J., & Olavarria, E. (2020). Evidence for HIV-1 cure after CCR5Δ32/Δ32 allogeneic haemopoietic stem-cell transplantation 30 months post analytical treatment interruption: a case report. *The lancet. HIV*, *7*(5), e340–e347.

Ha K, Takeda Y, Dynan WS. Sequences in PSF/SFPQ mediate radioresistance and recruitment of PSF/SFPQ-containing complexes to DNA damage sites in human cells. DNA Repair (Amst) 2011, 10:252–259

Hadzopoulou-Cladaras M, Felber BK, Cladaras C, Athanassopoulos A, Tse A, Pavlakis GN (March 1989). "The rev (trs/art) protein of human immunodeficiency virus type 1 affects viral mRNA and protein expression via a cis-acting sequence in the env region". *Journal of Virology*. **63** (3)

Hall-Pogar T, Liang S, Hague LK, Lutz CS. Specific trans-acting proteins interact with auxiliary RNA polyadenylation elements in the COX-2 3'-UTR. RNA 2007, 13:1103–1115.

Halper-Stromberg, A., Lu, C.-L., Klein, F., Horwitz, J.A., Bournazos, S., Nogueira, L., Eisenreich, T.R., Liu, C., Gazumyan, A., Schaefer, U., et al. (2014). Broadly Neutralizing Antibodies and Viral inducers decrease rebound from HIV-1 latent reservoirs in humanized mice. Cell *158*, 989–999.

Han Y, Lin YB, An W, et al. Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. *Cell Host Microbe*. 2008;4(2):134-146. doi:10.1016/j.chom.2008.06.008

Hayashi, T., Jean, M., Huang, H., Simpson, S., Santoso, N.G., and Zhu, J. (2017). Screening of an FDA-approved compound library identifies levosimendan as a novel anti-HIV-1 agent that inhibits viral transcription. Antiviral Res. *146*, 76–85.

He, J., Chen, Y., Farzan, M., Choe, H., Ohagen, A., Gartner, S., Busciglio, J., Yang, X., Hofmann, W., Newman, W., Mackay, C. R., Sodroski, J., Gabuzda, D. (1997). CCR3 and CCR5 are co-receptors for HIV-1 infection of microglia. *Nature*, *385*(6617), 645-649.

He G, Margolis DM. Counterregulation of chromatin deacetylation and histone deacetylase occupancy at the integrated promoter of human immunodeficiency virus type 1 (HIV-1) by the HIV-1 repressor YY1 and HIV-1 activator Tat. *Mol Cell Biol*. 2002;22(9):2965-2973.

Heng, X.; Kharytonchyk, S.; Garcia, E.L.; Lu, K.; Divakaruni, S.S.; LaCotti, C.; Edme, K.; Telesnitsky, A.; Summers, M.F. Identification of a minimal region of the HIV-1 5′-leader required for RNA dimerization, NC binding, and packaging. *J. Mol. Biol.* **2012**, *417*, 224–239

Henrich, T.J., Hu, Z., Li, J.Z., Sciaranghella, G., Busch, M.P., Keating, S.M., Gallien, S., Lin, N.H., Giguel, F.F., Lavoie, L., et al. (2013). Long-term reduction in peripheral blood HIV type 1 reservoirs following reduced-intensity conditioning allogeneic stem cell transplantation. J.Infect. Dis. *207*, 1694–1702.

Heyd F, Lynch KW. Phosphorylation-dependent regulation of PSF by GSK3 controls CD45 alternative splicing. Mol Cell 2010, 40:126–137.

Hindmarsh, P., Ridky, T., Reeves, R., Andrake, M., Skalka, A. M., & Leis, J. (1999). HMG protein family members stimulate human immunodeficiency virus type 1 and avian sarcoma virus concerted DNA integration in vitro. *Journal of virology*, *73*(4), 2994-3003.

Hirano K, Erdodi F, Patton JG, Hartshorne DJ. Interaction of protein phosphatase type 1 with a splicing factor. FEBS Lett 1996, 389:191–194

Ho, D. D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M., & Markowitz, M. (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature*, *373*(6510), 123-126.

Ho, Y. C., Shan, L., Hosmane, N. N., Wang, J., Laskey, S. B., Rosenbloom, D. I., Lai J., Blankson J., Siciliano J., & Siliciano, R. F. (2013). Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell*, *155*(3), 540-551.

Honeycutt, J. B., Thayer, W. O., Baker, C. E., Ribeiro, R. M., Lada, S. M., Cao, Y., Cleary, R. A., Hudgens, M. G., Richman, D. D., & Garcia, J. V. (2017). HIV persistence in tissue macrophages of humanized myeloid-only mice during antiretroviral therapy. *Nature medicine*, *23*(5), 638–643

Hu, W. S., & Hughes, S. H. (2012). HIV-1 reverse transcription. *Cold Spring Harbor perspectives in medicine*, *2*(10), a006882.

Hu SB, Yao RW, Chen LL. Shedding light on paraspeckle structure by super-resolution microscopy. *J Cell Biol*. 2016;214(7):789-791. doi:10.1083/jcb.201609008

Hutchinson, J. N., Ensminger, A. W., Clemson, C. M., Lynch, C. R., Lawrence, J. B., & Chess, A. (2007). A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC genomics*, *8*, 39.

Imamura K, Imamachi N, Akizuki G, Kumakura M, Kawaguchi A, Nagata K, Kato A, Kawaguchi Y, Sato H, Yoneda M, et al. Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL8 expression upon immune stimuli. Mol Cell 2014, 53:393–406.

Iradi, M., Triplett, J. C., Thomas, J. D., Davila, R., Crown, A. M., Brown, H., Lewis, J., Swanson, M. S., Xu, G., Rodriguez-Lebron, E., & Borchelt, D. R. (2018). Characterization of gene

regulation and protein interaction networks for Matrin 3 encoding mutations linked to amyotrophic lateral sclerosis and myopathy. *Scientific reports*, *8*(1), 4049.

Izquierdo-Useros, Nuria & Mar, Naranjo & Erkizia, Itziar & Puertas, Maria & Borràs, Francesc & Blanco, Julià & Martinez-Picado, Javier. (2010). HIV and Mature Dendritic Cells: Trojan Exosomes Riding the Trojan Horse?. PLoS pathogens. 6. e1000740. 10.1371/journal.ppat.1000740.

Jan, E., Mohr, I., & Walsh, D. (2016). A cap-to-tail guide to mRNA translation strategies in virus-infected cells. *Annual review of virology*, *3*, 283-307.

Jang, M.K.; Mochizuki, K.; Zhou, M.; Jeong, H.S.; Brady, J.N.; Ozato, K. The Bromodomain Protein Brd4 Is a Positive Regulatory Component of P-TEFb and Stimulates RNA Polymerase II-Dependent Transcription. Mol. Cell 2005, 19, 523–534

Johnson, Silas F.; Telesnitsky, Alice (2010). Madhani, Hiten D (ed.). "Retroviral RNA Dimerization and Packaging: The What, How, when, Where, and Why". *PLoS Pathogens*. **6** (10)

Jones, Katherine A., and B. Matija Peterlin. "Control of RNA initiation and elongation at the HIV-1 promoter." *Annual review of biochemistry* 63.1 (1994): 717-743

Jordan, A., Defechereux, P., & Verdin, E. (2001). The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. *The EMBO journal*, *20*(7), 1726-1738.

Kaneko S, Rozenblatt-Rosen O, Meyerson M, Manley JL. The multifunctional protein p54nrb/PSF recruits the exonuclease XRN2 to facilitate pre-mRNA 3' processing and transcription termination. Genes Dev 2007, 21:1779–1789.

Karn, J. (2011). The molecular biology of HIV latency: breaking and restoring the Tat-dependent transcriptional circuit. *Current opinion in HIV and AIDS*, *6*(1), 4.

Kauder SE, Bosque A, Lindqvist A, Planelles V, Verdin E. Epigenetic regulation of HIV-1 latency by cytosine methylation. *PLoS Pathog*. 2009;5(6):e1000495.

Kaul, M., Garden, G. A., & Lipton, S. A. (2001). Pathways to neuronal injury and apoptosis in HIV-associated dementia. *Nature*, *410*(6831), 988-994.

Ke, Y., Dramiga, J., Schütz, U., Kril, J. J., Ittner, L. M., Schröder, H., & Götz, J. (2012). Tau-mediated nuclear depletion and cytoplasmic accumulation of SFPQ in Alzheimer's and Pick's disease. *PloS one*, *7*(4), e35678.

Kennedy EM, Bogerd HP, Kornepati AV, 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. 2016 May 11;19(5):675-85.

Kessing, C. F., Nixon, C. C., Li, C., Tsai, P., Takata, H., Mousseau, G., Ho, P. T., Honeycutt, J. B., Fallahi, M., Trautmann, L., Garcia, J. V., & Valente, S. T. (2017). In Vivo Suppression of

HIV Rebound by Didehydro-Cortistatin A, a "Block-and-Lock" Strategy for HIV-1 Treatment. *Cell reports*, *21*(3), 600–611.

Kleiman, Lawrence. "tRNALys3: The Primer tRNA for Reverse Transcription in HIV-1." *IUBMB life* 53.2 (2002): 107-114.

Knott,G.J., Bond,C.S. and Fox,A.H. (2016) The DBHS proteins SFPQ, NONO and PSPC1: a multipurpose molecular scaffold. Nucleic Acids Res., 44, 3989–4004.

Kogan, M., & Rappaport, J. (2011). HIV-1 accessory protein Vpr: relevance in the pathogenesis of HIV and potential for therapeutic intervention. *Retrovirology*, *8*(1), 1-20.

König, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D. J., Luscombe, N. M., & Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature structural & molecular biology*, *17*(7), 909–915

Kominami R. (2012). Role of the transcription factor Bcl11b in development and lymphomagenesis. *Proceedings of the Japan Academy. Series B, Physical and biological sciences*, *88*(3), 72–87.

Kondo, E., Mammano, F., Cohen, E. A., & Göttlinger, H. G. (1995). The p6gag domain of human immunodeficiency virus type 1 is sufficient for the incorporation of Vpr into heterologous viral particles. *Journal of virology*, *69*(5), 2759-2764.

Kotov, A., Zhou, J., Flicker, P., & Aiken, C. (1999). Association of Nef with the human immunodeficiency virus type 1 core. *Journal of virology*, *73*(10), 8824-8830.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell*, *128*(4), 693-705.

Kula, A., Guerra, J., Knezevich, A., Kleva, D., Myers, M. P., & Marcello, A. (2011). Characterization of the HIV-1 RNA associated proteome identifies Matrin 3 as a nuclear cofactor of Rev function. *Retrovirology*, *8*(1), 60.

Kula, A., Gharu, L., & Marcello, A. (2013). HIV-1 pre-mRNA commitment to Rev mediated export through PSF and Matrin 3. Virology, 435(2), 329-340.

Kuwahara S, Ikei A, Taguchi Y, Tabuchi Y, Fujimoto N, Obinata M, Uesugi S, Kurihara Y. PSPC1, NONO, and SFPQ are expressed in mouse Sertoli cells and may function as coregulators of androgen receptor-mediated transcription. Biol Reprod 2006, 75:352–359.

Lama, J., & Planelles, V. (2007). Host factors influencing susceptibility to HIV infection and AIDS progression. *Retrovirology*, *4*(1), 52.

Landeras-Bueno S, Jorba N, Perez-Cidoncha M, Ortin J. The splicing factor proline-glutamine rich (SFPQ/PSF) is involved in influenza virus transcription. PLoS Pathog 2011, 7:e1002397.

