Understanding the molecular pathogenesis of HIV-associated Burkitt Lymphoma – the impact of HIV-1 protein Tat on lymphoma driver genes

By

Leonardo Alves de Souza Rios

Thesis presented for the Degree of

DOCTOR OF PHILOSOPHY

In the Department of Pathology,
Faculty of Health Sciences, UNIVERSITY OF CAPE TOWN

October 2020

Declaration

I, Leonardo Alves de Souza Rios, hereby declare that the work on which this thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other University. I empower the University to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signed by candidate

October 2020

Acknowledgements

Nothing in life that is worthwhile doing is easy and this Ph.D. thesis has greatly exemplified this to me. It would have been an even greater task to overcome if I had to do it alone. This work is dedicated to my wonderful parents, Gessé and Iolanda, who never once made me doubt myself or allowed me to conceive of giving this up, always encouraging me to keep going. To my older brothers, Guilherme and Philipe, thank you for being the examples that you are. Both extremely hardworking individuals who inspire me to thrive.

To my friends and fellow lab mice Beatrice, Lungile, Aaliyah and Riyaadh, thanks for always lightening the mood around the lab with jokes and all-round foolery that helped me cope with the disappointing nature of experimental molecular biology. I'm looking forward to following all your bright careers in science. Many thanks to the division of Haematology staff members, Associate Professors Jessica Opie and Karen shires for always being willing to help with expert advice. To Jean, my lab aunty, thank you for always looking after us students and for enabling the lab to run smoothly. To Rygana, Celyne and Coleen, thanks for always spoiling me with food and for your caring advice during these last few years, you made this lab feel a little bit more like home. Thanks to the Prince lab, especially Dr Shanel Swart, Victoria Damerell and Dr Jade Peres for their indispensable help during these last few years. I would also like to thank the National Research Foundation for funding this Ph.D. degree.

Lastly, but by no means least, to Dr Shaheen Mowla my mentor and friend. I cannot stress enough how important you have been to me during this Ph.D. Thank you for guiding me through this journey and for facilitating the conceptualization and execution of this research. Your calmness and positivity kept me going during the tough times. I could not have asked for a better supervisor; I hope to keep working together in whatever capacity it may be.

Table of contents

List of figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of tables

Chapter 2

Chapter 3

Abbreviations

ABC-DLBCL	Activated B-cell like diffuse large B-cell lymphoma
AICDA/AID	Activation-induced cytidine deaminase
AIDS	Acquired immunodeficiency syndrome
ANOVA	Analysis of variance
AP-1	Activator protein 1
APE1	apyrimidinic endonuclease 1
ATCC	American type culture collection
ATP	Adenosine triphosphate
BACH2	BTB domain and CNC homolog 2
BCA	Bicinchoninic acid assay
BCL2	B-cell lymphoma 2
BCL6	B-cell lymphoma 6
BCL7A	B-cell lymphoma 7A
BER	Base excision repair
b-FGF	Basic fibroblast growth factor
BL	Burkitt Lymphoma
BTG2	BTG anti-proliferation factor 2
bp	Base pairs
CARD11	Caspase recruitment domain family member 11
CH	Constant heavy chain
CHOP	Cyclophosphamide, doxorubicin hydrochloride, vincristine sulfate (Oncovin) and prednisone

CMV	Cytomegalovirus
CREB	cAMP response element-binding protein
CREBBP	CREB-binding protein
CRM-1	Chromosome maintenance 1
CSR	Class switch recombination
CXCR1	C-X-C motif chemokine receptor 1
CXCR2	C-X-C motif chemokine receptor 2
CXCR4	C-X-C motif chemokine receptor 4
DEPC	Diethyl pyrocarbonate
DF	Deletion fragment
DLBCL	Diffuse large B-cell lymphoma
DMEM	Dulbecco's modified Eagle's media
DMNT1	DNA methyltransferase 1
DMNT3a/b	DNA methyltransferase 3 a/b
DNA	Deoxyribonucleic acid
DSB	Double strand break
DZ	Dark zone
EBNA2	Epstein-Barr nuclear antigen 2
EBNA3C	Epstein-Barr nuclear antigen 3C
EBV	Epstein-Barr virus
EC	Empty control
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
EMSA	Electrophoretic mobility shift assay

EMT	Epithelial-to-mesenchymal transition
ERK	Extracellular signal-regulated kinase
EZH2	Enhancer of zeste homolog 2
FANCA	Fanconi anaemia complementation group A
FBS	Foetal bovine serum
FISH	Fluorescent *in situ* hybridization
FL	Full length
FOXO1	Forkhead box protein O1
GSK3-b	Glycogen synthase kinase 3 beta
GCB-DLBCL	Germinal centre B-cell like diffuse large B-cell lymphoma
HAART	Highly active antiretroviral therapy
HAND	HIV-associated neurocognitive disorder
HHV	Human herpes virus
HIV	Human immunodeficiency virus
HIV-BL	HIV-associated Burkitt Lymphoma
HRP	Horseradish peroxidase
HSP90	Heat shock protein 90
IG	Immunoglobulin
IGH	Immunoglobulin heavy locus
IGLL5	Immunoglobulin lambda like polypeptide 5
INK4/p16	Cyclin-dependent kinase 4 inhibitor A
IP	Immunoprecipitation
IRF8	Interferon regulatory factor 8
JNK	c-Jun N-terminal kinases
KS	Kaposi sarcoma

KSHV	Kaposi sarcoma associated herpes virus
LB	Lysogeny broth
LMP1	Latent membrane protein 1
LTR	Long terminal repeat
MAX	MYC associated factor X
MCP-1	Monocyte chemoattractant protein 1
MEK	Mitogen-activated protein kinase kinase
MLX	MAX-like protein X
MMR	Mismatch repair
mRNA	Messenger RNA
MT	Mutation
MYB	V-myb avian myeloblastosis viral oncogene
MYC	V-myc avian myelocytomatosis viral oncogene homolog
MYD	Myeloid differentiation primary response 88
NES	Nuclear export signal
NHL	Non-Hodgkin lymphoma
NIH	National institute of health
NLS	Nuclear localization signal
NFAT	Nuclear factor of activated T-cells
PAX5	Paired box 5
PBL	Plasmablastic lymphoma
PBMC	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDGF	Platelet derived growth factor

PIM1	Proto-oncogene serine/threonine-protein kinase
PKA	Protein kinase A
PLWH	People living with HIV
PTEN	Phosphatase and tensin homolog
P-TEFb	The positive transcription elongation factor b
RDG	Arginine, glycine, and aspartate
RE	Restriction enzyme
REG-γ	REG-gamma
RHOH	Ras homolog family member h
RIPA	Radioimmunoprecipitation assay
RL	Renilla
RLU	Relative luciferase units
RNA	Ribonucleic acid
RNAi	Ribonucleic acid interference
RNU6B	U6 small nuclear RNA
ROS	Reactive oxygen species
RPM	Revolutions per minute
RPMI	Roswell park memorial institute
RT	Room temperature
SDM	Site-directed mutagenesis
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
SHM	Somatic hypermutation
SP-1	Specificity protein 1
STAT6	Signal transducer and activator of transcription 6
ssDNA	Single-stranded DNA

S1PR2	Sphingosine-1-phosphate receptor 2
TAR	Transactivation response element
TCL1A	T-cell leukaemia/lymphoma protein 1A
TET2	Tet methylcytosine dioxygenase 2
TF	Transcription factor
TFBS	Transcription factor binding site
TSS	Transcription start site
UK	United Kingdom
UNAIDS	The Joint United Nations Programme on HIV and AIDS
UNG	Uracil DNA glycosylase
USA	United States of America
UTR	Untranslated region
UV	Ultraviolet
VEGF	Vascular endothelial growth factor
VEGFR	Vascular endothelial growth factor receptor
WT	Wild type
ZFP36L1	ZFP36 ring finger protein like 1

Abstract

Burkitt Lymphoma (BL) is a B cell non-Hodgkin lymphoma that occurs as three distinct subtypes, namely: endemic, sporadic, and immunodeficiency/HIV-associated. This cancer represents a frequent cause of mortality among HIV+ people in Southern Africa which has the highest incidence of HIV/AIDS worldwide. Recent reports associate a direct oncogenic function of HIV in BL development. However, the molecular mechanisms underlying this HIV-associated malignancy are not well understood. This study explores the oncogenic potential of HIV-1 protein Tat in BL *via* its ability to manipulate the expression of c-MYC and activation-induced cytidine deaminase (AID), two key drivers of BL progression.

Using dual-luciferase reporter assays, HIV-1 Tat was shown to enhance the activity of the *c-MYC* promoter (-2324 bp - +537 bp), which corresponded with elevated c-MYC protein levels in BL cells (Ramos) expressing HIV-1 Tat. By generating sequential promoter deletions, the minimal promoter region mediating HIV-1 Tat induced activation was identified. Site-directed mutagenesis indicated that this response was mediated by AP-1 binding elements, and co-immunoprecipitation assays revealed that HIV-1 Tat and the AP-1 factor JunB interacted within the same complex. Chromatin immunoprecipitation assays confirmed that JunB bound the c-MYC promoter *in vivo* under the influence of HIV-1 Tat.

The effect of HIV-1 Tat on the expression of the DNA editing enzyme AID was also investigated. Dual-luciferase assays revealed that HIV-1 Tat could enhance the activity of the three regulatory regions of the *AICDA* gene, namely R1, R2 and R4. This translated into

elevated AID protein expression in Ramos cells expressing HIV-1 Tat, which was also reflected in an increase in genomic instability as shown by enhanced phosphorylated H2AX expression. Sequential promoter deletions of the R1 promoter did not lead to a loss in HIV-1 Tat-mediated activation, pointing to potential post-transcriptional regulation. Indeed, HIV-1 Tat was found to downregulate the expression of hsa-miRNA-181b-5p, a known repressor of the murine *Aicda* gene. Furthermore, using reporter assays, we show that an hsa-miRNA-181b-5p mimic could repress *AICDA* via the full-length 3'UTR which contains three putative binding elements for the miRNA.

To date, this study reveals that HIV-1 Tat induced *c-MYC* promoter activation is mediated by two AP-1 binding sites. HIV-1 Tat was shown to couple with JunB and bind to both AP-1 sites on the *c-MYC* promoter inducing promoter activation. Furthermore, this study reveals a novel mechanism of AID deregulation via HIV-1 Tat-mediated miRNA perturbation. Lastly, we show that HIV-1 Tat interferes with hsa-miRNA-181b-5p expression in B cells, alleviating *AICDA* 3'UTR repression.

Chapter 1: Introduction and literature review

1.1 Introduction

This literature review provides a broad introduction to Human Immunodeficiency Virus (HIV) associated Non-Hodgkin Lymphoma (NHL) and gives detailed descriptions about the molecular features behind Burkitt Lymphoma (BL), a common HIV associated malignancy. It describes the recent advances in the field, specifically, the novel findings elucidating the mutagenic landscape of the two most common HIV-associated NHLs, namely Diffuse Large B-Cell Lymphoma (DLBCL) and BL. It describes the role of *c-MYC* chromosomal rearrangements and mutations in carcinogenesis and highlights the involvement of the DNA modifying enzyme Activation-Induced cytidine Deaminase (AID) in generating these genetic lesions. It also provides an overview of the known oncogenic roles of HIV and its viral proteins in both lymphoid and non-lymphoid cancers highlighting the gaps in knowledge of the molecular mechanisms behind HIV-associated Burkitt Lymphoma (HIV-BL) development. Parts of Section 1.3 of this literature review have been published as a review article in the Journal of Cancer Research and Clinical Oncology (Rios, L.A.d., Cloete, B. & Mowla, S. Activation-induced cytidine deaminase: in sickness and in health. *J Cancer Res Clin Oncol* **146,** 2721–2730 (2020). https://doi.org/10.1007/s00432-020-03348-x)

1.2 Background

Infection caused by the Human Immunodeficiency Virus (HIV) and the development of Acquired Immunodeficiency Syndrome (AIDS) remains a global health care problem with an estimated 38 million people living with HIV/AIDS in 2019 [1]. Although new infections per year have decreased worldwide, from the 1996 high of 3.4 million to 1.7 million estimated for 2019, efforts are still required to meet the 2020 UNAIDS goal of reducing this number to 0.5 million. According to UNAIDS, Eastern and Southern Africa are the regions with the highest burden

of HIV, with 20.7 million people living with HIV/AIDS, followed by Asia and the Pacific with 5.8 million [1]. South Africa is the country with the largest HIV positive population in the world with 13% of the population infected with the virus (7.8 million) as reported in 2020 [2]. Though access to antiretroviral therapy has improved significantly in South Africa since the start of the public sector Highly Active Antiretroviral Therapy (HAART) programme in 2004, many still do not receive treatment - almost 30% of HIV positive individuals are currently still not accessing treatment [1].

The introduction of HAART in South Africa, beginning in the early 2000s, has significantly improved the life expectancy of people living with HIV (PLWH), however, this has led to a dramatic increase in the incidence of HIV-associated co-morbidities including HIV-associated NHL [3,4]. While HAART allows for the reconstitution of the immune system, it does not eliminate the virus and latent reservoirs of the virus remain [5–7]. These, as well as other concomitant and chronic infections, are responsible for B-cell hyperactivation, leading to the transformation of these cells [8,9].

HIV positive individuals are at a 60-200-fold greater risk of developing NHLs compared to the general population [10,11]. Two main B-cell derived NHLs are associated with HIV infection, namely DLBCL and BL and notably, the prevalent subtype of DLBCL which occurs in PLWH is Activated B-Cell like DLBCL (ABC-DLBCL), which is highly aggressive and less responsive to current treatment regimens [12]. Another aggressive subset of DLBCLs known as Plasmablastic Lymphoma (PBL) is also highly prevalent in HIV positive individuals [13]. Interestingly this subset shares similar molecular characteristics with BL which are discussed later in this chapter (Section 1.3.1). The introduction of HAART combined with

chemotherapeutic treatment (combination therapy) has resulted in an improved survival outcome [14]. However, the prognosis of HIV-associated NHL remains poor with a 2-year overall survival (OS) rate of 55.6% (48.5 to 63.8) reported for HIV-associated DLBCL and 53.1% (40.2 to 70.1) for HIV-associated BL[15–17]. Therefore, there is an urgent need to further define the molecular pathology of HIV-DLBCL and HIV-BL to develop better treatment approaches. Several studies found the incidence of HIV-BL to still be relatively high post the HAART era with some reporting no improvement at all [16,18,19]. It is well known that infection with the Epstein-Barr Virus (EBV) is associated with the development of a specific subtype of BL [20]. However, this is observed to a lesser extent in HIV-associated BL indicating that HIV may play a more direct role in lymphomagenesis.

Intriguingly, BL seems to develop in HIV positive patients with a relatively high CD4 lymphocyte count [21]. This indicates that the immunocompromised status of these HIV positive individuals is not the only factor predisposing them to develop BL. Therefore, a better understanding of the role of HIV in the development of BL, and the molecular mechanisms behind this interaction is required for the improved treatment strategies of HIV-BL patients. This project is aimed at defining the relationship between HIV infection and BL, with a focus on specific host lymphoma driver genes, and therefore the sections below elaborate on the molecular pathology of this cancer.

1.3 Burkitt Lymphoma

Burkitt Lymphoma is an aggressive Non-Hodgkin lymphoma of germinal or post-germinal centre B cell origin. In 1958, Denis Burkitt was the first physician to report that children with

jaw tumours in Uganda had similar tumours at multiple additional sites, particularly in the abdomen [22]. The cancer was initially described as a sarcoma by Burkitt; however, it was later reclassified as a lymphoma. BL occurs in three main forms: endemic BL (eBL), sporadic BL (sBL) and HIV-associated BL (HIV-BL). The eBL form occurs predominantly along the malaria belt including equatorial Africa, Papua New Guinea and Northern Brazil. It is the most prevalent cancer among children in these regions, and presents with a swollen jaw and/or abdomen due to the formation of large tumours [23,24]. Notably, eBL is highly associated with Epstein Barr Virus (EBV) infection (detected in >90% of cases), with the virus being first described in eBL derived cells by Epstein et al in 1964 [25–28]. The sBL subtype occurs outside of the malaria belt, and it is less frequent but presents a similar morphology to the endemic form, although it is not as highly associated with EBV infection (20% of cases) [24,27]. A recent genome-wide sequencing study by Grande et al (2019) revealed that EBV status has a significant impact on the genetic landscape of paediatric BL, potentially superseding clinical variant status. Here it was shown that both EBV positive eBL and sBL had elevated levels of *AICDA* expression, correlating with a higher somatic hypermutation burden compared to EBV negative cases albeit with lower rates observed in coding regions. [29]. Of interest to this research project is HIV-BL which has lower associations with EBV (30-40% of cases) and malaria infection as compared to eBL [21,27]. The molecular characteristics and pathobiology of this subtype remain undefined and form the basis of this project.

BL is highly aggressive, characterised by its rapid tumour growth, with a doubling rate of 24-48 hours and under the microscope, the characteristic appearance of a BL tumour section upon Haematoxylin and Eosin (H&E) staining is the presence of a "starry sky" pattern, imparted by scattered macrophages phagocytizing cell debris and apoptotic cells (**Figure 1.1**). A hallmark of BL is the translocation and overexpression of oncogenic *c-MYC* which is detected through

Fluorescent In-Situ Hybridization (FISH). Although treatment regimens vary according to geographical setting, a common treatment approach for BL utilizes CHOP (Cyclophosphamide, Hydroxydaunorubicin (also called doxorubicin or adriamycin) Oncovin (vincristine) and Prednisone) or R(rituximab)-CHOP chemotherapy as well as a surgical intervention if necessary [21,30–32]. In the case of HIV-BL, the introduction of HAART before or during chemotherapeutic intervention has been shown to improve survival as compared to chemotherapy only regimens [33,34].

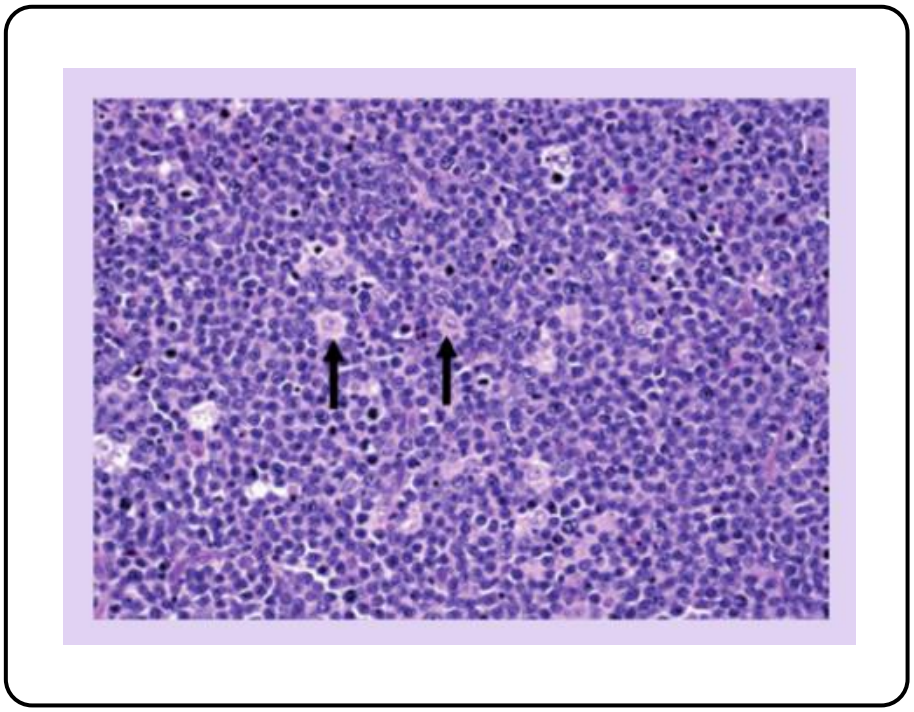

Figure 1.1: H&E stain of a BL tumour biopsy. Staining with Haematoxylin and Eosin results in a starry sky appearance with macrophages (denoted by black arrows) containing apoptotic tumour B cells surrounded by darker tumour cells (adapted from Molyneux et al., 2012).

1.3.1 Molecular hallmarks of BL

The pathogenesis of BL is strongly related to the overexpression of c-MYC, a Basic-Helix-Loop-Helix-Leucine Zipper transcription factor, which is an important regulator of apoptosis, cell growth and cellular metabolism [35]. c-MYC acts broadly in the regulation of mammalian gene expression forming part of the Extended MYC Network [36]. The c-MYC protein forms heterodimers with either MAX or MLX to bind to E-box elements (5'-CACGTG-3') thereby promoting gene transcription [37–39]. Under normal conditions, *c-MYC* expression is governed by mitogenic signals which activate transcription factors responsible for *c-MYC* promoter/enhancer activation. Importantly, c-MYC can immortalize cells and is one of the most overexpressed proteins in cancer and is, therefore, involved in the carcinogenesis of many human cancers [34]. Approximately 28% of the data sets from The Cancer Genome Atlas (TCGA), which includes 33 tumour types, harbour *c-MYC* genetic alterations [40]. These mutations are often non-coding, preserving c-MYC's potent pro-transcriptional ability but abrogate the regulatory mechanisms that govern its activity. The involvement of c-MYC in lymphoma has been well described. Importantly, in the context of BL, the translocation of the *c-MYC* gene is a hallmark of the disease. This translocation results in the overexpression of c-MYC allowing for the aggressive growth of the mutant cells [41]. Three main translocations are known to occur, the most frequent (70-80% of patients) being t(8;14)(q24;q32), involving the juxtaposition of *c-MYC* in chromosome band 8q24 to the immunoglobulin heavy chain gene (IGH) in chromosome band 14q32. The other two are t(2;8)(p12;q24), which involves the IG kappa-light chain locus in the 2p12, and can be detected in 15% of BL patients; and t(8;22)(q24;q11), involving the IG lambda-light chain genes in the 22q11, and is detected in 5% of BL patients [21]. Although the translocated *c-MYC* gene is the main contributor to the c-MYC overexpression in BL, c-MYC can still be expressed from its loci on chromosome 8

although at a much lower rate[42]. Furthermore, mutations that affect c-MYC activity and stability have been reported in both BL and DLBCL [43].

Recently, next-generation sequencing of sporadic BL has shown frequent genetic alterations to many genes further exemplifying the molecular complexity and heterogeneity of BL. Notably, *TCF3 (E2A)* and its repressor *ID3* which are both involved in the differentiation of B cells and cell cycle progression, often carry mutations in BL [44,45]. A whole-genome sequencing study analysing BL tumour samples compared to matched germline controls revealed recurring mutations in the *ID3* gene in 34% of samples. These mutations were shown to increase cell cycle progression from G1 to S phase *in vitro* [46]. Alterations to these genes have also been reported in HIV-associated BL with at least 67% of cases presenting with *ID3* and/or *TCF3* mutations [47,48]. An important feature of this malignancy is its dependency on the germinal centre (GC) dark zone (DZ) cell survival programme, relying on *BCL6, FOXO1, MYB* and *TCF3* expression, and the PTEN-PI3K-AKT pathway which is maintained at similar levels to normal DZ B cells [49–51].

The main driver of the *c-MYC* translocation has been identified as the DNA deaminating enzyme AID [52]. AID is a DNA modifying enzyme which plays an essential role in humoral immunity. Under normal conditions, the expression of AID is largely restricted to activated B cells and ensures that a diverse array of antibodies can be produced. However, when deregulated it can lead to genetic lesions including translocations [52–55]. The normal function of AID, as well as its involvement in cancer, are discussed further in the sections below.

1.4 The DNA modifying enzyme AID

The Human body is constantly challenged by a myriad of antigens and therefore requires a highly efficient way of protecting itself from these assaults. There are three lines of defence that aid in preventing pathogenic disease, namely, physical barriers, non-specific innate immunity, and specific adaptive immunity. Of the three lines of defence of the immune system, adaptive immunity is the most complex and sophisticated, being able to rapidly adapt to meet new challenges. Adaptive immunity is mediated through B and T lymphocytes working together to recognise foreign antigens and generate specific antibodies able to eliminate these antigens. The first line of diversification of antibodies occurs during B cell maturation when V(D)J recombination takes place [56]. After naïve B cells encounter antigens in secondary lymphoid tissues two additional processes occur namely somatic hypermutation (SHM) and class switch recombination (CSR). The AID enzyme plays a central role in these diversification mechanisms [53]

1.4.1 Somatic Hypermutation and Class Switch Recombination

Mutations are introduced into the variable region of the Immunoglobulin (Ig) gene at a rate of up to 10^{-3}–10^{-4} mutations per base/division, which is much higher when compared to the background 10^{-9} rate of mutation per base/division within the rest of the genome [57]. AID targets the "hot spot motif" RGYW/WRCY with R= A/G, Y= C/T and W=A/T at a higher frequency compared to the rest of the genome [58]. During SHM, AID converts deoxycytosine (dC) in the variable region to deoxyuridine (dU) mutations occur when replication proceeds before the correction of the mismatch or through the mismatch repair (MMR) and base excision repair (BER) pathways (**Figure 1.2**) [59,60].

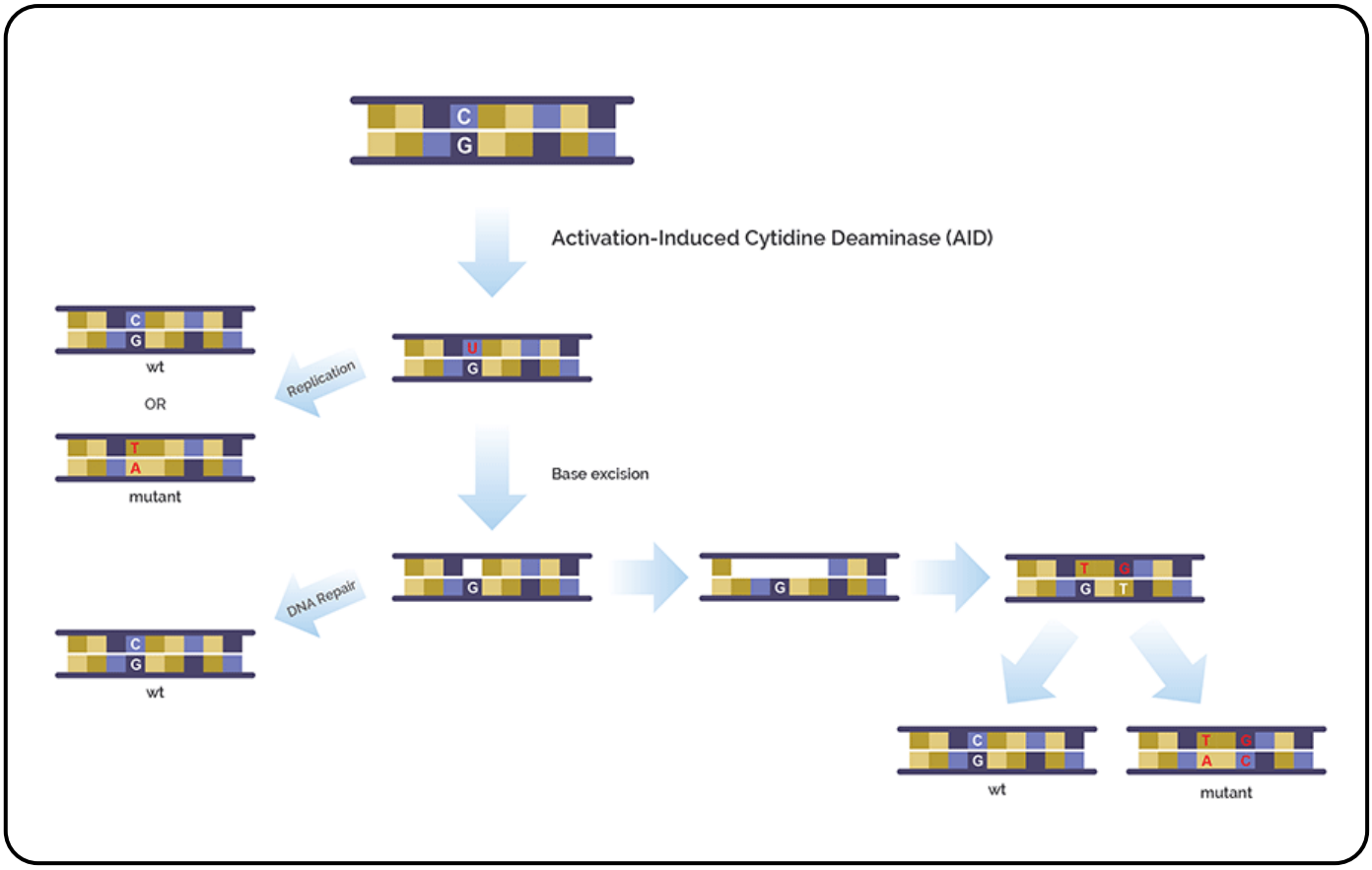

Figure 1.2: The role of AID in somatic hypermutation. In the variable region of *Ig* loci AID causes a change from cytidine to uracil. Mutations can occur if replication is carried out before DNA repair or through aberrant BER or MMR DNA repair mechanisms. The resulting mutated sections of DNA allow for the diversification of the hypervariable region of antibodies (image adapted from https://www.anaptysbio.com/technology/.).

