PREDICTING HBZ INTERACTION WITH *BATF3* ENHANCER REGIONS THROUGH CHIP-SEQ

by

Mikayla Luke

A Senior Honors Project Presented to the

Honors College

East Carolina University

In Partial Fulfillment of the

Requirements for

Graduation with Honors

by

Mikayla Luke

Greenville, NC

May, 2021

Approved by:

Nicholas Polakowski

Department of Microbiology/Immunology, Brody School of Medicine

Human Lymphotropic Virus Type 1 (HTLV-1) is a complex retrovirus that causes Adult T-cell Leukemia (ATL), an often-fatal form of cancer characterized by uncontrolled proliferation of infected CD4+ T-cells. The viral protein HTLV-1 basic leucine zipper factor (HBZ) has been proven as essential for HTLV-1 proliferation and is believed to contribute to progression to ATL and maintenance of the disease. HBZ functions as a transcriptional regulator. The knock-down of the hbz gene suppresses ATL proliferation, supporting that HBZ is essential for leukemogenesis. Recently, the basic leucine zipper transcription factor ATF-like 3 (BATF3) and interferon regulatory factor 4 (IRF4) were shown to help drive the ATL-specific transcription program. Both BATF3 and IRF4 promote T-cell differentiation and proliferation. HBZ was found to control BATF3 transcription by binding an enhancer region of the BATF3 gene. However, the molecular mechanism used by HBZ to control BATF3 transcription was not resolved. This review utilizes data analyses of chromatin immunoprecipitation specimens evaluated by next generation sequencing (ChIP-seq) to predict how HBZ interacts with the BATF3 enhancer region and then produces the elevated transcription level.

Introduction

Worldwide, 10-20 million people are infected with HTLV-1. HTLV-1-endemic regions include Japan, Sub-Saharan Africa, the Caribbean basin and South America. Viral transmission occurs primarily through blood transfusion, breastfeeding and sexual intercourse. About 5% of those infected will develop an HTLV-1-associated disease (Futsch, 2017). A major disease caused by the virus is Adult T-Cell Leukemia (ATL). ATL develops from a single infected CD4+

T-cell that accumulates genetic alterations through the actions of viral proteins. These alterations transform the cell, leading to its monoclonal expansion. Clinical features of aggressive subtypes of ATL include skin lesions, hypercalcemia, bone lesions, organ dysfunction and failure, and opportunistic infections (Futsch, 2017). ATL is often fatal with few effective treatment options.

HTLV-1 is a complex retrovirus. Retroviruses contain an RNA genome that, upon infection, is converted into DNA by the viral enzyme, reverse transcriptase. The viral DNA genome is then incorporated into one of the host cell's chromosomes. This "provirus" then is used to produce viral proteins. Retroviruses share the genes organization from 5' to 3' of *gag, pro, pol* and *env* flanked by long terminal repeats (LTRs). HTLV-1 is classified as a complex retrovirus, as it contains additional regulatory genes, *tax, rex*, and *hbz*. HBZ is uniquely positioned on the minus RNA strand. Transcription of *hbz* is regulated by a bidirectional promoter in the 3' LTR (Gaudray et al. 2002). This promoter has been found to remain intact and functional even in defective proviruses which arise during ATL development and progression (Matsuoka, Mesnard, 2020). Defective proviruses lose the ability to express sense strand viral genes, leaving *hbz* as the only viral gene expressed, however leukemogenesis still occurs in these defective cell lines, suggestingHBZ is essential for leukemogenesis in ATL. HBZ contains three distinct domains. The activation domain (AD) at the N-terminus directly interacts with the cellular coactivators CBP/p300. The basic regions in the center mediate nuclear localization of HBZ in ATL cells . The 3' bZIP region is subdivided into a basic region, which is potentially involved in DNA-binding, and a leucine zipper (ZIP) domain that forms coiled-coil interactions with certain cellular bZIP transcription factors, such as c-Jun (Cook et al. 2011).