Learmont, J. C., Geczy, A. F., Mills, J., Ashton, L. J., Raynes-Greenow, C. H., Garsia, R. J., ... & Deacon, N. J. (1999). Immunologic and virologic status after 14 to 18 years of infection with an attenuated strain of HIV-1—a report from the Sydney Blood Bank Cohort. *New England Journal of Medicine*, *340*(22), 1715-1722.

Lecossier, D., Bouchonnet, F., Clavel, F., & Hance, A. J. (2003). Hypermutation of HIV-1 DNA in the absence of the Vif protein. *Science*, *300*(5622), 1112-1112.

Le Douce V, Janossy A, Hallay H, et al. Achieving a cure for HIV infection: do we have reasons to be optimistic?. *J Antimicrob Chemother*. 2012;67(5):1063-1074.

Leid, M., Ishmael, J. E., Avram, D., Shepherd, D., Fraulob, V., & Dollé, P. (2004). CTIP1 and CTIP2 are differentially expressed during mouse embryogenesis. *Gene expression patterns : GEP*, *4*(6), 733–739.

Lessel, D., Gehbauer, C., Bramswig, N. C., Schluth-Bolard, C., Venkataramanappa, S., van Gassen, K., Hempel, M., Haack, T. B., Baresic, A., Genetti, C. A., Funari, M., Lessel, I., Kuhlmann, L., Simon, R., Liu, P., Denecke, J., Kuechler, A., de Kruijff, I., Shoukier, M., Lek, M., … Kubisch, C. (2018). BCL11B mutations in patients affected by a neurodevelopmental disorder with reduced type 2 innate lymphoid cells. *Brain : a journal of neurology*, *141*(8), 2299–2311.

Leth, S., Schleimann, M.H., Nissen, S.K., Højen, J.F., Olesen, R., Graversen, M.E., Jørgensen, S., Kjær, A.S., Denton, P.W., Mørk, A., et al. (2016). Combined effect of Vacc-4x, recombinant human granulocyte macrophage colony-stimulating factor vaccination, and romidepsin on the HIV-1 reservoir (REDUC): a single-arm, phase 1B/2A trial. Lancet HIV *3*, e463-472.

Li Y, Ndjango JB, Learn GH, Ramirez MA, Keele BF, Bibollet-Ruche F, Liu W, Easlick JL, Decker JM, Rudicell RS, Inogwabini BI, Ahuka-Mundeke S, Leendertz FH, Reynolds V, Muller MN, Chancellor RL, Rundus AS, Simmons N, Worobey M, Shaw GM, Peeters M, Sharp PM, Hahn BH. Eastern chimpanzees, but not bonobos, represent a simian immunodeficiency virus reservoir. J Virol. 2012 Oct;86(19):10776-91

Li, L., Dahiya, S., Kortagere, S., Aiamkitsumrit, B., Cunningham, D., Pirrone, V., Nonnemacher, M. R., & Wigdahl, B. (2012). Impact of Tat Genetic Variation on HIV-1 Disease. *Advances in virology*, *2012*, 123605.

Li, A., Chen, Y. S., Ping, X. L., Yang, X., Xiao, W., Yang, Y., Sun, H. Y., Zhu, Q., Baidya, P., Wang, X., Bhattarai, D. P., Zhao, Y. L., Sun, B. F., & Yang, Y. G. (2017). Cytoplasmic m^6A reader YTHDF3 promotes mRNA translation. *Cell research*, *27*(3), 444–447.

Liang, Xiaoping, Danilo R. Casimiro, William A. Schleif, Fubao Wang, Mary-Ellen Davies, Zhi-Qiang Zhang, Tong-Ming Fu, Adam C. Finnefrock, Larry Handt, Michael P. Citron, Gwendolyn Heidecker, Aimin Tang, Minchun Chen, KeithA. Wilson, Lori Gabryel ski, Michael McElhaugh, Anthony Carella, Cheryl Moyer, Lingyi Huang, Salvatore Vitelli, D eepa Patel, Jing Lin, Emilio A. Emini, John W. Shiver "Vectored Gag and Env but Not Tat Show Efficacy against Simian-Human Immunodeficiency Virus 89.6P Challenge in Mamu-A*01-Negative Rhesus Monkeys." *Journal of Virology* 79, no. 19 (October 1, 2005): 12321.

Liang S, Lutz CS. p54nrb is a component of the snRNP-free U1A (SF-A) complex that promotes pre-mRNA cleavage during polyadenylation. RNA 2006, 12:111–121.

Liu L, Xie N, Rennie P, Challis JR, Gleave M, Lye SJ, Dong X. Consensus PP1 binding motifs regulate transcriptional corepression and alternative RNA splicing activities of the steroid receptor coregulators, p54nrb and PSF. Mol Endocrinol 2012, 25:1197–1210.

Lowery,L.A., Rubin,J. and Sive,H. (2007) Whitesnake/sfpq is required for cell survival and neuronal development in the zebrafish. Dev. Dyn., 236, 1347–1357.

Lu, M., Blacklow, S. & Kim, P. A trimeric structural domain of the HIV-1 transmembrane glycoprotein. *Nat Struct Mol Biol* **2,** 1075–1082 (1995).

Lu, Jia; Kadakkuzha, Beena M.; Zhao, Liang; et al. (2011). "Dynamic Ensemble View of the Conformational Landscape of HIV-1 TAR RNA and Allosteric Recognition". *Biochemistry*. **50** (22): 5042–5057.

Luban, J., Bossolt, K. L., Franke, E. K., Kalpana, G. V., & Goff, S. P. (1993). Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B. *Cell*, *73*(6), 1067-1078.

Lukong KE, Huot ME, Richard S. BRK phosphorylates PSF promoting its cytoplasmic localization and cell cycle arrest. Cell Signal 2009, 21:1415–1422.

Maertens, G., Cherepanov, P., Pluymers, W., Busschots, K., De Clercq, E., Debyser, Z., & Engelborghs, Y. (2003). LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *Journal of Biological Chemistry*, *278*(35), 33528-33539.

Maison, C., Bailly, D., Roche, D., Montes de Oca, R., Probst, A. V., Vassias, I., Dingli, F., Lombard, B., Loew, D., Quivy, J. P., & Almouzni, G. (2011). SUMOylation promotes de novo targeting of HP1α to pericentric heterochromatin. *Nature genetics*, *43*(3), 220–227.

Maison, C., Bailly, D., Quivy, J. P., & Almouzni, G. (2016). The methyltransferase Suv39h1 links the SUMO pathway to HP1α marking at pericentric heterochromatin. *Nature communications*, *7*, 12224.

Mak, J., Khorchid, A., Cao, Q., Huang, Y., Lowy, I., Parniak, M. A., Prasad, V., R., Wainberg, M., A., & Kleiman, L. (1997). Effects of mutations in Pr160gag-pol upon tRNALys3 and Pr160gag-pol incorporation into HIV-1. *Journal of molecular biology*, *265*(4), 419-431.

Maldarelli, F., Martin, M., & Strebel, K. (1991). Identification of posttranscriptionally active inhibitory sequences in human immunodeficiency virus type 1 RNA: novel level of gene regulation. *Journal of Virology*, *65*(11), 5732-5743.

Mallon, P. W. (2007). Antiretroviral therapy-induced lipid alterations: in-vitro, animal and human studies. *Current Opinion in HIV and AIDS*, *2*(4), 282-292.

Mallon, P. W. (2014). Aging with HIV: osteoporosis and fractures. *Current Opinion in HIV and AIDS*, *9*(4), 428-435.

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*, *424*(6944), 99-103.

Marban, C., Redel, L., Suzanne, S., Van Lint, C., Lecestre, D., Chasserot-Golaz, S., Leid, M., Aunis, D., Schaeffer, E., & Rohr, O. (2005). COUP-TF interacting protein 2 represses the initial phase of HIV-1 gene transcription in human microglial cells. *Nucleic acids research*, *33*(7), 2318–2331.

Marshall, N. F., Peng, J., Xie, Z., & Price, D. H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. *Journal of Biological Chemistry*, *271*(43), 27176-27183.

Martin-Serrano, J., Zang, T., & Bieniasz, P. D. (2003). Role of ESCRT-I in retroviral budding. *Journal of virology*, *77*(8), 4794–4804.

Martinez-Picado, J., & Deeks, S. G. (2016). Persistent HIV-1 replication during antiretroviral therapy. *Current Opinion in HIV and AIDS*, *11*(4), 417.

Marx, Preston A., Phillip G. Alcabes, and Ernest Drucker. "Serial Human Passage of Simian Immunodeficiency Virus by Unsterile Injections and the Emergence of Epidemic Human Immunodeficiency Virus in Africa." Edited by W. Hamilton, R. A. Weiss, and S. Wain–Hobson. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 356, no. 1410 (June 29, 2001): 911–20.

McNamara, L. A., & Collins, K. L. (2011). Hematopoietic stem/precursor cells as HIV reservoirs. *Current opinion in HIV and AIDS*, *6*(1), 43–48.

Michels, A. A., Fraldi, A., Li, Q., Adamson, T. E., Bonnet, F., Nguyen, V. T., Sedore, S. C., Price, J. P., Price, D. H., Lania, L., & Bensaude, O. (2004). Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *The EMBO journal*, *23*(13), 2608–2619.

Milburn, J., Jones, R., & Levy, J. B. (2017). Renal effects of novel antiretroviral drugs. *Nephrology Dialysis Transplantation*, *32*(3), 434-439.

Miller, M. D., Farnet, C. M., & Bushman, F. D. (1997). Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. *Journal of virology*, *71*(7), 5382-5390.

Morozumi Y, Takizawa Y, Takaku M, Kurumizaka H. Human PSF binds to RAD51 and modulates its homologous-pairing and strand-exchange activities. Nucleic Acids Res 2009, 37:4296–4307.

Mothe, B., Moltó, J., Manzardo, C., Coll, J., Puertas, M.C., Martinez-Picado, J., Hanke, T., Clotet, B. and Brander, C., 2017, February. Viral control induced by HIVconsv vaccines & Romidepsin in early treated individuals. In *The Conference on Retroviruses and Opportunistic Infections*.

Mousseau, Guillaume, Sonia Mediouni, and Susana T. Valente. "Targeting HIV transcription: the quest for a functional cure." *The Future of HIV-1 Therapeutics*. Springer, Cham, 2015. 121-145.

Muniz, L., Kiss, T., & Egloff, S. (2012). Perturbations de la transcription liées à une dérégulation de P-TEFb : cancer, Sida et hypertrophie cardiaque [Misregulation of P-TEFb activity: pathological consequences]. *Medecine sciences : M/S*, *28*(2), 200–205.

Murray, A. J., Kwon, K. J., Farber, D. L., & Siliciano, R. F. (2016). The latent reservoir for HIV-1: how immunologic memory and clonal expansion contribute to HIV-1 persistence. *The Journal of Immunology*, *197*(2), 407-417.

Muthumani K, Choo AY, Zong WX, Madesh M, Hwang DS, Premkumar A, Thieu KP, Emmanuel J, Kumar S, Thompson CB, Weiner DB (February 2006). "The HIV-1 Vpr and glucocorticoid receptor complex is a gain-of-function interaction that prevents the nuclear localization of PARP-1". *Nature Cell Biology*. **8** (2): 170–9

Naganuma T, Nakagawa S, Tanigawa A, Sasaki YF, Goshima N, Hirose T. Alternative 3'-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J*. 2012;31(20):4020-4034.

Nakagawa S, Hirose T. Paraspeckle nuclear bodies—useful uselessness? Cell Mol Life Sci 2012, 69:3027–3036.

Nakayama J, Rice JC, Strahl BD, Allis CD, Grewal SI. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science*. 2001;292(5514):110-113.