Effector function of antibodies is changed through CSR, a process that occurs between two switch regions (S) found upstream of each constant heavy chain (CH) gene and results in the switching of antibody isotype expression from the initially expressed IgM to IgD, IgG, IgA, and IgE isotypes. AID converts dC to dU and these are removed by the actions of Uracil-DNA Glycosylase (UNG) and apyrimidinic endonuclease (APE1) leaving ssDNA breaks, dsDNA breaks are formed when ssDNA breaks are close to each other on opposite DNA strands or by MMR repair pathway. The intervening DNA is excised as a switch circle, resulting in the

variable region being juxtaposed to the downstream CH gene, these are then joined by either non-homologous end-joining (NHEJ) or alternative end-joining (A-EJ) processes (**Figure 1.3**) [61].

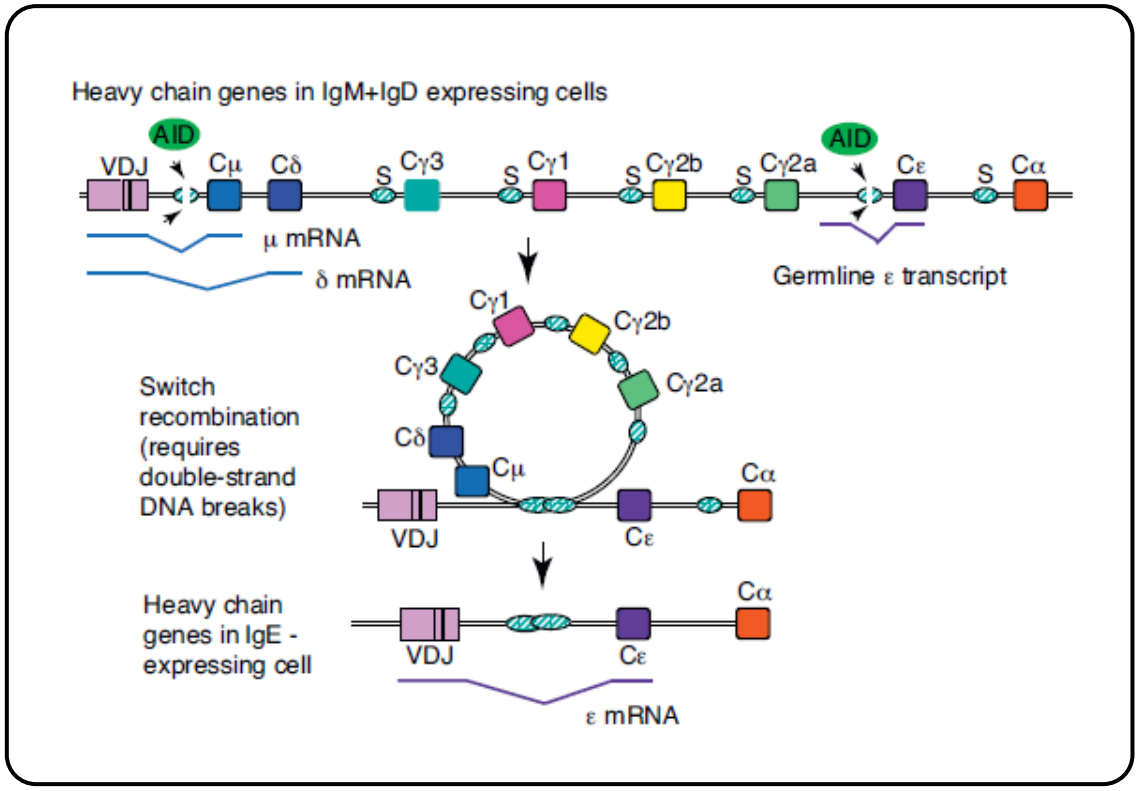

Figure 1.3: CSR mediated via the AID enzyme. CSR is a gene rearrangement process that allows for the change of expressed Ig constant regions (C) resulting in the change from IgM to IgG, IgA, IgE or IgD isotypes. AID deaminates dCs at S regions (S) causing DSBs. The intermediate DNA is excised as a switch circle and the DSB are joined through either NHEJ or A-EJ, resulting in the expression of a different Ig isotype. (Figure adapted from Stavnezer. J., 2011) [62].

1.4.2 AID leads to DNA demethylation

More recently, AID has become recognised as a DNA demethylator, linked to its ability to deaminate cytidine [63]. The conversion of deoxycytosine (dC) to deoxyuridine (dU) can lead to the disruption of CpG methylation sites in DNA, thus altering gene expression. This role for AID has only recently been recognised, with studies showing both genome-wide and gene-specific demethylation [64]. Although further investigations are needed, there is sufficient data to support the notion that AID can alter gene expression in B cell malignancy through active DNA demethylation [64–66].

1.4.3 The physiological regulation of AID

Due to its DNA mutagenic function, the expression of AID is strictly regulated within normal B cells and occurs at the transcription, post-transcription, and post-translation levels, involving nuclear/cytoplasmic distribution and stability, as well as enzymatic function. A comprehensive summary of these regulatory mechanisms is displayed in **Figure 1.4** [67,68]. Four conserved regulatory regions have been described for the human *AICDA* gene containing a variety of transcription factor binding sites (TFBS) including Sp1, Sp3, HoxC4-OCT, STAT 6 and others. Region 4 is primarily responsible for *AICDA* expression upon stimulation of B cells while region 1 seems to be responsible for basal expression.

Regulation of AID also occurs post-transcriptionally with the repressive action of miRNAs. To date, four miRNAs have been implicated in the repression of AID expression namely: miR-

155, miR-93, miR-29b and miR181b [69–72]. However, most of the research has focused on murine *Aicda* and there is limited research done on the repression of human *AICDA* by miRNA.

Post-translational regulation of AID is mediated by proteins like the Protein Kinase A (PKA) [73], with phosphorylation of S38, T140, T27 and S3 being important for AID function during SHM and CSR [74–77]. AID is predominantly localised to the cytoplasm of B cells where its stabilization is dependent on Heat shock protein 90 (Hsp90) coupling [78]. However, AID acts on ssDNA and is therefore shuttled into and out of the nucleus through its nuclear localization signal (NLS) and nuclear export signal (NES) located in the N and C termini respectively [79]. Furthermore, AID activity is limited by the cell cycle, in the early G1 phase when chromatin de-condense and transcription reinitiates, ssDNA is exposed allowing AID access to its preferred substrate [80]. Off-target AID activity is then limited when the nuclear envelope reforms and AID molecules are shuttled out of the nucleus by a Chromosome Maintenance 1 (CRM-1) dependant mechanism or rapidly degraded in the nucleus via a ubiquitin- and ATP-independent proteasomal degradation pathway mediated by REGgamma (REG-γ) or through ubiquitination by Cullin 7 E3 ubiquitin ligase [80–83] Together, these levels of regulation ensure that AID expression remains in check.

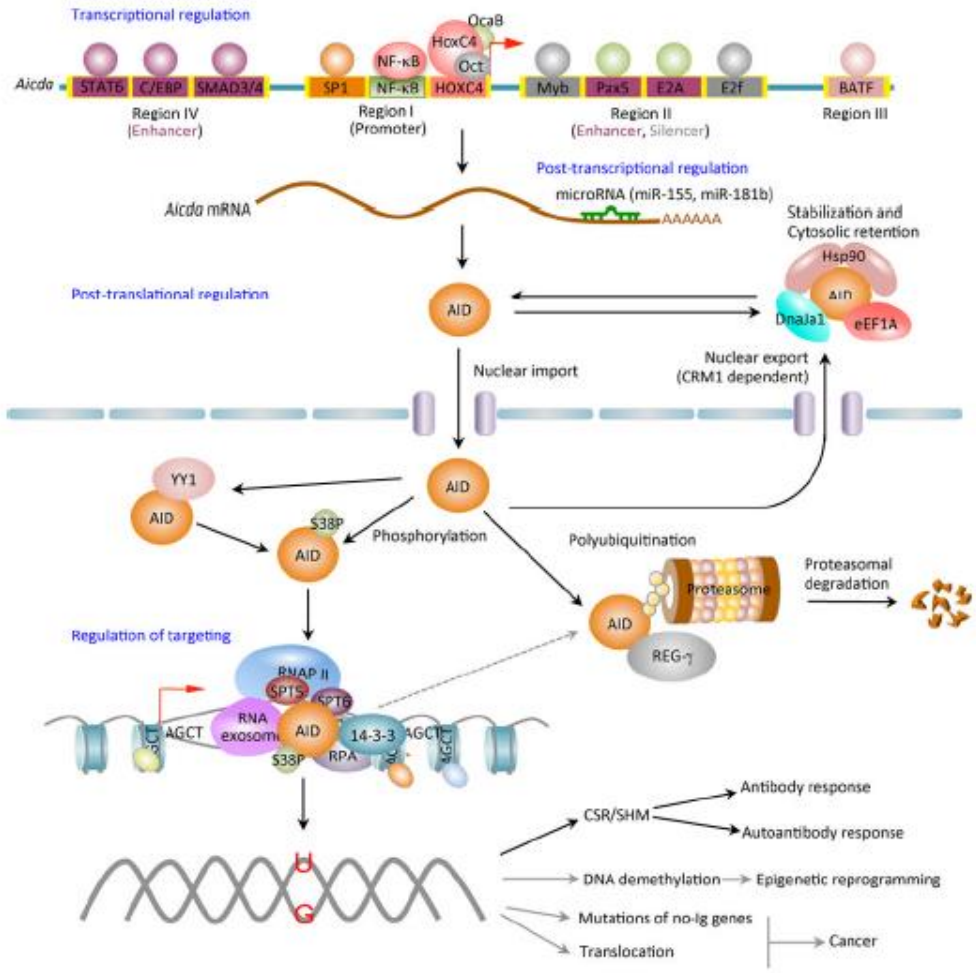

Figure 1.4: Summary of the *AICDA*/AID regulatory mechanisms. AID activity is regulated at multiple levels. Transcriptional regulation is mediated via four conserved regions which contain multiple TFBS. Posttranscriptional AID is regulated by several miRNAs. Binding partners like HSP90, promote AID stability in the cytoplasm. Posttranslational mechanism of regulation includes phosphorylation of the AID protein at select amino acids which govern its ability to induce SHM or CSR. Degradation of AID occurs by polyubiquitination, preventing its accumulation in the nucleus. (Figure adapted from Zan & Casali., 2013)

1.4.4 The role of AID in lymphoid cancers

The involvement of AID in lymphoid malignancies has been well described, with aberrant expression described in both B and T cell-derived lymphomas [67,84–90]. The main translocation event, t(8;14) in human BL and t(12;15) in murine plasmacytomas involving the *c-MYC* and *IgH* loci is driven by AID [52,54,86,91]. Importantly, Ramiro et al (2004) showed that AID expression was essential for the *c-MYC/IgH* translocation event in Balb/c-IL6tg transgenic mice, with $AID^{-/-}$ mice failing to develop the characteristic translocation events in B cells [52]. Duquette et al (2005) showed, using electron microscopy, that AID binds G-loops that form during transcription of the *c-MYC* gene as well as the immunoglobulin S region. These G-loop regions were mapped to the breakpoints related to *c-MYC* translocation [92].

Besides its link to *MYC* translocation, AID mutational activity has also been reported in other important B cell-specific genes like the *ID3* gene resulting in its inactivation [93]. A recent whole-genome sequencing study by Panea et al (2019) elucidated new driver mutations in BL including *IGLL5*, *BACH2*, *BTG2*, *BCL6*, *BCL7A*, *TCL1A*, *IRF8*, *CXCR4*, *ZFP36L1*, and *S1PR2* which all harboured somatic hypermutation signatures in promoter regions [94].

Further association between AID and the development of BL relates to its relationship with the human gammaherpesvirus Epstein-Barr virus (EBV), a virus that is closely associated with eBL. Studies suggest that EBV directly targets AID [95,96] and that this is reliant on the latency program of EBV infection that differs according to the lymphoproliferative malignancy [97]. Recent studies have shed light on the mechanisms whereby EBV proteins including the Latent Membrane Protein 1 (LMP1), Epstein-Barr nuclear antigen A2 (EBNA2) and Epstein-Barr virus nuclear antigen 3C (EBNA3C) induce AID expression [29,98–102]. These studies highlight

the complex relationship between EBV factors and AID expression and support a collaborative role between EBV and AID in lymphomagenesis.

Although AID is an essential enzyme of the adaptive immune system which, under normal conditions, enables antibody affinity and specificity diversification. It is also associated with a range of pathological conditions (**Figure 1.5 & 1.6**). Under these pathophysiological conditions including infection with EBV, malaria and HIV, AID can become deregulated, leading to off-target mutational and epigenetic changes to DNA, affecting the expression and activity of both tumour promoters and tumour suppressors. This has been described for both B-cell lymphomas, as well as several non-lymphoid cancers (**Figure 1.6**).

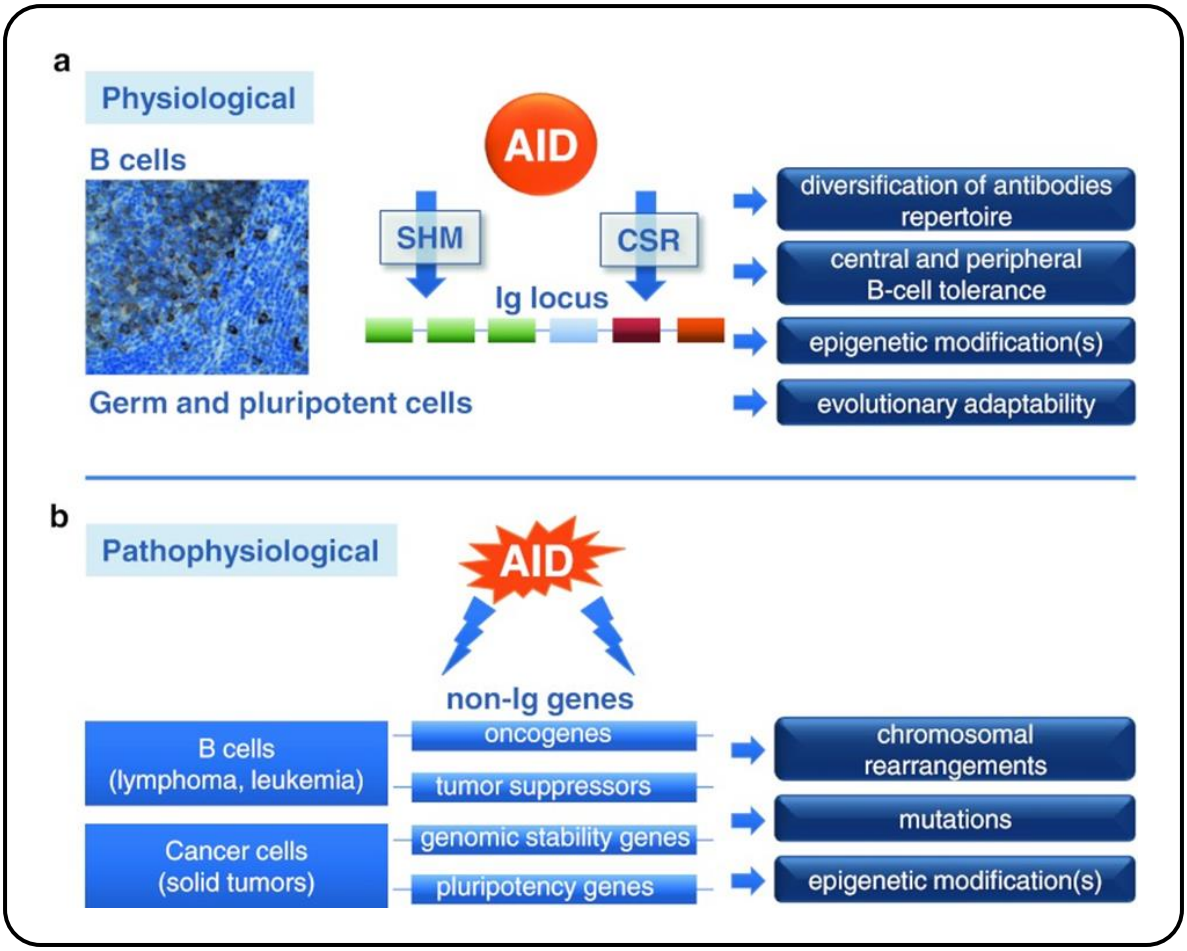

Figure 1.5: AID activity in physiological vs pathophysiological settings. A) normal physiological AID function, which occurs in activated B cells in germinal centres, where AID acts on the Ig locus, inducing SHM and CRS. B) aberrant activity under a pathological setting where off-target AID activity occurs at non-Ig genes leading to carcinogenesis (image adapted from Mechtcheriakova et al., 2012).

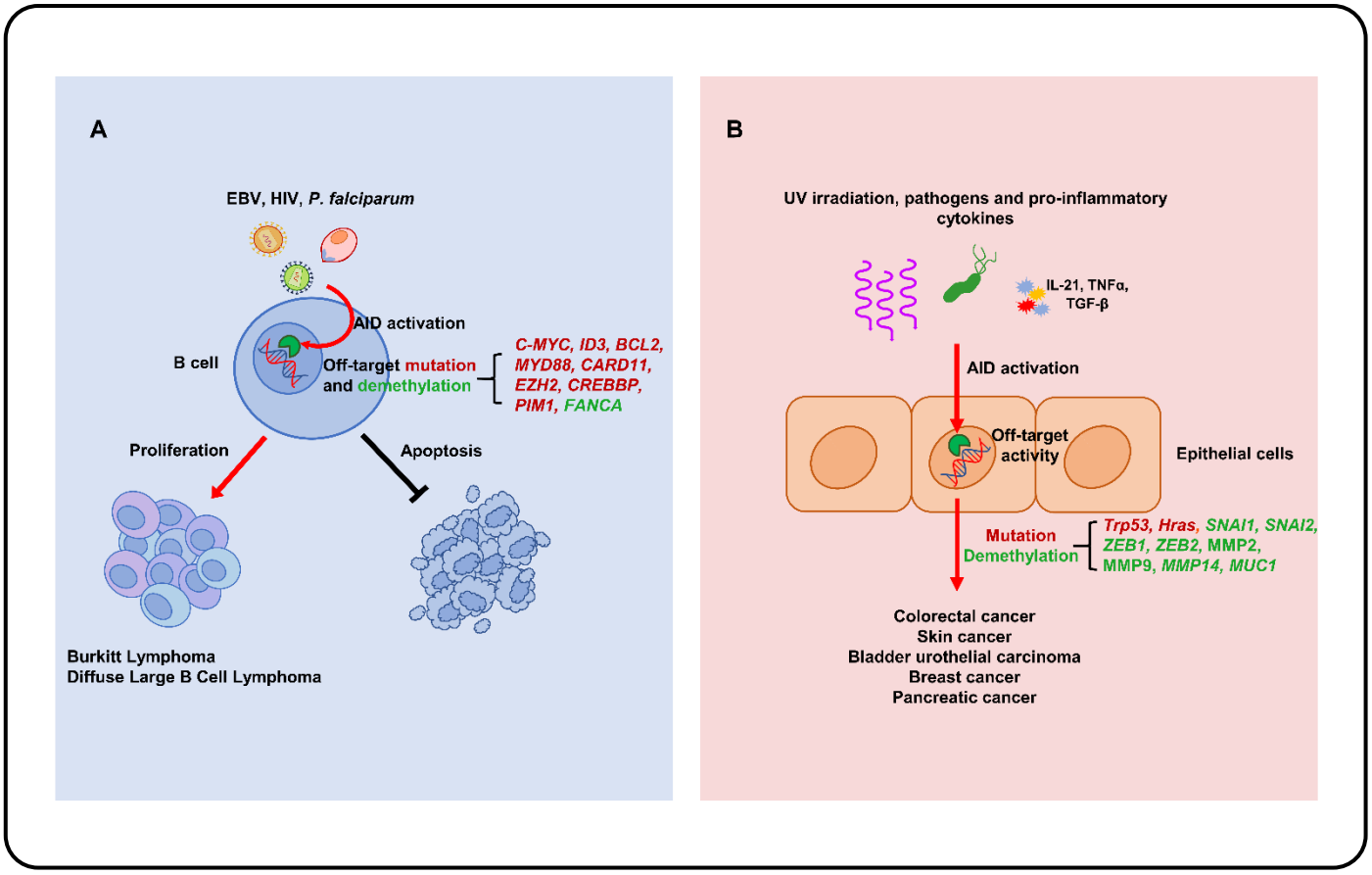

Figure 1.6: Illustration of AID deregulation, affected downstream genes and cellular consequences in lymphoid and non-lymphoid cancers. A) AID deregulation by pathogenic agents leads to off-target mutations in key oncogenes/ tumour suppressors inhibiting apoptosis and promoting proliferation, culminating in mature B cell lymphomas. B) AID dysregulation by UV irradiation, pathogenic infection and chronic inflammation leads to off-target mutations in genes involved in EMT and proliferation, ultimately leading to the development of carcinomas (Adapted from Rios et al., 2020) [103].

1.5 The role of HIV in carcinogenesis

1.5.1 HIV and cancer

As discussed earlier, people infected with HIV are predisposed to develop certain types of cancer both AIDS-defining, such as Kaposi Sarcoma (KS), NHL, and cervical cancer and non-AIDS-defining cancers including Hodgkin lymphoma, head and neck cancer, anal cancer, and lung cancer [3,104]. Although the introduction of HAART has decreased the incidence of AIDS-defining cancers, NHLs continue to present as a common cause of mortality in PLWH, especially in the Southern African context[15,105]. The most commonly occurring NHL in HIV positive individuals is BL followed closely by DLBCL and PBL (**Figure 1.7**). While previously this association was seen to be exclusively due to the immunodeficiency caused by HIV, recent evidence points to HIV playing a more direct and active role in cellular transformation and oncogenesis. This is discussed further in the sections below.

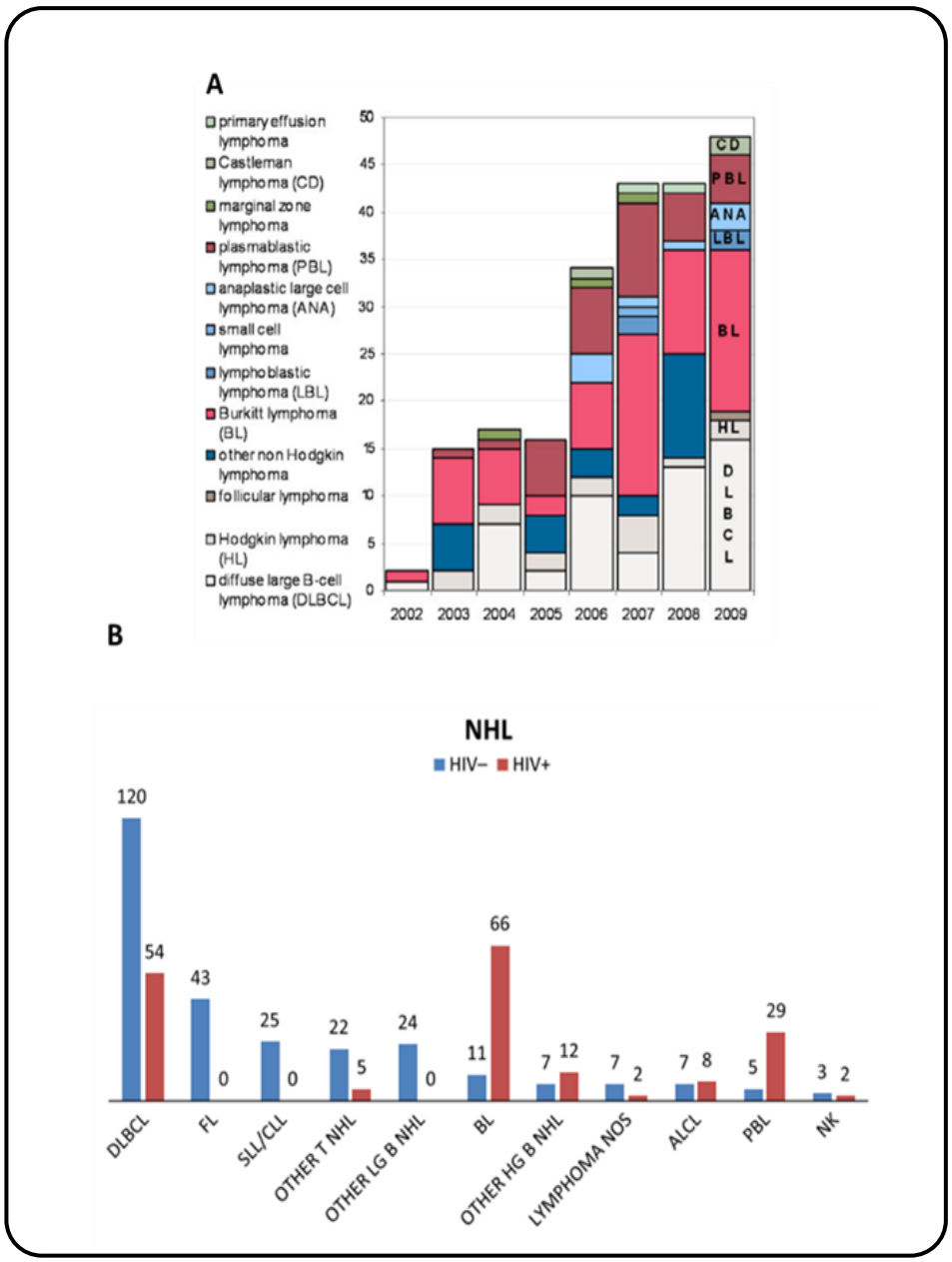

Figure 1.7: Graphs showing the distribution of HIV associated lymphomas from two independent studies conducted in the Western Cape, South Africa. A) Distribution of HIV associated lymphomas at the Tygerberg Hospital from 2002-2009 [106]. B) Retrospective analysis showing the distribution of HIV-associated lymphomas from Groote Schuur Hospital from 2005-2010 [107]. Both studies highlight the significant overrepresentation of BL in the HIV positive population.

1.5.2 Oncogenic role of HIV proteins

HIV proteins are essential for the survival and propagation of the virus. However, some of these viral proteins like HIV-1 Nef and Tat have also been recently implicated in carcinogenesis, which has been well described in the context of KS development. Importantly, HIV-1 proteins are actively secreted into the serum of infected individuals and are detectable even in patients undergoing HAART, who are virally suppressed [108–110]. HIV-1 Nef is one of the first transcribed proteins to accumulate in the early stages of macrophage infection [111]. Upon entering the cell it is thought to interfere with host gene expression, promoting the development of HIV related cancers [112]. A study by Xue et al (2014) showed that Nef and the HHV8 protein K1 could synergistically promote cellular proliferation and angiogenesis (two important processes in oncogenesis) through the PTEN/AKT/mTOR pathway mediated by the miRNA-718 in both primary Human Vascular Endothelial Cells (HUVECs) and EA.hy926 cells.[113]. Similarly, another study by Yao et al (2015) also utilizing the same cell lines showed that ectopic or soluble HIV-1 Tat and K1 could synergistically increase miRNA-891a-5p expression leading to elevated NFκB expression resulting in increased angiogenesis and cellular proliferation [114]. HIV-1 Tat can also accelerate KS tumorigenesis induced by another HHV8 protein, Kaposin A. Both these proteins activate multiple signalling pathways such as the MEK/ERK and PI3K/Akt pathways which together result in increased proliferation of NIH3T3 mouse fibroblast cells *in vitro* and increased tumour formation *in vivo* in both nude and immunocompetent mice [115].

A limited number of studies have reported on the oncogenic potential of HIV proteins in NHL. For instance, the HIV protein p17 has been shown to promote the DNA binding activity of c-MYC and CREB, both important regulators in cell proliferation, differentiation, and survival.

p17 increases the phosphorylation of these factors through time and dose-dependent activation of the cAMP/PKA and MEK/ERK signalling pathways. This deregulation is specifically important in the context of B cell lymphomas because c-MYC and CREB are both transcription factors involved in the pathogenesis of these diseases [116]. A study by Giagulli et al (2011) also showed that p17 variants differentially regulate PTEN expression with one variant (p17Δ36) leading to the repression of PTEN and the subsequent promotion of B cell proliferation [117]. HIV proteins can also mimic important growth factors inducing tumorigenic processes like angiogenesis. HIV-1 Tat can recognise and bind to vascular endothelial growth factor receptor (VEGFR) acting as a mimic of the VEGF protein, leading to the increased density of micro-vessels in BL associated tumours [118]. Furthermore, HIV-1 Tat has also been reported to induce epigenetic modifications in HIV related B cell lymphomas, with evidence showing that HIV-1 Tat leads to DNMT1 and DNMT3a/b overexpression by the repression of the miRNA148-152 cluster which are known regulators of DNMTs. This aberrant expression of DNMTs thus leads to hypermethylation of the *INK4/p16* gene promoting the oncogenic phenotype by alleviating cell cycle inhibition [119]. Recently, our research group has shown that BL cells exposed to recombinant HIV-1 Nef have increased transcription of both *c-MYC* and *AICDA* [120]. This was correlated with increased AID activity, demonstrated by exacerbated DNA damage within these cells. Taken together these studies highlight the importance of HIV-1 proteins as host genome regulators in the development of cancer.

1.5.3 Oncogenic role of HIV- 1 Tat

This project is focused on the role of HIV-1 Tat in oncogenesis, its normal function and interaction with the cellular host genome are discussed here. HIV-1 Tat is a small protein (86-101 amino acids) (**Figure 1.8**) that has a crucial role in HIV DNA transcription and survival.