HBZ was recently shown to promote the survival of ATL cells by activating the expression of BATF3. Survival of ATL cells was then sustained by transcriptional regulation

through enhancer complexes containing BATF3 and IRF4. BATF3 is essential for CD8+ and dendritic cell differentiation (Chandra, 2017). In HTLV1 infected cells, BATF3 binds to its own super-enhancer creating a positive autoregulatory loop after HBZ binds, upregulating itself and its downstream targets, such as oncogenic MYC. IRF4 is a transcription factor important to interferon inducible genes and differentiation from B cells to plasma cells. In normal T cells, an IRF4-BATF3 complex binds to AP-1 DNA sites (Nakagawa et al. 2018).

Methods

In order to look for further evidence of suspected protein-DNA interactions, existing data was examined from the Nakagawa et al. GEO DataSets (Figure 1). RNA-seq data sets were included if they showed consistent significant decrease in expression, moderate decrease, or increase in order to show varieties. Days and cell lines with multiple repeats were chosen and averages were taken. For ChIP-seq datasets, both KK1 and ST1 cell lines with IRF4, H3K27ac (a modification produced by the acetyltransferase activity of CBP/p300), HBZ, and IgG control treatments were included (Table 1).

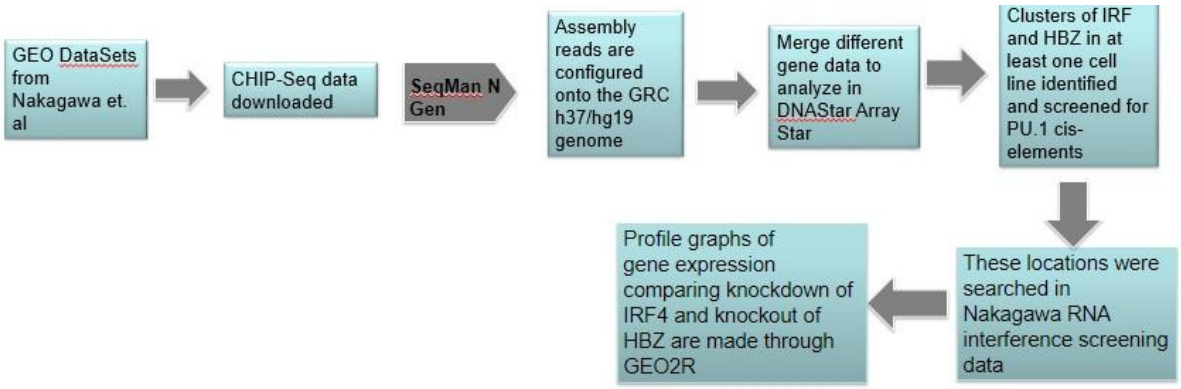


Figure 1. Flow chart of GEO data set analysis

KK1 shIRF4 - day 1 - repeat 1	KK1 shIRF4 - day 1 - repeat 2	KK1 sgHBZ_1 – day -7	KK1 sgHBZ_2 - day 7
GSM2474923	GSM2474925	GSM2474939	GSM2474942

KK1 IRF4	ST1 IRF4	KK1 DMSO Treated H3K27ac	KK1 BirA HBZ	ST1 BirA HBZ	KK1 IgG Control	ST1 IgG Control
GSM2481669	GSM2481671	GSM2481682	GSM2481679	GSM2481681	GSM2481668	GSM2481670

Table 1. GEO datasets used for RNA-seq (top) and ChIP-seq (bottom)

Results

Through ChIP-seq analyses (Nakagawa et al. 2018) IRF4, BATF3 and HBZ were found to associate with a super-enhancer that activates BATF3 transcription. IRF4 and BATF3 interact, and the super-enhancer contains a cis-element for this complex (Figure 2). HBZ does not interact with IRF4 or BATF3, and the super-enhancer does not contain potential cis-elements for HBZ DNA-binding. The super-enhancer contains a cis-element for a IRF4/PU.1 (SPI1). The basic region of the bZIP domain of c-Jun interacts with PU.1. Through these observations, we suggest that IRF4 and PU.1 form a complex that directly binds to the DNA, c-Jun and HBZ are recruited to the complex through protein-protein interactions, and HBZ recruits CBP/p300 to activate transcription (Figure 3). This mechanism controls the transcription of other genes.