Navia, B. A., Jordan, B. D., & Price, R. W. (1986). The AIDS dementia complex: I. Clinical features. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *19*(6), 517-524.

Ne, E., Palstra, R. J., & Mahmoudi, T. (2018). Transcription: insights from the HIV-1 promoter. In *International Review of Cell and Molecular Biology* (Vol. 335, pp. 191-243). Academic Press.

Neville, M., Stutz, F., Lee, L., Davis, L. I., & Rosbash, M. (1997). The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Current Biology*, *7*(10), 767-775.

Niu, Q.; Liu, Z.; Alamer, E.; Fan, X.; Chen, H.; Endsley, J.; Gelman, B.B.; Tian, B.; Kim, J.H.; Michael, N.L.; et al. Structure-guided drug design identifies a BRD4-selective small molecule that suppresses HIV. J. Clin. Investg. 2019, 129, 3361–3373.

Nostramo, R. T., & Hopper, A. K. (2019). Beyond rRNA and snRNA: tRNA as a 2'-O-methylation target for nucleolar and Cajal body box C/D RNPs. *Genes & development*, *33*(13-14), 739–740.

Olson, Erik D., William A. Cantara, and Karin Musier-Forsyth. "New structure sheds light on selective HIV-1 genomic RNA packaging." Viruses 7.8 (2015): 4826-4835.

Ong SE, Mittler G, Mann M. Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. Nat Methods 2004, 1:119–126.

Ono, A. (2010). Relationships between plasma membrane microdomains and HIV-1 assembly. *Biology of the Cell*, *102*(6), 335-350.

Osuna, C. E., Lim, S. Y., Kublin, J. L., Apps, R., Chen, E., Mota, T. M., Huang, S. H., Ren, Y., Bachtel, N. D., Tsibris, A. M., Ackerman, M. E., Jones, R. B., Nixon, D. F., & Whitney, J. B. (2018). Evidence that CD32a does not mark the HIV-1 latent reservoir. *Nature*, *561*(7723), E20–E28.

Pace, M., & O'Doherty, U. (2013). Hematopoietic stem cells and HIV infection.

Padilla-Parra, S., & Dustin, M. L. (2016). Actin dynamics and HIV-1 entry. *Trends in molecular medicine*, *22*(5), 354-356.

Paillart, J. C., Shehu-Xhilaga, M., Marquet, R., & Mak, J. (2004). Dimerization of retroviral RNA genomes: an inseparable pair. *Nature Reviews Microbiology*, *2*(6), 461-472.

Parada, C. A., & Roeder, R. G. (1996). Enhanced processivity of RNA polymerase II triggered by Tat-induced phosphorylation of its carboxy-terminal domain. *Nature*, *384*(6607), 375-378.

Passon DM, Lee M, Rackham O, Stanley WA, Sadowska A, Filipovska A, Fox AH, Bond CS. Structure of the heterodimer of human NONO and paraspeckle protein component 1 and analysis of its role in subnuclear body formation. Proc Natl Acad Sci USA 2012, 109:4846–4850.

Patton JG, Mayer SA, Tempst P, Nadal-Ginard B. Characterization and molecular cloning of polypyrimidine tract-binding protein: a component of a complex necessary for pre-mRNA splicing. Genes Dev 1991, 5:1237–1251.

Patton JG, Porro EB, Galceran J, Tempst P, Nadal-Ginard B. Cloning and characterization of PSF, a novel pre-mRNA splicing factor. Genes Dev 1993, 7:393–406

Pereyra F, Heckerman D, Carlson JM, Kadie C, Soghoian DZ, Karel D, et al. HIV control is mediated in part by CD8+ T-cell targeting of specific epitopes. J Virol (2014) 88:12937–48.

Perkins, N. D., Edwards, N. L., Duckett, C. S., Agranoff, A. B., Schmid, R. M., & Nabel, G. J. (1993). A cooperative interaction between NF-kappa B and Sp1 is required for HIV-1 enhancer activation. *The EMBO journal*, *12*(9), 3551-3558.

Permatasari, H.K., Nakahata, S., Ichikawa, T., and Morishita, K. (2017). BCL11B is frequently downregulated in HTLV-1-infected T-cells through Tax-mediated proteasomal degradation. Biochem. Biophys. Res. Commun. *490*, 1086–1092.

Perreau, M., Banga, R., and Pantaleo, G. (2017). Targeted Immune Interventions for an HIV-1 Cure. Trends Mol. Med. *23*, 945–961

Persaud, D., Gay, H., Ziemniak, C., Chen, Y.H., Piatak, M., Chun, T.-W., Strain, M., Richman, D., and Luzuriaga, K. (2013). Absence of detectable HIV-1 viremia after treatment cessation in an infant. N. Engl. J. Med. *369*, 1828–1835.

Peterlin, B. M., & Price, D. H. (2006). Controlling the elongation phase of transcription with P-TEFb. *Molecular cell*, *23*(3), 297-305.

Polymenidou, M., Lagier-Tourenne, C., Hutt, K. R., Bennett, C. F., Cleveland, D. W., & Yeo, G. W. (2012). Misregulated RNA processing in amyotrophic lateral sclerosis. *Brain research*, *1462*, 3–15.

Popov, S., Rexach, M., Zybarth, G., Reiling, N., Lee, M. A., Ratner, L., Bukrinsky, M. (1998). Viral protein R regulates nuclear import of the HIV-1 pre-integration complex. *The EMBO journal*, *17*(4), 909-917.

Prasanth, K. V., Prasanth, S. G., Xuan, Z., Hearn, S., Freier, S. M., Bennett, C. F., Zhang, M. Q., & Spector, D. L. (2005). Regulating gene expression through RNA nuclear retention. *Cell*, *123*(2), 249–263.

Punwani, D., Zhang, Y., Yu, J., Cowan, M. J., Rana, S., Kwan, A., Adhikari, A. N., Lizama, C. O., Mendelsohn, B. A., Fahl, S. P., Chellappan, A., Srinivasan, R., Brenner, S. E., Wiest, D. L., & Puck, J. M. (2016). Multisystem Anomalies in Severe Combined Immunodeficiency with Mutant BCL11B. *The New England journal of medicine*, *375*(22), 2165–2176.

Rajesh C, Baker DK, Pierce AJ, Pittman DL. The splicing-factor related protein SFPQ/PSF interacts with RAD51D and is necessary for homology-directed repair and sister chromatid cohesion. Nucleic Acids Res 2011, 39:132–145.

Rambaut, Posada, Crandall and E. C. Holmes, *Nat. Rev. Genet.*, 2004, **5**, 52–61

Raposo, G., Moore, M., Innes, D., Leijendekker, R., Leigh-Brown, A., Benaroch, P., & Geuze, H. (2002). Human macrophages accumulate HIV-1 particles in MHC II compartments. *Traffic*, *3*(10), 718-729.

Redel, L., Le Douce, V., Cherrier, T., Marban, C., Janossy, A., Aunis, D., Van Lint, C., Rohr, O., & Schwartz, C. (2010). HIV-1 regulation of latency in the monocyte-macrophage lineage and in CD4+ T lymphocytes. *Journal of leukocyte biology*, *87*(4), 575–588.

Rerks-Ngarm, S., Pitisuttithum, P., Nitayaphan, S., Kaewkungwal, J., Chiu, J., Paris, R., Premsri, N., Namwat, C., de Souza, M., Adams, E., Benenson, M., Gurunathan, S., Tartaglia, J., McNeil, J. G., Francis, D. P., Stablein, D., Birx, D. L., Chunsuttiwat, S., Khamboonruang, C., Thongcharoen, P., Kunasol P, Michael NL, Robb ML, Gilbert PB, Kim JH. MOPH-TAVEG Investigators (2009). Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *The New England journal of medicine*, *361*(23), 2209–2220

Ringrose, J.H.; Jeeninga, R.E.; Berkhout, B.; Speijer, D. Proteomic studies reveal coordinated changes in T-cell expression patterns upon infection with human immunodeficiency virus type 1. J. Virol. 2008, 82, 4320–4330

Robb, M. L., Rerks-Ngarm, S., Nitayaphan, S., Pitisuttithum, P., Kaewkungwal, J., Kunasol, P., Khamboonruang, C., Thongcharoen, P., Morgan, P., Benenson, M., Paris, R. M., Chiu, J., Adams, E., Francis, D., Gurunathan, S., Tartaglia, J., Gilbert, P., Stablein, D., Michael, N. L., & Kim, J. H. (2012). Risk behaviour and time as covariates for efficacy of the HIV vaccine regimen ALVAC-HIV (vCP1521) and AIDSVAX B/E: a post-hoc analysis of the Thai phase 3 efficacy trial RV 144. *The Lancet. Infectious diseases*, *12*(7), 531–537.

Rohr, O., Lecestre, D., Chasserot-Golaz, S., Marban, C., Avram, D., Aunis, D., Leid, M., and Schaeffer, E. (2003a). Recruitment of Tat to Heterochromatin Protein HP1 via Interaction with CTIP2 Inhibits Human Immunodeficiency Virus Type 1 Replication in Microglial Cells. J.Virol. *77*, 5415–5427.

Rojas-Araya, B., Ohlmann, T., & Soto-Rifo, R. (2015). Translational control of the HIV unspliced genomic RNA. *Viruses*, *7*(8), 4326-4351.

Rosonina E, Ip JY, Calarco JA, Bakowski MA, Emili A, McCracken S, Tucker P, Ingles CJ, Blencowe BJ. Role for PSF in mediating transcriptional activator-dependent stimulation of pre-mRNA processing in vivo. Mol Cell Biol 2005, 25:6734–6746

Sáez-Cirión, A., Bacchus, C., Hocqueloux, L., Avettand-Fenoel, V., Girault, I., Lecuroux, C., Potard, V., Versmisse, P., Melard, A., Prazuck, T., Descours, B., Guergnon, J., Viard, J. P., Boufassa, F., Lambotte, O., Goujard, C., Meyer, L., Costagliola, D., Venet, A., Pancino, G., … ANRS VISCONTI Study Group (2013). Post-treatment HIV-1 controllers with a long-term virological remission after the interruption of early initiated antiretroviral therapy ANRS VISCONTI Study. *PLoS pathogens*, *9*(3), e1003211.

Salton M, Lerenthal Y, Wang SY, Chen DJ, Shiloh Y. Involvement of Matrin 3 and SFPQ/NONO in the DNA damage response. Cell Cycle 2010, 9:1568–1576.

Sarracino, A., Gharu, L., Kula, A., Pasternak, A. O., Avettand-Fenoel, V., Rouzioux, C., Bardina M., De Wit S., Benkirane M., Berkhout B., Van Lint, C., Marcello A.,(2018). Posttranscriptional regulation of HIV-1 gene expression during replication and reactivation from latency by nuclear matrix protein MATR3. *MBio*, *9*(6).

Sasaki, Y.T., T. Ideue, M. Sano, T. Mituyama, and T. Hirose. 2009. MENε/β noncoding RNAs are essential for structural integrity of nuclear paraspeckles. Proc. Natl. Acad. Sci. USA. 106:2525–2530

Sattentau, Q. J., & Stevenson, M. (2016). Macrophages and HIV-1: an unhealthy constellation. *Cell host & microbe*, *19*(3), 304-310.

Schnittman, SM, MC Psallidopoulos, HC Lane, L Thompson, M Baseler, F Massari, CH Fox, NP Salzman, and AS Fauci. "The Reservoir for HIV-1 in Human Peripheral Blood Is a T Cell That Maintains Expression of CD4." *Science* 245, no. 4915 (July 21, 1989): 305.