HIV-1 Tat is a highly versatile protein with many functions attributed to its basic domain, an essential structural component for its interactions with both viral and host factors [121]. Viral HIV-1 Tat activity requires the transactivation-responsive region (TAR) [122], a stem-loop structure present at the 5' end of all HIV mRNAs. The TAR region is recognised by the basic arginine-rich domain [123,124]. HIV-1 Tat increases the rate of gene transcription, central to this role is HIV-1 Tat's ability to bind to P-TEFb which is then recruited to RNA Polymerase II, releasing stalled mRNA synthesis [125–127]. HIV-1 Tat also recruits other host TFs to the LTR region of the integrated viral genome promoting the initiation of mRNA synthesis [128–130]. However, HIV-1 Tat does not only target the viral DNA/mRNA motifs; it has also been shown to bind to the host genome with many reports indicating that it influences host gene expression.

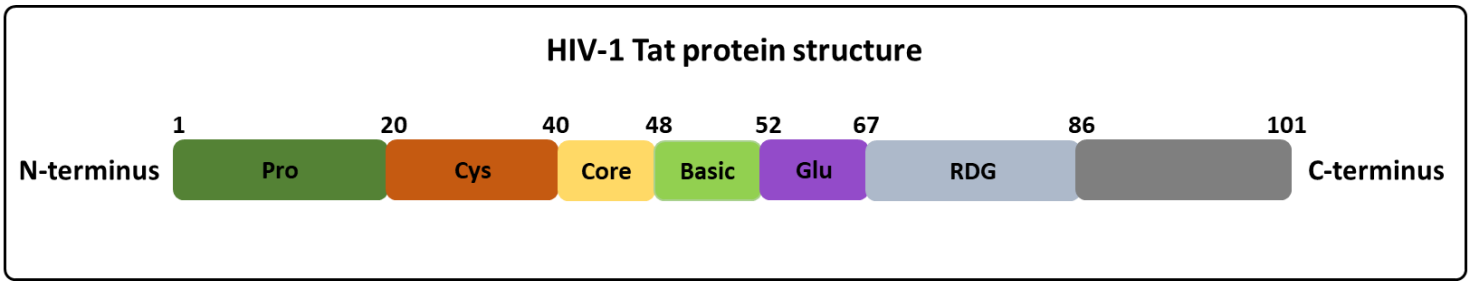

Figure 1.8: Structure of HIV-1 Tat protein showing the specific amino acid-rich domains. HIV-1 Tat is a small protein of approximately 14 kDa. It is composed of five distinct amino acid domains: a proline-rich domain (Pro), cysteine-rich domain (Cys), core region, basic domain (basic), a glutamine-rich domain (Gln) and an RDG motif. Numbers annotated represent amino acid positions.

A human genome-wide study by Marban et al (2011) revealed that HIV-1 Tat binds to a variety of regions within the genome, notably they showed that there was significant binding occurring within repetitive DNA elements, with the majority being Alu elements, which are the most common transposable elements in the human genome [131,132]. They also reported that in Jurkat

cells (acute T cell leukaemia cell line), 7% of HIV-1 Tat-bound regions are near transcription start sites (TSS) within gene promoters. They also showed that HIV-1 Tat binding within the human genome coincided with genes involved in T cell biology and immune response. A subsequent study to the work done by Marban et al (2011), using an improved methodology, further demonstrated HIV-1 Tat's preference for the promoter and 5'-UTR regions of genes (24% enrichment) and, importantly, it was also shown that these associations were conserved at the gene expression level, which was assessed through RNA-Seq [133]. Another human genome-wide study conducted by Dhamija et al (2015) showed that the HIV-1 Tat protein acts as a repressor of the *c-Rel* gene by interacting with specific NFkB sites on the *c-Rel* promoter, which they proposed aids in the persistent viral infection of the T cells [134]. There is also evidence that HIV-1 Tat indirectly influences the transcription of certain genes. Lim & Garzino-Demo, (2000) demonstrated that HIV-1 Tat increased binding activity of transcription factors SP1, AP-1 and NFkB to their DNA elements on the *Monocyte Chemoattractant Protein 1 (MCP-1)* promoter resulting in increased MCP-1 expression in astrocytes. HIV-1 Tat has also been shown to deregulate pRb2/p130 mRNA expression and inactivates its ability to arrest cell proliferation by binding to its pocket region, positively affecting the growth of HIV-associated NHL tumours [136,137]. Overall, these studies showed that the extent of HIV-1 Tat binding to the human genome was greater than previously thought and that the regulatory role of HIV-1 Tat is not only observed in the viral genome but also in the host genome.

As previously discussed, HIV proteins including HIV-1 Tat promote KS development and aggressiveness in HIV positive patients [113]. Zhou et al (2013) showed that HIV-1 Tat promotes vIL-6 induction of factors involved in angiogenesis and tumorigenesis (VEGF, b-FGF, CD31 and cyclin D1) through the regulation of the PI3K/PTEN/AKT/GSK-3b Pathway [138]. Importantly Germini et al (2017) recently showed that exposing healthy B cells to HIV-1 Tat

promotes the co-localization of the *c-MYC* locus to that of the *IgH* locus [139]. In a more recent publication, the same group demonstrated that HIV-1 Tat induces *AICDA* gene expression in healthy peripheral blood B cells resulting in DNA damage located at both the *c-MYC* and *IgH* loci [140]. Taken together, these studies indicate that HIV-1 Tat and AID can work synergistically to promote *c-MYC* translocation and overexpression.

HIV-1 Tat has been shown to alter cellular gene expression through altering cellular miRNA expression [114,141]. This was recently shown in Kaposi sarcoma, where HIV-1 Tat was shown to act synergistically with Orf-K1, a KSHV protein, to promote the expression of miRNA-891a-5p. The authors showed that this cooperation led to the promotion of angiogenesis through the NFkB pathway [114]. Work focusing on HIV Associated Neurocognitive Disorder (HAND) has also shown that HIV-1 Tat can bind to multiple miRNAs preventing them from accessing and repressing the translation of target mRNAs. Importantly, one of the miRNAs which were identified in the study was miRNA-181-b, a known regulator of AID [141,142]. The role of HIV-proteins in the deregulation of miRNAs involved in B-cell biology has not yet been elucidated, therefore, research in this area is needed to further explore the molecular pathology of HIV-NHLs.

1.6 Aims of this study

HIV-1 associated NHLs remain a clinical challenge in the South African context. To develop better treatment strategies a more in-depth knowledge of the molecular pathobiology of these malignancies is required. Previous preliminary work in our group had investigated the effect of HIV-1 Tat on key lymphoma driver genes and found that it could enhance the transcription of both *c-MYC* and *AICDA*. This study aimed to (i) confirm this preliminary finding through a more comprehensive investigation, and (ii) define the mechanisms via which HIV-Tat

enhanced the expression of these genes. The results of the investigations on the *c-MYC* oncogene are presented in Chapter 2, and that of *AICDA*/AID is presented in Chapter 3. In Chapter 4, the overall findings of the study are discussed and placed within the context of our current understanding of how HIV confers an increased risk of lymphoma development and progression.

Chapter 2: Investigating the mechanisms of HIV-1 Tat-induced activation of the *c-MYC* promoter

2.1 Introduction

The main function of the HIV-1 Tat protein is to stimulate and enhance transcription of the integrated HIV-1 genome, thus, the protein is essential for the virus's survival and propagation. Key to this role is HIV- 1 Tat's ability to bind to the transcription elongation factor P-TEFb which is then recruited to RNA Polymerase II releasing stalled mRNA synthesis [125–127]. HIV-1 Tat also recruits other host transcription factors (TFs) to the LTR region of the integrated viral genome which promotes initiation of mRNA synthesis [128–130]. Importantly, HIV-1 Tat is actively released from infected cells into circulation in HIV positive individuals and can be detected in the sera of patients undergoing HAART [110,143] In recent years, numerous studies have found that HIV-1 Tat does not exclusively bind to and interact with HIV-1 genes, but also to cellular genes and factors, as discussed in Chapter 1. Preliminary data from our research group has shown that HIV-1 Tat enhances the promoter activity of *c-MYC* in a dose-dependent manner[144]. Here, we aimed to confirm this activation and to uncover the mechanisms driving it. In silico analyses was employed to identify putative transcription factor (TF) binding elements within the ~2 kbp region upstream of the transcription start site (TSS) referred to as the full-length (FL) *c-MYC* promoter. Based on the location of these TF binding sites (TFBS), promoter deletion constructs were produced and used in dual-luciferase reporter assays to identify the minimal promoter region mediating the activation by HIV-1 Tat. Thereafter, site-directed mutagenesis (SDM) was used to interrogate the role of specific TFBS. This was followed by co-immunoprecipitation (Co-IP) assays, to identify protein-protein interactions, and Chromatin Immunoprecipitation (ChIP) assays, to assess protein-DNA interactions.

2.2. Methods

2.2.1 Cell lines and culture conditions

The human HT1080 cell line was cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Sigma Aldrich, USA) supplemented with 10% foetal bovine serum (FBS) (Invitrogen, USA) and 1% penicillin/streptomycin (Sigma Aldrich, USA). The human BL cell line Ramos was used in this study. Ramos is an EBV negative cell line derived from sporadic BL that has been widely used in the literature to study this cancer. Ramos cells were obtained from ATCC (Manassas, VA) and grown in Roswell Park Memorial Institute (RPMI-1640) media (Sigma Aldrich, USA) supplemented with 10% FBS (Sigma Aldrich, USA) and 1% penicillin/streptomycin (Sigma Aldrich, USA) or 20% FBS post electroporation. Cells were incubated in a 5% CO_2 humidified incubator at 37°C. Mycoplasma testing was done periodically to ensure cells were not contaminated. Testing was done using bisBenzimide H 33342 trihydrochloride (14533, Sigma Aldrich, USA) DNA staining.

2.2.2 Electroporation of plasmid DNA into BL cells

Ramos cells were counted at the log phase of growth and 4 x 10^6 cells were transferred to 0.4 cm electroporation cuvettes at a total volume of 450 µL RPMI-1640 media without supplementation. 15 µg of either pcDNA-Tat or pcDNA3.1 empty control (both plasmids were kindly donated by Professor Mitra from the National Center for Cell Sciences in India) was added to the cuvettes and electroporated using the Gene Pulser Xcell system (Bio-Rad, USA) with the following conditions: exponential wave, 0.28kV and 950 µF. Cells were incubated on ice for 10 minutes followed by gentle resuspension and transfer to prewarmed RPMi-1640

media supplemented with 20% FBS (Sigma Aldrich, USA) and 1% (Sigma Aldrich, USA) penicillin/streptomycin. Cells were incubated for 24-48 hrs before protein was harvested for western blot analysis or RNA for qRT-PCR.

2.2.3 Western blot analyses

Electroporated cells were centrifuged at 1000rpm for 5 minutes followed by two washes in 1x PBS (Appendix A). Pelleted cells were resuspended in 2x Laemmli buffer (Appendix A) and boiled for 5 min at 95°C. Protein was separated using 12% SDS-PAGE gels (Appendix A), transferred to nitrocellulose membranes (Bio-Rad, USA) using the Mini-PROTEAN 3 casting apparatus (Bio-Rad). The following antibodies were used to detect target proteins: primary antibodies used were anti-c-MYC (SC-764, Santa Cruz Biotechnology, USA; 1:1000) and anti-p38 (M0800, Sigma-Aldrich, USA; 1:5000). Secondary antibodies used were Goat Anti Rabbit (H+L) HRP conjugate (170-6515, Bio-Rad, USA; 1:5000) and Goat Anti Mouse (H+L) HRP conjugate (170-6516, Bio-Rad, USA; 1:5000). Densitometric analysis of the signal intensity of bands was done using the ImageJ software (NIH, USA).

2.2.4 *c-MYC* Promoter sequence analysis

The FL promoter (-2324 bp - +537 bp) of the *c-MYC* gene cloned within the pGL3-Basic luciferase reporter vector (kind gift of Professor Xing Dai from the University of California USA) was sequenced for verification and analysed using the online tool PROMO (Algorithmics and Genetics Group, Universitat Politecnica de Catalunya http://alggen.lsi.upc.es/cgibin/prom-v3/promo/promoinit.cgi?dirDB=TF_8.3) to identify potential TF binding sites [145,146].

2.2.5 Generation of *c-MYC* promoter deletion constructs

To determine the TF sites responsible for mediating the observed increased promoter activity under the effect of HIV-1 Tat, deletion constructs of the *c-MYC* promoter were designed to gradually remove promoter regions using primers with incorporated restriction enzyme sites.

2.2.5.1 Generation of deletion fragments using PCR

PCR was performed with the MyTaq^{TM} DNA polymerase kit (Bioline, USA) using forward and reverse primers (with incorporated restriction enzyme (RE) sites) using the plasmid containing the full-length promoter of the *c-MYC* gene (pGL3-cMYC kindly donated by Professor Hiroyuki Suzuki at the University of Tsukuba, Japan) as the template, to generate sequential deletions. Primers used to generate deletion constructs are listed in **Table 2.1** with inserted RE sites underlined and in bold. PCR conditions are listed in **Table 2.2**.

Table 2.1: c-MYC promoter cloning primers.

Primer	Sequence	Tm	GC%	RE site
Forward 1	5'-GGGAGCA**GAGCTC**TCATGTGTGGG-3'	58.2°C	62.5%	**SacI**
Forward 2	5'-GGCGCAAA**GAGCTC**TTGTCTCTTCTG-3'	62.2 °C	53.8%	**SacI**
Forward 3	5'-CTAGAGC**GAGCTC**GCTCGGCTGCC-3'	62.3°C	62.5%	**SacI**
Reverse	5'-CC**AAGCTT**GCTACTCTGCAGGTCG-3'	60.6°C	58.3%	**HindIII**

Table 2.2: PCR cycling conditions.

Program	Temperature	Time	Cycles
Initial denaturation	95°C	1 min	1
Denaturation	95°C	15 sec	
Annealing	54°C	15 sec	30
Extension	72°C	40 sec	
Final extension	72°C	8 min	1

2.2.5.2 Restriction enzyme digest

To incorporate the deletion fragments into the pGL3 Basic plasmid, the PCR products were first ligated into the p-GEM-T-Easy (Promega, USA) plasmid to facilitate restriction enzyme digestion. The pGEM-T-Easy plasmids with the shortened promoter inserts and the pGL3 Basic plasmid were digested with the SacI (GAGCTC) and HindIII (AAGCTT). Digested fragments were then electrophoresed on a 1% agarose gel (Appendix A) at 70V for 1 hour. The bands of interest were cut out of the gel, purified using the MinElute Gel extraction kit (Qiagen, Germany) and quantified using the Nanodrop1000 (Thermo Fisher Scientific, USA).

2.2.5.3 Promoter fragment ligation within pGL3 Basic vector

Promoter fragments were ligated within the linearized pGL3 Basic plasmid using T4 ligase (Promega, USA). A molar ratio of 4:1 (insert to vector) was used in the ligation and the reaction was set up using the formula below.

$$ng\ insert = \frac{ng\ vector\ \times Kbp\ insert}{Kbp\ vector}$$

2.2.5.4 Competent bacterial cells

E. coli (strain DH5α) cells were made competent using the $CaCl_2$ method [147]. Briefly, bacterial cells were grown in 200 mL lysogeny broth (LB) (Appendix A) for 3 hrs until the culture reached an O.D of 0.5-0.8. Cells were then pelleted at room temperature (RT) and resuspended in 30 mL of ice-cold 100 mM $CaCl_2$ (Appendix A), followed by incubation on ice for an hour. The cells were then pelleted at 4°C and resuspended with 1 mL of ice-cold 100 mM $CaCl_2$ and stored at 4°C for up to two weeks or stored at -80°C for long term storage with the inclusion of 25% sterile glycerol.

2.2.5.5 Bacterial transformation and colony PCR

The propagation of ligated pGL3 plasmids was done through bacterial transformation and expansion according to protocols outlined in Sambrook et al (2001). Briefly, the ligation reaction was added directly to 100 µL of competent cells and incubated on ice for 30 min followed by heat-shocking at 37°C for 5 min. Transformed bacterial cells were then allowed to grow in 500 µL of LB broth 2 hrs at 37°C with gentle agitation. The transformed cells were thereafter plated on LB Agar plates (Appendix A) containing ampicillin (100mg/ml) and incubated overnight at 37°C. To test for positive transformants, colony PCR was performed using MyTaq™ DNA polymerase (Bioline, USA) and vector-specific primers (**Table 2.3**) using the conditions outlined in **Table 2.4**.

Table 2.3: pGL3 Basic specific primers used in colony PCR.

Primer	Sequence (5'-3')	Tm
RVprimer3	5'-CTAGCAAAATAGGCTGTCCC-3'	53°C
GLprimer2	5'-CTTTATGTTTTTGGCGTCTTCCA-3'	55°C

Table 2.4: Colony PCR cycling conditions.

Program	Temperature	Time	Cycles
Initial denaturation	95°C	1 min	1
Denaturation	95°C	15 sec	
Annealing	47°C	15 sec	30
Extension	72°C	40 sec	
Final extension	72°C	8 min	1

2.2.5.6 Plasmid isolation from bacterial cells

The purifications of plasmids were done using the PureYield™ Plasmid Maxiprep System (Promega, USA). Briefly, 200 mL of LB broth was inoculated with a 1 mL overnight DH5α culture carrying the plasmids of interest. The 200 mL of LB broth (Appendix A) was incubated at 37°C overnight in gentle agitation. Following incubation, cells were pelleted at 3000 rpm for 10 min at RT and the PureYield™ Plasmid Maxiprep System (Promega, USA) was used to isolate pure and endotoxin-free plasmids. These were quantified using a Nanodrop1000 and the plasmid integrity assessed using gel electrophoresis.

2.2.6 Dual-luciferase reporter assays

Dual-luciferase reporter assays were conducted in HT1080 cells seeded at a density of 5x 10^4 cells/mL in 6-well cell culture dishes. 400 ng of each luciferase reporter construct (full length (pGL3-cMYC WT) or deletions of the *c-MYC* promoter) were co-transfected, using X-tremeGene^{TM} HP transfection reagent (Sigma Aldrich, USA) with up to 500 ng of pcDNA-Tat or the corresponding empty vector (pcDNA3.1 empty) as well as 50 ng the pRL-TK vector, kindly provided by Professor Sharon Prince from the University of Cape Town, driving the expression of a Renilla reporter (used as a control for transfection efficiency). All transfections were done in duplicates and Luciferase assays were repeated at least three times (n=3). Thirty (30) hrs post-transfection, whole-cell extracts were assayed for firefly and Renilla luciferase activity using the dual-luciferase reporter assay system (Promega, USA). Luciferase activity was measured using the GloMax®-Multi+ Luminescence Module (Promega, USA). Luciferase values were expressed relative to empty vector control values.

2.2.7 Site-directed mutagenesis

Forward and reverse primers containing the desired DNA base pair modifications (underlined and in bold in **Table 2.5**) were generated to disrupt AP-1 binding sites (**Figure 2.1**). The KAPA HiFi HotStart ReadyMix system (Roche, Germany) was used to amplify the pGL3-cMYC plasmid with the following cycling conditions outlined in **Table 2.6.**

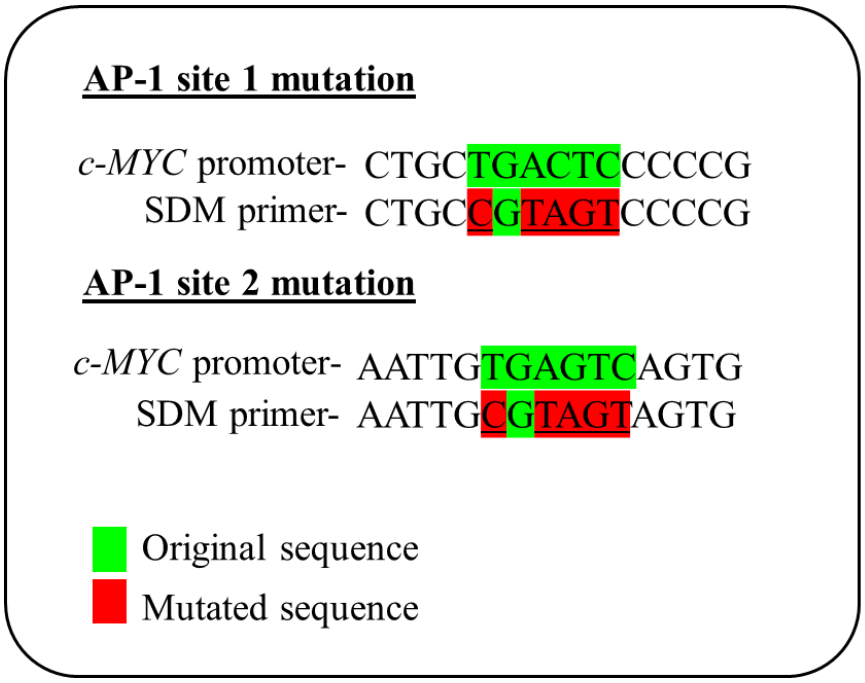

Figure 2.1: AP-1 sites on the *c-MYC* promoter targeted for mutation by SDM. Primers were designed to abrogate the two AP-1 sites located at positions -1128 bp and -1375 bp relative to the TSS present on the pGL3-cMYC WT plasmid. The resulting constructs were used in luciferase assays to determine mutating these sites resulted in the loss of promoter activity in the presence of HIV-1 Tat.

Table 2.5: SDM primers used to disrupt AP-1 binding sites on the c-MYC promoter.

Primer	Sequence	Tm	GC%
AP-1 Mut1 F	5'-CACAAGGGTCTCTGC**C**G**TAGT**CCCCGGCTCGGTCCACAAG-3'	78.8 °C	65%
AP-1 Mut1 R	5'-CTTGTGGACCGAGCCGGGG**ACTA**C**G**GCAGAGACCCTTGTG-3'	78.8 °C	65%
AP-1 Mut2 F	5'-CAGAAAAAATTG**C**G**TAGT**AGTGAACTAGGAAATTAATGCCTGGAAGGC-3'	73.2 °C	39.6%
AP-1 Mut2 R	5'-GCCTTCCAGGCATTAATTTCCTAGTTCACT**ACTA**C**G**CAATTTTTTCTG-3'	73.2 °C	39.6%

Table 2.6: SDM PCR cycling conditions.

Program	Temperature	Time	Cycles
Initial denaturation	95°C	3 min	1
Denaturation	98°C	20 sec	
Annealing	70°C	15 sec	15
Extension	72°C	8 min	
Final extension	72°C	8 min	1

2.2.8 Co-immunoprecipitation assay

HT1080 cells (5 x 10^4 /mL) were plated in 6-well cell culture plates. After 24 hrs, the cells were co-transfected with 500 ng of pcDNA-Tat and 500 ng pCMV-JunB plasmids with the XtremeGene-HP reagent (Sigma Aldrich, USA) using a 1:3 ratio of DNA (ng) to transfection reagent (µL) respectively. Cells were harvested after 24 hrs and lysed using RIPA buffer (Appendix A) with protease combined with 7x complete™, Mini, EDTA-free, Protease Inhibitor (Roche, Switzerland). Cell lysates were incubated overnight at -80 °C to optimize lysis. The following day cell lysates were thawed on ice and the protein fraction was isolated by removal of cell debris using centrifugation at 12000 xg for 10 min at 4°C. Protein samples were precleared with 50 µL protein A agarose beads (Roche, Switzerland) for 3 hrs at 4°C. the Pierce® BCA Protein Assay kit (Thermo Fisher Scientific, USA) was used to quantify the protein and 200 µg of protein was used per pulldown. 2 µg of primary antibody (anti-JunB (SC-8051, Santa Cruz Biotechnology, USA), anti-HIV-1 Tat (NIH AIDS reagent program (2A4.1 4373) or ab43014, Abcam, UK) and negative control IgG (170-6515, Bio-Rad, USA) were each added to individual tubes containing the 200 µg of the protein and incubated overnight at

4°C. The following day 50 µL protein A agarose beads were added to the protein-antibody complexes and incubated for 4 hrs at 4°C. Beads were collected by centrifugation at 2000 rpm and washed twice in ice-cold 1x PBS (Appendix A) containing protease inhibitors. Beads were resuspended in 2x Laemmli buffer (Appendix A) and boiled for 5 min at 95°C followed by SDS-PAGE (Appendix A) separation and western blot analysis as outlined in section 2.2.3, with appropriate antibodies.

2.2.9 Chromatin Immunoprecipitation assays

HT1080 cells (5×10^4/mL) were plated in 6-well cell culture plates. After 24 hrs, the cells were co-transfected with 500 ng of pcDNA-Tat and 500 ng of pCMV-JunB plasmids with the XtremeGene-HP reagent (Sigma Aldrich, USA) using a 1:3 ratio of DNA (ng) to transfection reagent (µL) respectively. The DNA was cross-linked using 1% formaldehyde for 10 min at RT, quenched with 125 mM glycine (Appendix A) for 5 min at RT. Cells were centrifuged at 2000 xg for 5 min at 4°C and resuspended in 15 mL ice-cold 1x PBS (Appendix A), centrifugation was repeated as before and cells were resuspended in 1.5 mL ice-cold 1x PBS. Cells were centrifuged at 1500 xg for 5 min at 4°C and washed in 1 mL NCP buffer I (Appendix A), centrifuged again followed by a wash in 1 mL NCP buffer II (Appendix A). Cells were pelleted as before and resuspended in ChIP lysis buffer with protease inhibitors (Appendix A) and allowed to lyse on ice for 10 min. The lysate was sonicated to obtain chromatin fragments between 300-500 bp in length (as assessed by gel electrophoresis). 480 µL of Sonicated DNA was added to 1 mL of RIPA buffer containing protease inhibitors (Appendix A) and cleared with 90 µL protein A agarose beads (Roche, Germany) for 4 hrs at 4°C in an orbital shaker. The precleared lysate was split into equal amounts in individual 1.5 mL Eppendorf tubes each containing 2 µg of the appropriate primary antibody (anti-JunB (Santa Cruz Biotechnology,

USA), anti-HIV-1 Tat (Abcam, UK) and negative control IgG (Bio-Rad, USA) and were incubated overnight at 4°C in an orbital shaker. 50 µL of protein A agarose beads were added to each tube and incubated at 4°C in an orbital shaker for 3 hrs. Beads were collected by centrifugation at 2000 xg for 5 min and washed with wash buffers 1-4 (Appendix A), twice for each buffer, in 4°C in an orbital shaker for 10 min. DNA bound by antibodies was extracted by incubation in Extraction buffer (Appendix A) for 15 min at RT followed by overnight incubation at 65°C. DNA was then purified using Phenol-chloroform extraction[148]. DNA enrichment was analysed through qRT-PCR on the Rotor Gene-Q (Qiagen, Germany) using the 2x QuantiTect SYBR Green PCR Master Mix (Qiagen, Germany). *c-MYC* promoter-specific primers were used (outlined in **Table 2.7**) with the PCR cycling conditions in **Table 2.8**.

Table 2.7: ChIP primers.

Primer	Sequence	Tm	GC%
Set 1 F	5'-GGAATTAAACGTCCGGTTTGTC-3'	55.9°C	52.4%
Set 1 R	5'- GGCAAGTGGAGAGCTTGTG-3'	55.4°C	55.6%
Set 2 F	5'- GCAACTAGCTAAGTCGAAGCG-3'	54.4°C	45.5%
Set 2 R	5'- GGCAAGTGGAGAGCTTGTG-3'	56.3°C	57.9%

Table 2.8: ChIP qRT-PCR cycling conditions.

Program	Temperature	Time	
Initial denaturation	95°C	3 min	1
Denaturation	95°C	5 sec	40
Annealing/extension	55°C	20 sec	
Final extention	55°C	8 min	1

2.2.10 Statistical analyses

Student's t-tests and one-way analysis of variance (ANOVA) tests were used. Significance was accepted at $p < 0.05$, Tukey's *post hoc* test was done for multiple comparisons. Statistical significance was determined using the GraphPad PRISM version 7 for Windows, GraphPad Software, San Diego California USA.

2.3 Results

2.3.1 HIV-1 Tat enhances the activity of the FL *c-MYC* promoter and increases c-MYC protein expression in the BL cell line Ramos

Previous preliminary work done in the laboratory had shown that HIV-1 Tat can increase the activity of the full-length *c-MYC* promoter in a dose-dependent manner [144]. To confirm this preliminary finding, the luciferase reporter assays were repeated. A significant increase, of more than 3 folds, in the activity of the promoter was observed in the presence of HIV-1 Tat when compared to the control (**Figure 2.2A)** confirming the previous results. To determine if the observed increase in *c-MYC* promoter activity, due to HIV-1 Tat, translates to an increase in the protein expression level, western blot analysis was performed in the BL cell line Ramos in which HIV-1 Tat was ectopically expressed. Due to the recalcitrant nature of these cell lines to conventional transfection techniques using commercially available transfection reagents, electroporation was utilised to deliver plasmid DNA into the cells. Ramos cells were electroporated with either pcDNA-Tat or pcDNA3.1 empty control vectors and protein was harvested 24 hrs post-electroporation. Western blot analysis showed that Ramos cells expressing HIV-1 Tat had elevated c-MYC protein expression compared to control (**Figure 2.2B**). To ensure that plasmid delivery was being achieved, the cells were also independently electroporated with the pEGFP-N3 vector, a plasmid that constitutively expresses EGFP, as a positive transfection control using the same parameters and plasmid concentrations as in the pcDNA-Tat electroporation. The expression of EGFP was indeed detected in the cells, albeit at low transfection efficiency (**Figure 2.2 C&D**).