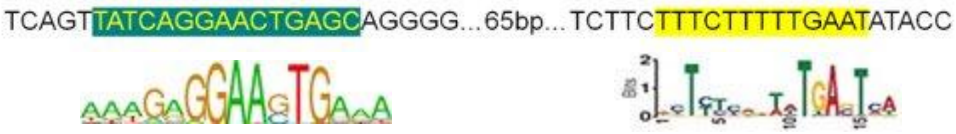


Fig 2. **A portion of the super-enhancer with the cis-elements for IRF4/PU.1-binding (green) and IRF4/BATF3-binding (yellow).**

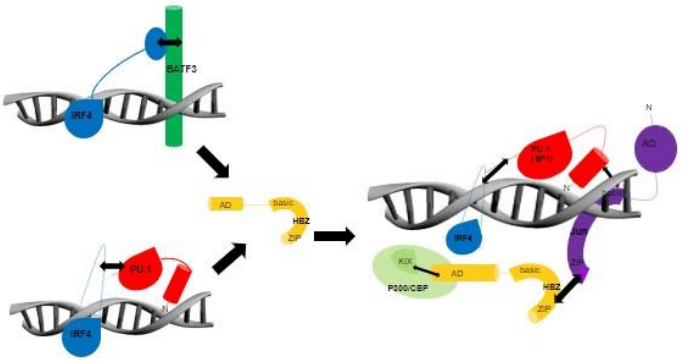


Figure 3: Proposed mechanism of HBZ mediated BATF3 enhancer region expression. The top left shows the IRF4-BATF3 complex that binds to AP-1 sites as previously mentioned. Bottom left is suspected interaction between IRF4 and PU.1. Right shows proposed mechanism of how HBZ is recruited to the BATF3 enhancer region to control transcription.

Evidence of such a complex existing may be found by reduced expression of HTLV-1 infected cell lines at gene locations other than the BATF3 super-enhancer when each of the proteins in the proposed complex are knocked down. RNA-seq found that multiple of these gene locations do exist (Figure 4). Further evidence exists in increased expression for these same gene sites in ChIP-seq data for these same gene sites (Figure 5).

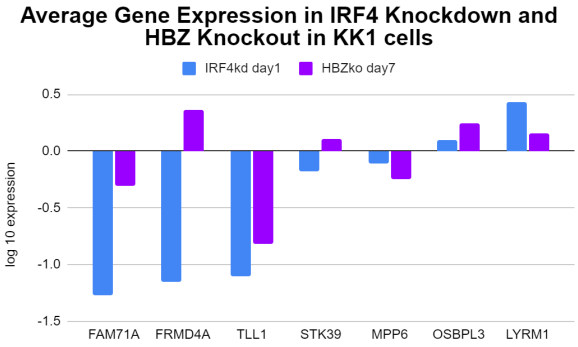
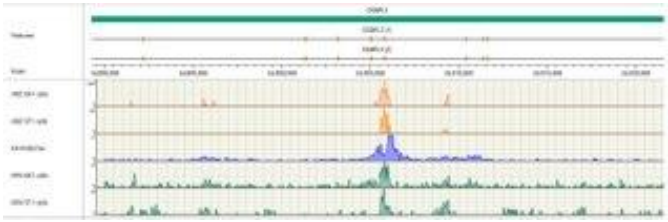


Fig. 4: Gene expression levels in KK1 infected cell lines when IRF4 is knocked down (blue) and when HBZ is knocked out (purple). The specific genes are shown along the bottom of the graph on the x-axis and expression is shown as log 10 values. Decreased expression suggests these genes are potentially regulated by the IRF4/PU.1/Jun/HBZ complex