Schö nichen, A., Bigalke, J. M., Urbanke, C., Grzesiek, S., Dames, S. A., & Geyer, M. (2010). A flexible bipartite coiled coil structure is required for the interaction of Hexim1 with the P-TEFB subunit cyclin T1. *Biochemistry*, *49*(14), 3083-3091.

Schröder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R., & Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*, *110*(4), 521-529.

Shultz, L.D., Ishikawa, F. & Greiner, D.L. Humanized mice in translational biomedical research. *Nat. Rev. Immunol.* **7**, 118–130 (2007).

Schwartz, J. C., Wang, X., Podell, E. R., & Cech, T. R. (2013). RNA seeds higher-order assembly of FUS protein. *Cell reports*, *5*(4), 918–925.

Sewald, X., Ladinsky, M. S., Uchil, P. D., Beloor, J., Pi, R., Herrmann, C., Motamedi N., Murooka T., Brehm M., Greiner D., Mempel T., Bjorkman P., Kumar P., Mothes W., Shultz, L. D. (2015). Retroviruses use CD169-mediated trans-infection of permissive lymphocytes to establish infection. *Science*, *350*(6260), 563-567.

Shan L, Deng K, Shroff NS, Durand CM, Rabi SA, Yang HC, et al. Stimulation of HIV-1-specific cytolytic T lymphocytes facilitates elimination of latent viral reservoir after virus reactivation. *Immunity* (2012) 36:491–501.

Shav-Tal Y, Zipori D. PSF and p54(nrb)/NonO— multi-functional nuclear proteins. FEBS Lett 2002, 531:109–114.

Shi, H., Wang, X., Lu, Z., Zhao, B. S., Ma, H., Hsu, P. J., Liu, C., & He, C. (2017). YTHDF3 facilitates translation and decay of N^6-methyladenosine-modified RNA. *Cell research*, *27*(3), 315–328.

Silva, Zacarias; Oliveira, Inês; Andersen, Andreas; Dias, Francisco; Rodrigues, Amabélia; Holmgren, Birgitta; Andersson, Sören; Aaby, Peter Changes in prevalence and incidence of HIV-1, HIV-2 and dual infections in urban areas of Bissau, Guinea-Bissau: is HIV-2 disappearing?, AIDS: June 19th, 2008 - Volume 22 - Issue 10 - p 1195-1202

Simon JH, Gaddis NC, Fouchier RA, Malim MH. Evidence for a newly discovered cellular anti-HIV-1 phenotype. *Nat Med*. 1998;4(12):1397-1400.

Solbak, S. M. Ø., Reksten, T. R., Hahn, F., Wray, V., Henklein, P., Henklein, P., Halskau Ø., chubert U., Fossen, T. (2013). HIV-1 p6—A structured to flexible multifunctional membrane-interacting protein. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, *1828*(2), 816-823.

Souquere, S., G. Beauclair, F. Harper, A. Fox, and G. Pierron. 2010. Highly ordered spatial organization of the structural long noncoding NEAT1 RNAs within paraspeckle nuclear bodies. *Mol. Biol. Cell.* 21:4020–4027.

Snijders AP, Hung ML, Wilson SA, Dickman MJ. Analysis of arginine and lysine methylation utilizing peptide separations at neutral pH and electron transfer dissociation mass spectrometry. J Am Soc Mass Spectrom 2010, 21:88–96.

Stauber, R. H., & Pavlakis, G. N. (1998). Intracellular trafficking and interactions of the HIV-1 Tat protein. *Virology*, *252*(1), 126-136.

Steinman RM (1991) The dendritic cell system and its role in immunogenicity. Annu Rev Immunol 9: 271–296.

Sumanasekera C, Kelemen O, Beullens M, Aubol BE, Adams JA, Sunkara M, Morris A, Bollen M, Andreadis A, Stamm S. C6 pyridinium ceramide influences alternative pre-mRNA splicing by inhibiting protein phosphatase-1. Nucleic Acids Res 2012, 40:4025–4039.

Sunwoo, H., M.E. Dinger, J.E. Wilusz, P.P. Amaral, J.S. Mattick, and D.L. Spector. 2009. MEN ε/β nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. Genome Res. 19:347–359

Suzuki, Y., & Craigie, R. (2007). The road to chromatin—nuclear entry of retroviruses. *Nature Reviews Microbiology*, *5*(3), 187-196.

Suzuki, K.; Juelich, T.; Lim, H.; Ishida, T.; Watanebe, T.; Cooper, D.A.; Rao, S.; Kelleher, A.D. Closed chromatin architecture is induced by an RNA duplex targeting the HIV-1 promoter region. J. Biol. Chem. 2008, 283, 23353–23363.

Takehisa, Jun, Matthias H. Kraus, Ahidjo Ayouba, Elizabeth Bailes, Fran Van Heuverswyn, Julie M. Decker, Yingying Li, et al. "Origin and Biology of Simian Immunodeficiency Virus in Wild-Living Western Gorillas." *Journal of Virology* 83, no. 4 (February 15, 2009): 1635–48.

Tang, C., Louis, J. M., Aniana, A., Suh, J. Y., & Clore, G. M. (2008). Visualizing transient events in amino-terminal autoprocessing of HIV-1 protease. *Nature*, *455*(7213), 693-696.

Thorlund, K., Horwitz, M. S., Fife, B. T., Lester, R., & Cameron, D. W. (2017). Landscape review of current HIV 'kick and kill'cure research-some kicking, not enough killing. *BMC infectious diseases*, *17*(1), 595.

Tyagi, M., Rusnati, M., Presta, M., & Giacca, M. (2001). Internalization of HIV-1 tat requires cell surface heparan sulfate proteoglycans. *Journal of Biological Chemistry*, *276*(5), 3254-3261.

van Kooyk Y, Geijtenbeek TB (2003) DC-SIGN: escape mechanism for pathogens. Nat Rev Immunol 3: 697–709.

Van Lint, C., Emiliani, S., Ott, M., & Verdin, E. (1996). Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation. *The EMBO journal*, *15*(5), 1112–1120.

Van Lint, C., Bouchat, S., & Marcello, A. (2013). HIV-1 transcription and latency: an update. *Retrovirology*, *10*(1), 1-38.

Van Maele, B., & Debyser, Z. (2005). HIV-1 integration: an interplay between HIV-1 integrase, cellular and viral proteins. *AIDS reviews*, *7*(1), 26-43.

Van Maele, B., Busschots, K., Vandekerckhove, L., Christ, F., & Debyser, Z. (2006). Cellular co-factors of HIV-1 integration. *Trends in biochemical sciences*, *31*(2), 98-105.

van Opijnen, T., Jeeninga, R. E., Boerlijst, M. C., Pollakis, G. P., Zetterberg, V., Salminen, M., & Berkhout, B. (2004). Human immunodeficiency virus type 1 subtypes have a distinct long

terminal repeat that determines the replication rate in a host-cell-specific manner. *Journal of virology*, *78*(7), 3675-3683.

Vasudevan, A., Smits, S..J., Höppner, A., Häussinger, D., Koenig, B. W., & Münk, C. (2013). Structural features of antiviral DNA cytidine deaminases, *Biological Chemistry*, *394*(11), 1357-1370.

Vázquez, N., Greenwell-Wild, T., Marinos, N.J., Swaim, W.D., Nares, S., Ott, D.E., Schubert, U., Henklein, P., Orenstein, J.M., Sporn, M.B., et al. (2005). Human immunodeficiency virus type 1-induced macrophage gene expression includes the p21 gene, a target for viral regulation. J. Virol. *79*, 4479–4491.

Verdin, E., Paras Jr, P., & Van Lint, C. (1993). Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. *The EMBO journal*, *12*(8), 3249-3259.

Vrolijk, M. M., Ooms, M., Harwig, A., Das, A. T., & Berkhout, B. (2008). Destabilization of the TAR hairpin affects the structure and function of the HIV-1 leader RNA. *Nucleic acids research*, *36*(13), 4352-4363.

Virus Taxonomy: 2018b Release". *International Committee on Taxonomy of Viruses (ICTV)*. March 2019. Retrieved 16 March 2019.

Wakabayashi Y, Watanabe H, Inoue J, Takeda N, Sakata J, Mishima Y, Hitomi J, Yamamoto T, Utsuyama M, Niwa O, Aizawa S, Kominami R (2003) Bcl11b is required for differentiation and survival of alphabeta T lymphocytes. ***Nat Immunol*** **4**: 533–539

Wallet, C., De Rovere, M., Van Assche, J., Daouad, F., De Wit, S., Gautier, V., Mallon, P., Marcello, A., Van Lint, C., Rohr, O., & Schwartz, C. (2019). Microglial Cells: The Main HIV-1 Reservoir in the Brain. *Frontiers in cellular and infection microbiology*, *9*, 362. https://doi.org/10.3389/fcimb.2019.00362

Wan, Z.; Chen, X. Triptolide inhibits human immunodeficiency virus type 1 replication by promoting proteasomal degradation of Tat protein. Retrovirology 2014, 11, 88

Wang G, Cui Y, Zhang G, Garen A, Song X. Regulation of proto-oncogene transcription, cell proliferation, and tumorigenesis in mice by PSF protein and a VL30 noncoding RNA. Proc Natl Acad Sci USA 2009, 106:16794–16798

Wang, Z., Fan, P., Zhao, Y., Zhang, S., Lu, J., Xie, W., Jiang, Y., Lei, F., Xu, N., & Zhang, Y. (2017). NEAT1 modulates herpes simplex virus-1 replication by regulating viral gene transcription. *Cellular and molecular life sciences : CMLS*, *74*(6), 1117–1131.

Wang X, Arai S, Song X, et al. Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature*. 2008;454(7200):126-130.

Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G., Ren, B., Pan, T., & He, C. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature*, *505*(7481)

Wang, X., Zhao, B. S., Roundtree, I. A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., & He, C. (2015). N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell*, *161*(6), 1388–1399.

Weiss, R. A. (2002). HIV receptors and cellular tropism. *IUBMB life*, *53*(4-5), 201-205.

Wensing, A. M., van Maarseveen, N. M., & Nijhuis, M. (2010). Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance. *Antiviral research*, *85*(1), 59–74. https://doi.org/10.1016/j.antiviral.2009.10.003

West, J.A., M. Mito, S. Kurosaka, T. Takumi, C. Tanegashima, T. Chujo, K. Yanaka, R.E. Kingston, T. Hirose, C. Bond, et al. 2016. Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization. *J. Cell Biol.*

Wilk, T., Gowen, B. E., & Fuller, S. D. (1999). Actin associates with the nucleocapsid domain of the Gag polyprotein in the human immunodeficiency virus (HIV-1). *J. Virol*, *73*, 1931-1940.

Wilson, Nancy A., Jason Reed, Gnankang S. Napoe, Shari Piaskowski, Andy Szymanski, Jessica Furlott, Edna J. Gonzalez Levi J. Yant, Nicholas J. Maness, Gemma E. May, Taeko Soma, Matthew R. Reynolds, Eva Rakasz, Richard Rudersdorf, Adrian B. McDermott, David H. O'Connor, Thomas C. Friedrich, David B. Allison, Amit Patki, Louis J. Picker, Dennis R. Burton, Jing Lin, Lingyi Huang, Deepa Patel, Gwendolyn Heindecker, Jiang Fan, Michael Citron, Melanie Horton, Fubao Wang, Xiaoping Liang, John W. Shiver, Danilo R. Casimiro, David I. Watkins. "Vaccine-Induced Cellular Immune Responses Reduce Plasma Viral Concentrations after Repeated Low-Dose Challenge with Pathogenic Simian Immunodeficiency Virus SIVmac239." *Journal of Virology* 80, no. 12 (June 15, 2006): 5875.