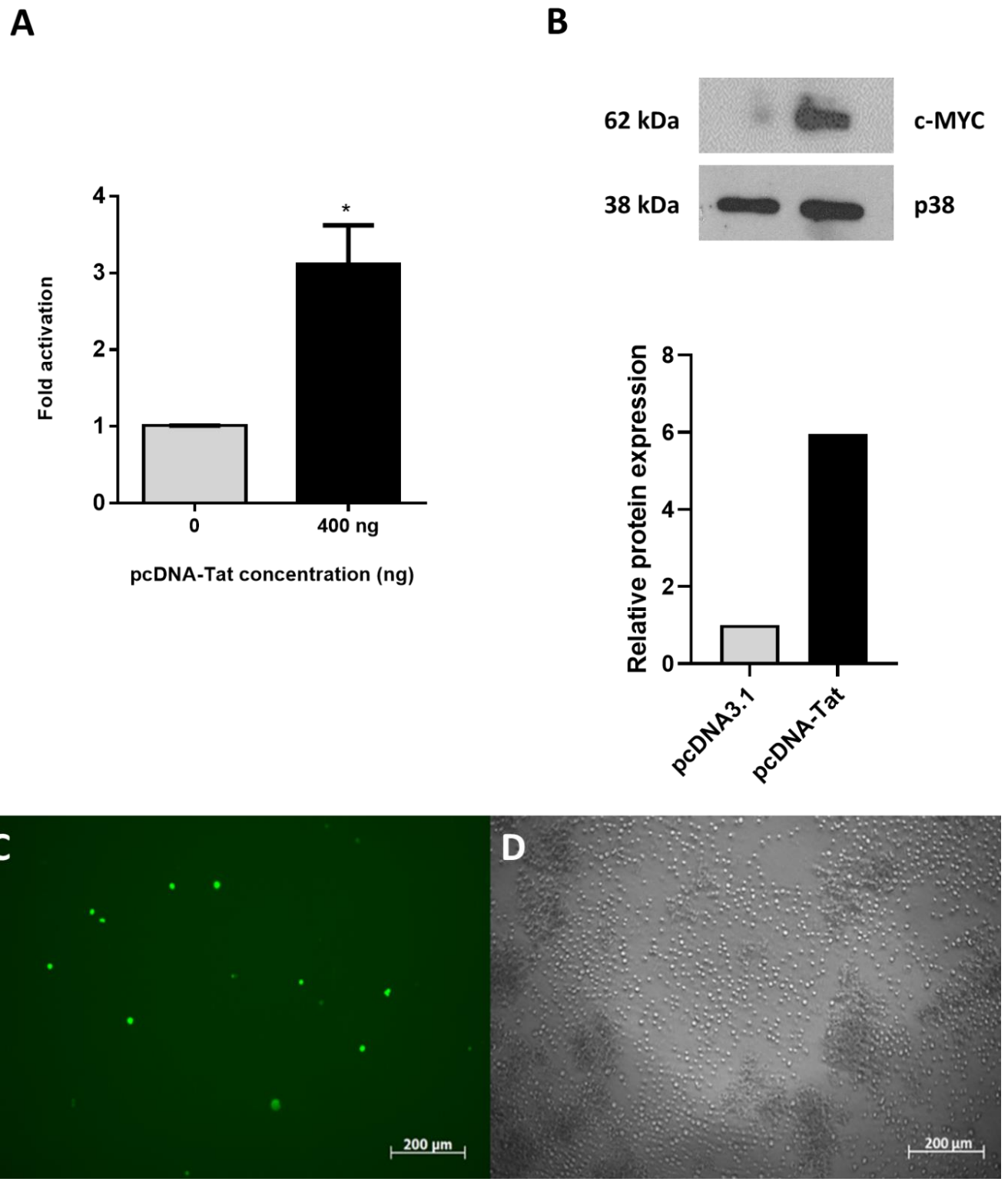
A
Fold activation
4
3
2
1
0
*
0
400 ng
pcDNA-Tat concentration (ng)
B
62 kDa
c-MYC
38 kDa
p38
Relative protein expression
8
6
4
2
0
pcDNA3.1
pcDNA-Tat
C
200 µm
D
200 µm

Figure 2.2: Luciferase reporter assay confirming that HIV-1 Tat enhances the activity of the *c-MYC* WT promoter and western blot showing increased c-MYC protein in BL cells expressing HIV-1 Tat. A) HT1080 cells were transfected with 500 ng of the *c-MYC* WT promoter (pGL3-c-MYC) along with 400 ng of the HIV-1 Tat expression vector (pcDNA-Tat) or 400 ng of the pcDNA3.1 empty vector as control. Fold activation was obtained by comparing the relative luciferase units (RLU) of the HIV-1 Tat transfected groups to their empty transfected controls. Statistical analysis was done using GraphPad PRISM 7 with a two-tailed Student's t-test. This experiment was performed at least in triplicate. B) Ramos cells were electroporated with either 15 µg pcDNA-Tat or 15 µg pcDNA3.1 empty vector control. Protein was harvested and western blot was done using anti-c-MYC antibody and the anti-p38 antibody as a loading control. The bar graph represents densitometric analysis (ImageJ). Fold increase was obtained by comparing the relative densitometric units of the HIV-1 Tat electroporated cells to the negative control cells, which were set to 1, after normalising to loading control (p38). Western blot is representative of at least two independent experiments. C) EGFP expressing cells. Ramos cells were electroporated with 15 µg of the pEGFP-N3 plasmid to serve as a positive control for electroporation. D) Phase contrast of the same field of vision as C. Scale bars represent 200 µm.

2.3.2 PROMO analysis identifies multiple putative TFBS within the *c-MYC* promoter

Using the online software PROMO, a comprehensive analysis of the FL *c-MYC* promoter was performed to identify putative transcription factor binding sites (TFBS) [145,146]. A dissimilarity rate was set at 15% to not dismiss potential binding sites. TFs with previously reported associations, directly or indirectly, with HIV-1 Tat, or which have been implicated in BL pathobiology were annotated. Sites were mapped onto the promoter sequence and the result is displayed in **Figure 2.3**. This analysis was done to identify TFBS which may potentially mediate the response observed in **Figure 2.2A** and to guide in the development of the sequential

promoter deletion constructs. These deletion constructs were then produced and used in luciferase assays.

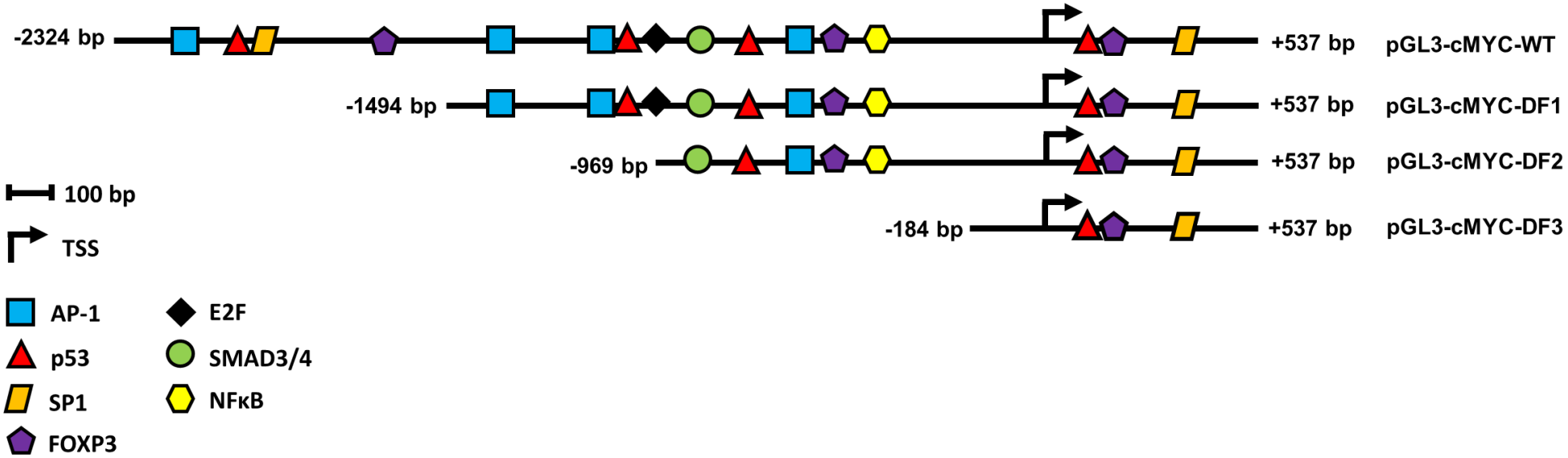

Figure 2.3: Putative transcription factor binding sites located within the *c-MYC* FL promoter, as identified using the PROMO software. The FL *c-MYC* promoter (-2324 bp to +537 bp) construct was analysed for TFBS using PROMO. Sites are represented as different coloured shapes which are labelled in the key and the transcription start site (TSS) is denoted by a black arrow. Sequential deletion constructs were designed to remove specific TFBS. Shortened promoter constructs are labelled accordingly on the right side of the diagram (pGL3-cMYC-DF1, pGL3-cMYC-DF2 and pGL3-cMYC-DF3)

2.3.3 HIV-1 Tat enhances the *c-MYC* promoter activity via AP-1 binding sites.

HIV-1 Tat's ability to directly or indirectly interact with the promoter regions of host genes has been described in the literature [132,133]. To identify the minimal promoter region required for HIV-1 Tat-mediated promoter activation three sequential promoter deletion constructs were generated using the full-length *c-MYC* promoter as a template, with each deletion resulting in loss of specific TF binding sites (**Figure 2.3**). Luciferase assays were performed using the full-length and deletion constructs and it was observed that the loss of the region spanning -1494 bp to -969 bp (pGL3-cMYC-DF2 construct) led to a significant reduction in promoter activation in the presence of HIV-1 Tat (**Figure 2.4A**). A loss of 35% in activity was observed from 2.6-fold in WT to 1.7-fold in the DF2 construct. Additionally, when compared to the empty control, the promoter activities of both the DF2 and DF3 constructs were not significantly higher than the WT and DF1 constructs.

The region spanning -1494 bp to -969 bp contains two AP-1 sites as well as one p53 and one E2F site. Further inspection revealed that the two AP-1 sites (AP-1 site 1 at position -1128 and AP-1 site 2 at position -1375) had favourable dissimilarity scores indicating that these sites are potentially functional *in vivo*, with site 2 being previously reported to induce *c-MYC* transcription upon Platelet-Derived Growth Factor (PDGF) receptor stimulation [149]. Furthermore, the AP-1 factor c-Jun has been shown to enhance the HIV-1 Tat-mediated transcription of the viral LTR [150]. These observations provided confidence that AP-1 sites 1 and/or 2 could be relevant in mediating the promoter activation. To test this, SDM was used to sequentially mutate these sites, as well as in combination. Luciferase assays were then repeated using FL, WT and mutated constructs. A significant loss in promoter activity of ~20% was observed when AP-1 site 1 was mutated (**Figure 2.4B**) and ~25% when AP-1 site 2 was

mutated (**Figure 2.4C**). Furthermore, when both sites were simultaneously abolished, a decrease of ~29% was observed (**Figure 2.4C**). These results indicated that indeed, HIV-1 Tat-mediated activation of the *c-MYC* promoter via these sites, albeit partially.

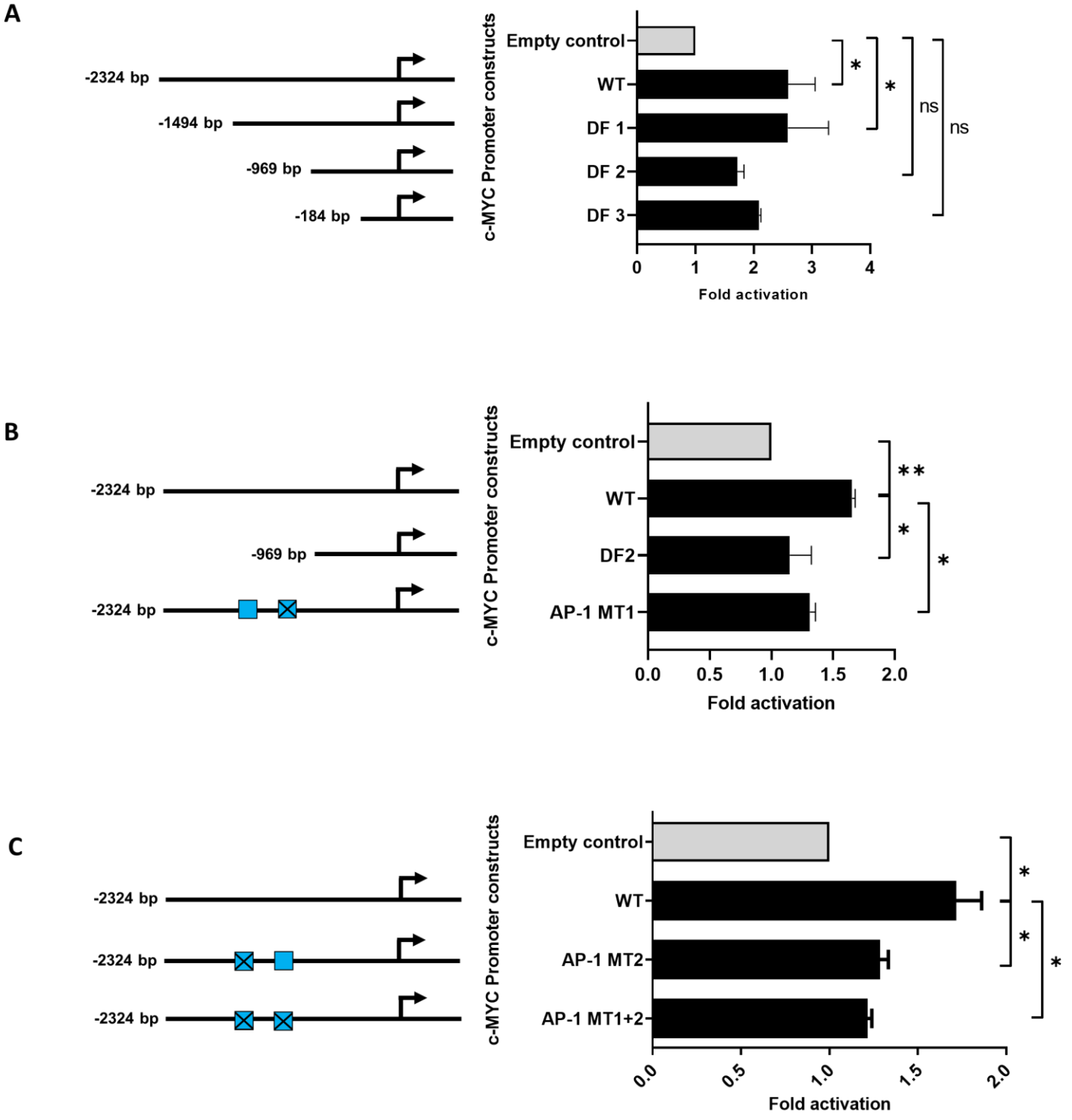
A
-2324 bp
-1494 bp
-969 bp
-184 bp
c-MYC Promoter constructs
Empty control
WT
DF 1
DF 2
DF 3
*
*
ns
ns
0
1
2
3
4
Fold activation
B
-2324 bp
-969 bp
-2324 bp
c-MYC Promoter constructs
Empty control
WT
DF2
AP-1 MT1
**
*
*
0.0
0.5
1.0
1.5
2.0
Fold activation
C
-2324 bp
-2324 bp
-2324 bp
c-MYC Promoter constructs
Empty control
WT
AP-1 MT2
AP-1 MT1+2
*
*
*
0.0
0.5
1.0
1.5
2.0
Fold activation

Figure 2.4: The AP-1 sites located at positions -1128 bp and -1375 bp partially mediate the HIV-1 Tat enhancement of *c-MYC* promoter activity. A) Luciferase reporter assay comparing the fold activation of the *c-MYC* WT promoter to the deletion constructs (DF1-DF3) in the presence of HIV-1 Tat. HT1080 cells were transfected with 500ng of the *c-MYC* WT promoter (pGL3-cMYC) or deletion constructs (pGL3-cMYC-DF1, pGL3-cMYC-DF2 and pGL3-cMYC-DF3) along with 400ng of the HIV-1 Tat expression vector (pcDNA-Tat) or 400ng of the empty pcDNA3.1 as control. B) Luciferase assay comparing fold activation in the presence of HIV-1 Tat between the WT promoter, DF2 and AP-1 MT1 constructs. C) Luciferase assay comparing fold activation in the presence of HIV-1 Tat between the WT promoter, AP-1 MT2 and AP-1 MT1+2 constructs. Fold activation was obtained by comparing the relative luciferase units (RLU) of the pcDNA-Tat transfected groups to their empty transfected controls, which were set to 1. AP-1 mutated sites are represented as striated blocks. Statistical analysis was done using one-way ANOVA in GraphPad PRISM 7. Significance *($p<0.05$), **($p<0.005$) and ns (not significant). The graphs are representative of at least three separate repeats.

2.3.4 Co-immunoprecipitation assay demonstrates strong protein to protein interaction between HIV-1 Tat and the AP-1 factor JunB

Previous studies have shown that HIV-1 Tat can form complexes with cellular proteins including transcription factors. For instance, HIV-1 Tat was found to form a complex with NFAT and c-Jun at an NFAT/AP-1 composite site [151]. In the context of this study, whether HIV-1 Tat mediates its effect through direct binding and/or interaction with cellular transcription factors, is not known. However, there is evidence in the literature that supports both scenarios [135,152]. To investigate whether HIV-1 Tat can form a complex with AP-1 factors, a co-immunoprecipitation assay was performed, where HT1080 cells were transfected with pcDNA-Tat and pCMV-JunB expression constructs. The Jun family of proteins are a major component of AP-1 complexes and have been associated with HIV-1 Tat [150,153,154]. A pull-down assay was performed using antibodies specific to HIV-1 Tat or JunB followed by western blot analyses. The result indicates that HIV-1 Tat can form strong associations with JunB as

the protein was detected in an immunoprecipitation assay where an HIV-1 Tat specific antibody was used to pull down proteins followed by a western blot with JunB antibody (**Figure 2.5A**). Similarly, HIV-1 Tat was detected in an immunoprecipitation assay where the JunB specific antibody was used to pull down proteins followed by a western blot with an HIV-1 Tat antibody. However, in this case, HIV-1 Tat was detected in the negative control sample (**Figure 2.5B**). The Co-IP was repeated with a different HIV-1 Tat antibody (ab43014, Abcam, UK) to eliminate the positive HIV-1 Tat signal observed in the IgG pulldown in **Figure 2.5B**. The HIV-1 Tat pulldown with the ab43014 antibody successfully enriched the JunB protein (**Figure 2.5C)**, however when performing the pulldown with JunB antibody and a western blot with the ab43014 antibody, no HIV-1 Tat protein could be detected (**Figure 2.5D**). These results indicate that the JunB antibody, while competent in western blot assays, is not compatible with immunoprecipitation. Nevertheless, we found confidence in this result by the fact that two independent anti-Tat antibodies were able to successfully pull down JunB in separate assays and can therefore conclude that HIV-1 Tat and JunB can exist within the same complex.

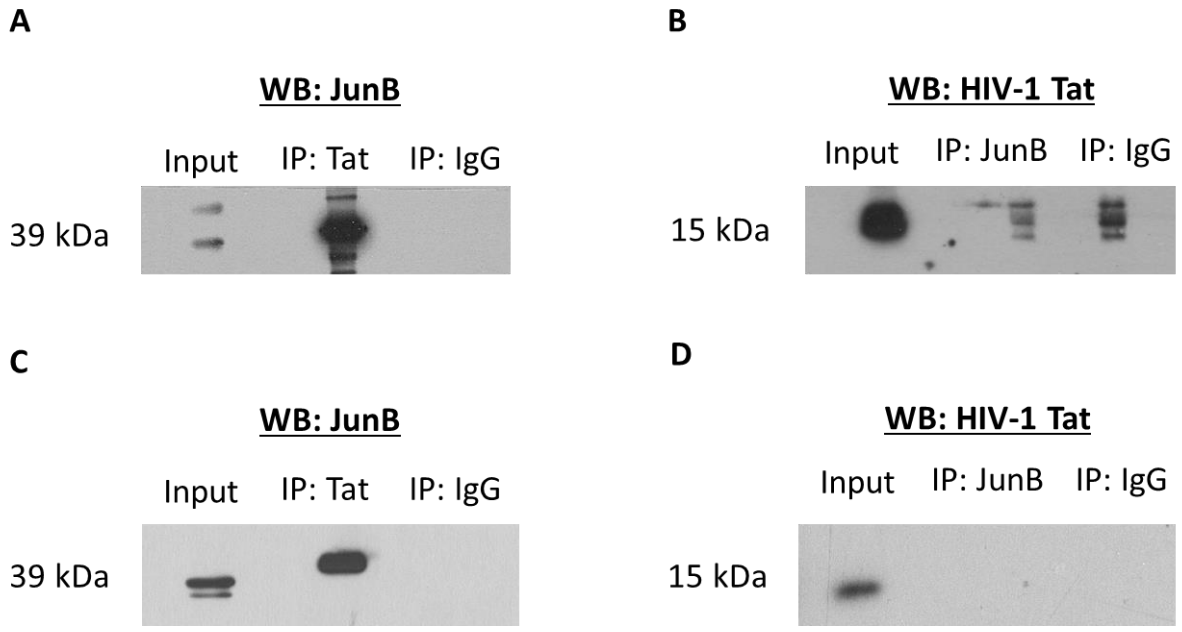

Figure 2.5: Co-immunoprecipitation assay shows binding between HIV-1 Tat and AP-1 factor JunB. HT1080 cells were transfected with both 500 ng HIV-1 Tat expression vector (pcDNA-Tat) and 500 ng JunB expression vector (pCMV-JunB). A&C) The cell lysate was incubated with anti-HIV-1 Tat antibody (Aids reagent program, NIH, USA used in A or ab43014, Abcam, UK used in C) and pulled down using protein A agarose beads (Roche, Germany). Western blot was carried out using an anti-JunB antibody (Santa Cruz, USA) to detect JunB in the HIV-1 Tat pull-down. B&D) The cell lysate was incubated with anti-JunB antibody (Santa Cruz, USA) and pulled down using protein A agarose beads (Roche, Germany). Western blot was carried out using an anti-HIV-1 Tat antibody (Aids reagent program, NIH, USA used in B or ab43014, Abcam, UK used in D) to detect HIV-1 Tat in the JunB pull down. In both western blots, the goat-anti-rabbit secondary antibody (Bio-Rad, USA) was used as a negative control (lane 3 in all blots). The western blots representative of at least three separate repeats.

2.3.5 Chromatin immunoprecipitation reveals JunB and HIV-1 Tat occupancy at the AP-1 sites on the *c-MYC* promoter

To determine if the identified AP-1 sites were indeed functional and if JunB was able to bind these sites *in vivo* in the presence of HIV-1 Tat, ChIP assay with primers designed to amplify regions of the promoter that harbour the previously identified AP-1 sites were performed. The results show that both sites were enriched when pulling down with JunB antibody with a significantly pronounced enrichment found in AP-1 site 2. HIV-1 Tat binding to site 2 was also detected but at a lower level, and although not statistically significant, the result was reproducible, while site 1 showed no enrichment for HIV-1 Tat (**Figure 2.6**).

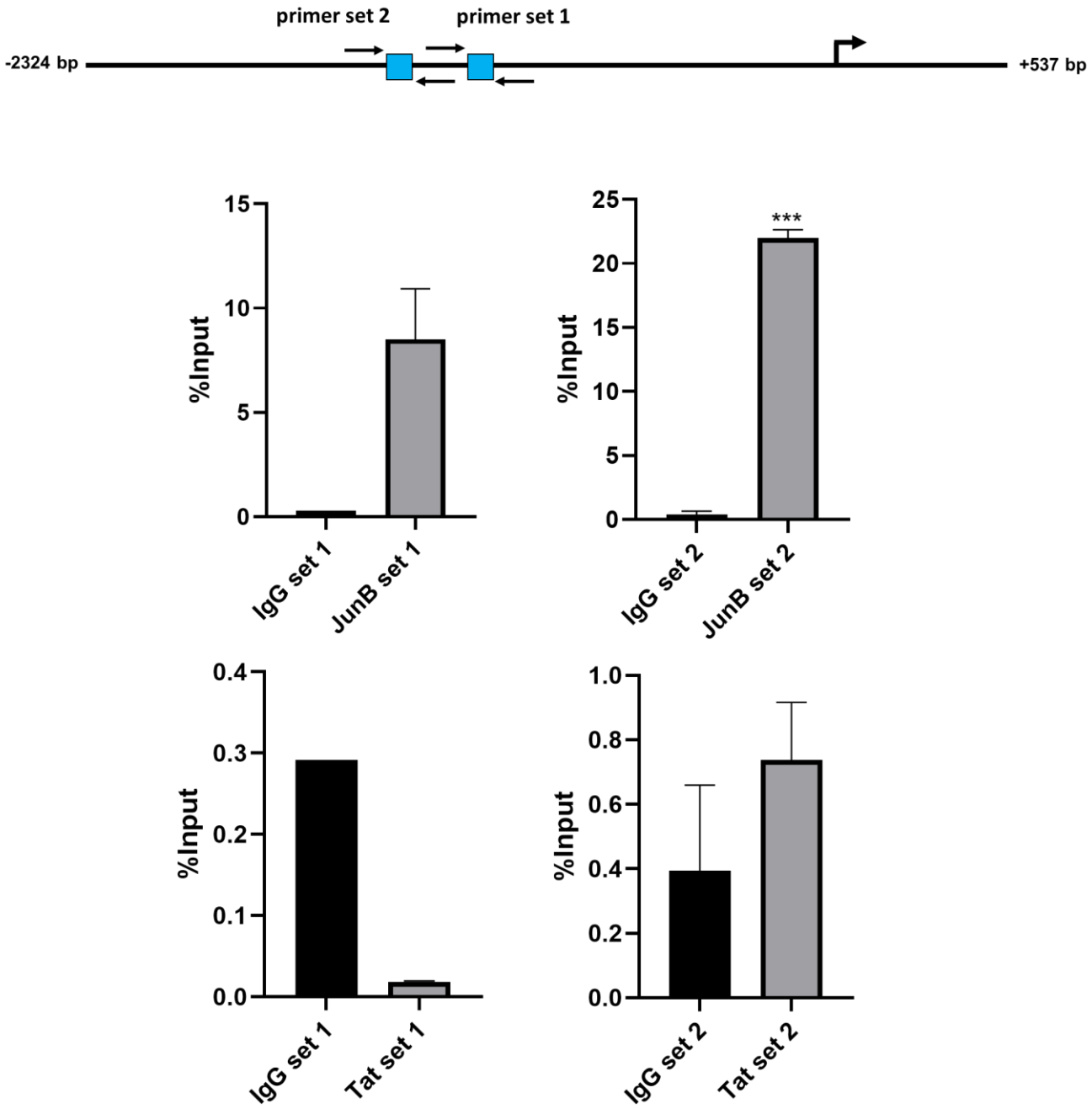

Figure 2.6: HIV-1 Tat and JunB proteins associate with the *c-MYC* promoter at AP-1 binding sites. Diagrammatic representation of the *c-MYC* promoter depicting the location of the two AP-1 sites in square blocks and the positions of the forward and reverse primers used for the ChIP assay are denoted by black arrows. Bar graphs show qRT-PCR data using immunoprecipitated DNA obtained from ChIP with either anti-JunB, anti-HIV-1 Tat or anti-IgG (negative control) antibodies and amplified with the primers depicted in the diagram Statistical significance determined using Student's t-test in GraphPad PRISM 7, ***($p<0,0005$). The graphs representative of at least two separate repeats.

2.4 Discussion

HIV-1 infection has been associated with the development of multiple cancers. Amongst these BL and DLBCL are particularly enriched in the HIV positive population [155,156]. Recent studies have brought attention to the oncogenic potential of HIV-1 and its associated proteins in promoting these cancers, however, the molecular mechanism driving these phenomena is as yet poorly defined. In this aspect of the study, we sought to establish the mechanism via which HIV-1 Tat influences the promoter activity of *c-MYC*. While overexpression of c-MYC in BL is primarily ascribed to its translocation to the *IgH* loci, hardly any studies have been performed on assessing any changes in expression of the intact *c-MYC* allele. Within the context of HIV-associated BL, this is particularly relevant since this subtype is more aggressive than the eBL and sBL subtypes, despite all three being considered as the same pathological entity.

The results presented here confirm preliminary assessments from our research group that the activity of a ~2 kbp promoter fragment of the *c-MYC* gene is enhanced in the presence of HIV-1 Tat. These findings are supported by work done by Germini et al (2017), where the authors showed that B cells exposed extracellularly to recombinant HIV-1 Tat protein had increased *c-MYC* mRNA expression[139]. The affinity of HIV-1 Tat for promoter regions of host genes has been described in the context of T-cell infection. Genome-wide studies have revealed that HIV-1 Tat binds to a variety of regions within the human T cell genome [132–134]. Notably, in Jurkat cells (acute T cell leukaemia cell line), HIV-1 Tat was found to preferentially occupy promoter and 5'-UTR regions in Jurkat cells. Importantly, these associations were conserved in RNA-Seq data showing that HIV-1 Tat not only binds to but also alters host gene expression, specifically genes involved in the cellular immune response. Other studies have shown that HIV-1 Tat forms protein-to-protein interactions with host TFs, including AP-1 factors, and

augments their binding to promoters of host genes [122,135,157]. Using promoter deletion luciferase reporter assays and SDM, we found that two AP-1 binding sites within the *c-MYC* promoter are important in mediating the HIV-1 Tat-dependent promoter increase. Of note, site 2 has previously been implicated in *c-MYC* promoter activation via observed binding of c-Jun and JunD, however, JunB demonstrated weak binding by EMSA and no binding by ChIP [149], contrary to what was observed in the dual-luciferase and ChIP assays performed in this study (**Figure 2.4** and **Figure 2.6** respectively). Furthermore, Co-IP showed that JunB and HIV-1 Tat form strong protein-to-protein interactions and ChIP assays demonstrated that JunB, in the presence of HIV-1 Tat, occupied the *c-MYC* promoter at both AP-1 binding sites with the highest enrichment shown at site 2. It is important to note that a complete decrease in fold activation was not observed with the promoter deletions or AP-1 site mutations, this could be due to other active sites compensating for the loss of the AP-1 sites, coupled with the potential promiscuous binding of HIV-1 Tat to the promoter. Future work is needed to clarify whether HIV-1 Tat utilizes other sites within the *c-MYC* promoter besides the two identified AP-1 sites.