A

B



C

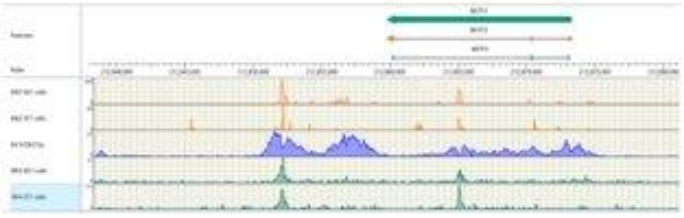


Fig. 5: ChIP-seq peaks for genes in infected cell lines for HBZ (top), H3K27ac (middle), as it's a modification produced by the acetyltransferase activity of CBP/p300 , and IRF4 (bottom). (A) TLL1 peaks on far-left (B) OSBPL3 (C)BATF3. Increased interactions with these proteins suggest further evidence this complex exists.

Discussion

The presence of other locations in the HTLV-1 infected cell genome that allow the proposed p300/CBP-HBZ-c-JUN-PU.1-IRF4 complex to bind and regulate transcription of these genes supports the hypothesis that such a complex exists. At least four of these locations were found, including FAM71A, FRMD4A, LTT1, and MPPG.

Gene	**Function**
FAM71A	Possibly essential for Golgi body integrity.
STK39	A serine/threonine kinase, thought to function in response to hypotonic stress, activating the p38 MAP kinase pathway.
FRMD4A	FERM domain containing protein that regulates epithelial cell polarity.
MPP6	Member of the MAGUK family function in tumor suppression and receptor clustering by forming multiprotein complexes containing sets of different signaling proteins.

Table 3. Functions of identified genes with suggested interactions with proposed complex in non-infected cells

Future research should be done to confirm this proposed mechanism in order to further develop and understand treatments. Experiments include verification of PU.1 association with IRF4 through CHIP-seq and reporter assays. Recently, treatments for ATL have been proposed utilizing BET-domain inhibitors, which utilizes JQ1 treatment to inhibit super-enhancer function, thus inhibiting cancer progression through BATF3 targeting. Further understanding of this mechanism will further this treatment course.

References

Bangham CR, Ratner L. How does HTLV-1 cause adult T-cell leukaemia/lymphoma (ATL)?. Curr Opin Virol. 2015;14:93-100. doi:10.1016/j.coviro.2015.09.004

Chandra J, Kuo PT, Hahn AM, Belz GT, Frazer IH. Batf3 selectively determines acquisition of CD8+ dendritic cell phenotype and function. Immunol Cell Biol. 2017;95(2):215-223. doi:10.1038/icb.2016.83

Cook P R, Polakowski N, Lemasson I, HTLV-1 HBZ Protein Deregulates Interactions between Cellular Factors and the KIX Domain of p300/CBP, Journal of Molecular Biology, Volume 409, Issue 3, 2011, Pages 384-398, https://doi.org/10.1016/j.jmb.2011.04.003.

Futsch N, Mahieux R, Dutartre H. HTLV-1, the Other Pathogenic Yet Neglected Human Retrovirus: From Transmission to Therapeutic Treatment. Viruses. 2018; 10(1):1. https://doi.org/10.3390/v10010001

Gaudray G, Gachon F, Basbous J, Biard-Piechaczyk M, Devaux C, Mesnard J-M. The Complementary Strand of the Human T-Cell Leukemia Virus Type 1 RNA Genome Encodes a bZIP Transcription Factor That Down-Regulates Viral Transcription. Journal of Virology. 2002;76(24):12813-12822. doi:10.1128/jvi.76.24.12813-12822.2002

Glasmacher E, et al. Science. Genomic Regulatory Element That Directs Assembly and Function of Immune-Specific AP-1–IRF Complexes. 2012 Nov 16; 338(6109): 975–980

Matsuoka, M., Mesnard, JM. HTLV-1 bZIP factor: the key viral gene for pathogenesis. Retrovirology 17, 2 (2020). https://doi.org/10.1186/s12977-020-0511-0

Nakagawa M, Shaffer AL, Ceribelli M, et al. Targeting the HTLV-I-Regulated BATF3/IRF4 Transcriptional Network in Adult T Cell Leukemia/Lymphoma. Cancer Cell. 2018;34(2). doi:10.1016/j.ccell.2018.06.014