Winkler, R., Gillis, E., Lasman, L., Safra, M., Geula, S., Soyris, C., Nachshon, A., Tai-Schmiedel, J., Friedman, N., Le-Trilling, V., Trilling, M., Mandelboim, M., Hanna, J. H., Schwartz, S., & Stern-Ginossar, N. (2019). m^6A modification controls the innate immune response to infection by targeting type I interferons. *Nature immunology*, *20*(2), 173–182

Wherry EJ, Ahmed R. Memory CD8 T-cell differentiation during viral infection. *J Virol* (2004) 78:5535–45. doi:10.1128/JVI.78.11.5535-5545.2004

Whitcomb, Jeannette M., Wei Huang, Signe Fransen, Kay Limoli, Jonathan Toma, Terri Wrin, Colombe Chappey, Linda D. B. Kiss, Ellen E. Paxinos, and Christos J. Petropoulos. "Development and Characterization of a Novel Single-Cycle Recombinant-Virus Assay To Determine Human Immunodeficiency Virus Type 1 Coreceptor Tropism." *Antimicrobial Agents and Chemotherapy* 51, no. 2 (February 1, 2007): 566.

Wolf D, Witte V, Laffert B, Laffert B, Blume K, Stromer K, Trapp S, d'Aloja P, Schurmann A, Baur A,. "HIV-1 Nef associated PAK and PI3-kinases stimulate Akt-independent Bad-phosphorylation to induce anti-apoptotic signals. *Nat Med*. 2001;7(11):1217-1224.

Workman, J. L., & Kingston, R. E. (1998). Alteration of nucleosome structure as a mechanism of transcriptional regulation.

Wu, Yuntao (2004). "HIV-1 gene expression: Lessons from provirus and non-integrated DNA". *Retrovirology*. **1**: 13. doi:10.1186/1742-4690-1-13.

Yadav, Tejas, Jean-Pierre Quivy, and Geneviève Almouzni. “Chromatin Plasticity: A Versatile Landscape That Underlies Cell Fate and Identity.” *Science* 361, no. 6409 (September 28, 2018): 1332

Yang, L., Embree, L. J., & Hickstein, D. D. (2000). TLS-ERG leukemia fusion protein inhibits RNA splicing mediated by serine-arginine proteins. *Molecular and cellular biology*, *20*(10), 3345–3354.

Yarosh CA, Iacona JR, Lutz CS, Lynch KW. PSF: nuclear busy-body or nuclear facilitator?. *Wiley Interdiscip Rev RNA*. 2015;6(4):351-367. doi:10.1002/wrna.1280

Yukl, S.A., Boritz, E., Busch, M., Bentsen, C., Chun, T.-W., Douek, D., Eisele, E., Haase, A., Ho, Y.-C., Hütter, G., et al. (2013b). Challenges in detecting HIV persistence during potentially curative interventions: a study of the Berlin patient. PLoS Pathog. *9*, e1003347

Zaccara, S., & Jaffrey, S. R. (2020). A Unified Model for the Function of YTHDF Proteins in Regulating m6A-Modified mRNA. *Cell*.

Zack, J. A., Arrigo, S. J., Weitsman, S. R., Go, A. S., Haislip, A., & Chen, I. S. (1990). HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. *Cell*, *61*(2), 213-222.

Zhang, Q., Chen, C. Y., Yedavalli, V. S., & Jeang, K. T. (2013). NEAT1 long noncoding RNA and paraspeckle bodies modulate HIV-1 posttranscriptional expression. *mBio*, *4*(1), e00596-12.

Zhang, X., Zhou, Y., Mehta, K. R., Danila, D. C., Scolavino, S., Johnson, S. R., & Klibanski, A. (2003). A pituitary-derived MEG3 isoform functions as a growth suppressor in tumor cells. *The Journal of clinical endocrinology and metabolism*, *88*(11), 5119–5126.

Zhong N, Kim CY, Rizzu P, Geula C, Porter DR, Pothos EN, Squitieri F, Heutink P, Xu J. DJ-1 transcriptionally up-regulates the human tyrosine hydroxylase by inhibiting the sumoylation of pyrimidine tract-binding protein-associated splicing factor. J Biol Chem 2006, 281:20940–20948.

Zolotukhin AS, Michalowski D, Bear J, Smulevitch SV, Traish AM, Peng R, Patton J, Shatsky IN, Felber BK. PSF acts through the human immunodeficiency virus type 1 mRNA instability elements to regulate virus expression. Mol Cell Biol 2003, 23:6618–6630.

12. Résumé en français

Introduction

Depuis sa découverte au début des années 1980, le VIH-1 et sa maladie associée, le SIDA, ont fait plus de 34 millions de morts. Le dernier rapport de l'OMS en 2018 estime qu'un nombre compris entre 31,1 et 41,9 millions de personnes sont actuellement infectées par le virus VIH-1. Ce virus fait partie de la famille des Retroviridae, sous-classe Orthoretrovirinae et est classé dans les genres Lentivirus (Virus Taxonomy 2019). La structure et la morphologie du VIH-1 ont été décrites pour la première fois à la fin des années 1980 comme un virus à peu près sphérique avec un diamètre de 120 nm à 140 nm en moyenne, et avec des projections de protéines saillantes incluses dans l'enveloppe virale. Le noyau, visible par microscopie électronique, est en forme de cône.

Le virus VIH-1 est responsable de la maladie connue comme Syndrome d'Immunodéficience Acquise (SIDA) qui est causée par la lente décomposition de la population de lymphocytes CD4 +, la principale cible de l'infection du virus, jusqu'au point où l'organisme est incapable de se protéger des autres agents pathogènes opportunistes qui, s'ils ne sont pas traités, conduiront irrévocablement à la mort du patient.

Le plus grand succès dans la lutte contre le VIH-1 a été l'introduction de la thérapie antirétrovirale combinée, un schéma pharmacologique composé de plusieurs médicaments, chacun étant spécifiquement conçu pour entraver et bloquer les étapes essentielles du cycle de réplication du VIH-1. L'introduction du cART a été accueillie avec beaucoup d'enthousiasme dans la communauté médicale et scientifique, avec une conviction optimiste d'avoir enfin trouvé le remède à la maladie. Dans les années 80 et 90, être séropositif pour le VIH-1 équivaut à une condamnation à mort, de nos jours avec les médicaments antirétroviraux actuellement développés, un patient infecté peut espérer une espérance de vie similaire à une personne en bonne santé. En outre, un patient séropositif pour le VIH-1, sous un régime de thérapie antirétrovirale, a son niveau viral réduit à un point où il est à peine considéré comme infectieux et sa virémie ne peut être détectée qu'avec les outils les plus précis et les plus sensibles disponibles; néanmoins le traitement doit être envisagé tout au long de la vie car aucun remède n'a encore été trouvé.

Comme son nom l'indique, cette thérapie ne fait pas appel à une seule classe de médicaments,

mais à une combinaison de ceux-ci. Actuellement, cinq classes de médicaments sont utilisées, chacune d'entre elles ciblant une étape spécifique et cruciale du cycle de réplication du VIH-1:

- Les inhibiteurs de fusion (ou inhibiteurs d'entrée) empêchent le virus de fusionner avec la membrane cellulaire et de libérer son noyau dans le cytoplasme. Le maraviroc et l'enfuvirtide sont actuellement les deux médicaments utilisés, le premier ciblant le co-récepteur CCR5, le second bloquant l'entrée en liant la protéine virale gp41 et en formant un faisceau d'hétéro-six hélices inactif
- Les inhibiteurs nucléosidiques / nucléotidiques de la transcriptase inverse (INTI) tels que la zidovudine, l'abacavir, la lamivudine, l'emtricitabine et le ténofovir sont des terminateurs de chaîne. Une fois incorporés dans le nouveau brin d'ADN naissant, ils empêchent d'autres nucléosides d'être également incorporés dans la chaîne d'ADN en raison de l'absence d'un groupe 3 'OH
- Les inhibiteurs non nucléosidiques de la transcriptase inverse (INNTI) tels que la névirapine, l'éfavirenz, l'étravirine et la rilpivirine affectent la manipulation du substrat (nucléotides) par la transcriptase inverse en se liant à un site allostérique de l'enzyme près du site actif
- Les inhibiteurs de l'intégrase tels que le raltégravir, l'elvitégravir et le dolutégravir inhibent l'intégrase de la protéine virale.
- Les inhibiteurs de protéase tels que le lopinavir, l'indinavir, le nelfinavir, l'amprénavir, le ritonavir, le darunavir et l'atazanavir empêchent la formation de virions matures en bloquant l'action du protéasome viral. Ils bloquent notamment le clivage des protéines précurseurs gag et gag / pol

Malheureusement, le virus s'est révélé plus résilient que prévu, et bien que la thérapie antirétrovirale hautement active (HAART) désormais évoluée puisse accorder au patient séropositif une espérance de vie similaire à celle d'une personne en bonne santé, elle n'a pas été en mesure d'éradiquer complètement le virus de l'organisme. Cette petite population de virus qui résiste au ART est presque indétectable et a une faible capacité à propager la maladie, mais si le traitement est interrompu, la virémie rebondit aux niveaux pré-traitement.

Pour mieux comprendre cette sécurité intégrée du virus, nombreux groupes de recherche ont commencé à étudier la latence du VIH-1. Ceci décrit les cellules infectées, mais le virus intégré dans leur génome n'est pas productif et il est considéré comme silencieux. Cette caractéristique le rend réfractaire au jeu actionné par le système immunitaire et résistant à l’ART.

En raison de ces populations de cellules qui subsistent après un traitement ART complet, les communautés scientifiques ont commencé à chercher à développer différentes stratégies pour résoudre ce problème. Deux d'entre elles sont actuellement en cours de développement sous les noms de stratégie Shock-and-Kill et Block-and-Lock. Pour le Shock-and-Kill, la première cible, dans la partie «choc» de la thérapie, différents composés sont actuellement à l'étude comme la prostratine et la bryostatine qui sont des agonistes de la protéine kinase C (PKC) et conduisent à l'activation du NF-kB et provoquent une forte activation de la transcription virale. Les agonistes du récepteur TRL ont montré une bonne efficacité dans la réactivation

l'expression du gène viral avec la réactivation de suivi de l'immunité spécifique chez les patients sous cART.

Des agents pour remodeler l'environnement de la chromatine ont également été testés et parmi eux les inhibiteurs des HDAC sont les plus étudiés à la fois in vivo et in vitro.

L'étape de clairance ou «kill» est actionnée par le système immunitaire porté par les cellules NK et les cellules T CD8 + spécifiques du VIH et par l'effet cytopathique viral. Des études récentes ont montré comment l'utilisation d'un seul composé en tant que LRA a montré des effets sous-optimaux dans la réactivation des cellules latentes et une combinaison de LRA est nécessaire.De plus, un traitement prolongé par cART a pour effet secondaire de réduire le nombre de CD8 + spécifiques du VIH. Les cellules T et introduit la nécessité de stimuler également la population de ces cellules pour tuer efficacement les cellules réactivées. La stratégie actuelle a alors évolué pour inclure une stimulation des cellules T CD8 + qui peuvent détecter et tuer avec une sensibilité étonnante et peuvent également former des populations «mémoire» à vie longue capables d'agir rapidement en cas de rencontre future avec le virus. Pendant la phase aiguë de l'infection, l'émergence de ces cellules CD8 + coïncide avec la baisse de la charge virale. Ces cellules ciblent le protéome du VIH de telle sorte que le virus ne peut pas s'échapper sans évoluer avec un coût de remise en forme.