Given that the activation of AP-1 factors through phosphorylation occurs downstream of the JNK pathway, it is plausible that HIV-1 Tat induces the JNK pathway directly, thereby increasing AP-1 factor binding to the *c-MYC* promoter. The effect of HIV-1 Tat on the induction of the JNK pathway in both T and B cells has previously been described [153,158]. Additionally, B cells exposed extracellularly to recombinant HIV-1 Tat have increased ROS production which could potentially induce the JNK pathway and lead to increased AP-1 phosphorylation [110]. Interestingly, a positive feedback loop between c-MYC and AP-1 factors has been proposed in BL, with increased c-MYC levels leading to increased AP-1 phosphorylation and binding to the Igκ Ei and E3' enhancers, potentially increasing expression of translocated *c-MYC* in *Igκ/c-MYC*-BL and possibly that of the unaffected allele on

chromosome 8 [159]. The increased transcriptional activity of the *c-MYC* gene due to HIV-1 Tat, coupled to its increased localization to the *IgH* locus and aberrant AID activity, could lead to enhanced proliferation and more aggressive disease in HIV positive people.

HIV-1 Tat binding to the human genome is greater than previously thought, and the regulatory role of HIV-1 Tat is not only observed in the viral genome but also in that of the host, with these interactions not limited to cells that are actively infected by HIV-1. Therefore, the possibility that Tat can lead to oncogenic transformation via altering cellular gene expression is significant and needs investigation. This study has shown further mechanistic evidence for HIV-1Tat induced *c-MYC* activation in the context of BL. Further work needs to be done to validate these findings *in vivo*. An interesting avenue for future research would be to determine if the observed findings in this study are downstream of the JNK pathway, activated by HIV-1 Tat. Furthermore, uncovering other interactions between HIV-1 Tat and the B cell genome will provide novel insights into the unique pathobiology of HIV-associated NHLs.

Chapter 3: The regulation of *AICDA*/AID expression by HIV-1 Tat in B cells

3.1 Introduction

AID plays a central role in the development of mature B-cell NHLs and HIV-1 infection has been linked to AID deregulation. Epeldegui et al (2007) showed that elevated AID expression in the peripheral mononuclear cells (PBMC) of HIV positive patients predated an eventual diagnosis of lymphoma. Notably, the most pronounced increase in AID levels was observed in patients subsequently diagnosed with BL [160]. This, coupled to the fact that HIV-1 proteins like HIV-1 Tat and Nef become internalized by bystander cells, including B cells, suggests the possibility of direct interference of *AICDA/AID* regulation by HIV-1 [137,161,162]. Work done by our research group showed that treating the BL cell line Ramos with extracellular recombinant HIV-1 Tat induced AID expression at the transcriptional level [144]. More recently, Sall et al (2019) showed that peripheral blood-derived B cells, from HIV positive patients, incubated with HIV-1 Tat recombinant protein had increased H3 acetylation in the promoter region of *AICDA*, indicative of active transcription [163]. The mechanism via which HIV-1 Tat leads to this increase remains as yet unclear. In a recent study, where the authors investigated the effect of HIV-1 Tat on astrocytes in the context of HIV-associated neurocognitive disorders (HAND), they found that HIV-1 Tat could inhibit specific cellular microRNAs, one of which was hsa-miRNA-181b [141]. This miRNA has been shown to directly bind to and repress the murine *Aicda* 3'UTR, and the expression of hsa-miRNA-181b was shown to be lower in BL as compared to human normal primary B cells [142]. This study therefore aimed at identifying the molecular interactions between HIV-1 Tat and the regulatory mechanism governing *AICDA*/AID expression at the transcriptional and posttranscriptional levels. To this end, the R1 promoter region was analysed for potential TFBS and sequential deletion constructs of the promoter were generated. Dual-luciferase assays were done using the full length (2141 bp) and deletion constructs of the R1 promoter, cloned upstream of the pGL3 luciferase reporter vector, in the

presence of HIV-1 Tat. *AICDA* transcription is also regulated by the R2 and R4 regions, therefore, dual-luciferase assays were also performed using these regulatory regions to further test if HIV-1 Tat alters their activity. Western blot was done to determine if the reported HIV-1 Tat-mediated increase of *AICDA* transcription was conserved at the protein level in BL cells and if increased levels of AID protein resulted in alterations to genomic stability as a result of AID perturbation. Furthermore, to investigate whether HIV-1 Tat is also involved in the post-transcriptional regulation of AID, via hsa-miRNA-181b-5p perturbation, electroporation of the BL cell line Ramos was performed followed by qRT-PCR to evaluate changes in hsa-miRNA-181b-5p expression. Lastly, dual-luciferase reporter assays were done to validate the repressive role of hsa-miRNA181b-5p on the human *AICDA* 3'UTR.

3.2 Methods

3.2.1 Cell lines and culture conditions

The Ramos and the HT1080 cell lines were cultured using the same methodology as described in Chapter 2 section 2.2.1.

3.2.2 *AICDA* R1 Promoter sequence analysis

The FL promoter (2141 bp) of the *AICDA* gene was sequenced followed by analysis using the online tool PROMO (Algorithmics and Genetics Group, Universitat Politecnica de Catalunya http://alggen.lsi.upc.es/cgibin/prom-v3/promo/promoinit.cgi?dirDB=TF_8.3) to identify potential TF binding sites [145,146].

3.2.3 Generation of *AICDA* R1 promoter deletion constructs

To determine the TF sites responsible for mediating the observed increased promoter activity under the effect of HIV-1 Tat, deletion constructs of the *AICDA* R1 promoter were designed to gradually remove promoter regions using primers with incorporated restriction enzyme sites.

3.2.3.1 Promoter fragment PCR

PCR was performed using forward and reverse primers (with incorporated restriction enzyme sites) using the plasmid containing the full-length R1 promoter of the *AICDA* gene as a template, to generate sequential deletions. Primers designed are listed in **Table 3.1**, with

inserted RE sites underlined and in bold. PCR conditions used to generate the shortened promoter constructs are listed in **Table 3.2**.

Table 3.1: *AICDA* R1 promoter cloning primers.

Primer	Sequence	Tm	GC%	RE site
Forward 1	5'-GCTATATTCTGGA**ACGCGT**GATAACAGG-3'	62.1°C	46.4%	**MluI**
Forward 2	5'-CTCTTGCCT**ACGCGT**ATGTAAGATG-3'	58.2°C	48%	**MluI**
Forward 3	5'-GTACTG**ACGCGT**AGAGAGACTGTGG-3'	61.5°C	56%	**MluI**
Forward 4	5'-GAAGGGC**ACGCGT**ACTGTCAGACTAAG-3'	60.8°C	55.6%	**MluI**
Reverse	5'-CTGTCCA**AAGCTT**TGTCCAGAGTGTC-3'	60.7°C	50%	**HindIII**

Table 3.2: PCR cycling conditions.

Program	Temperature	Time	Cycles
Initial denaturation	95°C	1 min	1
Denaturation	95°C	15 sec	
Annealing	54°C	15 sec	30
Extension	72°C	40 sec	
Final extension	72°C	8 min	1

3.2.3.2 Restriction enzyme digest

To incorporate the deletion fragments into the pGL3 Basic plasmid, the PCR products were ligated into the p-GEM-T-Easy (Promega, USA) plasmid to facilitate restriction enzyme digestion. The pGEM-T-Easy plasmids with the shortened promoter inserts and the pGL3 Basic plasmid were digested overnight at 37°C using the MluI (ACGCTG) and HindIII (AAGCTT). Digested fragments were then electrophoresed on a 1% agarose gel (70V, 1 hour) stained with ethidium bromide (0.5 μg/mL). Fragments were then isolated from the gel and purified using the MinElute Gel extraction kit (Qiagen, Germany) and quantified using the Nanodrop1000.

3.2.3.3 Promoter fragment ligation within pGL3 Basic vector

The same methodology was used as described in Chapter 2 section 2.2.5.3.

3.2.3.4 Competent bacterial cells

The same methodology was used as described in Chapter 2 section 2.2.5.4.

3.2.3.5 Bacterial transformation and colony PCR

The same methodology was used as described in Chapter 2 section 2.2.5.5.

3.2.3.6 Plasmid isolation from bacterial cells

The same methodology was used as described in Chapter 2 section 2.2.5.6.

3.2.4 Dual-luciferase reporter assays

Dual-luciferase reporter assays were conducted in HT1080 cells seeded at a density of $5x10^4$ cells/mL in 6-well cell culture dishes. 400 ng of each luciferase reporter construct [full length (pGL3-AICDA-R1 – previously cloned in our laboratory), deletions of the *AICDA* R1 promoter, *AICDA* R2 (pGL3-AICDA-R2) or *AICDA* R4 (pGL3-AICDA-R4) regulatory region] were co-transfected, using X-tremeGene^{TM} HP transfection reagent (Sigma Aldrich, USA) with up to 500 ng of pcDNA-Tat or the corresponding control vector (pcDNA3.1 empty) as well as 50 ng of the pRL-TK vector, kindly provided by Professor Sharon Prince from the University of Cape Town, driving the expression of a Renilla reporter (used as a control for transfection efficiency). All transfections were done in duplicates and Luciferase assays were repeated at least three times (n=3). Thirty (30) hrs post-transfection, whole-cell extracts were assayed for firefly and Renilla luciferase activity using the dual-luciferase reporter assay system (Promega, USA). Luciferase activity was measured using the GloMax®-Multi+ Luminescence Module (Promega, USA). Luciferase values were expressed relative to empty vector control values.

3.2.5 Electroporation of plasmid DNA into B cells

The same methodology was used as described in Chapter 2 section 2.2.2.

3.2.6 miScript qRT-PCR

3.2.6.1 Total RNA isolation

Total RNA was isolated using the AllPrep DNA/RNA/miRNA universal kit (Qiagen, Germany). Electroporated cells were centrifuged at 1000 rpm for 5 min to pellet cells. Pelleted cells were resuspended in 1x PBS (Appendix A) and centrifuged again at 1000rpm for 5 min. Cells were resuspended and lysed in 600 µL RLT buffer (Qiagen, Germany) supplemented with β-Mercaptoethanol (10µL/mL). Homogenization of the cell lysate was performed by passing lysate through a sterile 20-gauge needle. The homogenized lysate was then transferred to the AllPrep DNA Mini spin column and centrifuged at 16000 x g for 30 sec. The flow-through, containing the RNA, was treated with 80 µL of Proteinase K (Qiagen, Germany) and 350 µL 99% ethanol for 10 min at RT. 750 µL 99% ethanol was added and the mixture was centrifuged through the RNeasy mini spin column at 16000 x g for 30 sec. Flow-through was discarded and 500 µL RPE buffer (Qiagen, Germany) was added to the column and centrifuged at 16000 x g for 30 sec. DNase I incubation mix containing 10 uL DNase I and 70 µL RDD buffer (Qiagen, Germany) was added to the column and incubated for 15 min at RT followed 500 µL FRN buffer (Qiagen, Germany) and centrifuged at 16000 x g for 30 sec. Flow-through was returned to the column and centrifuged again. 500 µL of buffer RPE was added and centrifuged as before. An ethanol wash was done with 500 µL of 99% ethanol and centrifuged as before. An additional centrifugation step was done to remove residual ethanol (centrifuged at 16000 x g for 2 min). Total RNA was eluted into diethyl pyrocarbonate (DEPC) treated 1.5mL Eppendorf tubes using 40µL nuclease-free H_2O. Nanodrop1000 was used to quantify RNA elution.

3.2.6.2 cDNA synthesis and qRT-PCR

cDNA was generated from 100ng isolated total RNA using the miScript II RT kit (Qiagen, Germany). Briefly, the 5x HiSpec buffer was used in a total reaction volume of 20 µL. Reaction mixes were incubated at 37°C for 60 min followed by 5 min at 95°C. The resulting cDNA was diluted in 100 µL with nuclease-free H_2O and aliquoted for storage at -80°C. qRT-PCR was done using the Rotor Gene-Q (Qiagen, Germany) machine with the 2x QuantiTect SYBR Green PCR Master Mix (Qiagen, Germany) with cycling conditions outlined in **Table 3.3**. The primers used are listed in **Table 3.4**. The human snRNA RNU6B was used as an internal PCR control. Comparative delta-delta CT method was used for data analysis and presentation.

Table 3.3: miScript qRT-PCR cycling conditions.

Program	Temperature	Time	Cycles
PCR initial activation step	95°C	15 min	1
Denaturation	94°C	15 sec	40
Annealing	55°C	30 sec	
Extension	70°C	30 sec	

Table 3.4: miScript qRT-PCR primers.

Primer	Qiagen catalogue number
miScript Universal Primer	218073
Hs_miR-181b_1 miScript primer assay	218300 MS00006699
Hs_RNU6-2_1(RNU6B) miScript Primer Assay	218300 MS00033740

3.2.7 *AICDA* 3'UTR dual-luciferase reporter assay

HT1080 cells were seeded at 0,8 x 10^5 cells/mL in wells of a 24-well cell culture plate and incubated until they reached 70% confluency. Per well, cells were transfected with 350 ng luciferase reporter construct containing the full-length *AICDA* 3'UTR sequence (2182 bp) cloned downstream of the luciferase gene in the pGL3-Promoter vector (Promega, Switzerland), 50 ng of the pRL-TK used as an internal control for transfection efficiency and 50nM of the miRNA-181b-5p mimic. The Dharmacon siRNA and Allstar siRNA negative controls were used at the same concentration as the miRNA-181b-5p mimic in separate experiments. Attractene Transfection Reagent (Qiagen, Germany) was used for the DNA-RNA dual transfection, using 1.5uL of Attractene per well. Cells were incubated for 24hrs post-transfection and whole-cell extracts were assayed for firefly and Renilla luciferase activity using the dual-luciferase reporter assay system (Promega, Switzerland). Luciferase activities were measured using the GloMax®-Multi+ Luminescence Module (Promega, Switzerland). Firefly luciferase values were normalized to the Renilla luciferase activity and expressed relative to empty vector control.

3.2.8 Western blot analyses

Electroporated cells were centrifuged at 1000rpm for 5 minutes followed by two washes in 1x PBS. Pelleted cells were resuspended in 2x Laemmli Blue buffer (Appendix A) and boiled for 5 min at 95°C. Protein was separated using 12% SDS-PAGE gels (Appendix A), transferred to nitrocellulose membranes (Bio-Rad, USA) using the Mini-PROTEAN 3 casting apparatus (Bio-Rad). Primary antibodies used were anti-AID (39-2500, Thermofisher, USA; 1:1000); anti-γH2AX (2577, Cell Signalling, US; 1:1000); and anti-p38 (M0800, Sigma-Aldrich;

1:5000). Secondary antibodies used were Goat Anti Rabbit (H+L) HRP conjugate (170-6515, Bio-Rad, USA; 1:5000) and Goat Anti Mouse (H+L) HRP conjugate (170-6516, Bio-Rad, USA; 1:5000) Densitometric analysis of the signal intensity of bands was done using the ImageJ software (NIH, USA).

3.3 Results

3.3.1 HIV-1 Tat induces the activity of the *AICDA* R1 promoter

Previous preliminary work done in the laboratory has shown that HIV-1 Tat can increase the activity of the *AICDA* R1 promoter in a dose-dependent manner [144]. To verify these experiments luciferase reporter assays were repeated. A significant increase of 1.6-fold was observed in promoter activity in the presence of HIV-1 Tat, compared to the control (**Figure 3.1**) hence confirming the previous results.

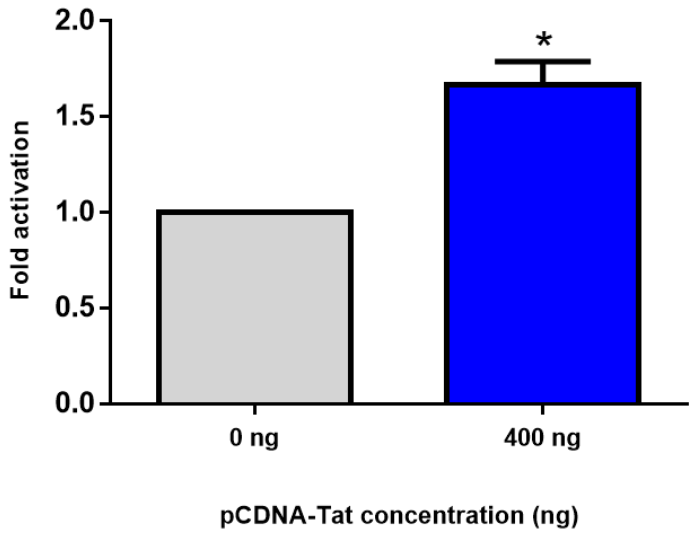

Figure 3.1: HIV-1 Tat enhances the activity of the *AICDA* R1 promoter. HT1080 cells were transfected with 500ng of the *AICDA* WT promoter (pGL3-AICDA-R1) along with 400ng of the HIV-1 Tat expression vector (pcDNA-Tat) or 400ng of the Empty pcDNA3.1 as control. Fold activation was obtained by comparing the relative luciferase units (RLU) of the HIV-1 Tat transfected groups to their empty transfected controls. Statistical analysis was done using GraphPad PRISM 7 with a two-tailed Student's t-test. Significance *($p<0.05$). The graph is representative of at least three separate repeats.

3.3.2 PROMO analysis identifies multiple putative TFBS within the *AICDA* R1 promoter

Using the online software PROMO, a comprehensive analysis of the *AICDA* R1 promoter was done to identify putative TFBS. Sites were mapped onto the promoter sequence of *AICDA* and the result is displayed in **Figure 3.2**. This analysis was done to identify TFBS which may potentially mediate the response observed in the result shown in **Figure 3.1**. To begin to address this, sequential deletion constructs of the R1 *AICDA* promoter was made. These deletion constructs were then used in repeat luciferase assays.

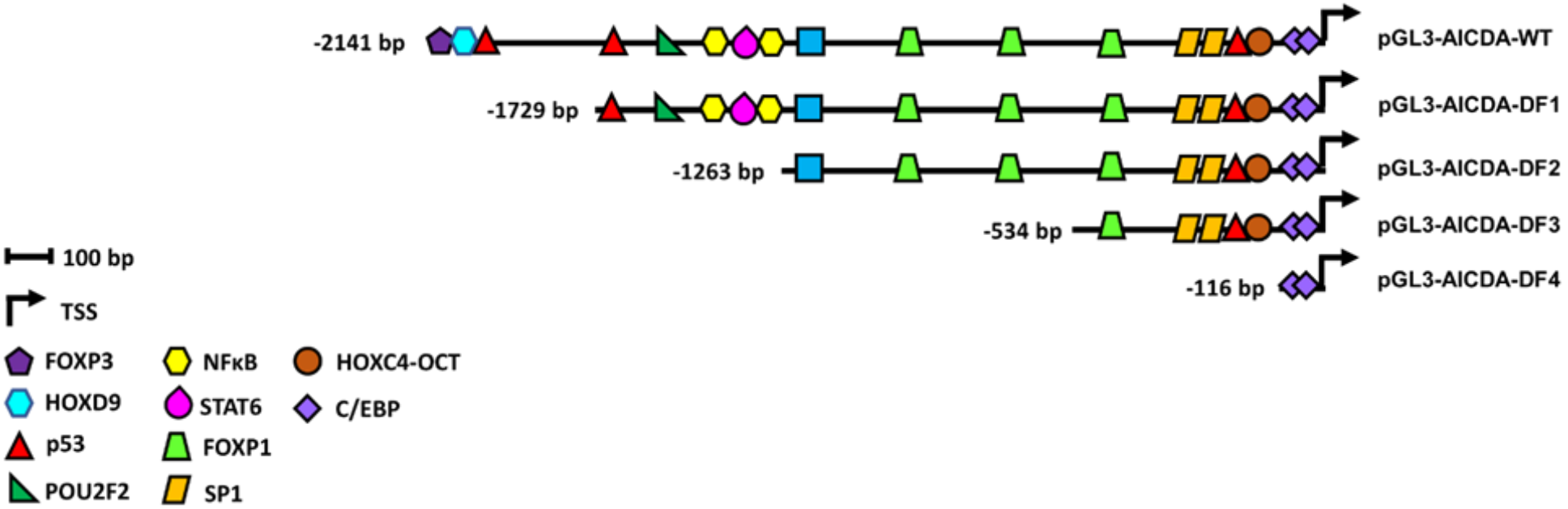

Figure 3.2: Transcription factor binding sites located on the *AICDA* R1 promoter as identified by PROMO. The 2141 bp FL *AICDA* promoter construct (pGL3- AICDA-WT) was analysed for TFBS using PROMO. Sites are represented as different coloured shapes which are labelled in the key and the transcription start site (TSS) is denoted by a black arrow. Sequential deletion constructs were designed to remove specific TFBS. Shortened promoter constructs are labelled accordingly on the right side of the diagram (pGL3-AICDA-DF1, pGL3- AICDA -DF2 and pGL3- AICDA -DF3 and pGL3-AICDA-DF4).

3.3.3 Sequential deletions of the *AICDA* R1 promoter do not lead to a decrease in activity in the presence of HIV-1 Tat

To identify the minimum promoter region which could be mediating this enhanced expression, a similar strategy was adopted as the one described for the *c-MYC* promoter in Chapter 2 in the creation of deletions construct for the *AICDA* promoter. This was followed by luciferase assays to assess whether the loss of specific promoter regions affected the observed activation. The use of sequential deletions constructs of the *AICDA* promoter in luciferase assays did not show a significant decrease in the promoter activity in the presence of HIV-1 Tat (**Figure 3.3**). While no loss in promoter activation was observed when the regions from -1729 bp to -534 bp (DF1 – 3) were deleted, intriguingly activation was enhanced when a region spanning -534 bp to -116 bp was deleted.

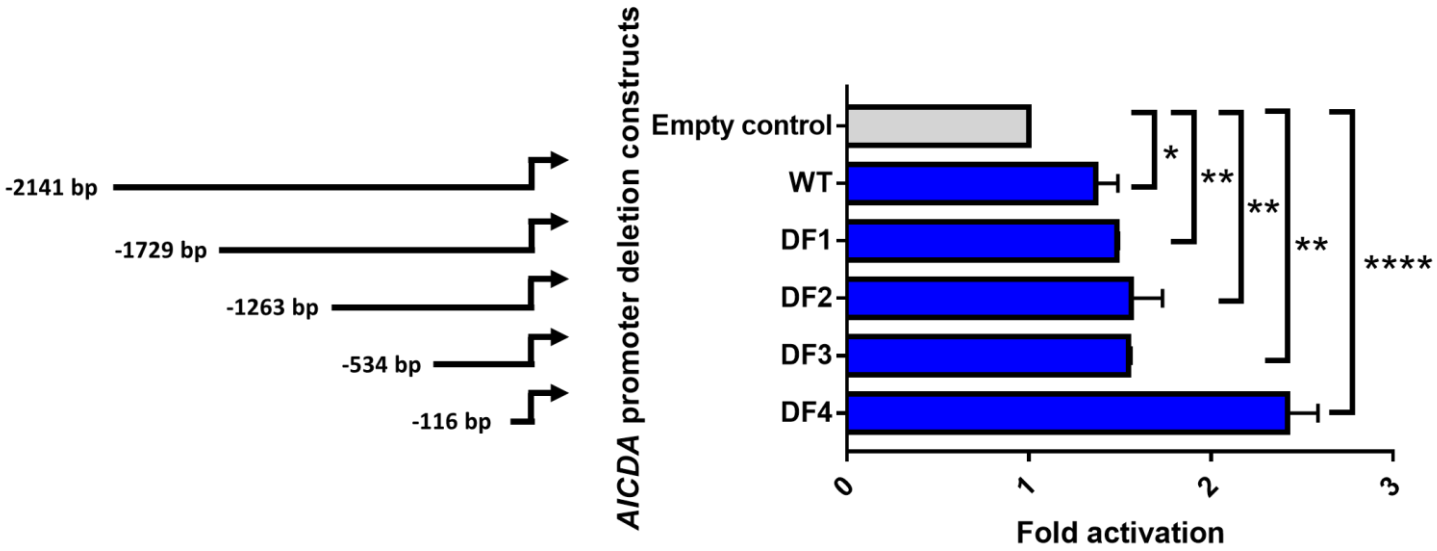

Figure 3.3: Sequential deletion of the *AICDA* R1 WT promoter does not abrogate the increased promoter activation under the influence of HIV-1 Tat. HT1080 cells were transfected with 500 ng of the *AICDA* R1 WT promoter (pGL3-AICDA-WT) or deletion constructs (pGL3-AICDA-DF1, pGL3-AICDA-DF2, pGL3-AICDA-DF3 and pGL3-AICDA-DF4) along with 400 ng of the HIV-1 Tat expression vector (pcDNA-Tat) or 400 ng of the Empty pcDNA3.1 as control. Fold activation was obtained by comparing the relative luciferase units (RLU) of the HIV- 1Tat transfected groups to their empty transfected controls which were set to 1. One-way ANOVA tests were done using GraphPad PRISM 7. Significance *($p<0.05$), **($p<0.005$) and ****($p<0,0001$). The graph is representative of at least three separate repeats.

3.3.4 HIV- 1 Tat elicits a dose-dependent increase in the activity of the R2 and R4 regulatory regions.

As discussed in Chapter 1, four regulatory regions have been described for the *AICDA* gene (see Figure 1.3) of which R1, R2 and R4 have defined roles in activating/repressing *AICDA* expression. To further assess the ability of HIV-1 Tat to affect *AICDA* transcription, luciferase assays were carried out using pGL3 constructs containing the R2 and R4 regulatory regions cloned upstream of the luciferase gene. Interestingly, both regions showed a dose-dependent increase in promoter activity with increasing concentrations of HIV-1 Tat (**Figure 3.4A & B**).

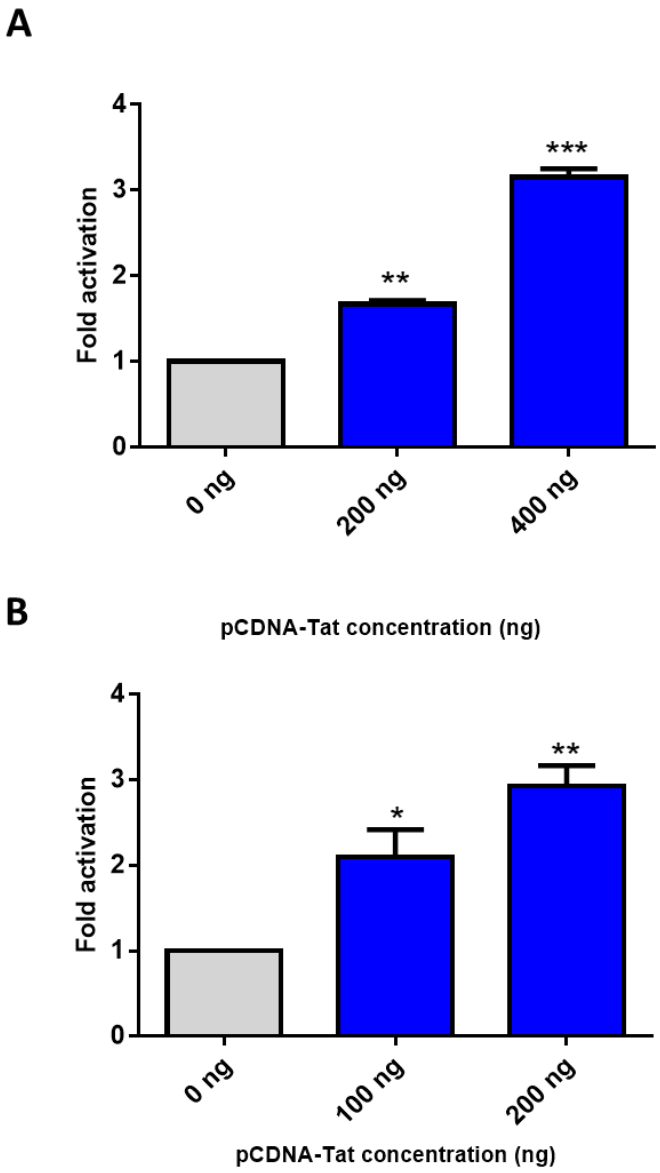

Figure 3.4: HIV-1 Tat induces activation of both the AID R2 and R4 regulatory regions. A) HT1080 cells were transfected with 500 ng of the *AICDA* R2 promoter (pGL3-AICDA-R2) along with 200-400 ng of the HIV-1 Tat expression vector (pcDNA-Tat) or 200-400 ng of the Empty pcDNA3.1 as control. B) HT1080 cells were transfected with 250 ng of the *AICDA* R4 promoter (pGL3-AICDA-R4) along with 100-200 ng of the HIV-1 Tat expression vector (pcDNA-Tat) or 100-200 ng of the Empty pcDNA3.1 as control. Fold activation was obtained by comparing the relative luciferase units (RLU) of the HIV-1 Tat transfected groups to their empty transfected controls which were set to 1. Statistical analysis was done using one-way ANOVA with GraphPad PRISM 7. Significance *($p<0.05$), **($p<0.005$) and ***($p<0,0005$). The graphs are representative of at least three separate repeats.