Une combinaison de Romidepsin avec un vaccin spécifique conçu pour stimuler les lymphocytes T CD8 + a rapporté, dans deux études différentes, un retard significatif du rebond viral après interruption du cART et une réduction de la taille du réservoir avec un niveau de virémie inférieur au niveau de détection.

En outre, l'utilisation d'anticorps neutralisants spécifiques est à l'étude et certains résultats intéressants ont été obtenus chez des souris humanisées traitées avec Varinostat qui ont montré un retard dans le rebond viral après interruption du cART.

Parallèlement à la stratégie «Shock and Kill», une nouvelle approche est actuellement testée dans le but d'arrêter complètement la réplication dans les cellules infectées de façon latente.

Comme nous l'avons décrit précédemment, les cellules infectées de manière latente présentent un niveau extrêmement faible de réplication virale, ce qui les rend responsables du rebond viral une fois le régime cART interrompu. La stratégie «Block and Lock» vise directement à utiliser des composés qui suppriment complètement cette réplication basale puis les verrouillent dans cet état. De nombreuses protéines virales et cellulaires sont impliquées dans la transcription et le silençage du VIH, et représentent donc des cibles potentielles pour les futures approches block-and-lock.

Notre groupe de recherche ainsi que mon projet de recherche, se concentrent sur une meilleure compréhension de la latence du VIH-1 dans l'un des types de cellules communément infectées par le virus, les cellules microgliales. Ces cellules sont les principaux réservoirs viraux au sein du système nerveux central, et en étant isolées dans ce sanctuaire anatomique, protégé par la barrière hémato-encéphalique, elles constituent un problème qui doit être abordé dans la voie de l'éradication complète du virus d'un organisme infecté.

Au cours des dernières années, notre groupe s'est concentré sur une protéine spécifique, CTIP2, et nous avons pu démontrer comment cette protéine joue un rôle crucial dans l'établissement et la persistance de la latence du VIH-1 dans les cellules microgliales.
Le facteur transcriptionnel Bcl11b (lymphome à cellules B / leucémie 11b) également connu sous le nom de CTIP2 (COUP-TF interacting protein 2) agit comme un répresseur transcriptionnel et il est principalement exprimé dans le cerveau et le système immunitaire.
CTIP2 a différentes fonctions. Nécessaire pour un développement correct du SNC, le CTIP2 joue un rôle clé dans le développement des projections axonales des motoneurones cortico-spinaux vers la moelle épinière et a également été lié à la formation de tumeurs cérébrales. Le CTIP2 est également nécessaire au développement du système immunitaire où il joue un rôle crucial dans la différenciation des thymocytes. De plus, le CTIP2 a récemment été associé pour jouer un rôle à la fois dans l'odontogenèse et dans la formation de la barrière protectrice perméable du derme
Un autre rôle dans la répression de la transcription qui a été lié à CTIP2 est sa capacité à saisir sous une forme inactive le facteur d'élongation P-TEFb. P-TEFb est un hétérodimère composé de la cycline T et de la kinase CDK9; ces deux protéines ont pour fonction de phosphoryler le domaine carboxy-terminal de l'ARN polymérase II lorsqu'elle en est aux premiers stades de transcription. Cette phosphorylation active complètement la polymérase permettant l'allongement des transcrits. La quantité de P-TEFb disponible pour réactiver l'ARNpolII est

contrôlée par un équilibre entre un complexe actif «libre» et un complexe inactif lié à l'ARN non codant 7sk. Ce complexe inactif comprend CTIP2 et HEXIM1 et il est dans un état catalytiquement inactif. L'équilibre entre la forme inactive et active du P-TEFb est sous contrôle cellulaire serré pour s'adapter précisément au besoin de la cellule au moment de la date. La perturbation de cet équilibre est souvent associée à l'apparition d'une maladie. Notre laboratoire a montré le rôle du CTIP2 dans la régulation de ce complexe. Une analyse globale de l'expression génique lors de la surexpression de CTIP2 a montré que des grappes significatives de gènes sensibles au PTEF-B étaient corégulées par CTIP2. De plus, nous avons démontré que le CTIP2 a un effet négatif direct sur la sous-unité catalytique du PTEF-B: le CDK9. Dans le contexte de la latence virale, notre laboratoire a décrit comment le complexe inactif P-TEFb est recruté sur le promoteur du VIH-1 grâce à l'interaction mutuelle de CTIP2 avec la protéine HMGA1 qui lie les facteurs de transcription sur le promoteur viral. Comme je l'ai mentionné précédemment, ce complexe est ciblé par Tat, dans la phase dépendante de Tat de la transcription. Tat est capable d'augmenter la partie active du complexe P-TEFb pour permettre la transcription des gènes viraux. On peut montrer alors comment le CTIP2 est un acteur majeur dans l'établissement de la latence virale, en créant un environnement hétérochromatine mais aussi en maintenant la latence en inactivant l'activité du complexe P-TEFb.

Nous avons démontré comment le CTIP2 joue un rôle crucial dans le remodelage de l'environnement de la chromatine autour du promoteur proviral VIH-1 intégré, en le faisant passer d'un état d'euchromatine ouvert, facile d'accès et de transcription à un état d'hétérochromatine emballé, réfractaire à l'accès des facteurs de transcription. CTIP2 le fait en recrutant et en interagissant avec différentes protéines qui agissent comme des enzymes de remodelage de la chromatine, et en séquestrant dans les états inactifs d'autres facteurs de transcription tels que le P-TEFb, comme indiqué sur la Figure 1. Pour cette raison, la relation entre le CTIP2 et le VIH-1 était et doit être examinée plus avant. Pour commencer, les patients infectés par le VIH-1 montrent une quantité élevée de CTIP2 circulant dans le liquide céphalo-rachidien (LCR) et l'analyse post-mortem a montré cette expression plus élevée également dans les cellules du SNC telles que les astrocytes et les microgliales. Cette augmentation de l'expression est directement corrélée à l'état de latence dans les cellules infectées du SNC par rapport aux patients asymptomatiques ou aux patients présentant des symptômes corrélés à l'inflammation du cerveau (encéphalite) causée par l'infection par le VIH-1. Cette augmentation des taux de CTIP2 s'accompagne également d'une augmentation des taux de HDAC1, HP1α et

d'une dérégulation des gènes pro-inflammatoires tels que le TNF-α et l'IL-6, qui, lorsqu'ils sont en fait régulés à la hausse, offrent une explication aux lésions cérébrales chez les patients. infecté par le VIH-1.

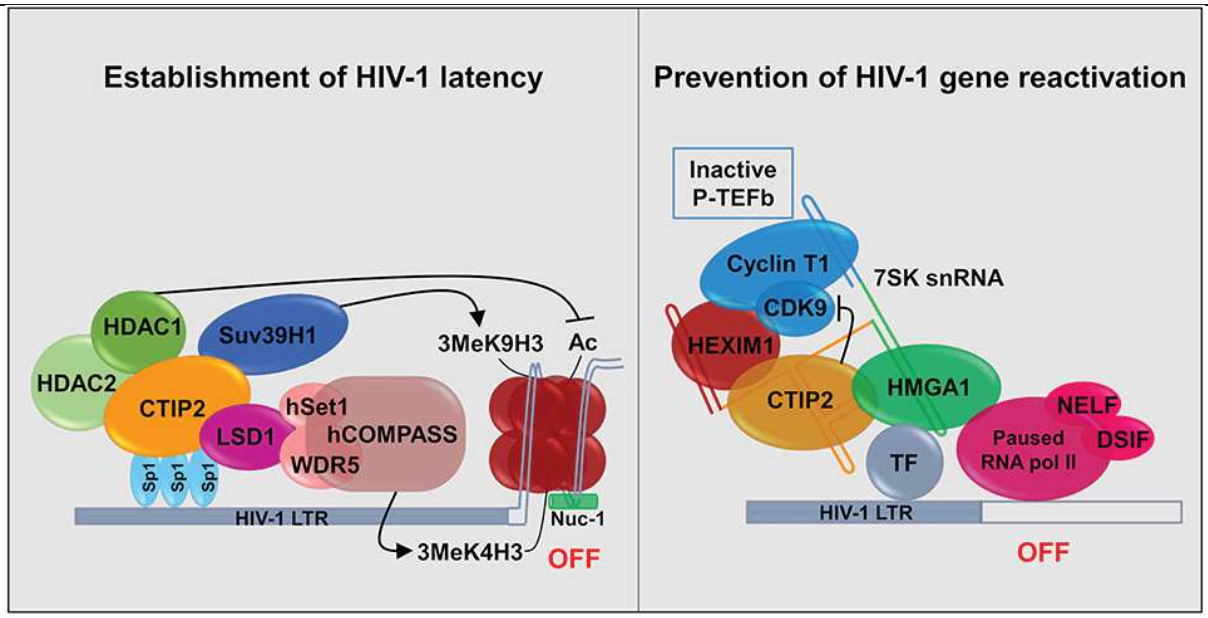


Figure 1: : **Rôles de CTIP2 dans l'établissement et la persistance de la latence du VIH-1 dans les cellules microgliales.** A gauche, la protéine CTIP2 est impliqué dans l'établissement de la latence du VIH-1 en recrutant un complexe modifiant la chromatine au niveau du promoteur viral. CTIP2 est ancré aux sites Sp1 sur le promoteur viral et agit comme un échafaudage pour le recrutement de plusieurs protéines modifiant la chromatine. À droite, CTIP2 empêche le redémarrage de la transcription en séquestrant le facteur d'élongation pTEF-B dans un complexe inactif où 7SK snRNA agit comme un complexe d'échafaudage et est recruté au promoteur via HMGA1 par interaction avec des facteurs de transcription (TF). La transcription est inhibée à deux niveaux. Premièrement, CTIP2 inhibe l'activité kinase de CDK9, une sous-unité pTEF-B qui empêche l'activation de l'ARN pol II en inhibant la phosphorylation de l'ARN pol II et des deux sous-unités NELF et DSIF. Deuxièmement, la persistance du pTEF-B inactif au niveau du promoteur empêche tout recrutement de pTEF-B actif. (Wallet et al., 2019)

Des recherches ultérieures ont été menées pour analyser différentes interactions de CTIP2 avec d'autres protéines proches de ces complexes ainsi que l'interaction de CTIP2 avec des protéines virales telles que Vpr pour mieux comprendre comment le virus lui-même peut ou non réguler l'expression et la présence de CTIP2, pour contrôler le état de latence de la cellule.

Objectifs et Résultats

Avec notre expérience précédente sur la façon dont le CTIP2 joue un rôle crucial dans l'établissement et la persistance de la latence du VIH-1 dans les cellules microgliales, nous avons consacré nos études à un projet plus large dans le but de répertorier et de décrire tous les partenaires du CTIP2 qui n'étaient pas déjà découverts, et d'évaluer leur fonction en ce qui concerne la latence du VIH-1. Ce projet était divisé en deux parties, l'une dans le but de lister toutes les protéines interagissant avec le CTIP2 et l'autre avec le même objectif concernant les ARN. Le projet d'interaction CTIP2-protéine a été lancé par un ancien étudiant de notre équipe, qui a réalisé une spectrométrie de masse aprè immunoprecipitation de CTIP2. Ces résultats ont ensuite été analysés et ont produit une liste de nouvelles protéines qui n'ont jamais été décrites comme des partenaires de CTIP2. Nous avons choisi d'étudier un groupe de facteurs cellulaires ayant été précédemment décrit comme des régulateurs du VIH-1 et/ou de l'expression des ARN cellulaires (Figure 2).