3.3.5 AID and yH2AX protein levels are elevated in HIV-1 Tat expressing BL cells

To determine if the increased *AICDA* expression previously shown was conserved at the protein level, electroporation of Ramos cells with an HIV-1 Tat expression construct was performed followed by western blot analysis. As shown in **Figure 3.5A**, an increase in AID protein levels was observed in cells expressing HIV-1 Tat compared to empty control. We and others have previously reported that an increase in AID is accompanied by upregulation of the double-strand DNA break marker, phosphorylated H2AX (γH2AX) [120,163]. Indeed, we found a corresponding increase in γH2AX in the same cells (**Figure 3.5B**).

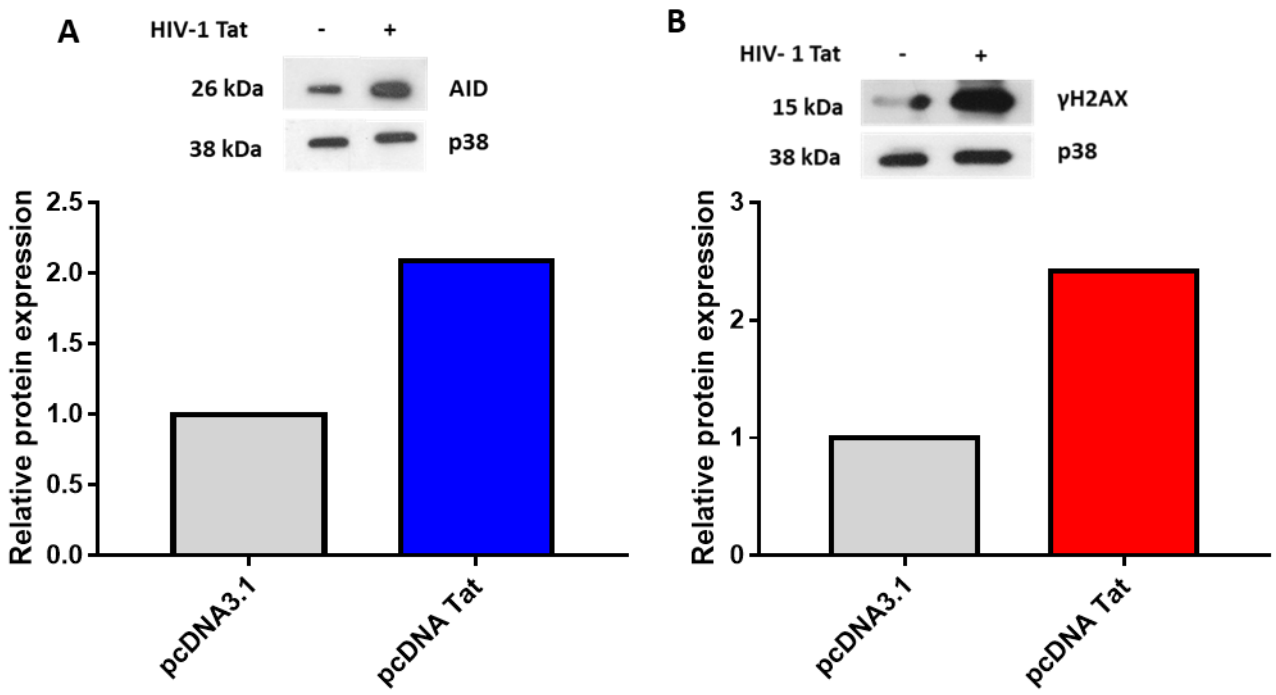

Figure 3.5: HIV-1 Tat expression in Ramos cells leads to increased AID and γH2AX protein levels. Cells were electroporated with 15 µg of the HIV-1 Tat expression vector (pcDNA-Tat) or 15 µg of the pcDNA3.1 plasmid as a control. A) AID western blot. B) γH2AX western blot. Bar graphs represent densitometric analysis (ImageJ). Fold increase was obtained by comparing the relative densitometric units of the HIV-1 Tat electroporated cells to the negative control cells, which were set to 1, after normalising to loading control (p38).

3.3.6 Expression of HIV-1 Tat in BL cells leads to a decrease in hsa-miRNA-181b-5p levels.

Our assays on the R1 promoter indicate that the observed increase in AID expression in the presence of HIV-1 Tat is unlikely to be via a direct mechanism involving this region but could potentially be regulated indirectly at the post-transcriptional level. AID has been reported to be the target of a handful of microRNAs. Furthermore, and as stated earlier, a study on HIV-associated neurocognitive disorders found that HIV-1 Tat could inhibit hsa-miRNA-181b in astrocytes [141]. This miRNA has been shown to directly bind to and repress the murine *Aicda* 3'UTR [142]. Analysis of the human *AICDA* 3'UTR revealed that it contains three predicted hsa-miRNA-181b-5p binding sites with two of the sites sharing significant homology between the two species (**Figure 3.6A&B**).

A

ID	Duplex structure	Position
1	miRNA 3' ugGGUGGCUGUCGUU------ACUUACAa 5' Target 5' aaCAATTGGAAGGAAGTTGCTTGAATGTt 3'	828 - 856
2	miRNA 3' ugggUGGCUGUCGUU----ACUUACaa 5' Target 5' gcagATCCACAGGAAAACTTGAATGca 3'	1679 - 1705
3	miRNA 3' ugGGUGGC--UGUCGUUACUUACaa 5' Target 5' tgCAATTGGTGCAGTTTTGAATGca 3'	203 - 227

B

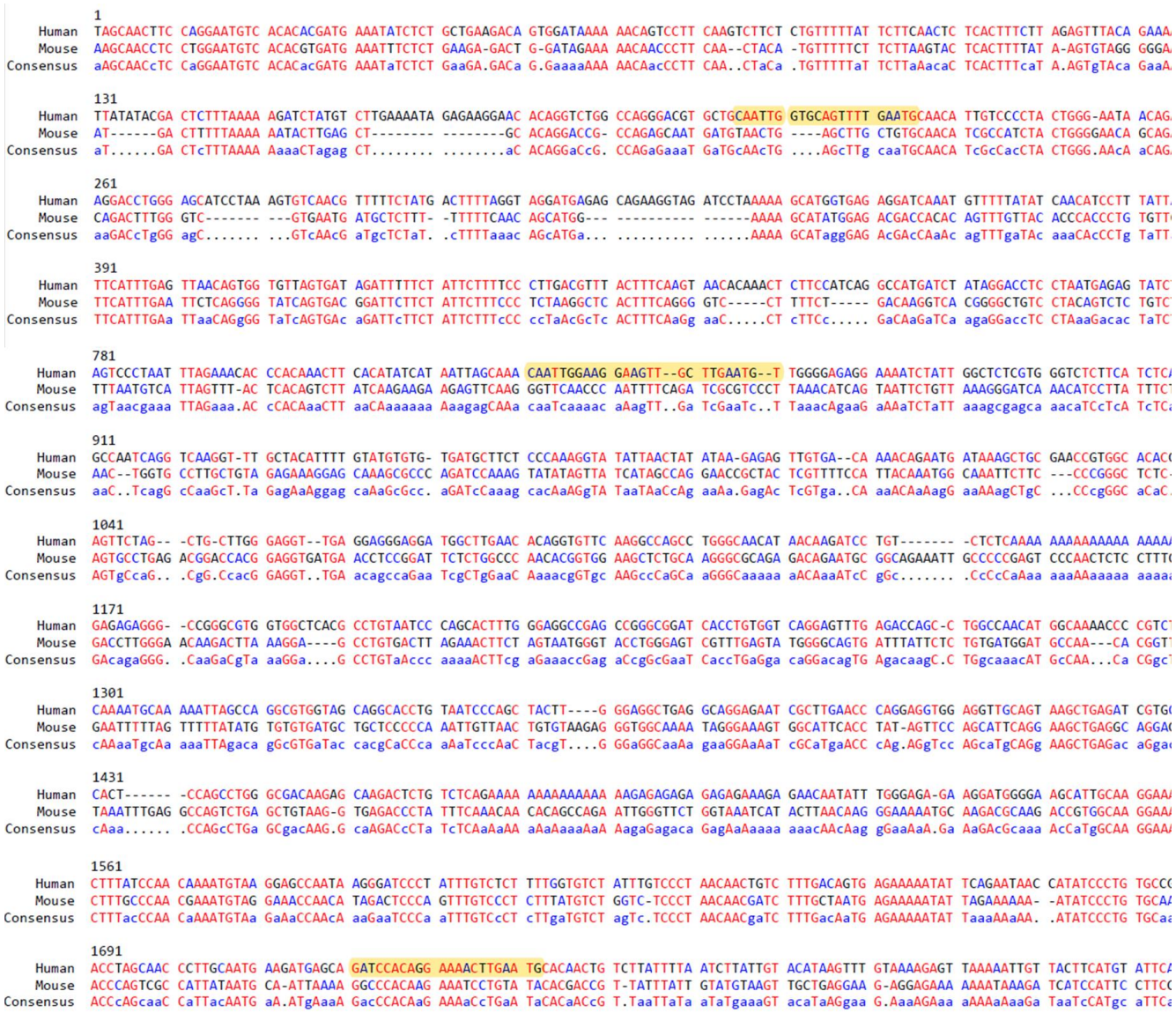

Figure 3.6: hsa-miR-181b-5p predicted binding sites within the 3'UTR of the *AICDA* mRNA. A) Putative sites for hsa-miR-181b-5p on the human *AICDA* 3'UTR were obtained using miRTarBase [164] and confirmed through searches on alternative bioinformatics tools: miRmap [165] and TargetScan [166]. B) The human and mouse *AICDA* 3'UTR was aligned using the online sequence alignment tool MultAlin[167]. Blue nucleotides indicate sequence differences, red nucleotides show sequence conservation and black nucleotides indicate sequences only present in one of the 3'UTRs. The three putative binding sites for hsa-miRNA-181b-5p are highlighted in orange.

To determine if HIV-1 Tat can alter the expression of hsa-miR-181b-5p in B cells, Ramos cells were electroporated with pcDNA-Tat or pcDNA3.1 empty control and RNA was isolated to quantify the levels of hsa-miRNA-181b-5b. Here it was found that hsa-miRNA-181b-5p is significantly downregulated when HIV-1 Tat is expressed 24hrs after electroporation, downregulation was maintained at 48hrs, although diminished and not statistically significant (**Figure 3.7**).

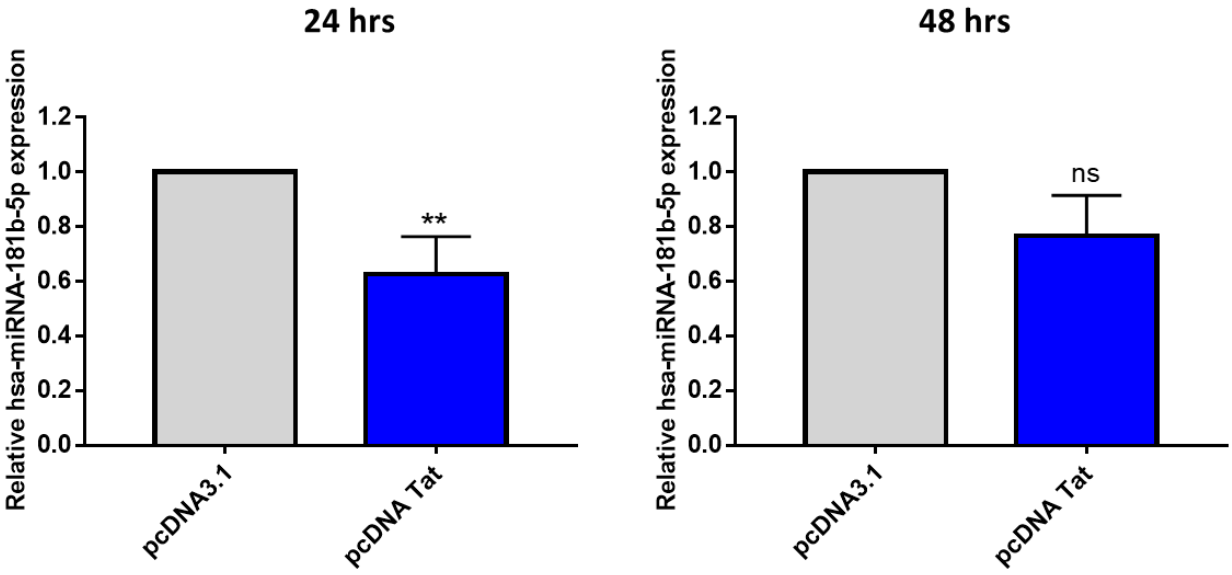

Figure 3.7: hsa-miRNA-181b-5p is downregulated in HIV-1 Tat expression BL cells. Ramos cells were electroporated with pcDNA-Tat or pcDNA3.1 control. miScript qRT-PCR results are expressed as fold change relative to pcDNA3.1 negative control which was set to 1. U6 primers were used as an internal control. Statistical analysis was done using GraphPad PRISM 7 with a two-tailed Student's t-test with a two-sided t-test, **($p<0.005$). The graph is representative of at least three separate repeats.

3.3.7 An hsa-miRNA-181b-5p mimic interacts with the *AICDA* 3'UTR, leading to translational repression.

To further investigate and confirm the HIV-1 Tat – miRNA181b – AID axis, HT1080 cells were transfected with the pGL3-AICDA-3'UTR construct along with 50 nM of a commercially available hsa-miRNA-181b-5p mimic or 50 nM of the Allstar siRNA negative control (**Figure 3.8)**. Significant repression of luciferase activity was observed in the hsa-miRNA-181b-5p mimic transfected cells compared to no RNA control and siRNA negative controls.

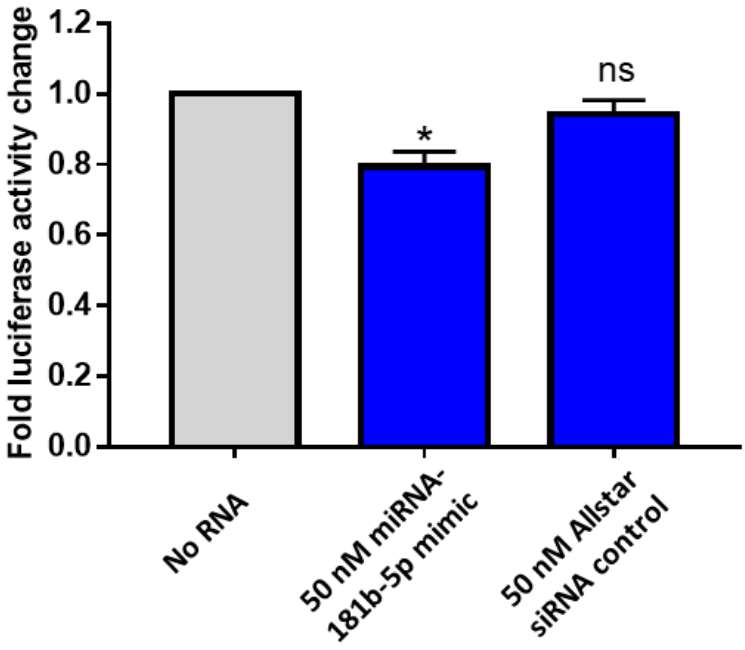

Figure 3.8: The hsa-miRNA-181b-5p mimic represses the *AICDA* 3'UTR. Luciferase reporter assay investigating the effect of hsa-miRNA-181b-5p on the *AICDA* full-length 3'UTR. HT1080 cells were transfected with 350 ng of the pGL3-AICDA 3'UTR, along with 50 nM of the hsa-miRNA-181b-5p mimic or Allstar siRNA negative control. Fold activation was obtained by comparing the relative luciferase units (RLU) of the hsa-miRNA-181b-5p or siRNA control transfected groups to the no RNA empty transfected controls, which were set to 1. The graph is representative of at least three separate repeats.

3.4 Discussion

The sequential deletion of the *AICDA* R1 promoter was unable to reduce the HIV-1 Tat-mediated promoter activation. This indicates that the increase in *AICDA* transcription via the R1 promoter is unlikely to occur as a direct effect of HIV-1 Tat on the promoter. A possible mechanism could be that HIV-1 Tat induces changes to cellular pathways which in turn activate the R1 promoter. Further interrogation is needed to decipher this, for instance, potential cellular pathways which are activated by HIV-1 Tat could be identified through the use of inhibitory molecules that target specific pathways, and assess if these inhibitory interventions abrogate the observed HIV-1 Tat-mediated R1 promoter activity. It is also possible that HIV-1 Tat may be mediating the activation via sites present within the minimal promoter region of 116 bp present in pGL3-AICDA-DF4. Of Note, the 116 bp region contains multiple C/EBP binding sites which could be abolished through SDM to determine if they are involved in this mechanism. A future avenue would be to perform ChIP assays with HIV-1 Tat specific antibody to determine whether HIV-1 Tat potentially interacts with the R1 promoter. If no region is enriched it would mean that there is less likelihood that HIV-1 Tat is directly interacting with this regulatory region and that the observed increase in promoter activity is not specific. Interestingly, the observed increase in promoter activity in the presence of HIV-1 Tat was further enhanced when only a small region of 116 bp upstream of the TSS was present. The region of the promoter lost within this deletion construct could harbour repressive elements, proximal to the core promoter, which keeps the *AICDA* R1 promoter in check. Therefore, the loss of these elements might be amplifying the observed HIV-1 Tat activation. Future work could be done to identify these sites using an SDM approach.

The effect of HIV-1 Tat on the *AICDA* regulatory regions R2 and R4 was also assessed and both regions were found to be significantly enhanced by HIV-1 Tat, more robustly than the R1 promoter. These regions contain multiple TFBS that regulate *AICDA* expression including STAT6, C/EBP and SMAD3/4 in R4 and MYB, PAX5, E2A and E2F in R2. Due to time constraints, further interrogation of this result could not be carried out, however, this does form part of future studies in the laboratory to continue to uncover the complete picture of HIV-1 Tat's role in the regulation of AID. It would be worthwhile to investigate whether HIV-1 Tat binds to and/or interacts with these TFs and TFBS given that HIV-1 Tat has previously been reported to interact with some of these proteins [168].

AID's physiological function is to induce changes in DNA which result in the activation of DNA repair response. Depending on the mechanistic responses that occur after deamination, there can be the formation of DSBs. An important marker for DNA DSB is the phosphorylation of Serine 139 on histone H2AX, producing γH2AX [169,170]. This modification to the H2AX histone can be used to evaluate the level of DNA damage occurring in the cell after specific stimuli[171]. This study shows that BL cells expressing HIV-1 Tat, not only have increased AID protein expression but also elevated phosphorylation of H2AX. This finding is in line with what has been reported recently by our group where HIV-1 Nef induced AID expression correlated resulted in increased genomic instability [120]. Taken together, this study, as well as other recent publications, give evidence of deleterious HIV-1 Tat modulation of AID activity within BL cells both at the transcriptional and posttranscriptional levels.

Our result of the R1 promoter suggested that the observed increase in AID expression could be due to interaction with other regulatory regions or potentially through events happening at the post-transcriptional level. AID is tightly regulated within B cells. These mechanisms of

regulation prevent off-target AID activity which could be oncogenic due to its ability to mutate DNA. One of the ways a cell regulates the expression of genes is through RNA interference [172]. miRNAs are short non-coding RNAs that target specific, often conserved, 3'UTR regions of mRNA targeting them for repression/degradation [173,174]. In the context of *AICDA*, a handful of miRNAs have been implicated in its repression. Of importance to this study is hsa-miRNA-181b-5p which has been implicated in the murine *Aicda* repression and has also been reported to be downregulated in astrocytes by HIV-1 Tat in the context of HIV associated neurocognitive disorder (HAND). In our study, we show that the presence of HIV-1 Tat led to a significant decrease in hsa-miRNA-181b-5p in BL cells. It is important to note that this decrease was consistently observed despite the relatively low transfection efficiency of Ramos cells. Within the context of HIV infection, this has an important implication for *AICDA* regulation, as it suggests that if hsa-miRNA-181b-5p is repressed by HIV-1 Tat, *AICDA*/AID levels would increase due to the loss of miRNA interference with mRNA translational processes, and this is indeed the case as shown in this study. The precise mechanism whereby HIV-1 Tat alters miRNA expression remains unclear. A potential mechanism behind this was proposed by the Jeang group, where HIV-1 Tat binds to Dicer protein thereby interfering with miRNA processing [175,176]. Future work needs to be done with a broader approach to investigate the extent of HIV-1 Tat and BL miRNA associations and the consequential changes in the miRNA targetome. A potential way to tackle this, which is currently being developed by our research group, would be to employ AGO2-RIP-ChIP followed by microarray or RNA-Seq analyses to identify actively targeted mRNA sequences as a result of this HIV-1 Tat-mediated miRNA deregulation [177].

The study that first implicated hsa-miRNA-181b-5p in the repression of AID was done using the mouse *Aicda* 3'UTR. Although the human and mouse *AICDA* 3'UTR share conserved

regions, it has not been shown if hsa-miRNA-181b-5p directly interacts with the human *AICDA* 3'UTR. The full-length sequence was cloned into the pGL3-Promoter vector and this construct was used in luciferase assays to assess the ability of hsa-miRNA-181b-5p to repress the full-length 3'UTR. Our result aligns with what was reported by Yébenes et al (2008) although their finding was in the context of the murine *Aicda* 3`UTR. Future work should focus on confirming that the putative hsa-miRNA-181b-5p binding sites within the 3'UTR do indeed mediate this repression. To this end, SDM can be employed to sequentially disrupt the three hsa-miRNA-181b-5p binding sites and assess if a loss in repression is observed in the luciferase assay.

Chapter 4: Collated discussion and concluding remarks

The last two decades have seen important advances in our understanding of HIV-1 biology, the progression of AIDS and the comorbidities associated with this life-long viral infection. Amongst the common comorbidities are the AIDS-defining cancers which include BL and DLBCL. These cancers are often reported to be aggressive in HIV infected patients, even those whose HIV viral load has been suppressed, with these patients experiencing poorer than average outcomes. In recent years the new concept of HIV having direct oncogenic effects in patients has emerged, supported by molecular studies showing the ability of components of the virus to induce cancer-driving changes in cellular pathways. The studies carried out and presented in this thesis contributes to this body of work, and serve to enhance our understanding of these phenomena [110,116,120,144,178].

This work presents a novel mechanism of how HIV-1 Tat interacts with the promoter region of *c-MYC* to induce its transcription, and how it enhances AID expression through two independent mechanisms; by inducing the activation of three regulatory regions of the *AICDA* gene and through the repression of hsa-miRNA-181b-5p-mediated translational interference. In **Figure 4.1** we summarise some of the pathways and mechanisms whereby HIV-1 proteins promote oncogenesis in B cells and demonstrate how the findings of this present study complement the growing body of evidence for the direct role of HIV-1 in B-cell lymphoma development.

Not only could HIV be driving oncogenic pathways in patients harbouring high viral loads, but also in those who are virally suppressed through effective antiretroviral therapy. Importantly, some studies suggest that latently infected memory CD4 T cells are not the only compartments of HIV persistence, but that myeloid cells, and possibly haematopoietic progenitors, can also serve as long-term viral reservoirs [179]. Furthermore, circulating HIV-1 proteins remain

detectable in the sera of HIV positive individuals including those that are receiving HAART and are virally-suppressed[110,180,181]. These proteins have been shown to be able to either translocate through the cell membrane or interact with cell surface receptors, culminating in the deregulation of cellular pathways, altering host gene expression, and disrupting genomic stability [112,137].

Previous work in our groups which formed part of a Master's thesis showed that BL cells incubated with recombinant HIV-1 Tat protein had elevated *c-MYC* mRNA and protein, and is in support of the published work by Germini et al (2017) demonstrating HIV-1 Tat's ability to promote the colocalization of the *c-MYC* and *IGH* loci in B cells from HIV positive patients. They further demonstrated that along with this increased co-localization there was also evidence of DNA DSBs in these loci [139,144]. Additionally, our research group and others have uncovered oncogenic properties associated with the HIV-1 protein Nef. Mdletshe et al (2020) found that BL cells exposed to HIV-1 Nef recombinant protein have significantly elevated levels of *c-MYC* mRNA and protein expression, with a concomitant increase in genomic instability[120]. Furthermore, Li et al (2010) demonstrated that HIV-1 p17 promotes c-MYC and CREB phosphorylation in B cells through the activation of the cAMP/PKA and MEK/ERK signalling pathways, thereby enhancing the DNA binding activity of these TFs [116]. The binding of p17 to the receptors CXCR1 and CXCR2 induces the activation of these signalling pathways, as well as the AKT/PI3K pathway, and has also been reported to enhance angiogenesis, a feature of more advanced stage disease [182–184]. Taken together these studies show that, as a result of HIV-1 infection, the expression and function of *c-MYC* is enhanced in B lymphocytes via complex direct and indirect mechanisms, hence promoting a more aggressive phenotype.

The direct perturbation of AID expression by pathogenic agents like EBV, *Helicobacter pylori*, *Plasmodium falciparum* and HIV-1 has been a focal point of research within the field of diseases caused by infectious agents[185–188]. In the context of HIV-1 infection, Sall et al (2019) demonstrated that within B cells of both healthy donors and HIV positive patients, recombinant HIV-1 Tat protein induces increased *AICDA* mRNA expression and H3 acetylation at the R1 promoter region, indicative of active chromatin. Concomitant with this, they observed γH2AX foci formation at the *IGH* locus through immuno-3D FISH. Our data also corroborates the Sall et al (2019) findings by showing that BL cells transiently transfected with HIV-1 Tat have increased AID protein expression and show increased DNA damage demonstrated by heightened γH2AX levels. AID targeting to DNA is complex and not fully understood. However, a proposed requirement for its targeting to DNA is active transcription, with AID activity often mapped to actively transcribed regions [91,92,189]. This phenomenon could be a result of ssDNA exposure within R-loops during transcription or due to AID's association with the transcriptional machinery, giving it access to DNA [190–193]. However, this is not enough to explain the selectivity of AID towards certain actively transcribed genes over others, suggesting the importance of other secondary DNA structures necessary for AID's enzymatic activity. For example, G-quadruplexes have been shown to be preferentially targeted by AID [194]. Notably, known targets of AID like *c-MYC, BCL2* and *IGH* contain such structures within their promoter regions (*c-MYC* and *BCL2)* and S region (IGH) [195–197]. Taken together, this could explain why HIV positive individuals have an increased risk of developing BL compared to the general population, despite the advent of HAART. Specifically, in B cells of HIV positive individuals, *c-MYC* translocations could be favoured due to i) HIV-1 Tat/Nef/p17-mediated increased *c-MYC* transcription and activity, ii) the increased *c-MYC* and *IGH* loci co-localization induced by HIV-1 Tat and iii) the concomitant elevation in AID protein levels and off-target activity promoted by HIV-1 Tat and Nef. Altogether, we propose that these features

potentiate a highly favourable environment for translocation and other mutational events to occur, thereby enhancing cell proliferation and other cellular phenotypes which support BL progression.

Ultimately, our increased understanding of the oncogenic features of HIV-1 will lead to novel avenues for therapy, disease monitoring and prevention. Potential therapies could be aimed at blocking the interactions between specific HIV-1 proteins and their host co-factors. For instance, the use of small-molecule inhibitors could be used to prevent p17's binding to the CXCR1/2 receptors on the B cell surface, therefore blocking the signalling pathways downstream of this interaction. Concurrently, other methods could be implemented to block HIV-1 Tat from inducing transcriptional activation of host genes. This could be achieved through the use of HIV-1 Tat inhibitors like Didehydro-Cortistatin A which blocks HIV-1 Tat function [198,199]. Furthermore, the monitoring of the levels of circulating viral proteins in the sera of HIV positive patients could be used as a factor to assess the risk of developing disease. These approaches would have to be optimized through continued research in this field.

Conclusion

HIV-BL is an aggressive disease with an often poor prognosis in the South African setting. Finding novel ways to tackle this malignancy is imperative to the improvement of treatment and outcome. This thesis describes three novel mechanisms which implicate HIV-1 Tat in the pathobiology of HIV-BL. (1) We show that HIV-1 Tat, through interacting with JunB, binds to the *c-MYC* promoter resulting in the increased transcription and protein levels in BL cells, (2) HIV-1 Tat activates the *AICDA* R1, R2 and R4 regulatory regions, upregulating transcription, and (3) HIV-1 Tat represses hsa-miRNA181b-5p, alleviating *AICDA* 3'UTR

repression further increasing AID protein levels in BL cells. Together, these findings not only highlight the complexity of HIV-1 Tat activity within BL cells but also the importance of targeting this viral protein to prevent its deleterious effects within these cells.