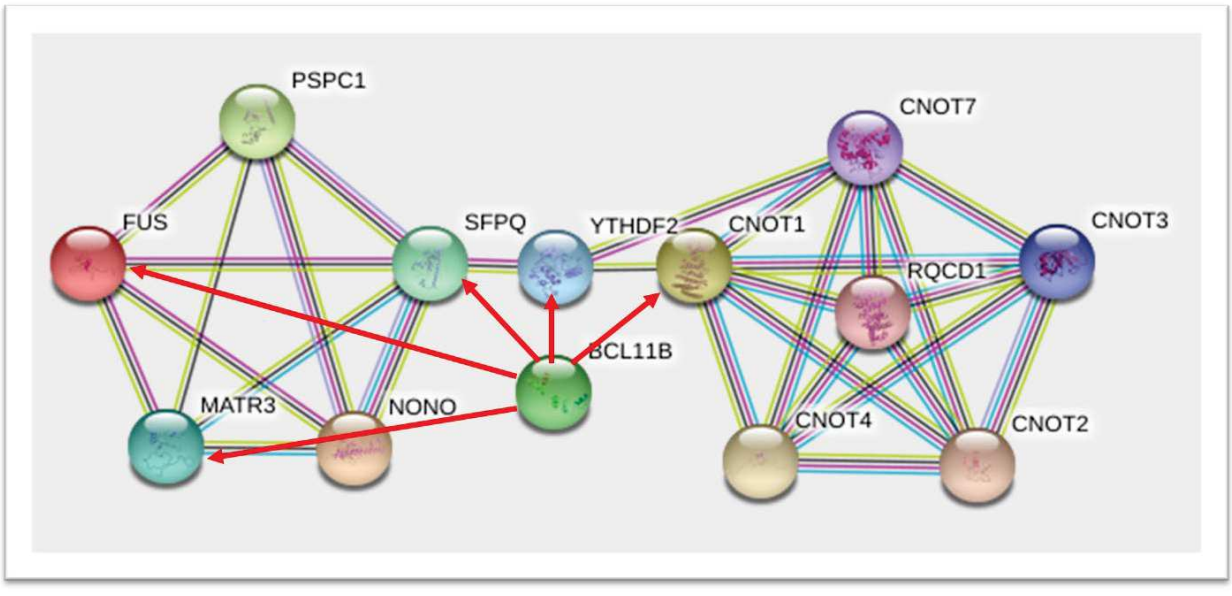


Figure 2: Parmi les résultats que nous avons obtenus à partir des résultats de spectroscopie de masse quantitative, nous avons observé, sans surprise, des protéines qui ont été précédemment décrites comme ayant des rôles dans le cycle de réplication du VIH-1 ou d'autres processus cellulaires impliqués dans l'infection par le VIH1.

Nous avons pu vérifer l'interaction avec certains de ces facteurs. Les flèches rouges montrent la protéine que nous avons réussi à immunoprécipiter avec CTIP2. Il est intéressant de noter que ces protéines ont été décrites comme ayant un rôle dans les processus d'édition et de dégradation de l'ARN:

- Le complexe CCR4-NOT, qui est l'une des principales deadénylases de l'ARNm cellulaire et est lié à divers processus cellulaires, y compris la dégradation de l'ARNm en masse, la répression médiée par les miARN, la répression de la traduction pendant l'initiation de la traduction et la régulation générale de la transcription.
- YTHDF2: reconnaît et lie spécifiquement les ARN méthylés en N6 sur l'Adénosine (m6A) et régule la stabilité de l'ARNm. m6A est une modification présente sur les sites internes des ARNm et de certains ARN non codants. La liaison de YTHDF aux ARNm contenant du m6A entraîne sa relocalisation vers des sites de dégradation, tels que les corps de traitement (corps P).
- MATR3 avec SFPQ s'est avéré nécessaire avec Rev pour exporter de l'ARN du VIH-1 partiellement non épissé et non épissé vers le cytoplasme.
- PSPC1, FUS et SFPQ font ensemble partie de la structure nucléaire nommée paraspeckle.

Sans surprise, certaines de ces protéines ont déjà été décrites dans d'autres recherches pour jouer un rôle dans le cycle de réplication du VIH-1 comme MATR3. MATR3 est une protéine de matrice nucléaire interne abondante largement exprimée dans divers tissus. Il peut se lier à l'ADN et à l'ARN et jouer un rôle dans la traduction des protéines, la réplication et la réparation de l'ADN, l'apoptose, la médiation de la mort des cellules neuronales en réponse au récepteur NMDA (N-méthyl-D-aspartate), la rétention nucléaire de l'ARN hyperedited et la stabilisation de l'ARN messager . Récemment, il a été décrit comme un régulateur positif du VIH-1 agissant au niveau post-transcriptionnel dans les lymphocytes primaires du sang périphérique (PBL). MATR3 forme un complexe avec SFPQ, qui était déjà impliqué dans l'exportation médiée par Rev des ARN du VIH-1. Alors que SFPQ et Rev se lient au pré-ARNm viral au site de transcription virale, MATR3 interagit à une étape ultérieure impliquée dans l'exportation nucléaire

Nous avons commencé à porter notre attention sur la protéine SFPQ, en particulier après qu'une immunoprécipitation ait confirmé l'interaction avec CTIP2. Ces expériences ont montré que l'interaction de CTIP2 se produisait non seulement lorsque les deux protéines étaient

surexprimées mais que CTIP2 était également capable de se lier à la protéine SFPQ endogène.

Ces résultats ont également été confirmés par des expériences de microscopie confocale, qui ont montré que CTIP2 relocalise SFPQ dans des structures nucléaires denses (figure 3).

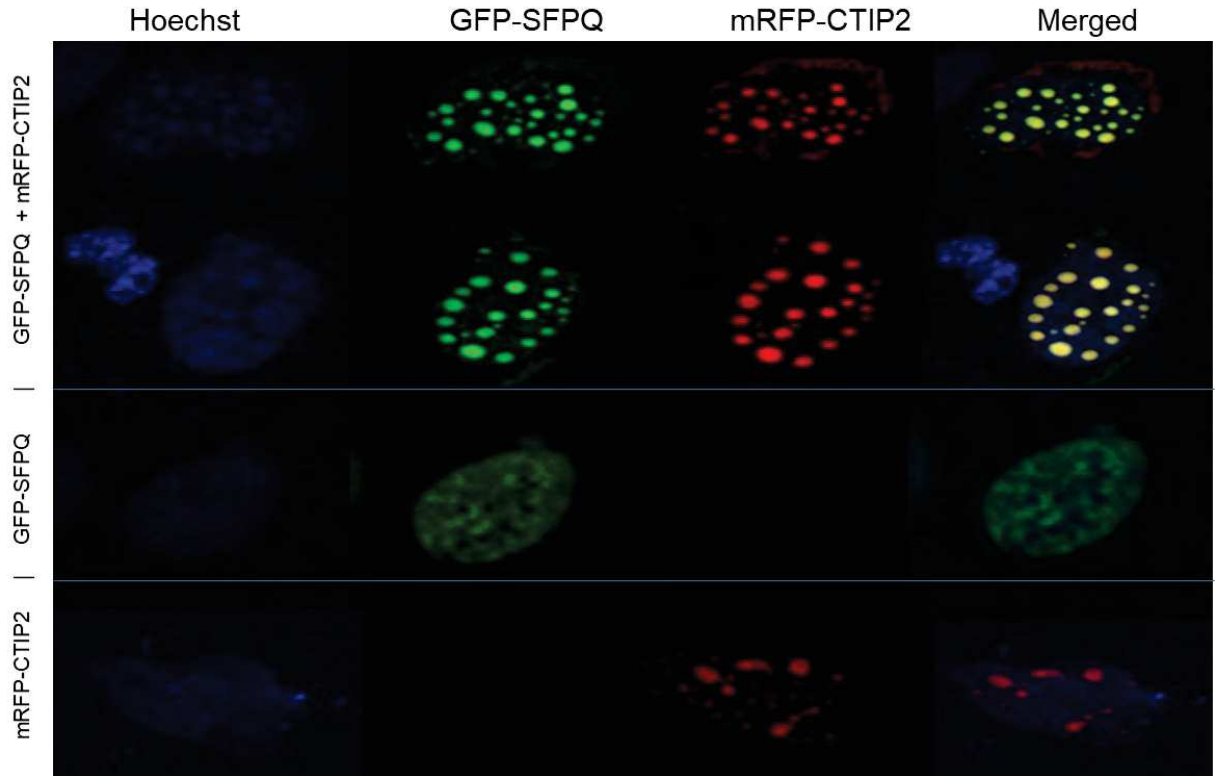


Figure 3 : Colocalisation de SFPQ et CTIP2 dans des structures de type sphère dans les noyaux de cellules microgliales.

Nous avons continué à étudier ces structures pour comprendre si elles pouvaient être connectées à des structures nucléaires décrites précédemment. Des expériences de microscopie confocale ont été réalisées dans le but de comprendre si ces corps étaient des corps PML (les corps PML varient en composition et sont impliqués dans des processus cellulaires tels que l'allongement des télomères et la réponse aux dommages à l'ADN) ou des paraspeckles qui sont des corps composés à la fois de protéines et d'ARN. Les paraspeckles réguleraient l'expression des gènes en séquestrant des protéines ou des ARNm. Si une certaine co-localisation avec la protéine PML a été constatée, les colocalisations quasi totale de CTIP2 avec SFPQ (un élément prépondérant des paraspeckles) et la protéine FUS (une autre protéine formant les paraspeckles) suggèrent que les structures denses formée par CTIP2 sont bien du type « paraspeckle ». Cette théorie est renforcée par notre analyse du RNAome associé à CTIP2. En effet, CTIP2 interagit avec différents ARN codant et non codant dont NEAT1, un lncRNA qui agit comme un

échafaudage de structure des paraspeckles. Nous avons également observé que la protéine virale Tat est séquestrée dans ces organelles lorsque CTIP2 est surexprimé. Bien que nos résultats nécessitent encore d'être confirmés, une hypothèse corroborée par d'autres groupes de recherche sur ces structures nucléaires, suggère que NEAT1_2 / Paraspeckle agit comme un promoteur du déclenchement de la défense cellulaire et aide le mécanisme de défense cellulaire. L'interaction avec FUS a également été étudiée plus en détail par immunoprécipitation, montrant une interaction plus faible par rapport à l'interaction CTIP2-SFPQ. Cette interaction s'est également avérée être dépendante de l'ARN et pourrait être expliquée par la relocalisation de CTIP2 dans les paraspeckles via un ARN (très probablement NEAT1).

Compte tenu de cette forte association avec la SFPQ, nous avons recherché si cette protéine pouvait avoir un effet sur la transcription et la réplication du VIH-1. La protéine CTIP2 a été précédemment décrite par notre groupe comme étant capable d'entraver la transcription des gènes du VIH-1. Nous montrons que SFPQ régule la transcription et la réplication du VIH1. De plus, nous montrons que le motif de liaison à l'ARN de SFPQ est important pour son activité. Nous pouvons proposer l'hypothèse que la délocalisation de la SFPQ par CTIP2 dans ceux décrits du compartiment nucléaire, pourrait affecter sa capacité à stimuler la transcription médiée par Tat. Nous avons ensuite comparé l'effet de surexpression unique de l'une (et de l'autre) protéines par rapport aux deux protéines combinées. Par rapport à la surexpression de CTIP2 seule, la co-expression des deux 159 protéines n'avaient aucun effet statistique sur la transcription mais si on les compare à l'effet SFPQ, les deux protéines ensemble avaient un effet répresseur plus fort sur le LTR et comme déjà décrit précédemment sur le LTR + Tat le phénotype a été déplacé d'un effet activateur à un répresseur une. CTIP2 et SFPQ ont également coopéré pour réprimer l'expression du provirus montrant une répression plus forte.