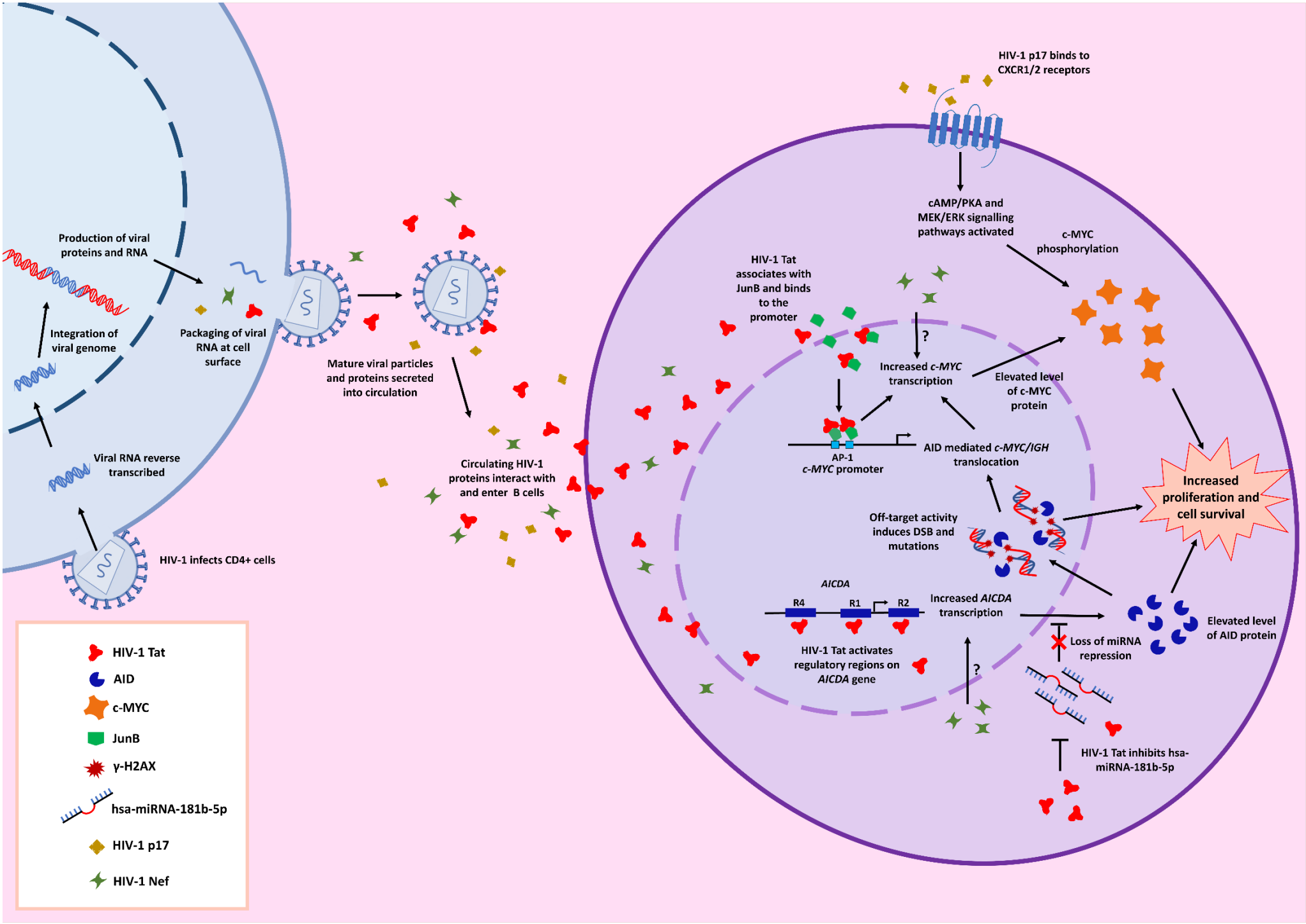
HIV-1 p17 binds to CXCR1/2 receptors
cAMP/PKA and MEK/ERK signalling pathways activated
c-MYC phosphorylation
HIV-1 Tat associates with JunB and binds to the promoter
?
Increased c-MYC transcription
Elevated level of c-MYC protein
AP-1 c-MYC promoter
AID mediated c-MYC/IGH translocation
Increased proliferation and cell survival
Off-target activity induces DSB and mutations
AICDA
R4
R1
R2
Increased AICDA transcription
Elevated level of AID protein
Loss of miRNA repression
HIV-1 Tat activates regulatory regions on AICDA gene
?
HIV-1 Tat inhibits hsa-miRNA-181b-5p
Production of viral proteins and RNA
Integration of viral genome
Packaging of viral RNA at cell surface
Mature viral particles and proteins secreted into circulation
Circulating HIV-1 proteins interact with and enter B cells
Viral RNA reverse transcribed
HIV-1 infects CD4+ cells
HIV-1 Tat
AID
c-MYC
JunB
γ-H2AX
hsa-miRNA-181b-5p
HIV-1 p17
HIV-1 Nef

Figure 4.1: Schematic diagram showing the different cellular mechanisms whereby HIV-1 viral proteins have been shown to promote B-cell lymphomagenesis. Infected CD4+ cells actively secrete viral proteins (including HIV-1 Tat, Nef and p17) into circulation. Circulating HIV-1 Tat and Nef proteins in the sera of HIV positive people can translocate into B cells [112,137]. Upon entering the nucleus HIV-1 Tat associates with host factors like JunB and binds to the *c-MYC* promoter (Chapter 2). HIV-1 Nef also induces *c-MYC* promoter activity resulting in increased transcription via a yet unidentified mechanism [120]. HIV-1 Tat can enhance the transcription of the *AICDA* gene via three of its regulatory regions (Chapter 3). Additionally, HIV-1 Tat inhibits the *AICDA* repressor hsa-miRNA-181b-5p, further increasing AID protein levels (Chapter 3). The HIV-1 Nef protein induces AID transcription via an as yet undescribed mechanism [120]. Increased AID protein in B cells promotes the occurrence of off-target DSBs and mutagenesis culminating in *c-MYC/IGH* translocations and other pro-cell survival/proliferative advantages [52,88,90]. Increased *c-MYC* transcription and/or translocations result in the overexpression of the c-MYC protein, coupled to HIV-1 p17 induced c-MYC protein phosphorylation via the cAMP/PKA and MEK/ERK signalling pathways, promotes c-MYC driven cell survival and proliferation[116].

References

1. UNAIDS. Global HIV and AIDS statistics 2019 Fact sheet. *Glob. HIV AIDS Stat. World AIDS day 2019 Fact Sheet* **1**, 1–6 (2019).

2. StatSA. *Mid-Year population estimates 2020*. http://www.statssa.gov.za/publications/P0302/P03022020.pdf (2020).

3. Guiguet, M. *et al.* Effect of immunodeficiency, HIV viral load, and antiretroviral therapy on the risk of individual malignancies (FHDH-ANRS CO4): a prospective cohort study. *Lancet. Oncol.* **10**, 1152–9 (2009).

4. Department of Health. Operational plan for comprehensive HIV and AIDS care, management and treatment for South Africa. 262 (2003).

5. Walker, C. M. *et al.* Identification of a Reservoir for HIV-1 in Patients on Highly Active Antiretroviral Therapy. *Science (80-.).* **278**, 1–7 (1997).

6. Chun, T. W. *et al.* In vivo fate of HIV-1-infected T cells: Quantitative analysis of the transition to stable latency. *Nat. Med.* **1**, 1284–1290 (1995).

7. Fischer-Smith, T. *et al.* CNS invasion by CD14+/CD16+ peripheral blood-derived monocytes in HIV dementia: Perivascular accumulation and reservoir of HIV infection. *J. Neurovirol.* **7**, 528–541 (2001).

8. Abudulai, L. N. *et al.* Chronic HIV-1 Infection Induces B-Cell Dysfunction That Is Incompletely Resolved by Long-Term Antiretroviral Therapy. *JAIDS J. Acquir. Immune Defic. Syndr.* **71**, 381–389 (2016).

9. Moir, S. *et al.* HIV-1 induces phenotypic and functional perturbations of B cells in chronically infected individuals. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10362–7 (2001).

10. Grogg, K. L., Miller, R. F. & Dogan, a. HIV infection and lymphoma. *J. Clin. Pathol.* **60**, 1365–1372 (2007).

11. Rubinstein, P. G., Aboulafia, D. M. & Zloza, A. Malignancies in HIV/AIDS. *Aids* **28**, 453–465 (2014).

12. Morton, L. M. *et al.* Molecular characteristics of diffuse large B-cell lymphoma in human immunodeficiency virus-infected and -uninfected patients in the pre-highly active antiretroviral therapy and pre-rituximab era. *Leuk. Lymphoma* **55**, 551–557 (2014).

13. Castillo, J. J., Bibas, M. & Miranda, R. N. The biology and treatment of plasmablastic lymphoma. *Blood* **125**, 2323–2330 (2015).

14. Carbone, A. *et al.* Diagnosis and management of lymphomas and other cancers in HIV-infected patients. *Nat Rev Clin Oncol* **11**, 223–238 (2014).

15. Gopal, S. *et al.* Temporal Trends in Presentation and Survival for HIV-Associated Lymphoma in the Antiretroviral Therapy Era. *JNCI J. Natl. Cancer Inst.* **105**, 1221–1229 (2013).

16. Shen, Y. *et al.* Clinical and prognostic analysis of 78 patients with human immuno-deficiency virus associated non-Hodgkin's lymphoma in Chinese population. *Infect. Agent. Cancer* **12**, 7 (2017).

17. Opie, J., Antel, K., Koller, A. & Novitzky, N. In the South African setting, HIV-associated Burkitt lymphoma is associated with frequent leukaemic presentation, complex cytogenetic karyotypes, and adverse clinical outcomes. *Ann. Hematol.* **99**, 571–578 (2020).

18. Diamond, C., Taylor, T. H., Aboumrad, T. & Anton-Culver, H. Changes in acquired

immunodeficiency syndrome-related non-Hodgkin lymphoma in the era of highly active antiretroviral therapy: Incidence, presentation, treatment, and survival. *Cancer* **106**, 128–135 (2006).

19. Gibson, T. M., Morton, L. M., Shiels, M. S., Clarke, C. A. & Engels, E. A. Risk of non-Hodgkin lymphoma subtypes in HIV-infected people during the HAART era. *AIDS* **28**, 2313–2318 (2014).

20. Brady, G., MacArthur, G. J. & Farrell, P. J. Epstein-Barr virus and Burkitt lymphoma. *J. Clin. Pathol.* **60**, 1397–402 (2007).

21. Donadoni, G., Bruno-Ventre, M. & Ferreri, A. Treatment of HIV-associated Burkitt lymphoma. *EMJ. Hema.* **1**, 38–52 (2013).

22. Burkitt, D. A sarcoma involving the jaws in african children. *Br. J. Surg.* **46**, 218–223 (1958).

23. Molyneux, E. M. *et al.* Burkitt's lymphoma. *Lancet* **379**, 1234–1244 (2012).

24. Orem, J., Mbidde, E. K., Lambert, B., De Sanjose, S. & Weiderpass, E. Burkitt's lymphoma in Africa, a review of the epidemiology and etiology. *Afr. Health Sci.* **7**, 166–175 (2007).

25. Crawford, D. H. Biology and disease associations of Epstein-Barr virus. *Philos. Trans. R. Soc. B Biol. Sci.* **356**, 461–473 (2001).

26. EPSTEIN, M. A., ACHONG, B. G. & BARR, Y. M. VIRUS PARTICLES IN CULTURED LYMPHOBLASTS FROM BURKITT'S LYMPHOMA. *Lancet (London, England)* **1**, 702–3 (1964).

27. Mbulaiteye, S. M. *et al.* Epstein-Barr virus patterns in US Burkitt lymphoma tumors from the SEER residual tissue repository during 1979-2009. *APMIS* **122**, 5–15 (2014).

28. Shannon-Lowe, C., Rickinson, A. B. & Bell, A. I. Epstein–Barr virus-associated lymphomas. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160271 (2017).

29. Grande, B. M. *et al.* Genome-wide discovery of somatic coding and noncoding mutations in pediatric endemic and sporadic Burkitt lymphoma. *Blood* **133**, 1313–1324 (2019).

30. De Witt, P., Maartens, D. J., Uldrick, T. S. & Sissolak, G. Treatment outcomes in AIDS-related diffuse large B-cell lymphoma in the setting roll out of combination antiretroviral therapy in South Africa. *J. Acquir. Immune Defic. Syndr.* **64**, 66–73 (2013).

31. Ozuah, N. W., Lubega, J., Allen, C. E. & El-Mallawany, N. K. Five decades of low intensity and low survival: adapting intensified regimens to cure pediatric Burkitt lymphoma in Africa. *Blood Adv.* **4**, 4007–4019 (2020).

32. Noy, A. Burkitt Lymphoma — Subtypes, Pathogenesis, and Treatment Strategies. *Clin. Lymphoma Myeloma Leuk.* **20**, S37–S38 (2020).

33. Rodrigo, J. A. *et al.* HIV-associated Burkitt lymphoma: Good efficacy and tolerance of intensive chemotherapy including CODOX-M/IVAC with or without rituximab in the HAART era. *Adv. Hematol.* **2012**, (2012).

34. Petrich, A. M., Sparano, J. A. & Parekh, S. Paradigms and Controversies in the Treatment of HIV-Related Burkitt Lymphoma. *Adv. Hematol.* **2012**, 403648 (2012).

35. Hoffman, B. & Liebermann, D. a. Apoptotic signaling by c-MYC. *Oncogene* **27**, 6462–72 (2008).

36. Carroll, P. A., Freie, B. W., Mathsyaraja, H. & Eisenman, R. N. The MYC transcription factor network: balancing metabolism, proliferation and oncogenesis.

Front. Med. **12**, 412–425 (2018).

37. Blackwood, E. & Eisenman, R. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science (80-.).* **251**, 1211–1217 (1991).

38. Billin, A. N., Eilers, A. L., Queva, C. & Ayer, D. E. Mlx, a novel Max-like BHLHZip protein that interacts with the max network of transcription factors. *J. Biol. Chem.* **274**, 36344–36350 (1999).

39. Meroni, G. *et al.* Mlx, a new Max-like bHLHZip family member: The center stage of a novel transcription factors regulatory pathway? *Oncogene* **19**, 3266–3277 (2000).

40. Schaub, F. X. *et al.* Pan-cancer Alterations of the MYC Oncogene and Its Proximal Network across the Cancer Genome Atlas. *Cell Syst.* **6**, 282-300.e2 (2018).

41. Jack, A. & Barrans, S. Recent advances in the understanding of aggressive B-cell lymphomas. *Curr. Diagnostic Pathol.* **10**, 360–373 (2004).

42. Eick, D. & Bornkamm, G. W. Expression of normal and translocated c-myc alleles in Burkitt's lymphoma cells: evidence for different regulation. *EMBO J.* **8**, 1965–1972 (1989).

43. Salghetti, S. E., Kim, S. Y. & Tansey, W. P. Destruction of Myc by ubiquitin-mediated proteolysis: Cancer-associated and transforming mutations stabilize Myc. *EMBO J.* **18**, 717–726 (1999).

44. Richter, J. *et al.* Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. *Nat. Genet.* **44**, 1316–1320 (2012).

45. Schmitz, R. *et al.* Burkitt lymphoma pathogenesis and therapeutic targets from

structural and functional genomics. *Nature* **490**, 116–120 (2012).

46. Love, C. *et al.* The genetic landscape of mutations in Burkitt lymphoma. *Nat. Genet.* **44**, 1321–1325 (2012).

47. Dunleavy, K., Little, R. F. & Wilson, W. H. Update on Burkitt Lymphoma. *Hematol. Oncol. Clin. North Am.* **30**, 1333–1343 (2016).

48. Bödör, C. & Reiniger, L. Catalog of genetic progression of human cancers: non-Hodgkin lymphoma. *Cancer Metastasis Rev.* **35**, 109–127 (2016).

49. Gehringer, F., Weissinger, S. E., Möller, P., Wirth, T. & Ushmorov, A. Physiological levels of the PTEN-PI3K-AKT axis activity are required for maintenance of Burkitt lymphoma. *Leukemia* **34**, 857–871 (2020).

50. Gehringer, F. *et al.* FOXO1 Confers Maintenance of the Dark Zone Proliferation and Survival Program and Can Be Pharmacologically Targeted in Burkitt Lymphoma. *Cancers (Basel).* **11**, 1427 (2019).

51. Kabrani, E. *et al.* Nuclear FOXO1 promotes lymphomagenesis in germinal center B cells. *Blood* **132**, 2670–2683 (2018).

52. Ramiro, A. R. *et al.* AID is required for c-myc/IgH chromosome translocations in vivo. *Cell* **118**, 431–438 (2004).

53. Chaudhuri, J., Evans, T., Kumar, R. & DiMenna, L. Biological function of activation-induced cytidine deaminase (AID). *Biomed. J.* **37**, 269 (2014).

54. Greisman, H. A. *et al.* IgH partner breakpoint sequences provide evidence that AID initiates t(11;14) and t(8;14) chromosomal breaks in mantle cell and Burkitt lymphomas. *Blood* **120**, 2864–2867 (2012).

55. Mechtcheriakova, D., Svoboda, M., Meshcheryakova, A. & Jensen-Jarolim, E. Activation-induced cytidine deaminase (AID) linking immunity, chronic inflammation, and cancer. *Cancer Immunol. Immunother.* **61**, 1591–1598 (2012).

56. Chandra, V., Bortnick, A. & Murre, C. AID targeting: Old mysteries and new challenges. *Trends Immunol.* **36**, 527–535 (2015).

57. Liu, M. *et al.* Two levels of protection for the B cell genome during somatic hypermutation. *Nature* **451**, 841–845 (2008).

58. Yu, K., Huang, F. T. & Lieber, M. R. DNA Substrate Length and Surrounding Sequence Affect the Activation-induced Deaminase Activity at Cytidine. *J. Biol. Chem.* **279**, 6496–6500 (2004).

59. Neuberger, M. S., Harris, R. S., Di Noia, J. & Petersen-Mahrt, S. K. Immunity through DNA deamination. *Trends Biochem. Sci.* **28**, 305–312 (2003).

60. Petersen-Mahrt, S. K., Harris, R. S. & Neuberger, M. S. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature* **418**, 99–103 (2002).

61. Fear, D. J. Mechanisms regulating the targeting and activity of activation induced cytidine deaminase. *Curr. Opin. Immunol.* **25**, 619–628 (2013).

62. Stavnezer, J. The complex regulation and function of activation-induced cytidine deaminase (AID). *Trends Immunol.* **32**, 194–201 (2011).

63. Rai, K. *et al.* DNA Demethylation in Zebrafish Involves the Coupling of a Deaminase, a Glycosylase, and Gadd45. *Cell* **135**, 1201–1212 (2017).

64. Ramiro, A. R. & Barreto, V. M. Activation-induced cytidine deaminase and active cytidine demethylation. *Trends Biochem. Sci.* **40**, 172–181 (2015).

65. Jiao, J. *et al.* AID and TET2 co-operation modulates FANCA expression by active demethylation in diffuse large B cell lymphoma. *Clin. Exp. Immunol.* **195**, 190–201 (2019).

66. Teater, M. *et al.* AICDA drives epigenetic heterogeneity and accelerates germinal center-derived lymphomagenesis. *Nat. Commun.* **9**, 222 (2018).

67. Nagaoka, H., Tran, T. H., Kobayashi, M., Aida, M. & Honjo, T. Preventing AID, a physiological mutator, from deleterious activation: Regulation of the genomic instability that is associated with antibody diversity. *Int. Immunol.* **22**, 227–235 (2010).

68. Zan, H. & Casali, P. Regulation of Aicda expression and AID activity. *Autoimmunity* **46**, 83–101 (2013).

69. de Yébenes, V. G. *et al.* miR-181b negatively regulates activation-induced cytidine deaminase in B cells. *J. Exp. Med.* **205**, 2199–2206 (2008).

70. Teng, G. *et al.* MicroRNA-155 Is a Negative Regulator of Activation-Induced Cytidine Deaminase. *Immunity* **28**, 621–629 (2008).

71. Recaldin, T., Hobson, P. S., Mann, E. H. & Ramadani, F. miR-29b directly targets activation-induced cytidine deaminase in human B cells and can limit its inappropriate expression in naïve B cells ☆. *Mol. Immunol.* **101**, 419–428 (2018).

72. Borchert, G. M., Holton, N. W. & Larson, E. D. Repression of human activation induced cytidine deaminase by miR-93 and miR-155. *BMC Cancer* **11**, 347 (2011).

73. Pasqualucci, L., Kitaura, Y., Gu, H. & Dalla-Favera, R. From The Cover: PKA-mediated phosphorylation regulates the function of activation-induced deaminase (AID) in B cells. *Proc. Natl. Acad. Sci.* **103**, 395–400 (2006).

74. McBride, K. M. *et al.* Regulation of hypermutation by activation-induced cytidine

deaminase phosphorylation. *Proc. Natl. Acad. Sci.* **103**, 8798–8803 (2006).

75. Demorest, Z. L., Li, M. & Harris, R. S. Phosphorylation directly regulates the intrinsic DNA cytidine deaminase activity of activation-induced deaminase and APOBEC3G protein. *J. Biol. Chem.* **286**, 26568–26575 (2011).

76. Mu, Y., Zelazowska, M. A. & McBride, K. M. Phosphorylation promotes activation-induced cytidine deaminase activity at the Myc oncogene. *J. Exp. Med.* **214**, 3543–3552 (2017).

77. Gazumyan, A. *et al.* Amino-Terminal Phosphorylation of Activation-Induced Cytidine Deaminase Suppresses c-myc/IgH Translocation. *Mol. Cell. Biol.* **31**, 442–449 (2011).

78. Orthwein, A. *et al.* Regulation of activation-induced deaminase stability and antibody gene diversification by Hsp90. *J. Exp. Med.* **207**, 2751–2765 (2010).

79. Ito, S. *et al.* Activation-Induced Cytidine Deaminase Shuttles between Nucleus and Cytoplasm like Apolipoprotein B mRNA Editing Catalytic Polypeptide 1. *PNAS* **101**, 1975–1980 (2003).

80. Wang, Q. *et al.* The cell cycle restricts activation-induced cytidine deaminase activity to early G1. *J. Exp. Med.* **214**, 49–58 (2017).

81. Uchimura, Y., Barton, L. F., Rada, C. & Neuberger, M. S. Reg-γ associates with and modulates the abundance of nuclear activation-induced deaminase. *J. Exp. Med.* **208**, 2385–2391 (2011).

82. McBride, K. M., Barreto, V., Ramiro, A. R., Stavropoulos, P. & Nussenzweig, M. C. Somatic Hypermutation Is Limited by CRM1-dependent Nuclear Export of Activation-induced Deaminase. *J. Exp. Med.* **199**, 1235–1244 (2004).

83. Luo, Y. *et al.* CUL7 E3 Ubiquitin Ligase Mediates the Degradation of Activation-

Induced Cytidine Deaminase and Regulates the Ig Class Switch Recombination in B Lymphocytes. *J. Immunol.* **203**, 269–281 (2019).

84. Bödör, C. *et al.* Aberrant somatic hypermutation and expression of activation-induced cytidine deaminase mRNA in mediastinal large B-cell lymphoma. *Br. J. Haematol.* **129**, 373–376 (2005).

85. Nakamura, M. *et al.* High levels of activation-induced cytidine deaminase expression in adult T-cell leukaemia/lymphoma. *Br. J. Dermatol.* **165**, 437–439 (2011).

86. Takizawa, M. *et al.* AID expression levels determine the extent of cMyc oncogenic translocations and the incidence of B cell tumor development. *J. Exp. Med.* **205**, 1949–57 (2008).

87. Lohr, J. G. *et al.* Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 3879–3884 (2012).

88. Chapuy, B. *et al.* Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat. Med.* **24**, 679–690 (2018).

89. Kasar, S. *et al.* Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. *Nat. Commun.* **6**, 1–12 (2015).

90. Álvarez-Prado, Á. F. *et al.* A broad atlas of somatic hypermutation allows prediction of activation-induced deaminase targets. *J. Exp. Med.* **215**, 761–771 (2018).

91. Robbiani, D. F. *et al.* AID Is Required for the Chromosomal Breaks in c-myc that Lead to c-myc/IgH Translocations. *Cell* **135**, 1028–1038 (2008).

92. Duquette, M. L., Pham, P., Goodman, M. F. & Maizels, N. AID binds to transcription-

induced structures in c-MYC that map to regions associated with translocation and hypermutation. *Oncogene* **24**, 5791–5798 (2005).

93. Richter, J. *et al.* Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. *Nat. Genet.* **44**, 1316–1320 (2012).

94. Panea, R. I. *et al.* The whole-genome landscape of Burkitt lymphoma subtypes. *Blood* **134**, 1598–1607 (2019).

95. Heath, E. *et al.* Epstein-Barr Virus Infection of Naïve B Cells In Vitro Frequently Selects Clones with Mutated Immunoglobulin Genotypes: Implications for Virus Biology. *PLoS Pathog.* **8**, e1002697 (2012).

96. Epeldegui, M., Hung, Y. P., McQuay, A., Ambinder, R. F. & Martínez-Maza, O. Infection of human B cells with Epstein-Barr virus results in the expression of somatic hypermutation-inducing molecules and in the accrual of oncogene mutations. *Mol. Immunol.* **44**, 934–942 (2007).

97. Hui, K. F., Yiu, S. P. T., Tam, K. P. & Chiang, A. K. S. Viral-targeted strategies against EBV-associated lymphoproliferative diseases. *Front. Oncol.* **9**, 1–18 (2019).

98. Rastelli, J. *et al.* LMP1 signaling can replace CD40 signaling in B cells in vivo and has unique features of inducing class-switch recombination to IgG1. *Blood* **111**, 1448–1455 (2008).

99. He, B., Raab-Traub, N., Casali, P. & Cerutti, A. EBV-Encoded Latent Membrane Protein 1 Cooperates with BAFF/BLyS and APRIL to Induce T Cell-Independent Ig Heavy Chain Class Switching. *J. Immunol.* **171**, 5215–5224 (2003).

100. Kim, J. H., Kim, W. S. & Park, C. Epstein-Barr virus latent membrane protein 1

increases genomic instability through Egr-1-mediated up-regulation of activation-induced cytidine deaminase in B-cell lymphoma. *Leuk. Lymphoma* **54**, 2035–2040 (2013).

101. Tobollik, S. *et al.* Epstein-Barr virus nuclear antigen 2 inhibits AID expression during EBV-driven B-cell growth. *Blood* **108**, 3859–3864 (2006).

102. Kalchschmidt, J. S. *et al.* Epstein-Barr virus nuclear protein EBNA3C directly induces expression of AID and somatic mutations in B cells. *J. Exp. Med.* **213**, 921–928 (2016).

103. Rios, L. A. de S., Cloete, B. & Mowla, S. Activation-induced cytidine deaminase: in sickness and in health. *J. Cancer Res. Clin. Oncol.* **146**, 2721–2730 (2020).

104. Dhokotera, T. *et al.* The burden of cancers associated with HIV in the South African public health sector, 2004–2014: a record linkage study. *Infect. Agent. Cancer* **14**, 12 (2019).

105. Achenbach, C. J. *et al.* Mortality after cancer diagnosis in HIV-infected individuals treated with antiretroviral therapy. *AIDS* **25**, 691–700 (2011).

106. Abayomi, E. A. *et al.* Transfusion and Apheresis Science Impact of the HIV epidemic and Anti-Retroviral Treatment policy on lymphoma incidence and subtypes seen in the Western Cape of South Africa , 2002 – 2009 : Preliminary findings of the Tygerberg Lymphoma Study Group. *Transfus. Apher. Sci.* **44**, 161–166 (2011).

107. Phillips, L. & Opie, J. The utility of bone marrow sampling in the diagnosis and staging of lymphoma in South Africa. *Int. J. Lab. Hematol.* **40**, 276–283 (2018).

108. Lee, J. H. *et al.* HIV-Nef and ADAM17-Containing Plasma Extracellular Vesicles Induce and Correlate with Immune Pathogenesis in Chronic HIV Infection.

EBioMedicine **6**, 103–113 (2016).

109. Ferdin, J. *et al.* Viral protein Nef is detected in plasma of half of HIV-infected adults with undetectable plasma HIV RNA. *PLoS One* **13**, 1–10 (2018).

110. El-Amine, R. *et al.* HIV-1 Tat protein induces DNA damage in human peripheral blood B-lymphocytes via mitochondrial ROS production. *Redox Biol.* **15**, 97–108 (2018).

111. Klotman, M. E. *et al.* Kinetics of expression of multiply spliced RNA in early human immunodeficiency virus type 1 infection of lymphocytes and monocytes. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 5011–5 (1991).

112. Lamers, S. L., Fogel, G. B., Huysentruyt, L. C. & McGrath, M. S. HIV-1 nef protein visits B-cells via macrophage nanotubes: a mechanism for AIDS-related lymphoma pathogenesis? *Curr. HIV Res.* **8**, 638–40 (2010).

113. Xue, M. *et al.* HIV-1 Nef and KSHV oncogene K1 synergistically promote angiogenesis by inducing cellular miR-718 to regulate the PTEN/AKT/mTOR signaling pathway. *Nucleic Acids Res.* **42**, 9862–9879 (2014).

114. Yao, S. *et al.* MiRNA-891a-5p mediates HIV-1 Tat and KSHV Orf-K1 synergistic induction of angiogenesis by activating NF-κB signaling. *Nucleic Acids Res.* **43**, 9362–9378 (2015).

115. Chen, X. *et al.* Human immunodeficiency virus type 1 Tat accelerates Kaposi sarcoma-associated herpesvirus Kaposin A-mediated tumorigenesis of transformed fibroblasts in vitro as well as in nude and immunocompetent mice. *Neoplasia* **11**, 1272–84 (2009).

116. Li, S., Bozzo, L., Wu, Z., Lu, W. & Romerio, F. The HIV-1 matrix protein p17

activates the transcription factors c-Myc and CREB in human B cells. *New Microbiol.* **33**, 13–24 (2010).

117. Giagulli, C. *et al.* Opposite effects of HIV-1 p17 variants on PTEN activation and cell growth in B cells. *PLoS One* **6**, (2011).

118. Nyagol, J. *et al.* HIV-1 Tat mimetic of VEGF correlates with increased microvessels density in AIDS-related diffuse large B-cell and Burkitt lymphomas. *J. Hematop.* **1**, 3–10 (2008).

119. Luzzi, A. *et al.* HIV-1 Tat induces DNMT over-expression through microRNA dysregulation in HIV-related non Hodgkin lymphomas. *Infect. Agent. Cancer* **9**, 41 (2014).

120. Mdletshe, N., Nel, A., Shires, K. & Mowla, S. HIV Nef enhances the expression of oncogenic c-MYC and activation-induced cytidine deaminase in Burkitt lymphoma cells, promoting genomic instability. *Infect. Agent. Cancer* **15**, 54 (2020).

121. Kurnaeva, M. A., Sheval, E. V., Musinova, Y. R. & Vassetzky, Y. S. Tat basic domain: A “Swiss army knife” of HIV-1 Tat? *Rev. Med. Virol.* **29**, e2031 (2019).

122. Southgate, C. D. & Green, M. R. The HIV-1 Tat protein activates transcription from an upstream DNA- binding site: implications for Tat function. *Genes Dev* **5**, 2496–2507 (1991).

123. Jeang, K. HIV-1 Tat : Structure and Function. *Hum. retroviruses AIDS 1996 Compend.* 3–18 (1996).

124. Baier-Bitterlich, G. *et al.* Structure and function of HIV-1 and SIV Tat proteins based on carboxy- terminal truncations, chimeric Tat constructs, and NMR modeling. *Biomed. Pharmacother.* **52**, 421–430 (1998).

125. Mancebo, H. S. Y. *et al.* P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro. *Genes Dev.* **11**, 2633–2644 (1997).

126. Fischer, W. H., Fang, S.-M., Jones, K. A., Wei, P. & Garber, M. E. A Novel CDK9-Associated C-Type Cyclin Interacts Directly with HIV-1 Tat and Mediates Its High-Affinity, Loop-Specific Binding to TAR RNA. *Cell* **92**, 451–462 (1998).

127. Raha, T., Cheng, S. W. G. & Green, M. R. HIV-1 Tat stimulates transcription complex assembly through recruitment of TBP in the absence of TAFs. *PLoS Biol.* **3**, 0221–0230 (2005).

128. Pereira, L. A. SURVEY AND SUMMARY A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter. *Nucleic Acids Res.* **28**, 663–668 (2000).

129. Loregian, A., Bortolozzo, K., Boso, S., Caputo, A. & Palù, G. Interaction of Sp1 transcription factor with HIV-1 Tat protein: Looking for cellular partners. *FEBS Lett.* **543**, 61–65 (2003).

130. Roebuck, K. A. & Saifuddin, M. Regulation of HIV-1 transcription. *Gene Expr.* **8**, 67–84 (1999).

131. Szmulewicz, M. N., Novick, G. E. & Herrera, R. J. Effects of Alu insertions on gene function. in *Electrophoresis* vol. 19 1260–1264 (1998).

132. Marban, C. *et al.* Genome-wide binding map of the HIV-1 Tat protein to the human genome. *PLoS One* **6**, (2011).

133. Reeder, J. E., Kwak, Y. T., McNamara, R. P., Forst, C. V. & D'Orso, I. HIV Tat controls RNA Polymerase II and the epigenetic landscape to transcriptionally reprogram target immune cells. *Elife* **4**, 1–44 (2015).

134. Dhamija, N., Choudhary, D., Ladha, J. S., Pillai, B. & Mitra, D. Tat predominantly associates with host promoter elements in HIV-1-infected T-cells - Regulatory basis of transcriptional repression of c-Rel. *FEBS J.* **282**, 595–610 (2015).

135. Lim, S. P. & Garzino-Demo, A. The Human Immunodeficiency Virus Type 1 Tat Protein Up-Regulates the Promoter Activity of the Beta-Chemokine Monocyte Chemoattractant Protein 1 in the Human Astrocytoma Cell Line U-87 MG: Role of SP-1, AP-1, and NF-κB Consensus Sites. *J. Virol.* **74**, 1632–1640 (2000).

136. De Falco, G. *et al.* Interaction between HIV-1 Tat and pRb2/p130: a possible mechanism in the pathogenesis of AIDS-related neoplasms. *Oncogene* **22**, 6214–6219 (2003).

137. Lazzi, S. *et al.* Expression of RB2/p130 tumor-suppressor gene in AIDS-related non-Hodgkin's lymphomas: Implications for disease pathogenesis. *Hum. Pathol.* **33**, 723–731 (2002).

138. Zhou, F. *et al.* HIV-1 Tat Promotes Kaposi's Sarcoma-Associated Herpesvirus (KSHV) vIL-6-Induced Angiogenesis and Tumorigenesis by Regulating PI3K/PTEN/AKT/GSK-3β Signaling Pathway. *PLoS One* **8**, e53145 (2013).

139. Germini, D. *et al.* HIV Tat induces a prolonged MYC relocalization next to IGH in circulating B-cells. *Leukemia* **31**, 2515–2522 (2017).

140. Sall, F. B. *et al.* HIV-1 Tat protein induces aberrant activation of AICDA in human B-lymphocytes from peripheral blood. *J. Cell. Physiol.* **234**, 15678–15685 (2019).

141. Sardo, L., Vakil, P. R., Elbezanti, W., El-Sayed, A. & Klase, Z. The inhibition of microRNAs by HIV-1 Tat suppresses beta catenin activity in astrocytes. *Retrovirology* **13**, 25 (2016).

142. de Yébenes, V. G. *et al.* miR-181b negatively regulates activation-induced cytidine deaminase in B cells. *J. Exp. Med.* **205**, 2199–2206 (2008).

143. Musinova, Y. R., Sheval, E. V., Dib, C., Germini, D. & Vassetzky, Y. S. Functional roles of HIV-1 Tat protein in the nucleus. *Cell. Mol. Life Sci.* **73**, 589–601 (2016).

144. Mdletshe, N. The role of HIV proteins Tat and Nef in HIV / AIDS-related lymphoma : effects on c-MYC and AID expression . (2014).

145. Messeguer, X. *et al.* PROMO: Detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics* (2002) doi:10.1093/bioinformatics/18.2.333.

146. Farre, D. Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic Acids Res.* **31**, 3651–3653 (2003).

147. J. Sambrook i D. W. Russell. *Molecular cloning : a laboratory manual, III. Red. New York: Cold Spring Harbor Laboratory Press* (2001). doi:10.3724/SP.J.1141.2012.01075.

148. Koh, C. M. *Isolation of genomic DNA from mammalian cells*. *Methods in Enzymology* vol. 529 (Elsevier Inc., 2013).

149. Iavarone, C. *et al.* The Platelet-derived Growth Factor Controls c-myc Expression through a JNK- and AP-1-dependent Signaling Pathway. *J. Biol. Chem.* **278**, 50024–50030 (2003).

150. van der Sluis, R. M. *et al.* Interplay between viral Tat protein and c-Jun transcription factor in controlling LTR promoter activity in different human immunodeficiency virus type I subtypes. *J. Gen. Virol.* **95**, 968–979 (2014).

151. Hidalgo-Estévez, A. M., González, E., Punzón, C. & Fresno, M. Human

immunodeficiency virus type 1 Tat increases cooperation between AP-1 and NFAT transcription factors in T cells. *J. Gen. Virol.* **87**, 1603–1612 (2006).

152. Carvallo, L. *et al.* HIV-Tat regulates macrophage gene expression in the context of neuroAIDS. *PLoS One* **12**, e0179882 (2017).

153. Kumar, L. J., Ghosh, M. R. & Tucker, P. W. HIV-tat protein activates c-jun N-terminal kinase and activator protein-1. *J. Immunol.* **161**, 776 (1998).

154. Hidalgo-Estévez, A. M., González, E., Punzón, C. & Fresno, M. Human immunodeficiency virus type 1 Tat increases cooperation between AP-1 and NFAT transcription factors in T cells. *J. Gen. Virol.* **87**, 1603–1612 (2006).

155. Abayomi, E. A. *et al.* Impact of the HIV epidemic and Anti-Retroviral Treatment policy on lymphoma incidence and subtypes seen in the Western Cape of South Africa, 2002–2009: Preliminary findings of the Tygerberg Lymphoma Study Group. *Transfus. Apher. Sci.* **44**, 161–166 (2011).

156. Phillips, L. & Opie, J. The utility of bone marrow sampling in the diagnosis and staging of lymphoma in South Africa. *Int. J. Lab. Hematol.* **40**, 276–283 (2018).

157. Kim, N., Kukkonen, S., Gupta, S. & Aldovini, A. Association of Tat with promoters of PTEN and PP2A subunits is key to transcriptional activation of apoptotic pathways in HIV-infected CD4+ T cells. *PLoS Pathog.* **6**, (2010).

158. Ju, S. M. *et al.* Extracellular HIV-1 tat induces human beta-defensin-2 production via NF-kappaB/AP-1 dependent pathways in human B cells. *Mol. Cells* **33**, 335–341 (2012).

159. Ding, X. *et al.* JNK/AP1 Pathway Regulates MYC Expression and BCR Signaling through Ig Enhancers in Burkitt Lymphoma Cells. *J. Cancer* **11**, 610–618 (2020).

160. Epeldegui, M. *et al.* Elevated expression of activation induced cytidine deaminase in peripheral blood mononuclear cells precedes AIDS-NHL diagnosis. *AIDS* **21**, 2265–2270 (2007).

161. Frankel, A. D. & Pabo, C. O. Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* **55**, 1189–1193 (1988).

162. Liu, Y. *et al.* Uptake of HIV-1 Tat protein mediated by low-density lipoprotein receptor-related protein disrupts the neuronal metabolic balance of the receptor ligands. *Nat. Med.* **6**, 1380–1387 (2000).

163. Sall, F. B. *et al.* HIV-1 Tat protein induces aberrant activation of AICDA in human B-lymphocytes from peripheral blood. *J. Cell. Physiol.* **234**, 15678–15685 (2019).

164. Huang, H.-Y. *et al.* miRTarBase 2020: updates to the experimentally validated microRNA–target interaction database. *Nucleic Acids Res.* **48**, D148–D154 (2019).

165. Vejnar, C. E. & Zdobnov, E. M. miRmap: Comprehensive prediction of microRNA target repression strength. *Nucleic Acids Res.* **40**, 11673–11683 (2012).

166. Agarwal, V., Bell, G. W., Nam, J.-W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *Elife* **4**, 1–38 (2015).

167. Corpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**, 10881–10890 (1988).

168. Abraham, S. *et al.* Cooperative interaction of C/EBPβ and Tat modulates MCP-1 gene transcription in astrocytes. *J. Neuroimmunol.* **160**, 219–227 (2005).

169. Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. & Bonner, W. M. Double-stranded Brekas Induce Histone H2AX phosphorylation on Serine 139. *J. Biol. Chem.* **273**, 5858–5868 (1998).

170. Rogakou, E. P., Boon, C., Redon, C. & Bonner, W. M. Megabase chromatin domains involved in DNA double-strand breaks in vivo. *J. Cell Biol.* **146**, 905–915 (1999).

171. Kinner, A., Wu, W., Staudt, C. & Iliakis, G. Gamma-H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin. *Nucleic Acids Res.* **36**, 5678–5694 (2008).

172. Rana, T. M. Illuminating the silence: understanding the structure and function of small RNAs. *Nat. Rev. Mol. Cell Biol.* **8**, 23–36 (2007).

173. Lee, R. C., Feinbaum, R. L. & Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843–854 (1993).

174. Wightman, B., Ha, I. & Ruvkun, G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell* **75**, 855–862 (1993).

175. Bennasser, Y. & Jeang, K.-T. HIV-1 Tat interaction with Dicer: requirement for RNA. *Retrovirology* **3**, 95 (2006).

176. Bennasser, Y., Le, S.-Y., Benkirane, M. & Jeang, K.-T. Evidence that HIV-1 Encodes an siRNA and a Suppressor of RNA Silencing. *Immunity* **22**, 607–619 (2005).

177. Dzikiewicz-Krawczyk, A. *et al.* Argonaute 2 RNA Immunoprecipitation Reveals Distinct miRNA Targetomes of Primary Burkitt Lymphoma Tumors and Normal B Cells. *Am. J. Pathol.* **188**, 1289–1299 (2018).

178. Mapekula, L., Ramorola, B. R., Hoosen, T. G. & Mowla, S. Critical Reviews in Oncology / Hematology The interplay between viruses & host microRNAs in cancer – An emerging role for HIV in oncogenesis. *Crit. Rev. Oncol. / Hematol.* **137**, 108–114

(2019).

179. Kuo, H.-H. & Lichterfeld, M. Recent progress in understanding HIV reservoirs. *Curr. Opin. HIV AIDS* **13**, 137–142 (2018).

180. Caccuri, F. *et al.* Cellular aspartyl proteases promote the unconventional secretion of biologically active HIV-1 matrix protein p17. *Sci. Rep.* **6**, 38027 (2016).

181. Popovic, M. *et al.* Persistence of HIV-1 structural proteins and glycoproteins in lymph nodes of patients under highly active antiretroviral therapy. *Proc. Natl. Acad. Sci.* **102**, 14807–14812 (2005).

182. Dolcetti, R. *et al.* Role of HIV-1 matrix protein p17 variants in lymphoma pathogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 14331–14336 (2015).

183. Caccuri, F. *et al.* HIV-1 matrix protein p17 promotes angiogenesis via chemokine receptors CXCR1 and CXCR2. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 14580–14585 (2012).

184. Basta, D. *et al.* Angiogenic, lymphangiogenic and adipogenic effects of HIV-1 matrix protein p17. *Pathog. Dis.* **73**, 1–8 (2015).

185. Matsumoto, Y. *et al.* Helicobacter pylori infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. *Nat. Med.* **13**, 470–476 (2007).

186. Kim, J. H., Kim, W. S. & Park, C. Epstein – Barr virus latent membrane protein 1 increases genomic instability through Egr-1-mediated up-regulation of activation-induced cytidine deaminase in B-cell lymphoma. **54**, 2035–2040 (2013).

187. Kalchschmidt, J. S. *et al.* Epstein-Barr virus nuclear protein EBNA3C directly induces expression of AID and somatic mutations in B cells. *J. Exp. Med.* **213**, 921–928

(2016).

188. Torgbor, C. *et al.* A Multifactorial Role for P. falciparum Malaria in Endemic Burkitt's Lymphoma Pathogenesis. *PLoS Pathog.* **10**, (2014).

189. Meng, F.-L. *et al.* Convergent Transcription at Intragenic Super-Enhancers Targets AID-Initiated Genomic Instability. *Cell* **159**, 1538–1548 (2014).

190. Yu, K., Roy, D., Bayramyan, M., Haworth, I. S. & Lieber, M. R. Fine-Structure Analysis of Activation-Induced Deaminase Accessibility to Class Switch Region R-Loops. *Mol. Cell. Biol.* **25**, 1730–1736 (2005).

191. Kenter, A. L., Kumar, S., Wuerffel, R. & Grigera, F. AID hits the jackpot when missing the target. *Curr. Opin. Immunol.* **39**, 96–102 (2016).

192. Pavri, R. *et al.* Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* (2010) doi:10.1016/j.cell.2010.09.017.

193. Chaudhuri, J., Khuong, C. & Alt, F. W. Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. *Nature* **430**, 992–998 (2004).

194. Xu, Y. *et al.* Activation-induced cytidine deaminase localizes to G-quadruplex motifs at mutation hotspots in lymphoma. *NAR Cancer* **2**, 9–12 (2020).

195. Siddiqui-Jain, A., Grand, C. L., Bearss, D. J. & Hurley, L. H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci.* **99**, 11593–11598 (2002).

196. Dexheimer, T. S., Sun, D. & Hurley, L. H. Deconvoluting the Structural and Drug-Recognition Complexity of the G-Quadruplex-Forming Region Upstream of the bcl-2 P1 Promoter. *J. Am. Chem. Soc.* **128**, 5404–5415 (2006).

197. Qiao, Q. *et al.* AID Recognizes Structured DNA for Class Switch Recombination. *Mol. Cell* **67**, 361-373.e4 (2017).

198. Kessing, C. F. *et al.* In Vivo Suppression of HIV Rebound by Didehydro-Cortistatin A, a "Block-and-Lock" Strategy for HIV-1 Treatment. *Cell Rep.* **21**, 600–611 (2017).

199. Mediouni, S. *et al.* The Tat inhibitor didehydro-cortistatin A suppresses SIV replication and reactivation. *FASEB J.* **33**, 8280–8293 (2019).

Appendix A: recipes and reagents

Bacterial culture

Lysogeny Broth (LB) (1 L)

Dissolve 10 g of LB powder media in 800 mL ddH_2O and adjust the volume to 1 L

Sterilise by autoclaving and store at room temperature

LB agar

Mix 1.5 g Agar and 2 g LB powder media in 100 mL ddH_2O

Sterilize by autoclaving and store at room temperature

100 mM $CaCl_2$ (50 mL)

Dissolve 0.735g $CaCl_2$ in 40 mL ddH_2O

Adjust volume to 50 mL with ddH_2O and filter sterilise using a 0.2 µm syringe filter

Make fresh before use each time and store at 4°C

25% glycerol stocks

Add 0.5 mL 50% sterile glycerol to 0.5 mL bacterial liquid culture and mix gently

Flash-freeze with liquid nitrogen and transfer to -80°C for long term storage

Tissue culture

Supplemented growth media for Ramos cell line

Add 5 mL (10%) of FBS and 0.5 mL of P/S to a 50 mL sterile conical tube

Fill up to 50 mL with RPMI-1640 media and mix

Store at 4°C until needed

Serum rich media for post-electroporation incubation

Add 10 mL (20%) of FBS and 0.5 mL of P/S to a 50 mL sterile conical tube

Fill up to 50 mL with RPMI-1640 media and mix

Store at 4°C until needed

Supplemented growth media for HT1080 cell line

Add 5 mL (10%) of FBS and 0.5 mL of P/S to a 50 mL sterile conical tube

Fill up to 50 mL with DMEM media and mix

Store at 4°C until needed

Freezing media for Ramos and HT1080 cell lines

10% DMSO and 10% FBS in RPMI-1640/DMEM media (as indicated above for each cell line)

Place on ice until use

0.1% diethyl pyrocarbonate- (DEPC) treated water

Fill a 5 L beaker with autoclaved water and add 1 mL of DEPC per 500 mL of ddH_2O

Add a magnetic stirrer bar and place on the stirrer for a few hours (3-5 hrs)

Place the beaker in the fume hood overnight, remove the stirrer bar and autoclave to inactivate the remaining DEPC

Store the H_2O at room temperature until needed

0.1% diethyl pyrocarbonate- (DEPC) to treat plasticware

Fill 5 L beaker with ddH_2O and add 1 mL of DEPC per 1 L of H_2O

Add stirrer bar and plasticware (pipette tips, centrifuge tubes)

Place on the magnetic stirrer for 3-5 hrs

Place the beaker in a fume hood overnight

Remove the pipette tips from the H_2O and allow to air dry before autoclaving

RNA extraction

10x Tris/Borate/EDTA (TBE) buffer (1 L)

Dissolve 108 g Tris and 55 g Boric acid in 900 mL ddH_2O using a magnetic stirrer

Add 40 mL EDTA (0.5M pH 8)

Adjust to 1 L

Store at room temperature

1x TBE (1 L)

Measure 100 mL of 10x TBE stock

Fill up to 1 L using ddH2O

2x RNA loading buffer (1 mL)

Add 900 µL formamide, 99 µL sucrose, 0.5 µL Bromophenol Blue and 0.5 µL Xylene Cyanol

Store at room temperature

10x DNA loading dye (20 mL)

Mix 0.025 g of Xylene cyanol, 0.025 g of Bromophenol Blue to 6.25 mL of ddH_2O and dissolve using a magnetic stirrer

Add 1.25 mL of 10% SDS and 12.5 mL of glycerol

Aliquot and store at room temperature

1%/1.5% agarose gel (100 mL)

Weigh 1 g/1.5 g agarose and add to 100 mL 1x TBE

Dissolve all the agarose by heating using a microwave

Allow to cool and then add 5 µL Ethidium Bromide (EtBr, 0.5 mg/mL)

Pour gel on a tray with a well comb and allow to set

Protein Extraction

RIPA Buffer (50 mL)

Mix 1.5 mL 150 nM NaCl; 0.5 mL 1% Triton X100; 0.5 mL 0.1% SDS; 0.5 mL 10 mM Tris pH 7.5; 0.5 g 1% deoxycholate powder

Adjust the volume to 50 mL with ddH_2O

7x protease inhibitor

Add 1 protease inhibitor tablet to 1.5 mL 1x PBS and allow to dissolve

Aliquot and store at -20°C

RIPA solution (1 mL)

Combine 858 µL RIPA buffer with 142 µL 7x protease inhibitor

Make fresh before each use and store at 4°C

2x Laemmli buffer (10 mL)

Mix 0.125 mM Tris-HCL pH 6.8, 4% SDS, 10% β-mercaptoethanol, 20% glycerol, and 0.005% of bromophenol blue in a final volume of 10 mL made up with ddH_2O, aliquot to 1.5mL Eppendorf tubes and store at -20°C

SDS-PAGE Reagents

1.5M Tris pH 6.8 (500 mL)

Dissolve 60.5 g Tris base in 300 mL deionized H2O

Adjust pH to 6.8 with concentrated HCl and volume to 500 mL with ddH_2O

Store at 4°C

1.5M Tris pH 8.8 (500 mL)

Dissolve 60.5 g Tris base in 300 mL ddH2O

Adjust pH to 8.8 with concentrated HCl and volume to 500 mL with ddH2O

Store at 4°C

10% Sodium dodecyl sulphate (SDS) (100 mL)

Dissolve 10 g sodium dodecyl sulphate crystals in 90 mL ddH_2O

Adjust volume to 100 mL with ddH2O

Store at room temperature

10% Ammonium persulfate (APS) (1 mL)

Dissolve 0.1 g ammonium persulfate in 1 mL ddH_2O

Store at 4°C

30% acryl-bisacrylamide mix (100 mL)

Dissolve 29 g acrylamide and 1 g N.N'-methylenebisacrylamde in 60 mL ddH_2O

Heat the solution to 37°C to dissolve the chemicals

Adjust volume to 100 mL with ddH_2O

Filter through 0.45 μM membrane

Store at 4°C protected from light

10x SDS-PAGE running buffer (1 L)

Dissolve 10 g SDS; 30.3 g Tris and 144.1 g glycine in 800 mL ddH_2O

Adjust volume to 1 L with ddH_2O

Store at room temperature

1x SDS-PAGE running buffer (1 L)

Add 100 mL 10X SDS-PAGE running buffer to 900 mL ddH_2O

Store at room temperature

10x SDS-PAGE transfer buffer (1 L)

Dissolve 144 g glycine and 38 g Tris in 800 mL ddH_2O

Adjust volume to 1 L with ddH_2O

Store at room temperature

1x SDS-PAGE transfer buffer (1 L)

Mix 100 mL 10X transfer buffer; 700 mL ddH_2O and 200 mL isopropanol

Make fresh on the day of running gel and store at 4°C

1x Ponceau S staining solution (0.1% (w/v) Ponceau S in 5% (v/v) acetic acid) (50 mL)

Dissolve 0.1g Ponceau S in 5 mL acetic acid

Adjust volume to 100 mL with deionized ddH_2O and protect from light

Store at room temperature

12% resolving gel for SDS-PAGE (7,5 mL)

Mix 2.4 mL ddH_2O; 3 mL 30% acryl-bisacrylamide mix; 1.95 mL 1.5M Tris (pH 8.8); 0.075 mL 10% SDS; 0.075 mL 10% APS and 0.003 mL tetramethylethylenediamine (TEMED) (inside a fume hood)

Pour into 1 mm glass plate, minimise exposure to air

Add 1% SDS to set the top of the gel

15% resolving gel for SDS-PAGE (7,5 mL)

Mix 1.65 mL ddH_2O; 3.75 mL 30% acryl-bisacrylamide mix; 1.95 mL 1.5M Tris (pH 8.8); 0.075 mL 10% SDS; 0.075 mL 10% APS and 0.003 mL tetramethylethylenediamine (TEMED) (inside a fume hood)

Pour into 1 mm glass plate, minimise exposure to air

Add 1% SDS to set the top of the gel

3% stacking gel for denaturing SDS-PAGE (3 mL)

Mix 2.1 mL ddH_2O; 0.5 mL 30% acryl-bisacrylamide mix; 0.38 mL 1.5M Tris (pH 6.8); 0.03 mL 10% SDS; 0.03 mL 10% APS and 0.003 mL tetramethylethylenediamine (TEMED) (inside a fume hood)

Pour into 1 mm glass plate which has the set resolving gel, add comb and leave to set

5x SDS loading dye (10 mL)

Dissolve 1 g SDS (10%) and 0.004 g Bromophenol Blue (0.04%) in 5.25 mL ddH_2O and add 1.25 mL 2 M Tris-HCl pH 6.8; 3 mL 100% glycerol and 0.5 mL β-mercaptoethanol

Aliquot and tore at room temperature

<u>Western Blotting</u>

10xPBS (1 L)

Weigh out 80 g NaCl, 2 g KCl, 14.4 g Na_2HPO_4 and 2.4 g KH_2PO_4 and transfer to a 1 L beaker

Dissolve in 800 mL ddH_2O using a magnetic stirrer

Adjust pH to 7.4 with HCl and volume to 1 L with ddH_2O.

Aliquot the 10x PBS and sterilize by autoclaving

Store at 4°C

1x PBS (1 L)

Measure 100 mL of 10x PBS and add to 900 mL ddH_2O.

Aliquot as needed and store at 4°C

1x PBS/0.1% Tween (PBST) (1 L)

Add 1 mL Tween-20 to 1x PBS made up to 1 L

Mix well with magnetic stirrer bar

Store at 4°C

Blocking Solution (50 mL)

Add 2.5 g non-fat powder milk (5%) and dissolve in 50 mL in PBST

Store at 4°C

1M Tris-HCl (pH 6.7) used in stripping buffer (200 mL)

Dissolve 24.22 g Tris base in 160 mL ddH_2O

Adjust pH to 6.7 with concentrated HCl and adjust the volume to 200 mL with ddH_2O

Store at 4°C

Stripping buffer (100 mL)

Combine 0.69 mL 100 mM β-mercaptoethanol, 10 mL 2% SDS, 6.25 mL 62.5 mM Tris HCl (pH 6.7) adjusted with ddH_2O to 100 mL

Store at room temperature

ChIP buffers and reagents

1% Formaldehyde

Add 0.548 mL of 36.5% Formaldehyde stock solution to 20 mL of media containing cells

1.25 M Glycine (100 mL)

Dissolve 9.384 g Glycine in 100 mL 1x PBS

Sterilise by autoclaving

Store at room temperature

NCP buffer I (50 mL)

Mix a final concentration of 10mM EDTA, 0.5 mM EGTA, 10 mM HEPES pH 6.5 and 0.25% Triton-X in a total of 50 mL ddH_2O

Store at 4°C

NCP buffer II (50 mL)

Mix a final concentration of 1mM EDTA, 0.5 mM EGTA, 10 mM HEPES pH 6.5 and 200mM NaCl in a total of 50 mL ddH_2O

Store at 4°C

ChIP lysis buffer (10 mL)

Mix a final concentration of 10 mM EDTA, 50 mM Tris pH 8.0, 1% SDS and 0.5% Triton-X in a total of 10 mL ddH_2O

Store at 4°C

ChIP lysis buffer + protease inhibitors (2.5 mL)

Add 0.0125 mL of 100 mM PMSF and 0.357 mL of 7x protease inhibitors to 2.13 mL of ChIP lysis buffer

Make fresh each time before using and store at 4°C

RIPA-IP buffer (20 mL)

Mix a final concentration of 2 mM EDTA, 150 mM NaCl, 20 mM Tris pH 8.0 and 1% Triton-X in a total of 20 mL ddH_2O

Store at 4°C

RIPA-IP buffer with protease inhibitors (1 mL)

Add 0.005 mL 100 mM PMSF and 0.143 mL 7x protease inhibitor to a total of 1 mL RIPA-IP buffer

Make fresh each time before using and store at 4°C

Wash buffer 1 (10 mL)

Mix a final concentration of 2 mM EDTA, 150 mM NaCl, 20 mM Tris pH 8.0, 0.1% SDS and 1% Triton-X in a total of 10 mL ddH_2O

Store at room temperature

Wash buffer 2 (10 mL)

Mix a final concentration of 2 mM EDTA, 500 mM NaCl, 20 mM Tris pH 8.0, 0.1% SDS and 1% Triton-X in a total of 10 mL ddH_2O

Store at room temperature

Wash buffer 3 (10 mL)

Mix a final concentration of 1 mM EDTA, 250 mM LiCl, 10 mM Tris pH 8.0, 0.1% sodium deoxycholate and 1% Triton-X in a total of 10 mL ddH_2O

Store at room temperature

Wash buffer 4 (10 mL)

Mix a final concentration of 1 mM EDTA and 10 mM Tris pH 8.0 in a total of 10 mL ddH_2O

Store at room temperature

Extraction buffer (10 mL)

Mix a final concentration of 1 % SDS and 1 M $NaHCO_3$ in a total of 10 mL ddH_2O

Store at room temperature