Une autre protéine trouvée dans le criblage par spectrométrie de masse était FUS. Notre attention sur cette protéine a été attirée par le fait que, avec la SFPQ (et d'autres protéines et ARN) est un composant clé des structures de la paraspeckle.
Pour ces raisons, nous avons commencé à réaliser des expériences d'immunoprécipitation en présence / absence d'expression de CTIP2 et avec ou sans traitement à la RNase A pour déterminer si leur interaction est médiée par l'ARN. Nous avons réussi à co-immunoprécipiter FUS et CTIP2 sans impact significatif du traitement RNase. Nous avons également pu colocaliser les deux protéines dans le noyau des cellules microgliales dans la même structure

dense que nous avons observée auparavant mais avec deux phénotypes plutôt distincts. Avec des quantités croissantes de CTIP2, le FUS semblait être exclu par ces structures, se concentrant sur sa jante.

En résumant ces résultats ensemble et en préparant un plan pour de futures expériences, nous devrions étudier plus avant l'interaction de CTIP2 avec d'autres protéines de paraspeckle. Nous devrions viser à relier CTIP2 à la machinerie post-transcriptionnelle responsable de l'édition de l'ARN de l'adénosine-à-inosine (A-vers-I). Il a été démontré que l'ADAR, l'une des protéines responsables de cette modification, interagit avec les protéines que nous avons déjà décrites en interaction avec CTIP2, comme SFPQ et MATR3. De plus, il faudrait tenter d'identifier cette modification des ARN liés au CTIP2, prouvant son rôle dans la rétention de ces ARN au sein des structures de la paraspeckle. En outre, un nouvel ensemble d'expériences de microscopie confocale et IP doit être effectué après la stimulation par l'interféron. Je crois qu'une surexpression plus naturelle (et plus faible) de CTIP2 maintiendra la paraspeckle dans un état plus stable, et une colocalisation plus précise de CTIP2, en particulier avec FUS, sera obtenue; si ma théorie s'avère correcte, je m'attends à voir des résultats similaires dans l'expérience FISH sur la surexpression de CTIP2 par stimulation interféron en ce qui concerne son interaction avec Neat1. Enfin, un test d'immunoprécipitation de la chromatine doit être effectué lors de la surexpression de Neat1 pour confirmer s'il aura le même effet sur la liaison de CTIP2 sur le promoteur du gène IL8 que sur la SFPQ.

D'autres protéines ont été testées pour confirmer les données de spéctométrie de masse. La famille de protéines YTHDF a été testée.

La famille YTHDF comprend trois paralogues (YTHDF1, YTHDF2 et YTHDF3; également appelés DF1, DF2 et DF3, respectivement), chacun ayant des fonctions rapportées différentes; DF1 améliore la traduction de l'ARNm, DF2 favorise la dégradation de l'ARNm et DF3 améliore la traduction et la dégradation. Ces familles de protéines exercent leur fonction en se liant à une modification spécifique hébergée par de nombreux ARN. La N6-méthyladénosine (m6A) est la modification nucléotidique de l'ARNm la plus abondante et régule les aspects critiques de la physiologie et de la différenciation cellulaires. Pour confirmer que les protéines YTHDF s'associent à CTIP2, nous avons effectué une co-immunoprécipitation en utilisant un α-YTHDF1, α-YTHDF2, α-YTHDF3 IP, contre un IP Mock (IgG non spécifique) comme contrôle négatif en présence ou non de Flag- CTIP2. Nous avons réussi à immunoprécipiter spécifiquement CTIP2 avec les protéines YTHDF respectives. Fait intéressant, YTHDF2 a été trouvé dans les échantillons surexprimant ou non CTIP2, ce qui suggère que cette interaction

est suffisamment forte pour être détectable avec le CTIP2 endogène également. Bien que cela se produise avec YTHDF2 spécifiquement ne soit pas une surprise étant donné que c'était spécifiquement YTHDF2 qui a été trouvé dans les résultats de la spécification de masse, il est également important de savoir que le niveau d'expression basal de CTIP2 a toujours été extrêmement faible pour être facilement détecté par Western Blot. Une analyse plus approfondie sera nécessaire. Quoi qu'il en soit, l'interaction de CTIP2 avec les trois membres de la famille YTHD a été confirmée. Connaissant le rôle des protéines YTHDF dans la liaison de la modification spécifique de l'ARN N6-méthyladénosine, nous avons voulu prouver que CTIP2 était capable de se lier à ces ARN spécifiquement modifiés. Nous avons effectué une immunoprécipitation d'ARN inverse et nous avons pu immunoprécipiter le CTIP2 lié à cet ARN modifié, confirmant son interaction avec cette machinerie de traitement de l'ARN. Alors que l'interaction entre les protéines CTIP2 et YTHDF est pontée par l'ARN (ou vice versa), il reste encore à explorer.

La famille CCR4-NOT n'a été étudiée que par l'interaction de CTIP2 avec CNOT1 et les résultats étaient difficiles à interpréter. Cela peut être dû au faible niveau d'expression du CNOT1 endogène. MATR3 a montré une interaction très variable avec CTIP2, indépendante de l'ARN, ne se produisant peut-être qu'en présence de SFPQ.

La deuxième partie du projet a consisté à définir les ARN cellulaires interagissant avec CTIP2. Une analyse pan-génomique a été réalisé à la fois dans les cellules HEK et les cellules microgliales en utilisant la technique de CLIP-Seq. Cette technique permet l'immunoprécipitation d'ARN après réticulation aux UV avec les protéines associées. Elle est suivie du séquençage des ARN liés à la protéine d'intérêt. Les données de séquençage, après avoir été traitées, ont livré de nombreux résultats intéressants, qui ont été confirmés par des expériences d'immunoprécipitation d'ARN. De plus, l'analyse du transcriptome a révélé que plus de 75% des lectures CLIP étaient mappées sur des exons et lorsque les lectures étaient mappées sur des régions introniques, l'emplacement des lectures était sur des isoformes d'ARN qui ne codent pas pour la protéine telle que les introns retenus (IR), la désintégration à médiation non-sens (NMD), la désintégration non-stop ou le transcrit traité.
Une autre découverte intéressante a été qu'en réticulant les résultats de Mass Spec avec ceux de CLIP-Seq, nous avons trouvé CTIP2 lié à 137 protéines et ARN codés par les mêmes gens. Parmi ces protéines, certaines d'entre elles ont été analysées et confirmées dans les chapitres précédents tels que FUS et SFPQ. De tous les ARN liés à CTIP2, environ 97% étaient des

protéines codantes. Bien qu'il s'agisse encore d'une hypothèse, il est intéressant de se rappeler que de nombreux ARNm sont retenus dans les structures de la paraspeckle, prêts à être traduits si le besoin s'en fait sentir. Un rôle possible du CTIP2 sur l'action de ce mécanisme pourrait expliquer l'interaction avec autant d'ARNm, alors que le fait qu'il interagisse spécifiquement avec les ARN et les protéines du même gène reste à expliquer. Nous avons décidé de jeter un œil aux 3% restants des ARN non codants. Dans ce groupe, nous avons décidé de nous concentrer sur cinq ARNnc qui ont montré un nombre élevé de hits (présence élevée dans l'immunoprécipitation CTIP2) et une signification biologique pertinente:
Nous avons confirmé l'association de CTIP2 avec une série d'ARN non codants d'intérêts dont : NEAT1.1, NEAT1.2, MALAT1, MEG3 et SNORD133.

L'ARN NEAT1 est un lncRNA bien décrit, qui forme l'échafaudage de la structure nucléaire paraspeckle et cela a confirmé notre hypothèse sur la localisation de CTIP2 au sein de ces structures. Malat1 est un ARN de structure pour différent complexes nucléoprotéiques. MEG3 semble avoir des fonctions de suppresseur de tumeur en association avec p53. Enfin, SNORD133 est l'un des nombreux ARNnc découverts dans les corps de Cajal.

Conclusion

Dans ce projet, nous avons étudier un large éventail de protéines et d'ARN interagissant avec CTIP2 pour évaluer leur fonction lors de l'infection par le VIH-1. Bien que nous n'avons pas encore une analyse complète d'un grand nombre de ces interactions, ce projet a ouvert de nombreuses voies d'investigation qui semblent prometteuses.
Nous avons démontré l'interaction de CTIP2 avec les protéines SFPQ et MATR3 impliquées dans la latence post-transcriptionelle du VIH-1. Ces résultats suggèrent une coopération des facteurs cellulaires impliqués dans le silencing transcriptionnel (CTIP2 et ses cofacteurs) et le silencing post-transcriptionnel du provirus (SFPQ/MATR3). Les analyses du protéome et du RNAome associés à CTIP2 suggèrent que CTIP2 serait une protéine constitutive des paraspeckles. Nous avons montré que SFPQ joue également un rôle dans la régulation de la transcription et de la réplication du VIH-.
Nous avons également confirmé l'interaction de CTIP2 avec FUS, une autre protéine de la structure de la paraspeckle.

Nous avons ensuite testé l'interaction de CTIP2 avec au moins un membre de la famille des protéines YTHDF. Ces protéines s'associent à l'ARN viral méthylé. Le rôle de CTIP2 dans ces mécanismes de régulation reste à déterminer.

L'interaction avec le complexe CCR4-NOT, qui est l'une des principales deadénylases de l'ARNm cellulaire reste à confirmer.

Données CLIP-Seq, nous avons sélectionné un petit nombre d'ARNc en fonction de leur importance biologique et du nombre de HITS de l'analyse de séquençage et nous avons réalisé une série d'expériences dans le but de confirmer nos résultats. En ce qui concerne Neat1, nous avons déjà décrit comment notre attention était noyée par le fait qu'il s'agit de l'échafaudage du paraspeckle structures, mais d'autres ARN ont été confirmés interagissant avec notre protéine.
MALAT1, un lncRNA connecté à l'épissage alternatif, à l'organisation nucléaire et à l'épigénétique la modulation de l'expression génique, tous les processus cellulaires dans lesquels les protéines et les complexes que nous avons pu trouver interagissant avec CTIP2, s'est avérée liée à CTIP2. Il a également été rapporté être régulé à la hausse lors de l'infection par le VIH-1 dans les lignées cellulaires Jurkat et MT4.

MEG3 avec son rôle de gène suppresseur de tumeur a également été trouvé lié à CTIP2, et encore une fois, nous voir un ARN avec des fonctions éprouvées reflétant celles décrites par CTIP2.
U11, un composant important du complexe d'épissage mineur, était un autre des ARNnc trouvés dans le CLIP-Seq, confirmé dans le RIP et confirmant en double les résultats de CTIP2 étant
associée à la machinerie des épissosomes dans les expériences de spectrométrie de masse quantitative.
SNORD133 un ncRNA trouvé dans les corps Cajal, avec une fonction liée à la méthylation des ARNt.

Pour conclure, ce projet de thèse a ouvert de nombreuses projets que mettront en lumière le rôle des CTIP2 dans de nombreux processus cellulaires qui n'ont jamais été liés à cette protéine auparavant.
D'ailleurs, nous n'oublierons pas que l'objectif principal est, à la fin, de trouver la clé de voûte dans le mécanisme de latence du VIH-1. Alors que nous continuerons dans un avenir perceptible à analyser les interaction SFPQ-CTIP2, apparemment la plus prometteuse, nous devrions penser avec optimisme que parmi toutes ces interactions CTIP2-protéine, et l'interaction CTIP2-ARN, nous pouvons avoir le point faible qu'une LRA ciblée ou une thérapie génétique pourrait viser pour éradiquer définitivement le VIH.